博士論文

Study on molecular genetic analysis of bovine leukemia virus in South America and Asia

（南米とアジアにおける牛白血病ウイルスの分子遺伝学的解析）

买热帕提　帕拉提　Mairepati Palati

Contents

Abbreviations

EBL : Enzootic bovine leukosis

SBL : Sporadic bovine leukosis

BLV : Bovine leukemia virus

PL : Persistent lymphocytosis

STLV : Simian T-lymphotropic virus

HTLV : Human T-cell leukemia virus

nt : Nucleotide

LTR : Long terminal repeats

CA : Capsid protein

NC : Nucleocapsid protein

MA : Matrix protein

RNA : Ribonucleic acid

ENV : Envelop protein

IN : Integrase

RT : Reverse transcriptase

SU : Surface unit

TM : Transmembrane protein

ITAM : Immunoreceptor tyrosine-based activation

TxRE : Tax-responsive element

CRE : Cycle-AMP responsive element

CREB : Cycle-AMP responsive element binding protein

REF : Rat embryo fibroblasts

Ha-ras : Hayvey rat sarcoma viral oncoprotein

mRNA : Messenger RNA

miRNA : MicroRNA

DNA : Deoxyribonucleic acid

dsDNA : Double strand DNA

cDNA : Complementary DNA

PCR : Polymerase chain reaction

AGID : Agar gel immunodiffusion test

ELISA : Enzyme linked immunosorbent assay

RIA : Radio immuno assay

PHA : Passive hemagglutination assay

CoCoMo : Coordination of common motifs

NJ : Neighbor-joining method

bp : Base pair

ND : Neutralizing domain

G- : Genotype-

NGS : Next generation sequencing

ML : Maximum-likelihood method

K2+G : Kimura-2 parameter with gamma distribution

BI : Bayesian Inference

GTL : General-time-reversible model

GRE : Glucocorticoid responsive element

DAS : Downstream of the transcription site

IRF : Interferon regulatory fector

PAS : Polyadenylation site

NLS : Nuclear localization signal

NES : Nuclear export signal

MYB : Myb-like motif

ARR : Arginine-rich region

RFLP : Restricted fragment length polymorphisms

Abstract

【Background】

Bovine leukemia virus (BLV) is an oncogenic member of *retroviridae* family belonging to the genes *deltaretrovirus*, and is the etiological agent of enzootic bovine leukosis (EBL), the most common neoplastic disease of cattle. BLV infects cattle worldwide, some of BLV-infected cattle suffer from lymphomas and/or B-lymphocyte proliferation (persistent lymphocytosis, PL), but the majority of infected cattle are healthy carriers of the virus.

BLV complete genome is constituted of 8714 nucleotides, including structural and enzymatic *gag*, *pro*, *pol*, and *env* genes which are indispensible in the synthesis of viral particles. BLV genome also contains a pX region which encodes regulatory proteins Tax and Rex, and accessory proteins R3 and G4. The structural genes and pX region are surrounded by two identical long terminal repeats (LTR). The Env gp51 glycoprotein plays an essential role in the viral life cycle, and is required for cell entry and the target of neutralizing antibodies. BLV gp51 contains both conformational and linear epitopes and is essential for viral infectivity and syncytium formation. Therefore, the gp51 region has been widely used for BLV genotyping studies and recent phylogenetic studies of this region from viral strains isolated worldwide demonstrate that BLV can be classified into at least eight genotypes.

【Objective】

Cattle in South America and Asia are occupied over than 50 % of cattle populations in the world. However, there are no or very few studies on the distribution of BLV in South America and Asia, and the genetic characteristics of BLV strains there remain to be unknown. Therefore, the aim of this study is to detect the spread of BLV infection and to investigate the molecular genetic variability of BLV strains in South America and Asia, and to confirm the existence of new genotypes-9 and -10 (G-9 and G-10).

【Materials and Methods】

Blood samples were obtained from a total of 3386 cattle which contained 2204 samples in five South American countries (328 from Peru, 139 from Paraguay, 507 from Bolivia, 420 from Argentina, and 810 from Chile) and 1116 cattle from the Philippines and the rest 66 cattle from

Myanmar. Genomic DNA was extracted from these samples. BLV infection was detected by amplification of BLV long terminal repeats (LTRs) using nested polymerase chain reaction (PCR) and/ or BLV-CoCoMo-qPCR-2, an assay developed for the highly sensitivity detection of BLV. BLV positive samples were used for amplification of BLV *env*-gp51 by nested PCR and sequenced for phylogenetic analyses. Phylogenetic trees were constructed by using the neighbor-joining (NJ) method, Bayesian Inference (BI) and/ or maximum-likelihood (ML) method. Full genome sequences of new genotypes were obtained by next generation sequencing (NGS)-based whole genome sequencing for G-9 or clone-sequencing for G-10, respectively.

【Results and discussion】

1. Detection of BLV provirus prevalence in South American and Asian cattle samples.

To investigate the spread of BLV infection, 3386 samples were screened for BLV infection. Among the 2204 cattle tested in the South America, all South American countries showed relatively high level of infection: 139 cattle out of 328 (42.3 %) were positive for the BLV provirus in Peru. In Paraguay, 76 cattle samples out of 139 (54.5 %) were BLV positive. Of 507 samples collected from Bolivia, 156 (30.7 %) were positive for BLV provirus. Argentina samples (n = 420) demonstrated extremely high levels of BLV prevalence (77.4 %). In Chile, of the 810 samples screened, 236 (29.1 % prevalence) were BLV provirus positive. In contrast to the BLV prevalence level in South America, Asia showed remarkable low level of BLV infection. In the Philippines, a total of 9.7 % of tested samples (108/1116) were determined as BLV positive, while only 5 out of 66 (7.6 %) cattle from Myanmar showed the prevalence of BLV.

2. Phylogenetic analysis of partial *env* gp51 sequences

Phylogenetic analysis based on *env* gp51 sequences of BLV positive samples from South America and Asia showed that total of 32 philippine BLV strains were assigned to G-1, while the rest 11 strains were grouped into G-6. In South America, Peru BLV strains were assigned to G-1, -2, and -6. Likewise, the majority of Paraguayan BLV strains were clustered into G-1 and -6, with a small number in G-2. Interestingly, the BLV strains collected from Bolivia clustered not only into G-1, -2, and -6, together with Peruvian and Paraguayan strains, but also into a unique clade, which was distinct from the eight previously known BLV genotypes as a novel G-9.

Interestingly, Myanmar strains were not clustered within any known genotypes, but separately located in a different branch indicating that Myanmar strains is a new genotype, termed as G-10.

3. Amino acid substitutions of BLV *env* gp51 partial sequences from strains isolated in South America and Asia

Alignment of predicted amino acid partial sequences of *env-gp51* of each genotype BLV strains in studied area revealed that BLV strains observed were highly conserved. Philippine strains assigned to G-6 showed G-6-specific common amino acid substitution of isoleucine by threonine at residue 144 of second neutralizing domain (2^{nd} ND). In addition, six single amino acid substitutions, such as substitutions Y108C at $CD4^+$ epitope, A119P at 2^{nd} ND and L202F in G-1 Philippine BLV strains, amino acid mutations K175E at $CD8^+$ epitope, T231S and S234N in the B-epitope in G-6 Philippine BLV strains, were also observed, respectively. South American BLV strains were strictly conserved. G-2 and -6 South American BLV strains all shared G-2 and -6 specific amino acid substitution of N141D and I144T at 2^{nd} ND. All South American new G-9 BLV strains showed common unique amino acid substitution of alanine by valine at residue 133 (A133V) in 2^{nd} ND. However, Myanmar BLV strains showed a variety of unique amino acid substitutions. One G-6-specific amino acid substitution from isoleucine to threonine at residue 144 (I144T), which is common to all G-6 BLV strains worldwide, is also detected in 2^{nd} ND of Myanmar BLV strains. Besides, four unique amino acid mutations, substitution of valine replaced by alanine at residue 106 (V104A) in 1^{st} ND and $CD4^+$-epitope, substitutions of serine to phenylalaine at residue 137 (S137F) and of glutamine to arginine in residue 143 (Q143R) at 2^{nd} ND, and substitution from proline to serine at residue 177 (P177S) in $CD8^+$- and E-epitope, were observed only in Myanmar strains.

4. BLV complete genome sequencing and phylogenetic analysis

To confirm the new G-9 and -10, ML phylogenetic tree were constructed based on complete-genome sequences of BLV strains. The phylogenetic tree clearly demonstrates stratification of BLV genotypes, including G-1, -2, -4, and -6, and the novel G-9, into separate clades (Bootstrap values 100 % for every clade). Thus, this analysis provides evidence for the existence of a novel G-9. Interestingly, complete genome sequences of two Myanmar strains were separately located in a different branch (Bootstrap values 100 % for every clade), indicating that Myanmar strain is

also a new genotype, G-10. The ML tree further confirms the existence of new G-9 and -10 in the current study.

5. Comparative analysis of amino acid sequences of each BLV gene among new genotypes and other known strains

To further determine how much new BLV genotype full genome sequences were distinguished from previously known BLV strains in database, G-9 and G-10 BLV strains were aligned with other known complete genome sequences. Ten unique amino acid substitutions in South American G-9 and 22 unique amino acid substitutions in Myanmar G-10 BLV strains were observed in full BLV genome sequences. G-9-specific amino acid substitutions were described as follows: (1) Two substitutions, E166D and D447G, in the Pol (RT) region; two substitutions, H644Y and A826T, in Pol (IN) region; and one substitution at residue 792 of the Pol (IN) region, restricted to only sequences of all 12 samples collected from Portachuelo, but not in other G-9 strains. (2) In the Env (gp51) protein, one significant substitution, A133V, was observed only in all BLV G-9 strains. (3) In the regulatory proteins Tax and Rex, substitutions at residue 108 (F108L) of Tax and residue 113 (A113E) of Rex were detected only in G-9 strains. In addition, a substitution at residue 100 (P100S) of the Tax protein was observed only in samples collected from Portachuelo. (4) Likewise, a substitution at residue 26 (N26H) in the R3 accessory protein was detected only in the sequences of all G-9 BLV strains, but not in other strains or known BLV genome sequences. 22 unique amino acid substitutions of G-10 Myanmar BLV strains are as follows: (1) In the structural gene-encoded proteins: two substitutions, T38A and T366A, in the Gag; one substitution, S52F, in the Pro; four substitutions, V205L, I409V, P480S and A826V, in the Pol; and four substitutions, V106A, S137F, Q143R and P177S, in the Env, restricted to only sequences of Myanmar strains. (2) In the regulatory proteins: six substitutions, N140K, V142E, I152T, D181N, E229D and L273F, in the Tax; three substitutions, S103F, L140P and T156N, in the Rex; one substitution of K27N in R3; one substitution, L66P, in the G4 protein, were observed only in Myanmar strains. These substitutions were firstly detected in the sequences of the new genotypes strains identified in this study. Comparative analysis of amino acid sequence of full genome showed that G-9 and -10 strains were significantly different from other genotypes, further supporting the result of full genome ML tree that G-9 and -10 are novel genotypes.

【Conclusion】

The present study provides the first evidence of the prevalence of BLV infection among cattle in Paraguay, Peru and Bolivia and Myanmar, and also confirms the widespread distribution of BLV infection in Argentina, Chile and in the Philippines. All of five South American countries showed higher level of BLV prevalence, and also indicated tendency of dramatic increase of BLV infection in South America. By contrast, low to medium level of BLV prevalence were observed in the Philippines and Myanmar in Asia. The phylogenetic analysis demonstrated that two different BLV G-1 and -6 are present in the Philippines. BLV genotyping analysis of South America showed that Peruvian and Paraguayan BLV strains are of G-1, -2, and -6. Of interest, this study findings indicate that Bolivian BLV strains are clustered into G-1, -2, -6, and a new G-9. In addition, Myanmar BLV strains were classified into another novel G-10. Amino acid comparison studies showed that ten unique amino acid substitutions in South American G-9 and 23 unique amino acid substitutions in Myanmar G-10 BLV strains were observed. These unique amino acid substitutions clearly distinguish novel G-9 and -10 from previously known genotypes, further indicating the existence of new genotypes. Thus, the detection of G-9 and -10 is a novel finding of this study. These results provide important information on BLV infection levels and will enable the implementation of appropriate cattle-management policies in addition to providing supplementary information for the development of more-effective methods of BLV eradication in South America and Asia.

General introduction

In 1871, Leisuring found yellowish nodules in the enlarged spleen of a cow which recorded as the first case of leukosis in cattle ((Leisering, 1871), cited by (Gillet *et al.*)). The most remarkable clinical sign of bovine leukemia is spleen disruption consecutive to tumour formation (Gillet *et al.*, 2007). Bovine leukemia as a lymphoproliferative disease in cattle herds is classified into two types, namely as enzootic bovine leukosis (EBL) and sporadic bovine leukosis (SBL), based on their epidemiological characters (Gillet *et al.*, 2007). The etiological cause of SBL is unknown, and most of the cases, SBL occurs in young cattle under one year of age. SBL is non-contagious and non-transmissible, and is normally seen as a single case in a herd (Bundza *et al.*, 1980). On the other hand, EBL is a lymphoproliferative disease of cattle distributed worldwide (Kirkland & Rodwell, 2005) (Figure 1), which is characterized by B-cell leukosis (Gillet *et al.*, 2007). In the case of EBL, cattle can be infected at any age including the embryonic stage (Kirkland & Rodwell, 2005) and has relatively long incubation period (Bundza *et al.*, 1980). The causative agent of the EBL is a retrovirus called bovine leukemia virus (BLV), which is isolated in culture in 1969 ((Miller *et al.*, 1969), cited by (Rodriguez *et al.*, 2011)).

BLV is a *deltaretrovirus* of *retroviridae* family that infects cattle worldwide (Figure 1) and induce the accumulation of B-lymphocytes in peripheral blood and lymphoid tissues of cattle, resulting in leukemia/lymphoma (Burny *et al.*, 1985; Florins *et al.*, 2008; Gillet *et al.*, 2007). The majority of BLV infected cattle neither show any clinical signs nor any change in lymphocyte count, known as asymptomatic carriers of the virus (Figure 2). Approximately, one-third of cattle infected with BLV develop a benign form of non-malignant proliferation of untransformed B-lymphocytes called persistent lymphocytosis (PL) (Figure 2). The typical character of PL is permanent and stable increase of B-lymphocyte number circulating in peripheral blood. By contrast, after 1-8 years of latency, 1-5 % of infected animals develop tumours, characterized as malignant $CD5^+$ B-cell lymphoma originated from mono- or oligo- clonal accumulation of B-cell (Figure 2). This malignant form of B cell lymphoma predominantly detected in cattle older than 4-5 years old (Aida *et al.*, 2013; Gillet *et al.*, 2007; Rodriguez *et al.*, 2011). Except from spleen disruption, BLV can also negatively effect on other tissues and organs: lymph nodes may be remarkably enlarged and easily observed under skin. BLV-induced neoplastic cells can also penetrate into the abomasums, right auricle of the heart, intestine, kidney, lung, liver, and uterus.

The clinical marks of BLV-induced tumours are different, and mostly involved in digestive disturbance, weight loss, weakness, reduction of milk production, loss of appetite, enlarged lymph nodes (OIE, 2012).

The *deltaretrovirus* genus also includes Simian T-lymphotropic virus 1, 2, 3 and 5 (STLV-1, -2, -3 and -5) and human T-lymphotropic virus 1, 2, 3 and 4 (HTLV-1, -2, -3 and -4) (Willems *et al.*, 2000). Even though BLV and HTLV infects B-lymphocytes and T-lymphocytes respectively, BLV is considered as a model of pathogenesis for HTLV-1 and -2 because of harboring a similar genomic organization, and sharing similar replication and gene expression strategies, (Willems *et al.*, 2000).

The full genome of BLV is comprised of 8714 nucleotides (nt) (Sagata *et al.*, 1985) which includes structural and enzymatic genes, namely *gag*, *pro*, *pol*, and *env*, and is surrounded by two identical long terminal repeats (LTRs) as shown in Figure 3a. The structural genes are essential and indispensable in the production of infectious virions (Figure 3b). BLV *gag* gene encodes precursor $Pr44^{gag}$ which further processed into mature proteins: p24 capsid protein (CA), p12 nucleocapsid protein (NC), and the p15 matrix protein (MA) (Hamard-Peron & Muriaux, 2011; Sagata *et al.*, 1984). Marix p15 ties to viral ribonuclei acid (RNA) and is associated with the lipid bilayer of viral membrane (Gillet *et al.*, 2007). Retrovirus RNA packaging involves recognition of genome-length viral RNA by viral Gag polyprotein. BLV has two encapsidation signal region: the primary region is constituted of two stable RNA stem-loop structures, located downstream of the gag start codon in MA domain, and these structures are required for RNA packaging and virus production. Another encapsidation signal region for BLV was constituted of one stable stem-loop structure located in the CA domain of Gag (Jewell & Mansky, 2000). BLV proteases (Pro) p14 and p80 are encoded by *pro* and *pol* genes, harboring integrase (IN) and reverse transcriptase (RT) activities (Gillet *et al.*, 2007; Sagata *et al.*, 1984), providing essential enzymatic functions. The *env* gene encodes precursor $Pr72^{env}$, which further cleaved into a mature surface Unit (SU) gp51 and a transmembrane protein (TM) gp30 (Sagata *et al.*, 1984), and is involved in viral infectivity (Callebaut *et al.*, 1993; Inabe *et al.*, 1999; Johnston & Radke, 2000), and plays important biological role in viral replication. SU gp51 and TM gp30 link through disulfide bonds. Gp51 might contain cell receptor-binding domain. The extracellular SU is very immunogenic (Gillet *et al.*, 2007) and cause massive expression of antibodies in BLV-infected cattle. These antibodies are cytolytic and neutralize virus infectivity and syncytium

formation ((Kettmann *et al.*, 1994) cited by (Willems *et al.*, 1995)). TM gp30 protein plays very essential role in cell to cell viral transmission. TM gp30 leads to destabilize cell membrane by oblique insertion of fusion peptides into the membrane (Voneche *et al.*, 1992). Gp30 protein also induces signal transduction through immunoreceptor tyrosine-based activation (ITAM) motifs located in cytoplasmic domain (Inabe *et al.*, 1999).

In addition to the above, BLV genome contains a pX region, located between the *env* sequence and the 3′ LTR (Aida *et al.*, 2013; Gillet *et al.*, 2007) (Figure 3a). This pX region encodes the regulatory proteins Tax and Rex, and the accessory proteins R3 and G4. The Tax protein has been extensively studied and is believed to play a key role in BLV-induced leukemogenesis (Willems *et al.*, 1990). Tax proteins activates viral transcription through acting on triplicate Tax responsive element (TxRE) in the Unique 3 (U3) region of the 5`LTR, which binds to cellular transcription factors such as cycle-AMP responsive element binding protein (CREB) (Tajima & Aida, 2000b; Willems *et al.*, 1992). Tax proteins also induces the immortalization of primary rat embryo fibroblasts (REF) (Willems *et al.*, 1990), and this function is destroyed by simultaneous mutations of both phosphorylation sites (Willems *et al.*, 1998). In addition, Tax leads to the full transformation of REF and form tumor in nude mice through cooperating with Harvey rat sarcoma viral oncoprotein (Ha-ras) (Willems *et al.*, 1990). Rex is indispensable for nuclear export of viral RNA and accelerates cytoplasmic accumulation and translation of viral messenger RNA (mRNA) in BLV-infected cells (Felber *et al.*, 1989). The R3 and G4 proteins contribute to the maintenance of high viral load (Florins *et al.*, 2007; Willems *et al.*, 1994). The G4 protein is specifically relevant to leukemogenesis because of immortalizing primary embryonic fibroblasts (Florins *et al.*, 2007). The R3 protein is located in the nucleus and cellular membranes (Florins *et al.*, 2007), and contributes to the maintenance of infectivity (Willems *et al.*, 1994). Besides the above, BLV RNA polymerase III (pol III)-encoded viral microRNAs (miRNA) are strongly expressed in preleukemic and malignant cells, in which structural and regulatory gene expression is repressed, indicating a potential role in tumor onset and progression (Kincaid & Sullivan, 2012; Rosewick *et al.*, 2013). In addition, Gillet et al., (2016) found that the BLV-miRNAs had significant effects on proviral load and consequently viral replication in the natural host (Gillet *et al.*, 2016).

The lifecycle of BLV is as following (Figure 4): firstly, viral infection starts with the interaction between the viral envelope proteins and the host cell surface receptors which

consecutively leads to fusion of the viral and cellular membranes. Then, the virus enters the host cell and releases viral RNA into the cytoplasm where viral deoxyribonucleic acid (DNA) is produced by reverse transcriptase. Following, the viral double strand DNA (dsDNA) is translocated into the host cell nucleus where it integrates into the host genome of cells as provirus. Proviral DNA is transcripted by host cellular machinery and messenger RNA (mRNA) molecules were transferred into the cytoplasm, and are translated into viral proteins. After assembly of viral proteins and RNA, viral particles bud from the plasma membrane. It was observed through using BLV genome-specific probes that BLV-induced tumors and BLV-infected cells contain provirus, and approximately four copies of proviral DNA can be found in each tumor (Burny *et al.*, 1988).

A number of transmission pathways and associated risk factors of BLV have been identified, and include both horizontal modes and vertical routes (Gillet *et al.*, 2007; Johnson *et al.*, 1985; Rodriguez *et al.*, 2011). The most efficient transmission and major infection of BLV happens through horizontal factors through the transfer of infected cells in blood or milk via direct contact (Ferrer & Piper, 1978; Gillet *et al.*, 2007; Rodriguez *et al.*, 2011). In addition, physical contact between healthy and infected animals within a herd had been considered a potential factor of BLV transmission (Kono *et al.*, 1983). Iatrogenic and cattle management procedures which involve exposure to biological fluid contaminated with infected lymphocytes, blood contaminated dehorning devices (DiGiacomo *et al.*, 1985; Lassauzet *et al.*, 1990), introduction of infected animals (Kobayashi *et al.*, 2014), common use of needles (Hopkins *et al.*, 1988; Lassauzet *et al.*, 1990) and plastic sleeves for rectal palpation (Van der Maaten *et al.*, 1981), and BLV infected cells in blood through blood sucking insects such as biting flies (Bech-Nielsen *et al.*, 1978; Ferrer & Piper, 1981; Ohshima *et al.*, 1981), vertical transmission of BLV might occur via milk ((Kenyon et al. 1982) cited by (Johnson *et al.*, 1985)), and dam to calf contact (Ferrer & Piper, 1978) and pre-naturally through utero infection of fetus (Van der Maaten *et al.*, 1981).

The natural hosts of BLV are domestic cattle (*Bos Taurus* and *Bos indicus*) and water buffaloes. Although several epidemiological studies have shown that BLV unlikely infects and replicates in humans (Perzova *et al.*, 2000), a recent study showed a significant association of the presence of amplified BLV DNA with breast cancer (Buehring *et al.*, 2015), suggesting a risk for the acquisition and proliferation of this virus in humans. Sheep are very susceptible to experimentally infection and lead to development of lymphoma more often and at a younger age

than cattle (Aida *et al.*, 2013; Gillet *et al.*, 2007) (OIE, 2012). Besides the above, BLV can also be experimentally transmitted to many species including chicken (Altanerova *et al.*, 1990), rabbits (Onuma *et al.*, 1990; Wyatt *et al.*, 1989), pigs (Olson *et al.*, 1981) and goats (Djilali *et al.*, 1987), rats (Altanerova *et al.*, 1989), deer, cats, dogs, rhesus monkeys, chimpanzees, antelopes (OIE, 2012), and successfully infects a serious of cell lines *in vitro* (Graves & Ferrer, 1976; Inabe *et al.*, 1998). BLV, however, induce leukemia only in cattle and sheep, and BLV leads to $CD5^+$ IgM^+ B-lymphocyte transformation in cattle (Aida *et al.*, 1993) while $CD5^-$ IgM^+ B-lymphocyte in sheep (Murakami *et al.*, 1994a; Murakami *et al.*, 1994b), indicating the different mechanisms of BLV-induced leukemia in cattle and sheep ((Djilali & Parodi, 1989; Graves & Ferrer, 1976) cited by (Aida *et al.*, 2013)). Another difference of BLV-induced leukemia/leukosis between cattle and sheep is that 20-100 % of B-lymphocytes are carrying provirus in sheep while BLV-infected cows have quite less number of infected cells (Florins *et al.*, 2008).

A number of BLV diagnostic methods have been developed and applied in the detection of BLV worldwide, including serological testes and polymerase chain reaction (PCR) based proviral genome detection methods as described in Table 1. Widely used serological diagnosis methods of BLV infection are agar gel immunodiffusion (AGID) test (Aida *et al.*, 1989; Kurdi *et al.*, 1999; Monti *et al.*, 2005b; Wang, 1991), enzyme-linked immunosorbent assay (ELISA) (Aida *et al.*, 1989; Kurdi *et al.*, 1999; Monti *et al.*, 2005b; Wang, 1991), radio immuno assay (RIA) (Levy *et al.*, 1977), passive hemagglutination Assay (PHA) (Fukai *et al.*, 1999). Both AGID and ELISA tests are commonly used serological tests and target to detect the antibodies in the screened samples. AGID test is economically less expensive and many samples can be screened at the same time, but not sufficiently sensitive (Naif *et al.*, 1990). The ELISA test is most commonly used in Animal Health Authorities for freedom of BLV, and very sensitive and easy to procedure, but ELISA fails to detect antibodies in serum of cattle at the early phage of infection, producing false negative result (Naif *et al.*, 1990). Besides, ELISA can also produce false positive result in the case of screening calves that have anti-BLV antibodies obtained from cow colostrums or parturition (Burridge *et al.*, 1982). Since BLV randomly integrate into host genome as provirus and remain in the cellular genome throughout the disease even if the absence of detectable antibodies, PCR-based method, which targets provirus, is also commonly used in BLV diagnosis analysis nowadays. A variety of PCR methods have been widely used, including

BLV-CoCoMo-qPCR (Jimba *et al.*, 2010; Jimba *et al.*, 2012), BLV-CoCoMo-qPCR-2 (Takeshima *et al.*, 2015), real-time PCR (Brym *et al.*, 2013; Tawfeeq *et al.*, 2013). In addition, some other tests, namely expression of viral proteins by western blot (Inabe *et al.*, 1998; Tajima & Aida, 2000b; Tajima *et al.*, 2003), syncytium formation assay (Inabe *et al.*, 1998), and the presence of BLV particles by electron microscope, are also used.

BLV might be originated from Europe, more specifically from the region located in the contemporary Lithuania, during the 19th century. The widespread of the BLV prevalence happened through the introduction of BLV-infected cattle from Europe to American continent during the first half of the 20th century. Then, BLV from the American continent have disseminated back into Europe, and BLV introduced into other countries free of disease by cattle import from North America ((Johnson & Kaneene, 1992) cited by (Rodriguez *et al.*, 2011)). Currently, Western Europe has successfully eradicated BLV infection. However, BLV prevalence is still widespread in other parts of the world, including countries of Eastern Europe (Acaite *et al.*, 2007), United States of America, South America, Asia and Middle East countries (Rodriguez *et al.*, 2011). The BLV infection levels varies from countries to counties, but a tendency of high BLV prevalence of dairy herd than that of beef were detected in United states of America (USDA, 2008) and Japan (Murakami *et al.*, 2011). Currently, BLV strains were classified into eight genotypes based on genetic characteristics of *env-gp51* region. Even though the genetic clustering of BLV are not always follow the country origin of the strain, the grouping of counties in each assigned genotype is associated with geographic affinity (Balic *et al.*, 2012; Lee *et al.*, 2015; Rola-Luszczak *et al.*, 2013) as shown in the Table 2.

In order to effectively eradicate BLV infection in cattle, it is necessary to determine infection levels. Estimating BLV prevalence will help to establish appropriate eradication programs that will hopefully prevent further spread of disease. Studies on the genetic structure of a virus are of fundamental biological and epidemiological interest, which help us to get better insight into endemic patterns and disease progression. The genetic structure of pathogen populations is important in many contexts, such as development of vaccines, resistance and treatments.

In this study, firstly, I examined the distribution of BLV in South America, including Peru, Paraguay, Bolivia, Argentina and Chile, and Asia, specifically the Philippines and Myanmar; and next, I evaluated the diversity of BLV strains, and investigated phylogenetic relationship of BLV strains circulating in the studied area; finally, I used complete genome sequences of BLV strains

to confirm the existence of Novel genotypes, genotype-9 and -10, circulating in South America and Asia, respectively.

CHAPTER I.

Detection and molecular characterization of bovine leukemia virus in Philippine cattle

1-1. Introduction

BLV is a RNA virus and recognised as the etiological agent of EBL, one of the neoplastic desease of cattle (Aida *et al.*, 2013). In *deltaretrovirus genes*, BLV is closely related to HTLV-1 and -2 (Aida *et al.*, 2013). Serological studies of BLV revealed that BLV infection is widely disseminated throughout the world (Rodriguez SM et al., 2011). The natural hosts of BLV are domestic cattle and water buffaloes (Aida *et al.*, 2013). Unfortunately, recent studies revealed the detection of BLV provirus in some human tissues, namely human breast tissue, indicating a potential risk for prevalence and proliferation of BLV in humans (Buehring *et al.*, 2014). Most of BLV infected cattle remain healthy carriers of the virus as asymtomatic, and one third of cattle with BLV become PL (Aida *et al.*, 2013; Mirsky *et al.*, 1996), and after up to 8 to 10 years of latency, less than 5 % of BLV infected animals develop malignant $CD5^+$ B-cell lymphoma (Aida *et al.*, 1993; Mirsky *et al.*, 1996).

After BLV infection, viral genome integrates into the host genome as a provirus (Kettmann *et al.*, 1979). It is known that BLV-infected cattle retain at least one copy of the full-length proviral genome throughout the course of the disease (Tajima *et al.*, 1998). Moreover, large or small deletions of the proviral genome are thought to be very rare events in BLV-infected cattle (Tajima *et al.*, 1998).

The Philippines is primarily an agricultural country. In this country, livestock is highly associated with the economic structure of farm and village life (Castillo). In the Philippines, cattle are one of the important large ruminants mainly used for livestock meat and milk production for home consumption and additional source of income, and draft power in agriculture (Bondoc, 2013; Castillo; Stanton, 2010). Livestock rising remains a backyard enterprise, and about 85 % of cattle in the Philippines are raised in small-holder (backyard) farms, while a very little portion is commercial farms. Philippine native cattle, also called Philippine cattle, are descendants of Chinese yellow cattle and Iberian cattle from Mexico and Spain brought into the Philippines by Chinese and Spaniard. Since many Chinese native cattle breeds are still found in different part of the Philippines, Philippine native cattle are believed to be a mixture of European and Indian cattle (Bondoc, 2013; Watanabe *et al.*, 1989)**.** Until now, to improve the native breeds and to increase dairy and beef production of cattle, a variety of breeds, including at least 13 diary or dual purpose breeds, 27 beef breeds, and 3 draft breeds, were introduced into the Philippines from America, Europe, and Asia (Bondoc, 2013).

BLV provirus remains integrated in cellular genomes (Jimba *et al.*, 2012; Tajima *et al.*, 1998), even in the absence of detectable BLV antibodies. Since routine diagnosis of BLV infection through serological testes, such as AGID and IELSA which shows false negative result in the case of low viral titer or in the early phage of infection when antibodies cannot be detected, diagnostic BLV PCR techniques that detect the integrated BLV proviral genome within the host genome are also commonly used (Tajima *et al.*, 1998). In addition to the above, our lab recently developed new quantitative real-time PCR method using coordination of common motifs (CoCoMo) primers to measure the proviral load of both known and novel BLV variants in BLV-infected animals from a number of international locations (Jimba *et al.*, 2010; Jimba *et al.*, 2012; Panei *et al.*, 2013). And Takeshima et al,. (2015) improved the original BLV-CoCoMo-qPCR by optimizing the primer degeneracy and PCR conditions and reconstructing the standard plasmid to BLV-CoCoMo-qPCR-2 which is highly effective in detecting BLV in cattle (Takeshima *et al.*, 2015).

The glycoprotein encoded by *env* gene is indispensible in viral life cycle and required for viral entry into the cell, and is subjected to neutralizing antibodies (Johnston & Radke, 2000; Mamoun *et al.*, 1990). The Env glycoprotein is also thought to harbor potential receptor binding domain involved in virus-host interactions (Ban *et al.*, 1992; Johnston *et al.*, 2002). The Env surface Unit (SU) gp51 plays key role in viral infectivity and syncytium formation (Callebaut *et al.*, 1993; Portetelle *et al.*, 1989) and contains the conformational epitopes F, G and H on its N-terminal region (Bruck *et al.*, 1982a; Mamoun *et al.*, 1990), whilst the C-terminal half of BLV gp51 contains the linear epitopes A, B, D and E (Mamoun *et al.*, 1990). Phylogenetic analysis based on sequences of BLV *env gp51* gene from different geographic regions demostrated the presence of different genetic groups that correlate with the geographic origin of the isolates (Balic *et al.*, 2012; Beier *et al.*, 2001; Camargos *et al.*, 2007; Camargos *et al.*, 2002; Felmer *et al.*, 2005; Molteni *et al.*, 1996; Monti *et al.*, 2005a; Moratorio *et al.*, 2010; Rodriguez *et al.*, 2009). Recent studies of the BLV *env* gene of strains isolated worldwide revealed that the virus can be assigned into eight different genotypes (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Rola-Luszczak *et al.*, 2013). Very few studies on the distribution of BLV prevalence in the Philippine are available. However, research focused on the genetic variability of BLV in the Philippines remains to be held. In this study, I observed the spread of BLV infection in the Philippines by amplification of BLV LTRs using a combination of nested PCR and BLV-CoCoMo-qPCR-2, and investigated the

genetic variability of Philippine BLV strains by DNA sequencing and phylogenetic analyses based on a 423 base pair (bp) sequence of the BLV *env* gene. This is the first study to identity the BLV genotypes in the Philippines.

1-2. Materials and Methods

1-2-1 Experimental samples

1116 cattle blood samples were taken from a number of different farms of 5 Philippine islands, including 72 cattle samples from Bohol, 102 samples from Cebu, 372 cattle obtained from Leyte, 159 cattle from Iloilo and the rest 411 samples from Luzon island as shown in the Figure 5. 40 µl of blood sample from each cattle were spotted on Whatman FTA elute cards (GE Healthcare Japan Corp., Tokyo, Japan), and stored as experimental material in humidity-control container at room temperature for further usage.

1-2-2 Animal handling and research Ethics

Veterinarians from RIKEN and the Philippine Carabao Center handled all the experimental animals in strict accordance with good animal practice following the Philippine Carabao Center institutional guidelines. Current study was approved by the Institutional Animal Care and Use Committee (IACUC) at the Philippine Carabao Center on the Ethics of Animals for Research.

1-2-3 Genomic DNA extraction methods

DNA was extracted from 40 µl of whole blood spotted onto Whatman FTA elute cards (GE Healthcare Japan Corp., Tokyo, Japan) based on the manufacturer's instructions. Briefly, 3 mm sample disk from the center of the blood spotted on FTA elude cards is removed and transferred into 1.5 ml microfuge tube which has 500 µl of sterile H_2O. Then, the sample disc is washed through vortex the tube. After that, the sample disc is transferred from 1.5 ml microfuge tube to 0.5 ml microfuge tube containing 30 µl of sterile H_2O. Then, the 0.5 ml microfuge tube containing sample disc is incubated at 95 °C for 25 min in a heating block (PCR machine is used). After incubation, sample is removed from the heating block and vortex 1 min continuously. The sample disc is separated through brief centrifuge and is discarded. The remaining eluate in the 0.5 ml tube contains purified genomic DNA. The eluted DNA is stored at -20 °C until required for further experiment.

1-2-4 Detection of BLV provirus

1-2-4-1 Detection of BLV provirus by CoCoMo-qPCR-2

CoCoMo-qPCR is a real-time PCR assay which quantitatively measures proviral load of a variety of BLV strains in BLV-infected cattle from different locations by taking advantage of CoCoMo primers (Jimba *et al.*, 2010; Jimba *et al.*, 2012; Panei *et al.*, 2013). BLV-CoCoMo-qPCR-2 is upgraded version of CoCoMo-qPCR, and the detection of BLV provirus by BLV-CoCoMo-qPCR-2 is described previously (Takeshima *et al.*, 2015). Briefly, 120-bp fragment of provirus LTR gene was amplified by using degenerate primer pairs of forward primer mixture (CoCoMo-FRW1: 5`-AATCCMNMYCYKDAGCTGCTGAYYTCACCT-3`; CoCoMo-FRW2: 5`-ATCCACACCCTGAGCTGCTGCACCTCACCT-3`) and reverse primer (CoCoMo-REV: 5`-TTGCCTTACCTGMCSSCTKSCGGATAGCCGA-3`). One single copy of host gene bovine leukocyte antigen (BoLA)-DRA was also amplified by the primer pairs of DRA-Forward (5`-CCCAGAGTATGAAGCTCCAGCCC-3`) and DRA-Reverse primer (5`-CCCTCGGCGTTCAACGGTGT-3`) to be used as an inner control for balancing the amount of DNA in the reaction. FAM probes (FAM-LTR probe: 5`-FAM-CTCAGCTCTCGGTCC-NFQ-MGB-3` for viral gene LTR; FAM-DRA probe: 5`-FAM-TGTGTGCCCTGGGC-NFQ-MGB-3` for host gene BoLA-DRA) were used to detect both the viral and host genes (Takeshima *et al.*, 2015).

1-2-4-2 Detection of BLV provirus by Nested PCR

Each tested sample in this study were screened for the existence of BLV provirus by nested PCR targeting the BLV LTR gene using specific primers designed in this study. Briefly, the first PCR reaction was performed with outer primer pairs: BLTR256F (5`-GAGCTCTCTTGCTCCCGAGAC-3`) and BLTR453R (5`-GAAACAAACGCGGGTGCAAGCCAG-3`). The BoLA-DRA gene as an internal control was amplified using primer pairs of BDRA488F (5`-ACAACACCCCAAACACCAAT-3`) and BDRA1145R (5`-AGGAAGGGGAGGTAGTGGAA-3`). Five picomoles of each primer, 2 µl of 10xrTaq PCR Buffer (Toyobo, Osaka, Japan), 2 µl of 25 mM MgCl2, 2 µl of 2 mM dNTP mix, 0.1 µl of 5U/µl rTaq, and 10.4 µl of nuclease-free water were added to each sample, which was amplified in a final volume of 20 µl for 45 cycles of 94 °C for 30 sec, 58 °C for 30 sec, and

72 °C for 30 sec, and then a final extension step of 4 min. The PCR products from the above reaction was purified by applying 2 µl of Exo-SAP IT (USB Corp., Cleveland, OH, USA) through incubating at 37 °C for 15 min, then at 80 °C for 15 min. The initial PCR amplicons were subsequently applied to the second PCR with primer pairs of BLTR306F (5`-GTAAGGCAAACCACGGTTT-3`) and BLTR408R (5`-AGGAGGCAAAGGAGAGAGT-3`). The nested PCR reaction contains 1 µl of each purified PCR product from the first PCR, and five picomoles of each primer, 2 µl of 10xrTaq PCR Buffer, 2 µl of 25 mM $MgCl_2$, 2 µl of 2 mM dNTP mix, 0.1 µl of 5U/µl rTaq, and 11.9 µl of nuclease-free water. Sterilized water was used as a PCR negative control. The nested PCR amplicons were amplified under the following thermal conditions of 35 cycles of 94 °C for 30 sec, 63 °C for 30 sec, and 72 °C for 30 sec, and then a final extension step of 4 min.

1-2-5 PCR amplification and sequencing of BLV *env-gp51* gene fragments

Samples detected as BLV positive were further used for amplification of BLV *env* partial gene by nested PCR targeting gp51 partial region. PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Otsu, Japan) were applied for the PCR amplification of gp51 by using primers described previously (Asfaw *et al.*, 2005; Moratorio *et al.*, 2010). Briefly, the first PCR amplification was performed using the primer pairs of external forward (5`-ATGCCYAAAGAACGACGG-3`) and external reverse (5`-CGACGGGACTAGGTCTGACCC-3`) primers (Moratorio *et al.*, 2010). The first PCR reaction contains 5 µl of 5 x PrimeSTAR GXL Buffer, 2 µl of 2.5 mM dNTP mixture, 0.5 µl of PrimeSTAR GXL polymerase and 13.5 µl of distilled water. Nested PCR was amplified using 1 µl of first PCR amplicons by primer pairs of Env5032 (5`-TCTGTGCCAAGTCTCCCAGATA-3`) and Env5608r (5`-AACAACAACCTCTGGGAAGGGT-3`) (Asfaw *et al.*, 2005). Both the first and nested PCR reaction were running under the thermal condition of 30 cycles of 98 °C for 15 sec, 60 °C for 20 sec and 68 °C for 60 sec. The external primers produce amplification of a 913-bp DNA fragment, and internal primers result in a 597-bp fragment of the gp51 partial region of the *env* gene. Successful amplified second-round nested PCR products purified by applying 2 µl of Exo-SAP IT (USB Corp., Cleveland, OH, USA) and sequenced using an ABI PRISM Big Dye Terminator v 3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on an ABI3730xl DNA Analyzer. Sanger sequencing produces a 423-bp partial sequence of the *env*

gene, corresponding to nucleotide positions 5126 to 5548 of the BLV cell line FLK-BLV subclone pBLV913 complete genome (GenBank accession number EF600696) (Rovnak *et al.*, 1993). Editing, alignment, identification of nucleotide sequences, and amino acid prediction were performed using MEGA 5.1 software (Tamura *et al.*, 2011).

1-2-6 Construction of a phylogenetic tree

Partial *env* sequences of 43 philippine BLV strains were aligned with 72 references *env-gp51* partial sequences of eight known BLV genotypes from GenBank by Align the Muscle tools of MEGA 5.1 software (Tamura *et al.*, 2011). Phylogenetic trees were constructed using the neighbor-joining algorithm (NJ) (Saitou & Nei, 1987) with the Tamura-Nei model of nucleotide substitution (Tamura & Nei, 1993) based on alignments of the 423-bp partial *env* sequence.

1-3. Results

1-3-1 Detection of BLV infection by BLV-CoCoMo-qPCR-2 and nested PCR

A total of 1116 samples collected from different Philippine Islands were tested for BLV infection. Among all tested samples, 72 samples from Bohol Island (Table 3) and 102 samples from Cebu Island (Table 4) were determined as BLV negative by BLV-CoCoMo-qPCR-2 and nested PCR. The cattle breeds tested in Bohol Island includes Philippine native cattle, Philippine native x Brahman cattle and Philippine native x Holstein cattle as shown in Table 3. In Cebu Island, Cattle were classified as Holstein x Jersey cattle, Holstein x Brahman x Sahiwal cattle, and Holstein x Sahiwal cattle (Table 4). Four cattle out of 372 collected from Leyte Island were determined as positive for BLV provirus by BLV-CoCoMo-qPCR-2, but showed BLV negative by nested PCR (Table 5). And the four BLV positive cattle in Leyte Island were Philippine native cattle. Among the screened 159 sample obtained from Iloilo Island, three samples were BLV provirus positive by both BLV-CoCoMo-qPCR-2 and nested PCR, while other six samples were detected as BLV positive by BLV-CoCoMo-qPCR-2, but negative by nested PCR as shown in Table 6. Eight of these nine BLV positive cattle were Philippine native cattle breed. In contrast to the above mentioned four Islands located in the South Philippines, of 411 samples obtained from northern Philippine Island, Luzon Island, 95 samples (23.1%) by BLV-CoCoMo-qPCR-2 and 51 (12.4%) by nested PCR were determined as BLV provirus positive, respectively (Table 7). Luzon Island showed the highest frequency of BLV prevalence among all sampling Islands. To my interest, most of the BLV infected cattle from Luzon Island were Holstein x Sahiwal (64.2%) and Brahman cattle (29.5%), but not Philippine native cattle (0.1%). In addition, cattle from beef farms raising Brahman breeds were the next most frequently infected on Luzon Island. To sum up, 54 cattle samples out of the 1116 screened samples were determined as positive for BLV detection by both BLV-CoCoMo-qPCR-2 and nested PCR. Current study results indicate that the prevalence of BLV in Philippine cattle differ between two populations: Southern (Bohol, Cebu, Iloilo, and Leyte Islands) and Northern (Luzon Island) Philippines.

1-3-2 Phylogenetic analysis based partial sequence of BLV *env gp51* region

To get better insight into the degree of genetic variability of Philippines BLV strains, 43 field strains representative of different breeds were used for phylogenetic analysis. 423-bp *env-gp51*

partial sequences (excluding primers) corresponding to nucleotide positions 5126 to 5548 of the full-length BLV genome (BLV isolate LS3, GenBank accession number HE 967303) were obtained, and then aligned with 72 reference sequences of reported BLV strains which represents all known eight BLV genotypes (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013). The nucleotide sequence similarity of the Philippine BLV strains were from 97.2 % to 100 % for each other, while they showed 94.8 % to 99.8 % similar to those sequences corresponding to all known genotypes deposited in GenBank. A NJ phylogenetic tree was then constructed using the Philippine BLV strains together with references through Tamura-Nei model of nucleotide substitution (Tamura & Nei, 1993). As shown in Figure 6, NJ phylogenetic tree results were in accordance with previous studies in which BLV strains were classified into eight genotypes (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Rola-Luszczak *et al.*, 2013) and an extra Iranian isolate cluster (Hemmatzadeh, 2007). Interestingly, the 43 BLV strains from the Philippines were grouped into either genotype-1 or genotype-6. Thirty two out of the 43 BLV strains showed the highest nucleotide sequence similarity to the reference sequence of genotype-1 (98.6 to 99.8 %). These BLV strains were identified in close proximity to each other and were assigned to genotype-1 as shown in the tree (Figure 6). By contrast, the rest 11 Philippine BLV strains showed 98.8 to 99.1 % similarity to other genotype 6 sequences, and were classified into genotype 6 and formed three subgroups, named G-6a, G-6b and G-6c. Among the all 32 Philippine genotype-1 BLV strains, 31 strains were isolated from ‘‘V’’ City, while the rest one strain was from ‘‘II’’ City, in Luzon Island. By contrast, 11 Philippine genotype-6 BLV strains were obtained from a number of different cities, including ‘‘II’’ City, ‘‘IV’’ City, ‘‘V’’ City and ‘‘VI’’ City, on Luzon Island.

1-3-3 Nucleotide and amino acid substitutions of BLV *env-gp51* from the Philippines strains

Many strains from the Philippines showed 100 % homology to each other. Therefore, one typical sequence among identical strains was submitted to GenBank (accession numbers KJ668809-KJ668819). Nucleotide sequences of the 11 distinctive strains out of 43 were aligned together with reference sequences of the Japanese K02120 strain (Figure 7A). As shown in the Figure 7A, all of genotype-1 Philippine BLV strains showed one unique silent substitution at nucleotide position 387, and all genotype-6 Philippine BLV strains showed seven unique silent

substitutions at nucleotide (nt) position 405, at 483 nt, at 525 nt, at 552 nt, at 555 nt, at 582 nt, and at 615 nt. By contrast, all Philippine BLV strains belonging to genotypes-1 and -6 showed two common substitutions at 363 nt and 399 nt. Interestingly, of the ten silent substitutions in this study, seven were located within epitope regions, including the viral G epitope (nucleotide 363), the second neutralizing domain (ND) (nucleotides 399 and 405), the $CD8^+$ T-cell epitope (nucleotides 483 and 525) and the E epitope (nucleotides 552 and 555), as shown in Figure 7B. Besides, seven nucleotide changes caused amino acid variations.

Eight deduced amino acid sequences for the 11 different Philippine BLV strains were aligned with the predicted amino acid sequence of K02120. The distributions of amino acid changes within the middle region of gp51, encompassing amino acid positions 101 to 241, were described as Figure 7B. This *env-gp51* region includes a portion of the first ND (residues 101-105), second ND (residues 131-149) and third ND (residues 210-225), a portion of the $CD4^+$ T-cell epitope (residues 101-113) and $CD8^+$ T-cell epitope (residues 154-182), and the viral G (residues 121), E (residues 175-194) and B (residues 228-238) epitopes (Balic *et al.*, 2012). Alignment of predicted amino acid sequences of partial *env-gp51* gene from BLV strains circulating in the Philippines showed highly conserved regions as shown in Figure 7B. Although Philippine BLV strains revealed high homology to that of the Japanese K02120 strain (accession number: AY151262), seven different amino acid substitutions were detected in BLV strain in the Philippines. Three amino acid substitutions were specifically detected in the genotype-1 Philippine BLV strains, such as substitutions of tyrosine to cysteine at residue 108 within the $CD4^+$ T-cell epitope, alanine to proline at residue 119, and leucine to phenylalanine at residue 202. By contrast, Philippine strains grouped in genotype-6 showed four specific amino acid substitutions. Interestingly, all of the Philippine genotype-6 strains showed genotype-6 specific amino acid substitution of isoleucine replaced by threonine at residue 144 within the second ND.

1-4. Discussion

The current study aims at to investigate the prevalence of BLV in the Philippine farms, and to molecularly characterize philippine BLV strains. The findings of this study are summarized as following. First, the current study revealed the distribution of BLV in the Philippines as determined by BLV-CoCoMo-qPCR-2 and nested PCR. A total of 1116 samples obtained from different farms located on five Philippine Islands showed relatively low levels of BLV infection. Among the five Islands investigated, Bohol and Cebu were negative for BLV provirus. This result is in accordance with previous studies of BLV in the Philippines (Uera *et al.*, 2012). Likewise, BLV was detected at extremely low levels, 1.1 % (4/372) and 5.6 % (9/159), in Leyte and Iloilo, respectively. By contrast, BLV infection was remarkably high (23.1 %) in Luzon compared to the other four Islands. This observation is similar to that of a previous study (Mingala *et al.*, 2009), where BLV proviral DNA was detected in samples obtained from water buffalos present in different parts of Luzon. Thus, current results suggest that the prevalence of BLV infection in Philippine cattle is different between the Southern Philippines, such as Bohol, Cebu, Iloilo and Leyte Islands, and the Northern Philippines, such as Luzon Island. Second, the present results clearly indicate that the sensitivity of BLV-CoCoMo-qPCR-2 was higher than that of nested PCR because I found a number of cattle on Leyte, Iloilo and Luzon Islands that were BLV positive determined by the BLV-CoCoMo-qPCR-2 but were negative for BLV provirus as determined by nested PCR. This result showed the same tendency as discussed in previous publications (Jimba *et al.*, 2010; Jimba *et al.*, 2012), showing that BLV-CoCoMo-qPCR is highly specific and sensitive and able to detect BLV in samples that are negative by the nested PCR assay. In addition, Jimba M et al., (2012) (Jimba *et al.*, 2012) demonstrated that the positive rate for the nested PCR in cattle correlated with the proviral load determined by BLV-CoCoMo-qPCR as follows: 1) positive rates for the nested PCR ranged from 62.9 % to 98.5 % among animals with proviral copy numbers ranging from 10^0 to 10^4 copies per 10^5 cells; and 2) the positive rate for nested PCR was 100 % in cattle with high proviral loads ($>10^4$ copies per 10^5 cells). Therefore, cattle that were BLV positive as determined by the BLV-CoCoMo-qPCR-2 but were negative for BLV provirus as determined by nested PCR may have a very low copy number of BLV, which cannot be detected by nested PCR. Indeed, only samples that were positive by both methods resulted in successful amplification of BLV *env-gp51*, and therefore these samples were used for further phylogenetic analysis. Interestingly, our previous study showed that the

original BLV-CoCoMo-qPCR using highly degenerate primers was able to detect various BLV strains from a broad geographical origin, including Japan, Peru, Bolivia, Chile and the U.S.A. (Jimba *et al.*, 2010). Here, I analyzed whether BLV-CoCoMo-qPCR-2, which uses optimized degenerate primers, allows the highly sensitive detection of BLV by using two methods, BLV-CoCoMo-qPCR-2 and nested PCR. This study results showed that BLV-CoCoMo-qPCR-2 is highly effective in detecting BLV in cattle from the Philippines. Third, phylogenetic analysis based on a 423-bp fragment of the *env* gene demonstrated that Philippine BLV strains were generally of genotypes-1 and -6, out of the eight distinct BLV genotypes worldwide (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Rodriguez *et al.*, 2009). Interestingly, a number of substitutions were found in epitope regions, such as the $CD4^+$ T-cell epitope, the $CD8^+$ T-cell epitope, the second ND, B and E epitopes, and the substitutions varied according to genotype.

The majority of BLV-positive samples detected in this study were collected from relatively large dairy farms where milking is the most common activity. A tendency towards high rates of BLV infection on dairy farms was also found in a Japanese study (Kobayashi *et al.*, 2014). Thompson and Miller (Thompson & Miller, 1974) demonstrated that pathogens can be transmitted during the milking process. Therefore, it is possible that the high level of BLV infection on dairy farms in Luzon is a consequence of transmission of infection from residual infected milk present in a milking machine to a healthy receptive cow during the milking process. One interesting finding in this study is that the BLV-infected farms in ‘‘V’’ City on Luzon Island (Table 7), where the highest rate of BLV infection was observed, belong to a totally confined setting. By contrast, cattle that were negative for the BLV provirus were raised on small farms. Previous studies focusing on risk factors for BLV transmission revealed that loose housing systems, the presence of hematophagous insects, blood-contaminated dehorning devices (Lassauzet *et al.*, 1990), physical contact (Kono *et al.*, 1983), common use of needles and/or the introduction of infected animals (Kobayashi *et al.*, 2014) are the predominant risk factors. Therefore, I believe that the extremely high level of BLV present in farms in ‘‘V’’ City may be the result of cattle management procedures that involve transfer of infected blood, such as common use of needles, ear tattooing, physical contact and/or cattle exchange from one farm to another, as well as infection through milking.

The analysis of partial BLV *env* sequences resulted in the identification of 17 nucleotide substitutions, ten of which were silent substitutions, and seven were amino acid substitutions

(Figure 7). Nucleotide substitutions at residues 231 and 234 were located in the B epitope. The substitution at residue 108 was located in the $CD4^+$ T-cell epitope, and the substitution at residues 144 was located in the second ND. Another substitution at residue 175 was located between the $CD8^+$ T-cell epitope and the E epitope. Thus, most of the amino acid changes occurred within epitope regions. Of all the amino acid substitutions, substitutions at codons 108, 119 and 202 were a unique characteristic of the genotype-1 Philippine BLV strains, while those at residues 144, 175, 231 and 234 were unique to genotype-6 BLV strains. This result is in accordance with a previous finding that most substitutions in gp51 are found within epitopes rather than at random locations (Portetelle *et al.*, 1989).

In this study, two genotypes (G-1 and -6) were identified in BLV strains from the Philippines. The classification was based on the percentage of nucleotide sequence similarity, silent nucleotide substitution patterns, the phylogenetic tree, and amino acid substitutions within a partial *env* gene sequence, together with phylogenetic analysis. Most of the Philippine BLV strains clustered into genotype-1 together with two BLV strains from Japan, and these Philippine BLV strains were identified in samples located in close proximity. All of the Philippine BLV strains clustering into genotype-1 were isolated from dairy cattle, including cross breed of Holstein x Sahiwal, and were collected from ‘‘II’’ City and ‘‘V’’ City in Luzon Island. The rest of the Philippine BLV strains clustered into genotype-6, together with isolates from Brazil and Argentina. The Philippine BLV strains clustering into genotype 6 consisted of Brahman cattle from ‘‘VI’’ City and Holstein x Sahiwal cattle from ‘‘II’’ City, ‘‘IV’’ City and ‘‘V’’ City in Luzon Island. Notably, the Philippine National Dairy Authority imported Holstein x Sahiwal cattle from New Zealand and Australia to Luzon Island (http://www.nda.da.gov.ph/na123.htm; http://pcic.gov.ph/index.php/news/press-release/pcic-covers-13th-batch-of-nda-cattle importation/#). These two countries are known to have BLV-infected cattle (Coulston *et al.*, 1990; Lew *et al.*, 2004; Thompson *et al.*, 1993) (www.daff.gov.au/ animal-planthealth/pests-diseases-weeds/animal/ebl). This observation was also mentioned in a previous study (Uera *et al.*, 2012).

In addition to the above, farmers in the Philippines are also encouraged to cross Holstein cattle with Sahiwal cattle to increase milk productivity (http://www.nda.da.gov.ph/2013/PROGRAMS/programs2013.html). Collectively, current study is the first report to show that BLV strains isolated from Luzon Island were the most similar to

BLV strains from Japan, Brazil and Argentina. Another interesting finding was that both genotype-1 and -6 exist only on farms in ‘‘II’’ City and ‘‘V’’ City on Luzon Island, which belong to a totally confined setting, but they were not present on other farms, especially on small farms. The presence of more than one genotype in the same herd has previously been noted in open herds (Licursi *et al.*, 2002) and closed herds. The presence of two genotypes in ‘‘II’’ City and ‘‘V’’ City may be the consequence of separate infections of different viral origins in two cities and further transmission of the virus through exchange of infected cattle.

1-5. Conclusions

Current study result showed the different level of BLV prevalence among a number of farms on five Philippine Islands. Bohol and Cebu were tested as BLV negative, while Leyte and Iloilo Island showed low level of BLV infection. Luzon Island showed the highest level of BLV prevalence among all tested area. Two different BLV genotypes (G-1 and -6) are circulating in the Philippines. Most of the nucleotide substitutions are silent substitution. In addition, amino acid mutations were located in epitope regions. These results provide important information on BLV infection levels and will enable the implementation of appropriate cattle-management policies in addition to providing supplementary information for the development of more-effective methods of BLV eradication in the Philippines.

CHAPTER II.

A new genotype of bovine leukemia virus in South America identified by NGS-based whole genome sequencing and molecular evolutionary genetic analysis

2-1. Introduction

BLV is oncogenic member of *deltaretrovirus* genus, being considered a model of pathogens for HTLV-1 (Rodriguez *et al.*, 2011). BLV infects cattle worldwide and associated with EBL. BLV complete genome is constitutive of structural *gag, pol* and *evn* genes, and regulatory and accessory genes *tax, rex, R3* and *G4. gag* gene encodes structural proteins of *the virus, namely* MA p15, CA p24 and NC p12 (Hamard-Peron & Muriaux, 2011; Sagata *et al.*, 1984). Gag plays a key role in the assembly of virions at the plasma membrane (Qualley & Boleratz, 2014). Enzymatic genes *pro and pol* is involved in the reverse transcription of viral RNA into viral DNA and its integration in host cellular genomes (Gillet *et al.*, 2007; Sagata *et al.*, 1984). The *env* gene encodes surface glycoprotein (gp51) and transmembrane protein (gp30) (Sagata *et al.*, 1984). The envelop glycoprotein plays essential roles in viral lifecycle and infectivity (Callebaut *et al.*, 1993; Inabe *et al.*, 1999; Johnston & Radke, 2000). Regulatory gene encodes Tax and Rex proteins which are posttranscriptional regulator of viral expression (Rodriguez *et al.*, 2011). The R3 and G4 proteins encode by accessory genes contribute to the maintenance of high viral load (Florins *et al.*, 2007; Willems *et al.*, 1994). Besides the above, BLV RNA polymerase III (pol III)-encoded viral miRNAs are strongly expressed in preleukemic and malignant cells, in which structural and regulatory gene expression is repressed, indicating a potential role in tumor onset and progression (Kincaid *et al.*, 2012; Rosewick *et al.*, 2013). Gp51 is required for cell entry and the target of neutralizing antibodies (Johnston & Radke, 2000; Mamoun *et al.*, 1990) and is subjected to immune pressures and selection processes. Therefore, the gp51 region has been widely used for BLV genotyping studies and recent phylogenetic studies based on partial and full-length gp51 region from viral strains isolated worldwide demonstrate that BLV can be classified into at least eight genotypes (Balic *et al.*, 2012; Beier *et al.*, 2001; Camargos *et al.*, 2007; Camargos *et al.*, 2002; Coulston *et al.*, 1990; Felmer *et al.*, 2005; Licursi *et al.*, 2003; Matsumura *et al.*, 2011; Molteni *et al.*, 1996; Moratorio *et al.*, 2010; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013).

The Spanish conquerors introduced cattle to the American continent from 1493. Soon after that, these introduced cattle had spread all over the South America and their population size had sharply increased to several millions (Liron *et al.*, 2006). More than 300 years, Creole cattle were the only cattle bred in America. For the purpose of improving dairy and beef production through crossbreeding with local breeds or directly replacing of local breeds, cattle of British

breeds, such as Shorthorn, Hereford, Angus, and Holstein, and indicus breeds namely Nelore, Brahman, and Gir were introduced in temperate/cold and subtropical/tropical regions, respectively, during the end of 19th and the beginning of the 20th centuries.

BLV infects cattle worldwide, and the first case of EBL in South America has been detected in Brazil and confirmed the presence of EBL since 1943 (D'Angelino *et al.*, 1998). World Organization for Animal Health data (OIE 1999 and 2009) has confirmed the prevalence of BLV in most parts of South America. In Brazil, BLV infection rates between 17.1 and 60.8 % were detected, with individual prevalence rates varying considerably among states, and recent data indicated a sharp increase in the prevalence of BLV (Camargos *et al.*, 2007; Camargos *et al.*, 2002; Castro *et al.*, 1992; D'Angelino *et al.*, 1998; Molnár *et al.*, 1999; Samara *et al.*, 1997). Early phylogenetic analysis of the BLV *env* gene demonstrated that Brazilian BLV isolates were classified into genotypes-1, -2, -5, -6, and -7 (Camargos *et al.*, 2007; Camargos *et al.*, 2002). Up to 83.3 % of BLV prevalence level have been reported in Columbia, with average prevalence differing significantly among cattle breeds and virus detection methods used (Alfonso *et al.*, 1998; Hernández-Herrera *et al.*, 2011). In Argentina, previous studies have indicated BLV prevalence (Licursi *et al.*, 2003; Monti *et al.*, 2005a; Monti *et al.*, 2005b), up to 32.8% at individual level and 84 % at herd level (Trono *et al.*, 2001). In addition, the presence of BLV genotypes-1, -2, -4, and -6 were confirmed in Argentina (Licursi *et al.*, 2003; Monti *et al.*, 2005a; Rodriguez *et al.*, 2009). BLV infection has also been reported in Peru, Chile, and Uruguay (Ch, 1983; Felmer *et al.*, 2005; Moratorio *et al.*, 2010). Analysis of Chilean BLV strains showed the presence of genotypes-1 and -4 through restricted fragment length polymorphism (RFLP) (Felmer *et al.*, 2005); however, a study using cattle from Uruguay and Chile by Moratorio et al. (Moratorio *et al.*, 2010), aligning partial *env* gp51 sequences with data available from other South American BLV strains, found that Chilean BLV strains were assigned into genotypes-4 and -7, but not genotypes-1, while Uruguayan BLV strains were of genotype-1. Thus, it appears that at least six genotypes of BLV strains are circulating in South America. However, currently there are no studies of the molecular characteristics of BLV in Bolivia, Paraguay, or Peru (OIE, 2009).

Here, I investigated the distribution of BLV in Peru, Paraguay, Bolivia, Argentina, and Chile, and molecularly characterized BLV strains circulating in Peru, Paraguay and Bolivia, together with other South American BLV strains as follows: (1) first, the spread of BLV infection was

investigated by amplification of BLV LTRs through nested PCR from blood samples obtained from a total of 2204 cattle in Peru, Paraguay, Bolivia, Argentina, and Chile; (2) second, the genetic variability of BLV strains circulating in Peru, Paraguay, and Bolivia was examined by DNA sequencing and 35 distinct strains isolated from cattle collected from these countries were genotyping by phylogenetic analyses based on partial *env* (gp51) sequences, comparing these strains and isolates from other geographical locations worldwide; (3) third, whole genome sequences of 25 BLV strains from Peru, Paraguay and Bolivia, assigned to four different genotypes, were obtained by next generation sequencing (NGS) and compared with eight full-length BLV genome sequences, generated by Sanger sequencing, available in the NCBI database. Collectively, current studies provide strong evidence for the existence of a novel BLV genotype, which designated as genotype-9, in Bolivia.

2-2. Materials and Methods

2-2-1 Experimental samples

A Total of 2204 blood samples were collected from cattle from five individual countries in South America, including 328 cattle from three farms in Peru, 139 cattle obtained from five different farms in Paraguay, 507 from eight farms in Bolivia, 420 cattle samples from three farms in Argentina, and the rest 810 samples obtained from 19 farms in Chile (Figure 8; Table 8). In all South American countries, farms which were located in main-cattle raising area were selected for sampling in each country. 40 µl of blood from each cattle were spotted on Whatman FTA elute cards (GE Healthcare Japan Corp., Tokyo, Japan), and stored as experimental material in humidity-control container at room temperature for further usage.

2-2-2 Animal handling and research Ethics

All of the 2204 tested animals were handled by veterinarians from RIKEN, Universidad Austral de Chile and LAVET, in strict accordance with good animal practice following the Universidad Austral de Chile Institutional guidelines. This study was approved by the Committee on the Ethics of Animals for Research at the National University of LA PLATA (Certificate date May 26th, 2014) and by the Committee on the Ethics of Animals for Research at Universidad Austral de Chile (Certificate No. 153-2014).

2-2-3 Genomic DNA extraction methods

Genomic DNA was extracted from 40 µl of whole blood spotted onto Whatman FTA elute cards (GE Healthcare Japan Corp., Tokyo, Japan), according to the manufacturer's instructions (Brief explanations were given in changer 1). The concentrations and purity of each DNA sample was checked by ND-1000 Spectrophotometer (NanoDrop Technologies, Inc. Wilmington, USA) and was stored at -20 °C until required for PCR.

2-2-4 Detection of BLV provirus by nested PCR targeting BLV-LTR gene

The existence of BLV provirus in each tested sample was screened by nested PCR targeting two copies of BLV LTR gene through using specific primers designed in chapter 1 and previous

study (Tajima *et al.*, 1998). Briefly, the first PCR was performed with external primer pairs: BLTR256F (5`-GAGCTCTCTTGCTCCCGAGAC-3`) and BLTR453R (5`-GAAACAAACGCGGGTGCAAGCCAG-3`). Single copy of cattle host gene *BOLA-DRA* was amplified as internal control by using primer pairs: BDRA488F (5`-ACAACACCCCAAACACCAAT-3`) and BDRA1145R (5`-AGGAAGGGGAGGTAGTGGAA-3`). Final volume of 20 µl reaction mixture, containing 1 µl of five picomole of each primer, 2 µl of 10× rTaq PCR Buffer (Toyobo, Osaka,Japan), 2 µl of 25 mM $MgCl_2$, 2 µl of 2 mM dNTP mix, 0.1 µl of 5U/µl rTaq, and 10.4 µl of nuclease-free water, were amplified at 45 cycles of 94 °C for 30 sec, 58 °C for 30 sec, and 72 °C for 30 sec, and then a final extension step of 4 min. Two percent agar gel was used to visualize the PCR products. The external PCR produces 198-bp band for BLV-LTR and 658-bp band for host gene BoLA-DRA. Successful first PCR products were purified by applying 2 µl of Exo-SAP IT (USB Corp., Cleveland, OH, USA) through the process of incubating at 37 °C for 15 min, then at 80 °C for 15 min. The purified-external PCR amplicons were subsequently applied to the second PCR with inner primer pairs: BLTR306F (5′-GTAAGGCAAACCACGGTTT-3′) and BLTR408R (5′-AGGAGGCAAAGGAGAGAGT-3′). The nested PCR reaction contains 1 µl of purified external PCR product, 1 µl of five picomoles of each primer, 2 µl of 10× rTaq PCR Buffer, 2 µl of 25 mM $MgCl_2$, 2 µl of 2 mM dNTP mix, 0.1 µl of 5U/ µl rTaq, and 11.9 µl of nuclease-free water to final volume of 20 µl, and amplified through 35 cycles of 94 °C for 30 sec, 63 °C for 30 sec, and 72 °C for 30 sec, and then a final extension step of 4 min. Sterilized water was used as a PCR negative control. Three percent agar gel was used to visualize nested PCR products. The nested PCR produces 102-bp band for BLV-LTR. Only the sample, which showed both 658-bp amplicon of DRA in the first PCR and 102-bp LTR band in nested PCR, was determined as BLV positive sample.

2-2-5 PCR amplification and sequencing of BLV *env-gp51* gene fragments

One hundred and thirty-one BLV field strain samples including 41 from Peru, 34 from Paraguay, and 56 from Bolivia which represent 30 % of BLV positive samples from each country were randomly chosen for amplification of the partial BLV *env* gene by gp51 nested PCR. The first reaction were performed using PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Otsu, Japan) through the following external primer pairs: external forward (5`-

ATGCCYAAAGAACGACGG-3`) and external reverse (5`-CGACGGGACTAGGTCTGACCC-3`) described previously publication (Moratorio *et al.*, 2010). Nested PCR were amplified by inner primer pairs: Forward Env5032 (5`-TCTGTGCCAAGTCTCCCAGATA-3`) and Reverse Env5608r (5`-AACAACAACCTCTGGGAAGGGT-3`) as described in previous publication (Asfaw *et al.*, 2005). Final volume of 25 µl PCR reaction mixture (both first and nested PCR), containing 13.5 µl (initial PCR) and 14.5 µl (nested PCR) of distilled water, 5 µl of 5× PrimeSTAR GXL Buffer, 2 µl of 2.5 mM dNTP mix, 0.5 µl of PrimeSTAR GXL, 1 µl of each primer (10 µM), and 2 µl of genomic DNA (30 ng/µl) (for the first PCR) or 1 µl of first PCR product (for the nested PCR), were amplified through the following thermal condition: 98 °C for 2 min, followed by 30 cycles of denaturation at 98 °C for 15 sec, annealing at 60 °C for 20 sec, and extension at 68 °C for 60 sec. The external primers resulted in amplification of a 913-bp DNA fragment, and internal primers amplified a 597-bp fragment of the gp51 region of the *env* gene. Two µl of Exo-SAP IT (USB Corp., Cleveland, OH, USA) was applied to purify the successfully amplified nested-PCR products and sequenced on an ABI3730xl DNA Analyzer using an ABI PRISM Big Dye Terminator v 3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Sanger sequencing produces a 475-bp partial sequence of the *env-gp51* gene, corresponding to nucleotide positions 5090-5564 of the BLV cell line FLK-BLV subclone pBLV913 complete genome (GenBank accession number EF600696) (Rovnak *et al.*, 1993).

2-2-6 Construction of a phylogenetic tree based on gp51 partial sequences

BLV *env* partial sequences of 131 BLV positive cattle from Peru, Paraguay, and Bolivia were successfully amplified. Since strains from the same farm showed homology to each other, 35 distinctive sequences out of 131 were aligned together with 74 BLV *env* sequences from GenBank using MEGA 5.1 software (Tamura *et al.*, 2011), including sequences of previously known BLV strains from other South American countries and also of BLV isolates from worldwide, representative of the eight known BLV genotypes. The construction of phylogentetic tree used two different algorithms, namely Maximum-likelihood phylogenetic tree (ML) and Bayesian Inference tree (BI), for the purpose of accurate and robust phylogenetic analysis of the BLV *env* gp51 partial sequences. Firstly, phylogenetic tree was constructed by ML algorithm with Kimura-2 parameter model with gamma distribution (K2+G) of nucleotide substitution

(Kimura, 1980), which was chosen as the model with the best fit for analysis of the BLV *env* gp51 partial sequences, by MEGA 5.1 software (Tamura *et al.*, 2011). The reliability of the phylogenetic relationships determined by ML algorithm was evaluated using nonparametric bootstrap analysis with 1000 replicates. Second algorithm used in this study is BI. BI was performed using MrBayes v.3.2.5 (Ronquist *et al.*, 2012) with the evolutionary model set to lset nst = 6, rates = equal (corresponding to the general-time-reversible model (GTL)). In BI analysis, two runs with four Markov chains were carried out simultaneously for 100,000,000 generations, and the trees were sampled every 100 generations. The first 25 % of the BI trees were discarded as "burn-in". A consensus tree was constructed from the output file produced in the BI analysis using FigTree v. 1.4.2 (http: //tree. bio.ed.ac.uk/software/figtree/).

2-2-7 PCR amplification of the BLV provirus complete genome

To obtain deeper understanding into the genetic differences between the distinct BLV genotypes, PCR amplification of overlapping genomic fragments covering the complete BLV genome of South American was achieved using the PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Otsu, Japan) and specific primers designed for this study (Life Technologies Japan Ltd, Tokyo, Japan) (Figure 9). Briefly, four different sets of PCR were run to amplify overlapping genomic fragments as described in Figure 9. The 25 µl final reaction mixture contained 5 µl of 5× PrimerSTAR GXL Buffer, 2 µl of 2.5 mM dNTP mix, 1 µl of 10 pmol/µl each primer, total of 50 ng of template and 0.5 µl of PrimerSTAR GXL DNA Polymerase under the thermal condition of preheating at 98 °C for 2 min, followed by 33 cycles of denaturation at 98 °C for 15 sec, annealing at 60 °C for 20-30 sec, and extension at 68 °C for 1-6 min (1 min per kilobase), followed by a final extension at 72 °C for 4 min. Each successful PCR amplicon was quantified using the Qubit dsDNA BR Assay kit (Life Technologies Ltd, Oregon, OR, USA) according to manufacturer instructions, and four different BLV provirus PCR amplicons from each individual sample were pooled together at an equimolar ratio to a final concentration of 30 ng/µl as the starting material for whole genome sequencing library preparation.

2-2-8 BLV whole genome library preparation

Transposase-mediated library preparation method was used for preparation of DNA libraries of the overlapped genome fragment of South American pooled samples, using the SureSelect QXT Library prep for Illumina Multiplexed Sequencing Kit (Agilent Technologies, Santa Clara, CA, USA). Briefly, 30 ng of each pooled sample was treated with 1 µl of the SureSelect QXT enzyme mix ILM and 8.5 µl of QXT buffer, and incubated for 10 min at 45 °C, followed by addition of 16 µl of QXT stop solution, enabling enzymatic fragmentation and addition of adaptors to the ends of fragments in a single reaction. The adaptor-tagged library fragments were purified using Agencourt AMPure XP beads (Beckman Coulter, Inc., Brea, CA, USA) with the NGS Magna Stand Ch YS-Model (NIPPON Genetics Co., Ltd., Tokyo, Japan) according to the manual. Next, each purified adaptor-tagged library was amplified with 1 µl of each of the indexing primers, P5 and P7, added to the index in a reaction mixture containing 5 µl of 5× Herculase II reaction buffer, 0.25 µl of 100 mM dNTP mix, 1.25 µl of dimethyl sulfoxide (DMSO), 0.5 µl of Herculase II Fusion DNA Polymerase, and 6 µl of water to a final volume of 25 µl. Each sample was dual indexed. The amplified libraries were purified with AMPure XP beads (Beckman Coulter Inc., Brea, CA, USA) according to the manual. Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) and DNA 1000 Assay were used to determine the quality and quantity of amplified libraries. The final dual-indexed libraries were pooled into one at equimolar concentrations. The library was subjected to multiplex sequencing on the MiSeq sequencer (Illumina).

2-2-9 Next-generation sequencing (NGS) data analysis

All sequences, obtained by NGS platform, were analyzed through a framework described previously (Rabbani *et al.*, 2012). Briefly, Fastq files were generated as row sequences by the Illumina MiSeq, with 600 cycles of paired-end read, and validated by evaluation of the distribution of quality scores. Validated fastq files from each viral genome were aligned with the Burrows-Wheeler Aligner tool (BWA v. 0.7.8-r455) (Li & Durbin, 2009) against a reference sequence and the resulting Alignment Map (SAM) format output was suitable for analyses using SAMTools (Li *et al.*, 2009). Analysis of sequence quality, depth of coverage, short-read alignment, and variant identification were performed using SAMTools (Li *et al.*, 2009). The

indexed file was visualized using the Integrative Genomics Viewer (IGV) (Robinson *et al.*, 2011). For references sequences in alignment, BLV strain Arg41 (GenBank accession number: FJ914764) (Dube *et al.*, 2009), The BLV cell line FLK-BLV subclone pBLV913 (GenBank accession number: EF600696) (Rovnak *et al.*, 1993), and BLV (GenBank accession number: AF033818) complete genomes were used. Full genome consensus sequences of each individual BLV strain were saved for further analyses, including alignment using MAFFT v 7.123b (http://mafft.cbrc.jp/alignment/software/) and deduction of protein sequences by in silico translation of nucleotide to amino acid sequences using MEGA 5.1 (Tamura *et al.*, 2011).

BLV genome sequences (8714-bp) were aligned using MAFFT software, and the nucleotide and amino acid substitution per site was calculated using the Jukes and Cantor (Jukes & Canter, 1969), and *p*-distance models (Tamura *et al.*, 2011), respectively. A ML tree for full BLV genome sequences was constructed using MEGA 6.06 software (Tamura *et al.*, 2013), and 1000 replications were used to calculate bootstrap values.

2-3. Results

2-3-1 Investigation of the prevalence in South America

A total of 2204 blood samples obtained from cattle in five South American countries were screened for the prevalence of BLV by nested PCR targeting LTR region (Table 8; Figure 8). In Peru, 139 out of 328 cattle (42.3 %) were tested as BLV provirus positive, with infection level ranging from 0 to 58.6 % at the farm level. 139 cattle samples screened in Paraguay, 76 cattle samples showed BLV positive result. Infected cattle in Paraguay were classified as Holstein, Brown Swiss breed, and cross breed of Holstein X Brown Swiss. All of the breeds mentioned above showed high BLV prevalence rate, ranging from 28.6-71.4 % in the Holstein breed and 36.4 % in Brown Swiss. Among the 507 samples obtained from different farms in Bolivia, 156 samples were detected as BLV positive, which includes Holstein, Gir and Yacumeno breeds. Among the BLV infected breeds, Holstein cattle showed the highest BLV prevalence, with up to 100 % at the individual level, while Yacumeño showed moderate levels of BLV infection ranging from 4.5 to 24.3 %. Interestingly, I did not detect any positive samples from Brown Swiss, Brahman and Montana in Bolivia. 420 samples collected from Argentina showed remarkably high level of BLV prevalence as 77.4 %, with up to 90.9 % at the herd level. In Chile, 810 samples were obtained from different cattle breeds at 19 farms over a relatively wide geographical area (Table 8). 236 out of 810 samples were determined as BLV provirus positive. It is not easy to compare BLV infection levels between breeds in Chile, since a wide range of breeds were tested as BLV positive; however, remarkable differences were observed between farms, breeds, and locations, indicating variability of BLV infection levels in this country. When the BLV infection level of the dairy and beef cattle were compared, current study found that dairy cattle breeds, namely Holstein and Brown Swiss, have higher level of BLV prevalence than that of beef cattle breeds including Yacumeño and Nelore on average. Current study result agrees with previous studies (Kobayashi *et al.*, 2014; Ohno *et al.*, 2015). However, there is no statistically significant correlation of BLV prevalence among Breeds, because individually breeds are different from farm to farm and also from country to county.

2-3-2 Identification of a novel BLV genotype, genotype-9, by phylogenetic analysis based on partial *env* gp51 sequences

To gain insight into the degree of genetic variability of BLV strains circulating in South America, especially in Peru, Paraguay and Bolivia where no study about molecular characteristics of BLV available, *env* gp51 partial sequences were amplified from 131 field strains. Since strains from the same farm showed 100 % homology to each other, 475-bp gp51 partial sequences of 35 distinct strains from 131 field strains, corresponding to nucleotide positions 5090-5564 of the full-length BLV genome (BLV cell line FLK-BLV strain pBLV913, accession number EF600696) (Rovnak *et al.*, 1993), were aligned with 74 corresponding sequences of known BLV genotypes from other South American countries (Argentina, Brazil, Chile, and Uruguay) and different corner of the world (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013). Phylogenetic trees constructed by ML (Figure 10) and BI (Figure 11) algorithm showed congruent topologies, supported by moderate to high bootstrap values and high posterior probabilities. Previous phylogenetic studies of BLV strains revealed the presence of eight different BLV genotypes worldwide (Balic *et al.*, 2012; Beier *et al.*, 2001; Camargos *et al.*, 2007; Camargos *et al.*, 2002; Coulston *et al.*, 1990; Felmer *et al.*, 2005; Licursi *et al.*, 2003; Matsumura *et al.*, 2011; Molteni *et al.*, 1996; Monti *et al.*, 2005a; Moratorio *et al.*, 2010; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013). Interestingly, phylogenetic tree in this study assigned all aligned strains into nine clusters, designated genotypes 1-9, as shown in Figures 10 and 11. BLV strains circulating in Peru were grouped into genotypes-1, -2, and -6. Most of Paraguayan BLV strains were assigned into genotypes-1 and -6, with a small number in genotype-2, together with Argentine BLV strains. To my surprise, the BLV strains obtained from Bolivia were clustered not only into genotypes-1, -2, and -6, together with Peruvian and Paraguayan strains, but also into a unique clade, which was distinct from the eight previously known BLV genotypes. This novel genotype was supported by a high bootstrap value (92 %) in the ML tree (Figure 10) and a posterior probability of 100 % using the BI approach (Figure 11), indicating the presence of new genotype, genotype-9. In addition, the genotyping result of reference sequences of BLV strains isolated from Argentine, Brazil, Chile and Uruguay are consistent with previous studies (Camargos *et al.*, 2007; Camargos *et al.*, 2002; Felmer *et al.*, 2005; Licursi *et al.*, 2003; Monti *et al.*, 2005a; Rodriguez *et al.*, 2009). BLV strains from Argentina were mainly classified into genotypes-1 and -2, with relatively few samples were

grouped into genotypes-4 and -6. Reference sequences of Brazilian BLV strains were assigned into a wider variety of BLV genotypes, consisting of genotypes-1, -2, -5, -6, and -7. Chilean BLV strains were classified into genotypes-4 and -7. Furthermore, strains from Uruguay were assigned only into genotype-1. Thus, current study results conclude that at least seven genotypes of BLV strains (G-1, -2, -4, -5, -6, and -7, plus the novel G-9) are circulating in South America.

2-3-3 Whole genome sequencing and sequence comparison of strains of genotypes-1, -2, and -6, and novel genotype-9

To compare the divergence of the BLV complete genome, overlapping genomic fragments covering the full BLV genomic sequence (Figure 9) were amplified from strains grouped into the novel genotype-9 and those in other genotypes, such as genotype-1, -2 and -6, as determined by dual phylogenetic analyses of *env* gp51 partial sequences. A total of 25 samples, including one sample from Peru, seven from Paraguay, and 17 of Bolivian samples were amplified by PCR. Then, PCR amplicons were pooled to prepare DNA library for multiplex sequencing on the MiSeq system (Illumina). Short-read sequences were assembled by using three BLV reference sequences (Accession numbers; EF600696, FJ914764, and AF033818), and the average depth of genome coverage was between 143 and 1558 (Table 9). In addition, the *env-gp51* fragments of BLV complete genome sequences were 100 percent identical with those of partial *env-gp51* sequences produced by Sanger sequencing. Complete genome sequences of 25 South American BLV strains showed 98.29-100 % homology to each other, while these sequences were 95.63-99.42 % homologous to the FLK-BLV strain pBLV913 sequence. Next, I estimated the average substitution rate of paired sequences for each of viral gene between the 25 full sequences of novel BLV provirus and those of previously reported eight BLV complete genome sequences. The numbers of nucleotide and amino acid substitutions per site were calculated by the Jukes and Cantor method, and *p*-distance models, respectively. As I can see from Table 10 that the diversity of each structural genes of all BLV strains, including *gag* (*p15*, *p12*, and *p24*), *pro*, *pol*, and *env* (*gp30* and *gp51*) genes, and regulatory and accessory genes, including the *tax*, *rex*, *R3*, and *G4*, and LTRs, were very different. 0.023 was determined as the average nucleotide substitution rate per site for BLV complete genome. As far as each gene is concerned, *G4*, *rex*, and the LTR genes had the lowest average nucleotide substitution rates of 0.013, 0.016 and 0.013, while the *R3*, *p24*, and *gp30* genes fragment revealed the highest average rate of nucleotide

substitution as 0.027, 0.030 and 0.027 shown in Table 10. Amino acid substitution rates of each viral protein were compared, *tax* and *R3* revealed the most polymorphism as 0.053 and 0.057 respectively, whereas *p24* was significantly conserved.

2-3-4 Confirmation of the existence of novel BLV genotype, genotype-9 through phylogenetic analysis based on whole genome sequences

ML phylogenetic tree was constructed based on complete genome sequence of 25 South American BLV strains and of previously known BLV strains worldwide, which were classified into genotypes-1, -2 and -4 (Figure 12). Until now, full genome sequences of BLV strains classified into genotypes-3, -5, -6 and -8 were unknown. In this study, for the first time, I sequenced complete genome sequences of three genotype-6 BLV strains from South America, and included these sequences in phylogenetic analysis. ML full genome tree clearly demonstrated the classification of BLV genotypes. I can see from the Figure 12 that all strains were separately assigned into known genotypes-1, -2, -4 and -6 respectively, and a novel genotype-9, into distinct clades supported with bootstrap values 100 % for each clade. Interestingly, all of genotype-9 South American strains were close to each other than other strains, and grouped together to form the novel genotype-9 clade. Thus, full genome phylogenetic analysis provides strong evidence for the existence of the novel genotype-9 which was indicated as new by phylogenetic analysis based on gp51 partial sequences. Another interesting finding of this study is that BLV strains from two different geographical locations of Bolivia, named Portachuelo and Montero, were not mixed but separately located in two different sub-clades within genotype-9, indicating at least two distinctive genotype-9 BLV strains circulating in Bolivia. The presence of genotype-9 sub-strains is observed only by analysis through full BLV genome sequences, but not partial gp51, indicating the limitations of gp51 partial sequences. To sum up, ML tree confirmed the existence of novel BLV genotype, G-9.

2-3-5 Diversity of structural and non-structural genes of 16 complete genome sequence of novel BLV strains from Peru, Paraguay and Bolivia in South America

Since some of the full genome sequences of 25 South American BLV strains showed homology to each other, 16 unique sequences out of 25 sequences were chosen for further

studies. In order to get better insight into how the novel genotype-9 BLV strains different from other known BLV strains in South America, including genotype-1, -2, and -6, nucleotide sequences or predicted amino acid sequences of each viral gene or proteins, including structural gene *gag* (*p15*, *p12*, and *p24*), *pro*, *pol*, *env* (*gp30* and *gp51*), and non-structural genes *tax*, *rex*, *R3*, *G4*, LTRs and miRNAs, of 16 distinct strains were aligned to the FLK-BLV strain pBLV913 as a reference sequence (Figures 13, 14, 15, 16, 17, 18).

BLV transcription is initiated at the boundary of U3-R of 5`-LTR by regulatory protein Tax, and is terminated with polyadenylation at the 3`-LTR R region. An essential factor required for viral transcription is the presence of triplicate copy of 21-bp enhancer elements (TxRE) located in the 5`-LTR U3 region. Each TxRE has a cyclic AMP-responsive element (CRE) and an E-box sequence, which overlaps the CRE motif. The U3 region also includes binding sites of PU.1/ Spi-B and a glucocorticoid responsive element (GRE). Additionally, BLV expression is regulated by 5`-LTR sequences downstream of the transcription initiation site (DAS) at the 3` end of the R region and an interferon regulatory factor (IRF) binding site in the U5 region. LTR sequences comparison between 16 South American BLV strains, assigned as genotypes-1, -2, -6 and -9, and of FLK-BLV strain pBLV913 showed that entire LTR regions, including functional domains of CRE motifs, E Boxes, CAT Boxes, polyadenylation sites (PAS), and IRF binding sites, were well conserved in all strains (Figure 13). Interestingly, three nucleotide substitutions were detected downstream of the DAS, at position 374, 399, and 400 in sequences of genotype-9 strains and strains of other genotypes. Other individual nucleotide mutations were also detected outside of functional domains (Figure 13).

Gag is a polyprotein precursor constituted of p15 MA, p24 CA, and p12 NC proteins (Figure 14a). All of the above mentioned protein regions were conserved strictly among the aligned BLV full genome sequences. For the p15 MA protein, sequences of three genotype-1 South American BLV strains were identical to that of pBLV913, whereas BLV strains clustered into genotypes-2, -6, and -9 had amino acid substitutions at residues 48, 61, 63, 69, 88 and 90 (Figure 14a). The p24 CA protein region was also well conserved overall, with substitutions detected at residues 318 (V318 M/I/T), and 323 (V323I) (Figure 14a). The p12 NC protein is made up of a region rich in basic amino acid residues and zinc binding domains involved in RNA packaging, both of which are very well conserved. A study by Wang et al. concluded that substitutions affecting either the basic amino acid residues or the zinc finger domains might cause a reduction in RNA

packaging (Wang *et al.*, 2003). In this study, all basic residues and Zn finger domains were conserved in all the BLV strains analyzed (Figure 14a), but a substitution from Proline to Serine at amino acid residue 340 (P340S) was detected in the p12 NC protein region of South American genotype-1 strains. This substitution has previously been reported in genotype-1 BLV strains LS1, LS2, and SL3 (Moratorio *et al.*, 2013). In addition, another substitution at residue 365 (A365T) was detected in strains of genotypes-2, -6, and in novel genotype-9. However, the functional consequences of these amino acid substitutions are unknown.

The BLV Pro protein is an aspartic protease having the function in *gag* processing and also virion maturation (Moratorio *et al.*, 2013). As I can see from Figure 14b that although amino acid substitutions at residues 13, 81, 91, 134, 148 and 149 were detected among the Pro protein amino acid sequences of all South American strains, this region was generally conserved well, and any of new genotype-9-specific substitutions were not detected. The deduced amino acid sequences of BLV RT and IN protein encoded by the *pol* gene were also highly conserved among all of 16 South American BLV strains (Figure 15). The deduced amino acid sequences of polymerase (Pol) of the complete BLV genome sequences aligned were similar to the reference sequence; however, amino acid substitutions of glutamic acid by aspartic acid (E166D) and aspartic acid replaced by glycine (D447G) in the RT region, and mutations from histidine by tyrosine (H644Y) and alanine by threonine (A826T) in the IN region, were detected only in all sequences of novel genotype-9 BLV strains (Figure 15). In addition, an asparagine replaced by serine substitution (A792S) was detected only in the IN protein sequence of all 12 Portachuelo samples, which were classified into genotype-9, but this substitution was not detected in other genotype-9 strains, indicating that certain amino acid mutations could be limited to particular geographic regions (Figure 15). In addition, nine amino acid substitutions at residues 209, 231, 268, 350, 399, 437, 695, 814, and 839 were found in the Pol protein region of three sequences of genotype-6 Bolivian strains.

The deduced amino acid sequences of Env leader peptide and gp30 protein were highly conserved among all of South American BLV strains in this study, as were those of gp51 proteins (Figure 16). It is known through comparing of conformational (F, G, and H) and linear (A, B, D, and E) epitopes that A, B, and E linear apitopes, and F, H conformational epitopes were conserved well, while the G-epitope at residues 48, 74 and 82, and D-epitope at residues 254 and 267 showed variations (Figure 16). The K74R substitution at G-epitope has previously

been detected in *env* gp51 partial sequence deposited in the GenBank database (Rodriguez *et al.*, 2009). The S82F substitution was present in sequences of strains classified into genotypes-2, -6, and -9. Interestingly, the S82L substitution was detected only in all 12 sequences of genotype-9 Portachuelo-derived strains, but not other trains of genotype-9 Bolivian strains. As far as ND are concerned, the first and third NDs were conserved, while the second ND revealed some genotype-specific substitutions (Figure 16). Genotype-1 Paraguay strains showed substitutions of aspartic acid to asparagine at residue 134 (D134N) and of phenylalanine to serine at residue 146 (F146S), which have previously been detected in some *env* gp51 partial sequences of BLV trains isolated from Uruguay and Brazil (Moratorio *et al.*, 2010). Besides, genotypes-2 and -6 BLV strains in the current study also revealed specific substitutions, N141D and I144T, in the second ND, respectively. Interestingly, one significant amino acid substitution was observed at amino acid residue 133 of the gp51 protein, with alanine (A) substituted for valine (V) in all BLV genotype-9 strains, in consistent with results obtained by direct sequencing of the partial gp51 *env* region. Furthermore, both gp51 partial sequences and complete BLV genome sequences demonstrated the conservation of all of the eight N-linked glycosylation sites. In the gp30 transmembrane protein, the fusion peptide and GD 21 domain were highly conserved, while cytoplasmic domain of the gp30 protein showed divergence with some substitutions, including Q470E, L479F, T480A, H500R, and V504T, commonly detected in South American BLV strains of various genotypes (Figure 16).

One of the main functions of BLV Tax protein is to regulate viral expression. The functional domains located in the BLV Tax protein include putative zinc finger motif (residues 30-53) (Aida *et al.*, 2013; Gillet *et al.*, 2007), a leucine-rich activation domain (residues 157-197) (Willems *et al.*, 1991), two phosphorylation sites (residues 106 and 293) (Willems *et al.*, 1998), and a multi-functional domain (residues 240–265) (Aida *et al.*, 2013; Arainga *et al.*, 2012; Tajima & Aida, 2000a; 2002; Tajima *et al.*, 2003; Takahashi *et al.*, 2005; Takahashi *et al.*, 2004). The above mentioned domains of Tax proteins in South American BLV genotype-1, -2, -6 and -9 strains and reference sequences pBLV943 were compared as shown in Figure 17a. Substitutions at residue 40 (Q40R) in the zinc finger domain, mutations at residue 164 (S164P), at residue 169 (L169I) and at residue 183 (R183 K) in the leucine-rich activation domain, and residue 257 (C257N) in the multi-functional domain were detected in the deduced amino acid alignments of South American BLV strains (Figure 17a). By contrast, all South American BLV strains

included in this study showed the conservation of two phosphorylation sites in Tax protein. One of the most important findings in the alignment of the Tax protein were the detection of substitution at residue 108 (F108L) only in sequences of all of the South American novel genotype-9 BLV strains, and a substitution at residue 100 (P100S) observed only in samples collected from Portachuelo, indicating genotype-9 and/or Portachuelo region specific mutations in Tax.

Regulatory protein Rex is associated with nuclear pores and has a nuclear localization signal (NLS) and a nuclear export signal (NES) required for RNA-binding and nuclear localization (Figure 17b). Deduced amino acid sequences alignment of the 16 South American BLV strains showed that the NLS domain was conserved well, while one single mutation (A80S) was observed in the NES of the majority of samples. In addition, a variety of substitutions, including M39T, T72P, and L146S, were commonly detected in BLV strains grouped into genotypes-2, -6, and -9, and were also present in known sequences deposited in the GenBank database. Five substitutions at amino acid residues 12, 40, 53, 58, and 70 were observed only in BLV strains of genotype-6. Of interest, one amino acid mutation at residue 113 (A113E) of the Rex protein was detected only in all sequences of genotype-9 strain as described in Figure 17b.

The pX region of BLV genome also encodes two accessory proteins, such as G4 and R3. The amino acid sequence of G4 harbors an amino terminal stretch of hydrophobic residues (amino acids 1-24) followed by a potential proteolytic cleavage site, a myb-like motif (MYB; amino acids 39-44) and an arginine-rich region (ARR; amino acids 58-72) located in the middle of the protein (Alexandersen *et al.*, 1993) (Figure 17d). The MYB and ARR of G4 were conserved well; however, an amino acid change from phenylalanine to serine at residue 22 (F22S) was detected in one of the two cleavage sites in three South American BLV strains of genotypes-2 and -6. The biological functions of mutant F22S remains to be known. Other point mutations at residues 7, 25, 26, 27, 35, 45, 55, and 71 were found in some South American BLV strains. By contrast, the deduced amino acid sequences alignment of accessory protein R3 revealed that a single mutation at residue 26 (N26H) was detected only in the sequences of all genotype-9 BLV strains Figure 17c. In addition, four mutations, Q12E, K27E, H35Y and L40S, were observed only in three genotype-6 South American BLV strains, indicating genotype-6-specific mutations in the R3 protein.

BLV-encoded miRNAs might possibly play key roles in the regulation of gene expression and tumor development (Kincaid *et al.*, 2012; Rosewick *et al.*, 2013). In this study, ten unique BLV-miRNAs sequences derived from five predicted miRNA precursor hairpins, BLV-premiR-B1 to -B5, were also obtained (Figure 18). The BLV-miRNA sequences were corresponding to those described in previous studies (Kincaid *et al.*, 2012; Rosewick *et al.*, 2013). All of these miRNA sequences were conserved strictly as follows: BLV-miRNA-B2-5p, BLV-miRNA-B3-5p, and BLV-miRNAB3-3p were conserved in all four South American BLV genotypes in this study. In addition, the seed sequences of the predominant arms of seven out of the ten miRNAs were the same in all examined BLV strains. However, all BLV strains classified as genotype-9 had an A to G nucleotide change in the seed sequence of BLV-miRNA-B5-3p. Besides, nucleotide substitutions of A to G and C to T in BLV-miRNA-B1-3p, G to A in BLV-miRNA-B2-3p, A to T in BLV-miRNA-B4-3p, and T to C in BLV-miRNAB5-5p were detected only in BLV genotype-6 strains as shown in Figure 18.

2-3-6 Observation of novel genotype-9 specific mutations

To obtain better insight into the novel genotype-9 BLV strains, I compared each gene of genotype-9 to that of other genotypes worldwide. Sixteen distinctive novel BLV complete genome sequences of BLV strains obtained from Peru, Paraguay, and Bolivia were aligned with eight previously reported BLV whole genome sequences, including FLK-BLV pBLV913 (Rovnak *et al.*, 1993), strains LS1-LS3 from Uruguay (Wang *et al.*, 2003), strains Arg41 and Arg38 from Argentina (Dube *et al.*, 2009; Dube *et al.*, 2000), and two other strains from Japan (Sagata *et al.*, 1984) and the USA. As summarized in Figure 19, I detected ten amino acid substitutions unique to genotype-9 BLV strains as follows: (1) Two substitutions, E166D and D447G, in the Pol (RT) region; two substitutions, H644Y and A826T, in Pol (IN) region; and one substitution at residue 792 of the Pol (IN) region, restricted to only sequences of all 12 samples collected from Portachuelo, but not in other genotype-9 strains. (2) In the Env (gp51) protein, one significant substitution, A133V, was commonly observed in all BLV genotype-9 strains and not detected in eight known whole BLV genome sequences. Interestingly, another substitution in the Env (gp51) protein, S82L, was detected only in sequences of genotype-9 Portachuelo strain, but not in other genotype-9 sequences. (3) In the regulatory proteins Tax and Rex, substitutions at residue 108 (F108L) of Tax and residue 113 (A113E) of Rex were detected

only in genotype-9 strains, but not in other known BLV genome sequences. In addition, a substitution at residue 100 (P100S) of the Tax protein was observed only in samples collected from Portachuelo. (4) Likewise, one mutation at residue 26 (N26H) of the accessory protein R3 was detected only in the sequences of all genotype-9 BLV strains, but not in other strains or known BLV genome sequences. These ten substitutions were first detected in the sequences of the novel genotype-9 strains identified in this study and the impact of these substitutions on the function of the proteins will require further study. Besides, 15, 7, and 22 unique amino acid substitutions were observed in BLV strains of genotypes-1, -2, and -6, respectively (Figure 19).

2-4. Discussion

Recent phylogenetic studies based on BLV *env-gp51* gene sequences from strains isolated worldwide have classified the virus into eight genotypes (Balic *et al.*, 2012; Beier *et al.*, 2001; Camargos *et al.*, 2007; Camargos *et al.*, 2002; Coulston *et al.*, 1990; Felmer *et al.*, 2005; Licursi *et al.*, 2003; Matsumura *et al.*, 2011; Molteni *et al.*, 1996; Monti *et al.*, 2005a; Moratorio *et al.*, 2010; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013). In this study, I identified a novel BLV genotype, genotype-9, in samples from two Bolivian provinces, Montero and Portachuelo, by phylogenetic analyses of partial *env* gp51 and whole genome sequences (Figure 19). Thus, the identification of genotype-9 is most interesting and a novel finding of current study. The phylogenetic analysis of BLV complete genome was in agreement with the phylogenetic analyses of 109 partial *env gp51* sequences obtained by two independent methods (Figures 10 and 11). Notably, the bootstrap value of the genotype-9 cluster was remarkably increased in the ML phylogenetic tree constructed from complete genome sequences (bootstrap value = 100%), compared to that constructed using partial *env* sequences (bootstrap value = 92%). Therefore, the NGS analysis provided strong confirmatory evidence that genotype-9 is separate from the other eight genotypes.

ML phylogenetic analysis of BLV whole genome sequences also showed that genotype-9 sequences derived from Portachuelo were distinct from those collected in Montero (Figures 12). In addition to the above, comparison of BLV amino acid sequences showed that 10 unique amino acid substitutions were observed in the Pol, Env, Tax, Rex, and R3 regions of genotype-9 BLV strains, as shown in Figure 19. Importantly, two of these substitutions, including N792L in the *Pol* region and P100S in *Tax*, were observed only in samples collected from Portachuelo assigned to genotype-9, indicating the presence of at least two distinct genotype-9 strains in Bolivia, which is in consistent with ML phylogenetic analysis of BLV whole genome sequences. This result was obtained from BLV whole genome sequence data, but was not apparent from analyses of partial *env* gp51 sequences. Full BLV genome comparison also revealed that 15, 7, and 22 unique amino acid substitutions are encoded by BLV genotypes-1, -2, and -6, respectively.

Phylogenetic analyses of the *env gp51* gene also suggested the possibility that genotype-6 may be divided into three sub-genotypes, including clades for 1) the FJ808582/Argentina strain, 2) the DQ059415/Brazil and AY185360/Brazil strains, and 3) newly-identified sequences,

including Paraguay-91. By whole genome analysis, Paraguay-62, which was assigned to clade 2) in the gp51 *env* analysis, was clearly divided from Paraguay-91, with a bootstrap value of 100%. My previous work has also suggested that genotype-6 may be divided into three subgroups: G-6a, G-6b and G-6c (Chepter-1). Therefore, these two branches, including Paraguay-62, and Paraguay-89 and -91 may be designated as separate genotypes when additional whole genome sequences of genotype-6 are accumulated and analyzed in future.

This study revealed a widespread distribution and remarkably high levels of BLV in five South American countries, including Peru, Chile, Argentina, Paraguay, and Bolivia. In Peru, 139 of 328 cattle samples were positive for BLV provirus, with a prevalence rate of up to 58.6%. Of Peruvian samples, 58.6% of those from Holstein cattle in Lima were BLV positive, whereas 40.0% of samples from Pucallpa gave positive results. Current study results were consistent with those of Ch (Ch, 1983), who demonstrated that BLV infection is widespread in Peru; however, the percentages of BLV positive samples found in the Peruvian areas of Lima and Pucallpa were higher in my study than those reported by Ch (up to 31.0%) (Ch, 1983). In Argentina, I found that 325 of 420 samples were BLV positive, with an overall prevalence of 77.4% at individual level and up to 90.9% at the herd level. These results are consistent with previous reports demonstrating that BLV is widespread in Argentina (Ch, 1983; Gutierrez *et al.*, 2011; Juliarena *et al.*, 2013; Licursi *et al.*, 2003; Monti *et al.*, 2005a). Argentine samples were collected from farms in Buenos Aires, where a high individual BLV prevalence was reported by Trono *et al*., (Trono *et al.*, 2001) and this study results were broadly consistent with those of Trono *et al*.; however, I detected a higher BLV prevalence at the herd level, since Trono *et al*. reported infection levels of 84% (Trono *et al.*, 2001). In addition, at the individual level, the percentage (77.4%) of BLV positive Argentine samples in current study was slightly higher than that reported by Monti *et al*. (70%) (Monti *et al.*, 2005a), indicating a possible increased prevalence of BLV in Argentina. Compared to Peru and Argentina, relatively few BLV positive samples were identified from Chile. However, from the distribution of BLV presented in Table 8, it can be seen that more breeds and a wider geographical area were involved in BLV infection in Chile, indicating a tendency towards widespread infection, in agreement with a previous study of BLV prevalence in this country (Felmer *et al.*, 2005). By contrast, over 50% (76/139) of screened samples from Paraguay were BLV positive and all infected breeds in Paraguay demonstrated a high level of BLV prevalence, ranging from 28.6% to 71.4% in the Holstein breed, with 36.4%

in Brown Swiss cattle. In Bolivia, 156 of 507 samples were BLV positive and the BLV infection level differed from farm to farm. There are no previous data detailing the prevalence of BLV in Paraguay and Bolivia, and this study result is the first report confirming the presence of BLV in these countries.

Based on sequencing results from the BLV provirus, the predicted amino acid sequences of the partial and full gp51 *env* gene from Peruvian, Paraguayan, and Bolivian BLV strains were highly conserved. Amino acid substitutions in the gp51 protein were mostly located in the second neutralization domain, which is consistent with previous studies (Moratorio *et al.*, 2010; Rodriguez *et al.*, 2009). In this study, only two distinct genotype-1 BLV strains from Paraguay had amino acid changes at positions 134 (D134N) and 146 (F146S) of the second neutralizing domain. These two substitutions were also observed in Brazilian and Uruguayan BLV genotype-1 strain isolates (Moratorio *et al.*, 2010), indicating that these amino acid substitutions exist mainly in South American BLV strains, and suggesting a common origin for these virus strains. Interestingly, Moratorio *et al*. (Moratorio *et al.*, 2010) confirmed that D134N changes the net charge of a loop of the Env protein, indicating a potential impact of this substitution on virus-host interactions. BLV strains grouped into genotype-2 showed a common substitution, N141D. Interestingly, the V191I substitution in the B epitope, observed in only one distinct BLV strain from Peru in this study, has been reported in some Argentine BLV isolates (Moratorio *et al.*, 2010). Most importantly and interestingly, an A133V substitution in the second neutralization domain was observed in only genotype-9 BLV strains. Previous data have demonstrated amino acid changes of alanine to aspartic acid (A to D) or threonine (A to T) (Rodriguez *et al.*, 2009) at this residue, but there are no previous reports of alanine to valine substitutions, indicating that A133V is a novel finding of this study.

At present, there are eight full-length BLV genome sequences available in the GenBank database. The major limitations of Sanger sequencing compared to Next Generation sequencing methods are increased time and cost, and the limited amount of data generated by sequencing runs. The rapid development of NGS (DNA and RNA-seq) has accelerated sequencing analysis of new genomes and transcriptomes of complex organisms, allowing the identification of new virus species integrated within complete or draft genome sequences. Commonly, virus genomes are very compact and the sequences have multiple functions, such as protein encoding (sometimes for multiple proteins), transcription, regulation of host genes, and virus genome

packaging. Therefore, identification of complete viral genome sequences is becoming increasingly important. In this study, I identified both BLV partial *env* and complete viral genome sequences from the same samples. This study data indicate that the BLV *env* sequence is a good target for classification of BLV strains, and analysis of *env* sequences was helpful for selection of informative strains to include in the complete genome sequence analysis.

In this study, I sequenced complete genomes of 25 BLV strains, including 17 novel genotype-9 strains. This study results revealed that BLV genomes contain a number of unique genotype specific substitutions not only in the *env* region, but also in the LTR, Gag, Pro, Pol, Tax, Rex, R3, G4, and miRNA encoding regions. The Information about substitutions in viral genomes is important for investigating viral spread worldwide. Therefore, full understanding of the biological functions of these genotype specific substitutions is now essential. By contrast, I also found that BLV genome sequences of strains from different geographic origins, especially the important sites on the regulation of viral replication of BLV, are relatively stable and highly conserved as follows. Three copies of CRE sequences were imperfectly conserved in all strains of each genotype. Merezaki *et al*. (2001) found that introducing a perfect CRE sequences into TxRE increases the BLV LTR promoter activity, suggesting that imperfect conservation of CREs repress viral expression with escape from the host immune response (Merezak *et al.*, 2001). Furthermore, deep sequencing results of the current study were consistent with previous result (Rola-Luszczak *et al.*, 2013) that all of the eight N-linked glycosylation sites are conserved very well. The absence of mutation of N-linked glycosylation sites in naturally infected cattle is explained by the hypothesis as being favorable for the virus to coexist with its host for efficient replication and transmission (de Brogniez *et al.*, 2015). Moreover, the two phosphorylation sites on Tax protein, and NES and NLS regions in Rex Protein were also conserved well. Thus, the fact that limited sequence variations are compatible with the development of a vaccine (Bai *et al.*, 2015; Gutierrez *et al.*, 2014; LanLan *et al.*, 2015).

2-5 Conclusion

The present study provides the first evidence of the prevalence of BLV infection among cattle in Paraguay, Peru, and Bolivia, and also confirms the widespread distribution of BLV infection in Argentina and Chile. I identified a novel BLV genotype, genotype-9 in Bolivia. BLV genotyping studies concluded that Peruvian and Paraguayan BLV strains are of genotypes-1, -2,

and -6. Of interest, this study findings indicate that Bolivian BLV strains are clustered into genotypes-1, -2, -6, and -9. Complete genome sequences provide detailed information of nucleotide and amino acid sequences of novel genotype-9, and previously known genotype-6. Genotype-9 specific ten amino acid substitutions were firstly detected in this study. This information will be important for the control of BLV infection in South America, and the data regarding the prevalence of BLV in Bolivia and Paraguay is unique to this study.

CHAPTER III.

Confirmation of genotype-10 of bovine leukemia virus in Myanmar by full genome sequences

3-1. Introduction

BLV is the causative agent of EBL, the most common neoplastic disease of cattle (Aida *et al.*, 2013; Gillet *et al.*, 2007). Symptoms of lymphoma by BLV infection depend on the site of the tumors and may include digestive disturbances, weight loss, weakness, reduction of milk production and also neurological manifestations (OIE, 2012).

BLV complete genome includes structural and enzymatic *gag*, *pro*, *pol* and *env* genes, and regulatory genes *tax*, *rex*, *R3* and *G4* (Aida *et al.*, 2013; Gillet *et al.*, 2007). Like other structural genes, the BLV *env* gene is indispensable in the synthesis of virions. The *env* gene is transcribed as 5.1 kb mRNA coding for the $pr72^{env}$ precursor (Mamoun *et al.*, 1983; Zarkik *et al.*, 1997) which further cleaved into extracellular gp51 and transmembrane gp30. Among the Env glycoproteins, gp51 glycoprotein plays an essential role in the viral life cycle (Callebaut *et al.*, 1993; Johnston & Radke, 2000; Mamoun *et al.*, 1990), and is required for cell entry and the target of neutralizing antibodies (Johnston & Radke, 2000; Mamoun *et al.*, 1990; Portetelle *et al.*, 1989). The N-terminal half of BLV gp51 contains three conformational epitopes, F, G and H (Bruck *et al.*, 1982a), and plays an important role in viral infectivity and syncytium formation (Bruck *et al.*, 1982a; Portetelle *et al.*, 1989), while the C-terminal half of BLV gp51 contains the linear epitopes A, B, D, and E (Bruck *et al.*, 1982a; Mamoun *et al.*, 1983). Therefore, the gp51 region has been widely used for BLV genotyping studies.

Previously different clustering or genotyping were result from samples collected from different geographical locations. In the early studies of BLV genotyping, researchers used a variety of restriction enzymes to group BLV strains based on restricted fragment length polymorphisms (RFLP). They named BLV genotypes according to sample origin, such as Argentine type and Australian type, or based on clustering such as cluster 1 or group 1, and a total of seven BLV genotypes were identified by RFLP (Asfaw *et al.*, 2005; Coulston *et al.*, 1990; Fechner *et al.*, 1997; Felmer *et al.*, 2005; Licursi *et al.*, 2003; Mamoun *et al.*, 1990; Monti *et al.*, 2005a). Meanwhile, partial sequences of gp51 were also used in phylogenetic analysis in combination with or without RFLP, and BLV strains isolated from a number of regions worldwide were assigned to different groups (Camargos *et al.*, 2007; Camargos *et al.*, 2002; Felmer *et al.*, 2005; Licursi *et al.*, 2003; Mamoun *et al.*, 1990; Monti *et al.*, 2005a; Zhao & Buehring, 2007). Rodriguez et al (2009) combined his data with previously studied strains, which represented

different clusters or groupes by RFLP and gp51, and demonstrated at the first time the existence of seven BLV genotypes, gentoype-1 to -7 (Rodriguez *et al.*, 2009). Soon BLV genotype-8 were observed in a variety of geographical locations (Balic *et al.*, 2012; Lee *et al.*, 2015; Polat *et al.*, 2015; Rola-Luszczak *et al.*, 2013). Fortunately, more recent study found the presence of novel genotype based on phylogenetic analysis of BLV complete genome sequences (Polat *et al.*, 2016), and concluded that at least nine different genotypes, genotype-1 to -9, of BLV strains are circulating in the world (Ochirkhuu *et al.*, 2016; Polat *et al.*, 2016).

Myanmar is the largest country on the main land Southeast Asia, agriculture is the main economic sources in this country. Among all livestock in Myanmar, cattle occupies main portion of livestock population. Beef cattle farming are non-existing. Since Myanmar is agro-based country, draught cattle development is accorded the highest priority in livestock program. Dairy cattle plays a secondary role in ruminant production with some reasons such as demand for milk and milk products for urban population, improvement in the level of human nutrition and quality of the life of the rural population (National Consultative Committee, 2013). Until now, a variety of BLV genotypes have been detected in different parts of Asia, such as genotype-1, -2, -3 in Japan (Licursi *et al.*, 2003; Matsumura *et al.*, 2011), genotype-1, -6 in the Philippines (Polat *et al.*, 2015), genotype-1, -3 in Korea (Lee *et al.*, 2015), and genotype-1, -4, -7 in Mongolia (Ochirkhuu *et al.*, 2016). However, no information is available about the prevalence of BLV infection or the existence of EBL in this region (OIE, 2009).

Here, I investigated the distribution and molecular characteristics of BLV strains in Myanmar with other BLV strains worldwide as follows: (i) first, the spread of BLV infection was investigated in 66 cattle samples in Myanmar and five samples were confirmed as BLV positive by amplification of BLV LTRs through nested PCR and CoCoMo-qPCR, Takara real-time PCR targeting *tax* gene, and amplifying and sequence of *gp51* partial region; (ii) second, phylogenetic analysis of *env* (gp51) sequences of Myanmar BLV strains showed that Myanmar BLV strains were classified into a novel genotype, genotype-10; (iii) third, full genome sequence of two Myanmar BLV strains were obtained by cloning through pGEM-T easy vector, and the genetic variability and genotyping of these Myanmar strains were analyzed. Then, the existence of genotype-10 was confirmed by full-genome phylogenetic analysis. This is the first study to confirm genotype-10 by full genome sequences, and also to determine the prevalence and genotype classification of BLV in Myanmar.

3-2. Materials and Methods

3-2-1 Experimental samples

Blood samples were taken from 66 cattle from 23 farms in four different townships of Naypyidaw Union Territory of central Myanmar, namely Zeyathiri Township, Takkone Township, Pyinmana Township and Lewe Township. Buffy coats were separated from cattle whole blood and used for DNA extraction.

3-2-2 Animal handling and research Ethics

All animals were handled by veterinarians from the University of Veterinary Science in Yezin Myanmar in strict accordance with good animal practice following the University guidelines.

3-2-3 Genomic DNA extraction methods

Genomic DNA was extracted from the buffy coat separated from cattle whole blood by the standard phenol-chloroform method. Briefly: (1) Wash the buffy coat with 1.5 ml of 0.2 % NaCl, and centrifuge at 2200 rpm for 10 minutes at room temperature (about 25°C) (repeat this step 1 more time); (2) throw the supernatant and keep the precipitate; (3) add 0.5 ml of 2×SDS and 10 μl of Proteinase K to the precipitate, then incubate the mixture at 55 °C overnight; (4) Add 1/20 volume of 3M Na Acetate (pH 5.2) and same volume Phenol/Chloroform, and mix 5-10 min, then centrifuge at 2200 rpm for 10 minutes at 4 °C; (5) transfer the supernatant into new tube and add same volume of Phenol/Chloroform, then centrifuge at 2200 rpm for 10 min at room temperature (about 25 °C); (6) transfer aqueous layer (supernatant) into new tube and add 2 volumes of 100 % ethanol and mix, then centrifuge at 12000 rpm for 20 min at 4 °C, and discard the 100 % ethanol; (7) add 70 % of ethanol to wash the pellet, then centrifuge at 12000 rpm for 5 min at 4 °C; (8) discard the 70 % of ethanol carefully, and dry the DNA pellet for 10 min in air; (9) elute DNA by adding 50-100 μl of TE buffer. Finally, the eluted DNA was stored at -20 °C until experimental usage.

3-2-4 Detection of BLV provirus

3-2-4-1 Detection of BLV provirus by nested PCR

Each sample was screened for the presence of BLV provirus by the detection of BLV LTR using nested PCR. Detailed description of BLV detection by LTR nested PCR is provided in chapter 1 method section (BLV detection).

3-2-4-2 Confirmation of BLV positive result

Since some of the samples showed inconsistent result when screened by nested PCR targeting LTR region, I applied more sensitive and specific method for confirmation of BLV prevalence in those samples. The methods include CoCoMo-qPCR-2 targeting LTR, Takara BLV detection Probe/Primer/Positive control kit (real time-PCR BLV detection kit) (Takara Bio Inc., Otsu, Japan) targeting BLV tax *gene*, and nested PCR targeting BLV *env-gp51* region.

CoCoMo-qPCR-2 is highly sensitive and very specific in which two copies of BLV LTR were amplified and detected by using degenerated primers and probes as described in method section of chapter 1 of this thesis. Takara BLV detection Probe/Primer/Positive control kit was used to amplify BLV *tax* gene according to manufacturer's instructions.

Another method used for BLV prevalence confirmation is nested PCR in which BLV *env-gp51* partial region was amplified by using outer and inner reaction primers. Briefly, the first PCR amplification was performed using the primer pairs of external forward (5`-ATGCCYAAAGAACGACGG-3`) and external reverse (5`-CGACGGGACTAGGTCTGACCC-3`) primers (Moratorio *et al.*, 2010). The first PCR reaction contains 5 µl of 5 x PrimeSTAR GXL Buffer, 2 µl of 2.5 mM dNTP mixture, 0.5 µl of PrimeSTAR GXL polymerase (Takara Bio Inc., Otsu, Japan) and 13.5 µl of distilled water (10 lM). Nested PCR was amplified using 1 µl of first PCR amplicons by primer pairs of Env5032 (5`-TCTGTGCCAAGTCTCCCAGATA-3`) and Env5608r (5`-AACAACAACCTCTGGGAAGGGT-3`) (Asfaw *et al.*, 2005). Both the first and nested PCR reaction were running under the thermal condition of 30 cycles of 98 °C for 15 sec, 60 °C for 20 sec and 68 °C for 60 sec. The external primers produce amplification of a 913-bp DNA fragment, and internal primers result in a 597-bp fragment of the gp51 partial region of the *env* gene. Nested PCR results were visualized by agar gel to determine the prevalence of the BLV if the sample showed target band corresponding to the amplified *gp51* fragment. In each

method, cattle samples detected as positive by LTR nested PCR were also used as positive control together with the positive control or plasmid supported by each kit.

3-2-5 PCR amplification and sequencing of BLV *env-gp51* gene

3-2-5-1 PCR amplification and sequencing of BLV *env-gp51* partial fragments

Samples detected as BLV positive were further used for amplification of BLV *env* partial gene by nested PCR targeting gp51 partial region. The PCR amplification of partial gp51 region was described as the above paragraph and as the method section of chapter 1 of this thesis. Positive second-round PCR products were purified using Exo-SAP IT (USB Corp., Cleveland, OH) and sequenced on an ABI3730xl DNA Analyzer using an ABI PRISM Big Dye Terminator v 3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). 475-bp sequence (without primers) of the *env* gene, corresponding to nucleotide positions 5090 to 5564 of the BLV cell line FLK-BLV subclone pBLV913 complete genome (GenBank accession number EF600696) (Rovnak *et al.*, 1993), was obtained. Editing, alignment, and identification of nucleotide sequences were performed using MEGA 5.1 software (Tamura *et al.*, 2011).

3-2-5-2 PCR amplification, cloning and sequencing of BLV *env-gp51* gene full region

Full-length sequences of BLV *env-gp51* gene from four Myanmar BLV strains were amplified for phylogenetic analysis by PCR. PCR amplification was performed by using KOD FX Neo (TOYOBO CO., LTD. Tokyo, Japan) and the following primer pairs: Forward (pBLV-F0: 5'-AGATGGGAGCTACACCATTCA-3') and Reverse (pBLV-0R: 5'-GTCTGTAGAGACTCTTTGCGAG-3'). The reaction mixture contained 12.5 µl of 2 × PCR Buffer for KOD FX Neo, 5 µl of 2 mM dNTP mix, 0.5 µl of KOD FX Neo Polymerase (TOYOBO CO., LTD. Tokyo, Japan), 1 µl of each primer (10 µM), and added 2 µl of distilled water up to final reaction volume of 25 µl. Conditions for PCR amplification were as follows: 94°C for 2 min, followed by 30 cycles of denaturation at 98 °C for 10 sec, annealing at 60 °C for 30 sec, and extension at 68 °C for 45 sec. The successful amplification resulted in production of 1465-bp DNA fragment. PCR amplicons were purified by Centri-Sep Columns (Princeton Separations, Inc. USA) according to manufacturer instructions.

Because of the blunt-end termini produced by KOD FX New, purified *env-gp51* full length PCR products were treated with *rTaq* polymerase (TOYOBO CO., LTD. Tokyo, Japan) to add a single 3`-A overhang at both ends at 70 °C for 15 min and purified again by Centri-Sep Columns (Princeton Separations, Inc. USA) according to manufacturer. *rTaq* treated-purified PCR products were ligated to the pGEM-T easy vector with T4 ligase for cloning (Promega Corporation, USA), according to the manufacturer instructions. 10 µl of the ligation reaction was transformed into *E.coli* XL10-Gold Ultracompetent cells (Agilent Technologies Co., Ltd. US) and then incubated at 37 °C 150 rpm for 90 min. Then, the transformed *E.coli* were spread on agar plates containing 100 mg/ml ampicillin, 100 mM IPTG and 20 mg/ml X-Gal. At least 20 white colonies were selected to obtain positive clones. For plasmid DNA extraction, FastGene plasmid mini kit (Nippon Genetics Co., Ltd. Tokyo, Japan) was used according to manufacturer instructions. Each isolated plasmid was further confirmed by SalI and ApaI restriction enzyme (TOYOBO CO., LTD. Tokyo, Japan) digestion at 37 °C to check the correct insertion size. Three positive plasmids were confirmed by sequencing on an ABI3730xl DNA Analyzer using an ABI PRISM Big Dye Terminator v 3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and sequencing primers as shown in Table 11. 1230 bp sequences (including sequences of leader peptide, full-length gp51 and N-terminal region of gp30) corresponding to nucleotide positions 4826 to 6054 of the BLV cell line FLK-BLV subclone pBLV913 complete genome (GenBank accession number EF600696) (Rovnak *et al.*, 1993) were obtained, and deposited in the DDBJ database under accession numbers LC154064~LC154067. 807-bp gp51 full-length sequences, corresponding to nucleotide positions 4922 to 5728 of the BLV cell line FLK-BLV subclone pBLV913 complete genome, were used for phylogenetic analysis of full-length gp51.

3-2-6 PCR amplification of the BLV provirus complete genome

To gain better insight into the genetic differences between the Myanmar BLV strains and other available strains, the complete BLV genome of two Myanmar strains were achieved by PCR amplification of overlapping genomic fragments (Figure 20) through using the PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Otsu, Japan) and specific primers designed for this study (Life Technologies Japan Ltd, Tokyo, Japan) (Figure 20). The 25 µl final reaction mixture contained 5 µl of 5× PrimerSTAR GXL Buffer, 2 µl of 2.5 mM dNTP mix, 1 µl of each primer at

a concentration of 10 pmol, 2 µl of template (30 ng/ µl) and 0.5 µl of PrimerSTAR GXL DNA Polymerase. The cycles for the PCR amplification were as follows: 98 °C for 2 min, followed by 33 cycles of denaturation at 98 °C for 15 sec, annealing at 60 °C for 15 sec, and extension at 68 °C for 1 min/kb (1 min per kilobase). The successful full genome PCR products were purified by FastGene Gel/PCR extraction kit (Nippon Genetics Co., Ltd. Tokyo, Japan) according to manufacturer.

3-2-7 Cloning of the PCR products for sequencing of BLV full genome

Because of the blunt-end termini produced by PrimerSTAR GXL DNA, purified BLV full genome PCR products were treated with *rTaq* polymerase (TOYOBO CO., LTD. Tokyo, Japan) to add a single 3`-A overhang at both ends at 70 °C for 15 min and purified again by FastGene Gel/PCR extraction kit (Nippon Genetics Co., Ltd. Tokyo, Japan) according to manufacturer. *rTaq* treated-purified PCR products were ligated into the pGEM-T easy vector with T4 ligase for cloning (Promega Corporation, USA), according to the manufacturer instructions. 10 µl of the ligation reaction was transformed into *E.coli* XL10-Gold Ultracompetent cells (Agilent Technologies Co., Ltd. US) and then incubated at 37 °C 150 rpm for 90 min. Then, the transformed *E.coli* were spread on agar plates containing 100 mg/ml ampicillin, 100 mM IPTG and 20 mg/ml X-Gal. At least 20 white colonies were selected to obtain positive clones. For plasmid DNA extraction, FastGene plasmid mini kit (Nippon Genetics Co., Ltd. Tokyo, Japan) was used according to manufacturer instructions. Each isolated plasmid was further confirmed by SalI and ApaI restriction enzyme (TOYOBO CO., LTD. Tokyo, Japan) digestion at 37 °C to check the correct insertion size. BLV full genome sequences were obtained through using the primers listed in Table 12. Three positive plasmids were confirmed by sequencing on an ABI3730xl DNA Analyzer using an ABI PRISM Big Dye Terminator v 3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Myanmar BLV strain complete genome sequences obtained in this study were deposited in the DDBJ database under accession numbers LC154848-LC154849.

3-3. Results

3-3-1 Investigation of the spread of BLV infection in Myanmar

To investigate the spread of BLV infection in Myanmar, a total of 66 blood samples were collected from cattle in different farms of four regions of Naypyidaw Union Territory of central Myanmar. The samples were screened for BLV infection by different methods. Since the limitations of specificity and sensitivity of each methods used, inconsistent BLV detection results were obtained. Samples which showed BLV positive result by at least two methods were confirmed as BLV positive in the final result (Table 13). Five samples out of 66 screened were detected as BLV positive. As far as each township region is compared, Zeyathiri Township was negative for the BLV provirus. In Tat Kone Township, two cattle samples out of four were BLV positive. Of 20 samples collected from different farms of Pyinmana Township of Myanmar, only one sample (5.0 %) was positive for BLV provirus. Two out of 35 (5.7 %) samples collected from Lewe Township were positive for the BLV provirus.

3-3-2 Phylogenetic analysis of BLV strains based on *env*-gp51 partial fragment and full region

Since *env* gp51 partial region was commonly used in previous BLV genotyping studies (Balic *et al.*, 2012; Beier *et al.*, 2001; Camargos *et al.*, 2007; Camargos *et al.*, 2002; Coulston *et al.*, 1990; Felmer *et al.*, 2005; Lee *et al.*, 2015; Matsumura *et al.*, 2011; Monti *et al.*, 2005a; Polat *et al.*, 2015; Rola-Luszczak *et al.*, 2013), five BLV positive samples were used for phylogenetic analysis of *env* gene partial fragment and *gp51* full region. After direct sequencing, 475-bp *env* partial nucleotide sequences of these five field strains, corresponding to nucleotide positions 5090 to 5564 of the full-length BLV genome (BLV cell line FLK-BLV strain pBLV913, GenBank accession number EF600696) (Rovnak *et al.*, 1993), were aligned with 94 references sequences corresponding to known BLV genotypes. Phylogenetic tree were then constructed by ML approaches using the K2+G (Figure 21). The ML tree showed congruent topologies, supported by moderate to high bootstrap values. As shown, current study results were similar to those of previous studies where BLV strains were divided into eight genotypes (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013) and new genotype 9 (Polat *et al.*, 2016). Unexpectedly, the five Myanmar BLV strains were group close

to each other, and were assigned to a clade which shares common node with other genotype-6 BLV strains Figure 21.

To classify the Myanmar BLV strains clearly and also to observe whether there is any difference in phylogenetic analysis between partial and full length of *gp51* region, I constructed ML phylogenetic tree based on 807-bp full length sequence of Myanmar strains and 88 references strains, corresponding to nucleotide positions 4922 to 5728 of the full-length BLV genome (BLV cell line FLK-BLV strain pBLV913, GenBank accession number EF600696) (Rovnak *et al.*, 1993). Even though full-length gp51 tree clearly classified the previously known eight genotypes and newly identified novel genotype 9, which is similar to gp51 partial ML tree, Myanmar BLV strains were clearly separated from strains of genotype-6 and other strains, forming a distinctive cluster supported with bootstrap value of 99, indicating that Myanmar BLV strains might be a novel genotype termed as genotype-10 (G-10), but not a subgroup of genotype-6 as shown in Figure 22. To sum up, current study showed the difference in outcomes of phylogenetic analysis between partial and full-length gp51 sequences (Figure 21 & 22) that phylogenetic analysis based on full-length gp51 is needed in genotyping of BLV strains.

3-3-3 Nucleotide sequence analysis of partial *env* gp51 region of BLV strains in Myanmar

In order to compare the results of partial gp51 phylogenetic analyses with similarity relationships derived from sequences comparisons, I compared the possible pair of strains. The nucleotide sequence similarity of the 475-bp BLV *env* gene sequence ranged from 98.1 % to 100 % for the five Myanmar BLV strains, while these sequences were 94.1 % to 96.9 % similar to those corresponding to all known nine genotypes. As far as the nucleotide difference between Myanmar strains and representative references sequences within genotype-6 are concerned, strains from Myanmar genotype-10 and genotype-6 showed some genotype-specific nucleotide substitutions while having common nucleotide changes as shown in Figure 23A. Myanmar strains showed the highest nucleotide sequence similarity to Paraguay and Bolivian strains with homology of 97.3 % (Paraguay-39/Paraguay; Bolivia Ya-47/Bolivia), while they showed less identity of 94.01 % to Paraguay strains (Paraguay-96/Paraguay) (Figure 23A). Interestingly, eight distinctive nucleotide substitutions were detected only in all of the Myanmar strains, but not other strains grouped into genotype-6. Among them, four nucleotide substitutions at nt 282 (A->G), 561 (G->A), 564 (C->T) and nt 618 (C->T) were silent substitutions, while the rest four

nucleotide substitutions at nt 317 (T-> C), nt 410 (C->T), nt 428 (A->G) and nt 529 (C->T) lead to amino acid changes. Thus, nucleotide alignment showed the difference of Myanmar BLV strains from other genotype-6 strains.

3-3-4 Amino acid sequence analysis of partial gp51 Env protein of BLV strains in Myanmar

To get insight into amino acid changes observed in Myanmar BLV strains, and the difference with other genotype-6 known BLV strains, five partial gp51 *env* sequences were aligned with other representative sequences of genotype-6 strains and translated into amino acid sequence. Figure 23B shows the distribution of amino acid changes within the middle region of gp51, encompassing amino acid positions 89 to 246 (corresponding to FLK-BLV strains; GenBank accession number EF600696). Strains in genotype-6 showed strict conservation, and shared the common genotype-6 specific amino acid substitutions from Isoleucine to Threonine at residue 144 (Figure 23B). As expected, Myanmar strains showed different pattern. Although a comparison of the predicted amino acid sequences of partial gp51 sequences of Myanmar strains showed high homology to that of FLK-BLV (accession nom: EF600696) and genotype-6 strains, Myanmar strains showed a variety of unique amino acid substitutions, which were not detected in other strains as in Figure 23B. One genotype-6 specific amino acid substitution at residue 144 (I->T) which is common to all genotype-6 BLV strains worldwide is also detected in the 2^{nd} ND of Myanmar BLV strains. Besides that, four unique amino acid mutations were observed only in Myanmar strains, which clearly separate Myanmar strains from genotype-6 strains as a novel genotype. The four substitutions specific to Myanmar strains were all located in functional domains. For example, amino acid substitution of valine replaced by alanine at residue 106 (V->A) which is the common amino acid residue of both 1^{st} ND and $CD4^{+}$ epitope region. Two amino acid substitutions of serine to phenylalaine at residue 137 (S->F) and of glutamine to arginine at residue 143 (Q->R) were located at 2^{nd} ND. The last common amino acid substitutions from proline to serine at residue 177 (P->S) were located in $CD8^{+}$ epitope and E epitope region. In addition to the five common mutations, other three single mutations were also observed in Myanmar strains. One single substitution was detected at residue 121 which is corresponding to conformation epitope G-epitope. Glutamine to arginine mutation was detected at residue 181 (Q->R) in E-epitope region. The third single substitution from glutamic acid

replaced by aspartic acid was located at residue of 225 in 3rd ND. The impact of amino acid substitutions detected in Myanmar strains on virus functions need to be studied further. Thus, the distinctive nucleotide and amino acid mutation pattern of partial gp51 region might also indicate that Myanmar BLV strain is a novel genotype-10, but not a subgroup of genotype-6.

3-3-5 Maximum-likelihood phylogenetic tree based on BLV full genome sequences

To get full insight into the complete genome sequence diversity and into comprehensive phylogenetic characteristic of Myanmar BLV strains, complete BLV genome sequences of two Myanmar strains were obtained through clone-sequencing. Full genome sequences of Myanmar strains showed highest similarity of 96.06 % to genotype-6 BLV strains (Paraguay-89), while having less homology of 94.66 % to genotype-2 BLV strains (Arg38). Full genome ML phylogenetic tree was constructed based on two sequences in this study and 27 previously reported BLV complete genome sequences (Figure 24). The tree is strongly in consistent with previous clustering result of BLV genotypes (Polat *et al.*, 2016), including genotypes-1, -2, -4, -6 and -9 as shown in Figure 24. Interestingly, two Myanmar strains were not clustered within neither genotype-6, nor other known genotypes, but separately located in a different branch (Bootstrap values 100 % for every clade), confirming that Myanmar strains is a new genotype, termed as genotype-10 (G-10). This result is consistent with topology of full-length gp51 phylogenetic analysis.

3-3-6 Amino acid comparison of each BLV gene among Myanmar strains and other known strains

To further determine how much Myanmar BLV full genome sequences were distinguished from previously known BLV strains in database, two Myanmar provirus's full genome were aligned with 27 known complete genome sequences, and deduced amino acid sequences of each gene region were obtained by MEGA 5.1 (Tamura *et al.*, 2011). As summarized in Figure 25, 22 unique amino acid substitutions were observed in Myanmar strains L1 and S3 as follows: (1) In the structural gene-encoded proteins: two substitutions, T38A and T366A, in the Gag; one substitution, S52F, in the Pro; four substitutions, V205L, I409V, P480S and A826V, in the Pol; and four substitutions, V106A, S137F, Q143R and P177S, in the Env, restricted to only

sequences of Myanmar strains. (2) In the regulatory proteins: six substitutions, N140K, V142E, I152T, D181N, E229D and L273F, in the Tax; three substitutions, S103F, L140P and T156N, in the Rex; one substitution of K27N in R3; one substitution, L66P, in the G4 protein, were observed only in Myanmar strains. It can be seen from the Figure 25 that each genotype showed genotype-specific mutations which distinguish the genotype from others. The amino acid comparison clearly showed that Myanmar BLV strains were significantly different from genotype-6 BLV strains and other genotypes, further supporting the result of full genome ML tree. Thus, Myanmar BLV strains are a novel genotype, namely G-10.

3-4 Discussion

I draw three major conclusions from the results of this study of BLV in Myanmar cattle. First, the current study showed the distribution of BLV in the Myanmar. A total of 66 samples collected from Myanmar revealed relatively low level of BLV infection. BLV was detected in three out of the four townships, indicating the prevalence of BLV in Myanmar. Second and most important, phylogenetic analysis based on complete genome and amino acid comparative analysis strongly support the possibility that Myanmar BLV strains were clustered into a novel genotype, genotype-10 (G-10). This result is in consistent with phylogenetic analysis of full-length gp51. Thus the detection of genotype-10 is the novel finding of this study. Third, a number of nucleotide and amino acid substitutions were found both in full genome and *gp51* region of Myanmar strains. Most of the nucleotide substitutions were silent substitutions. As far as the amino acid substitutions in *env* partial region were concerned, all amino acid substitutions were located in epitope regions, such as the $CD4^+$ T-cell epitope, the 2^{nd} ND, the $CD8^+$ T-cell epitope, G, B and E epitopes.

A number of transmission pathways and associated risk factors of BLV have been identified and include both horizontal modes, such as physical contact (Kono *et al.*, 1983), exposure to biological fluid contaminated with infected lymphocytes, blood sucking insects (Bech-Nielsen *et al.*, 1978; Ohshima *et al.*, 1981), blood contaminated dehorning devices (DiGiacomo *et al.*, 1985; Lassauzet *et al.*, 1990), introduction of infected animals (Kobayashi *et al.*, 2014), iatrogenic factor including the common use of needles and plastic sleeves for rectal palpation (Hopkins *et al.*, 1988; Lassauzet *et al.*, 1990), and vertical routes through utero infection (Van der Maaten *et al.*, 1981). As far as the presence of BLV in Myanmar cattle is concerned, multiple potential risk factors have associated with the BLV positive result in the current study. Firstly, breeding strategy, to develop the draught, dairy cattle and also to upgrade the local breeds, cattle from most of the townships in Myanmar undertake insemination services using frozen semen exported from New Zealand and Australia, or crossed with many exotic breeds such as Holstein Frisian, Jersey, Shindi and Thari breeds (National Consultative Committee, 2013). It has been demonstrated that the introduction of semen from bulls did not cause BLV transmission (Monke, 1986), excluding the potential transmission of BLV by the artificial insemination in Myanmar. Animal introduction, however, is highly associated with BLV infection in Japan (Kobayashi *et al.*, 2014) and other countries. Therefore, the introduction of infected cattle to Myanmar and the

transfer of infected cattle from farm to farm within Myanmar, and also hybridizing or crossing of local breeds with exported breeds, might be one of the main reasons on BLV infection in Myanmar (National Consultative Committee, 2013). In addition to the above, embryo transfer technology had also practiced in Myanmar (National Consultative Committee, 2013), which might also cause BLV transmission through exposure to biological fluid contaminated with infected lymphocytes such as blood. Furthermore, there are special bulls kept for breeding purpose in some of the areas of Myanmar which used to breed with cows for selective breeding under payment (National Consultative Committee, 2013), indicating the potential routes of physical contact through breeding. Even though, 7.6 % (5/66) of total samples were detected as BLV positive, I can hypothesize from the risk factors associated with the presence of BLV in Myanmar that BLV infection might be wide spread in Myanmar, not only in the tested area, but also other Townships and territory of Myanmar.

The most interesting data in this study is that phylogenetic analysis of Myanmar BLV strains concurs with previously established clusters as well as discovering a novel genotype, namely G-10. The phylogenetic results of novel BLV genotype (G-10) in Myanmar confirmed by ML phylogenetic tree using two novel and 27 previously reported BLV whole genome sequences (Figure 24), which was further confirmed by comparative amino acid analysis of structural and regulatory proteins (Figure 25). Notably, the bootstrap value of the Myanmar novel genotype cluster (G-10) was supported by a value of 100 in the ML phylogenetic tree constructed from complete genome sequences. The genotype-10 Myanmar BLV strains have a close genetic relationship among themselves and a more comparatively distant relationship with other previously reported genotypes including genotype 6.

A study by Moratorio et al., (2010) (Moratorio *et al.*, 2010) reported the suitability of partial gp51 sequences for the assignment and establishment of phylogenetic relationship among BLV strains through comparing phylogenetic analysis obtained using partial or full-length *env-gp51* sequences. Current study result, however, showed difference in the outcomes of partial and full-length gp51 phylogenetic analysis. phylogenetic analyses based on partial gp51 *env* gene sequences supported previously established data indicating that BLV has nine genotypes (Balic *et al.*, 2012; Matsumura *et al.*, 2011; Polat *et al.*, 2016; Rodriguez *et al.*, 2009; Rola-Luszczak *et al.*, 2013). Unfortunately, gp51 partial sequences ML tree implies that Myanmar BLV strains might be a subgroup of genotypes-6, which is contrary to ML phylogenetic analysis based on

full-length gp51. This discrepancy of phylogenetic analysis might be caused by two different set of sequences (full-length and partial gp51) for the two datasets. Phylogenetic analysis of *env*-gp51 pairtial region uses very short sequences (475bp) of the BLV genome which has limited genetic variations. As variation among viruses of the same genus is not expected to differ much, I hypothesize that phylogenetic analysis based on partial *gp51* gene might be not sensitive enough to differ two distinctive subgroups or genotypes in the case of sharing a number of nucleotide substitutions. This has been proved by Polat M *et al*,. (Polat *et al.*, 2016) that two different sub-strains of genotype 9 circulating in Bolivia were observed by full genome phylogenetic analysis, but not by *gp51* partial sequence-based analysis. However, both alignments of nucleotide and amino acid sequences of *env* partial gene revealed the existence of novel genotype 10 in Myanmar. Therefore, phylogenetic trees based on *env* partial region might not be sufficient to determine new genotypes. It is necessary to use full-length gp51 sequences for phylogenetic genotyping studies, and nucleotide and amino acid sequences variation and diversity comparison of comple BLV genome sequences are supposed to be taken into consideration.

Lee *et al*., suggested that BLV gp51 sequences analysis of different geographical isolates showed specific sequences conservation depending on genotypes (Lee *et al.*, 2015). My study supports Lee et al., (Lee *et al.*, 2015) statement that even though nucleotide sequence showed higher variety (Figure 23A), deduced amino acid sequences conserved strictly based on genotypes (Figure 23B). Sequence alignment of partial *env* gp51 region revealed eight nucleotide substitutions and four amino acid substitutions unique for Myanmar strains (Figure 23). All amino acid substitutions were located in functional epitopes or neutralizing domains (Figure 23B), in agreement with previous molecular evolutionary analysis (Mamoun *et al.*, 1990; Portetelle *et al.*, 1989; Rodriguez *et al.*, 2009; Zhao & Buehring, 2007). Myanmar BLV strains had genotype-6 specific amino acid substitution at residue 144 (Q->T) of 2^{nd} ND which is common to all genotype-6 BLV strains. Besides that, I found substitutions at residue 137 (S->F) and at 143 (H->R) in the 2^{nd} ND only of Myanmar BLV strains. It had been reported that 2^{nd} ND could be involved in the interaction of gp51 with the receptor expressed on host cell members, suggesting that this region could have an impact on viral fusion and infectivity in vivo (Callebaut *et al.*, 1993; Gatot *et al.*, 2002). In addition to the above, one mutation at residue 106 (V->A) was located in 1^{st} ND and $CD4^{+}$ epitope, and the last one at residue 177 (P->S) was located on $CD8^{+}$

and E epitope region. It is interesting that two Myanmar strains (S3 and S5/Myanmar) have a substitution at residue 121 (R->H) of G-epitope which was previously described in genotype-4 BLV strains of Eastern European (Rola-Luszczak *et al.*, 2013) and of Chile (Moratorio *et al.*, 2010; Rodriguez *et al.*, 2009). This epitope was known as one of the conformational epitope, target for MAbs that cause virus neutralization and syncytium inhibition (Bruck *et al.*, 1982b; Bruck *et al.*, 1984). In addition, two amino acid out of 22 detected in comparison of amino acid sequences of each gene were also located in functional domains as the following: substitution of aspartic acid by asparagine at residue 181(D->N) is located in the leucin-reach domain (Willems *et al.*, 1991) of regulatory protein Tax; substitution of leucine by proline at residue 66 (L->P) of accessory protein G4 is located in arginine-rich nucleus (Alexandersen *et al.*, 1993) targeting RNA-binding region (ARR).

As previously explained, Holstei Frisian, Jersey, Shindi and Thari breeds were imported to Myanmar for upgrading the local breeds (National Consultative Committee, 2013). I hypothesize that these exotic breeds might previously infected by BLV outside of Myanmar and introduced this virus to Myanmar cattle. This is proved by this study result that I detected BLV in Friesian (sample ID: S3, S5, L1 and L2) and Pyarsein (p60) breeds. However, the exact geographical origin of the Myanmar BLV genotype-10 strains remains uncertain which requires further study to identify the diversity of BLV circulating in Myanmar and also the BLV genotypes circulating in the neighboring countries, including Thailand, Laos, Bangladesh, Bhutan, Eastern India, and in Southern China. Since the tested samples in this study were only collected from central Myanmar and also the limited number of samples collected, I only detected the novel genotype-10 in Myanmar. However, more genotypes of BLV strains might be observed through nationwide broad range of sampling.

3-5. Conclusions

The present study provides the first evidence of the prevalence of BLV infection among cattle in Myanmar. Current study phylogenetic analysis demonstrated that at least a new BLV genotype (G-10) is present in Myanmar. These results provide important information on BLV infection levels and will enable the implementation of appropriate cattle-management policies in addition to providing supplementary information for the development of more-effective methods of BLV eradication in Myanmar.

General Discussion

Even though previous studies showed the existence of eight different BLV genotypes, the prevalence of BLV and its phylogenetic analysis in many parts of the world still obscure. The objective of this study is to investigate the spread of BLV infection in South America and Asia, and to evaluate the molecular genetic variability of BLV strains circulating in study area by discussing the characterization and phylogenetic analysis of *env-gp51* gene and BLV full genome sequences.

First of all, I examined the BLV prevalence in five countries in South America, named Peru, Paraguay, Bolivia, Argentina and Chile, and two Asian countries such as the Philippines and Myanmar. The prevalence of BLV was observed in each of the countries where samples were collected. South America showed medium to high level of BLV infection. Comparing to previous studies, current study result revealed a tendency of increasing level of BLV infection in all of the South American countries. By contrast, in Asia, low level of BLV prevalence was observed in the Philippines and Myanmar. A variety of risk factors were associated with the spread of BLV in South America and Asia. The import of BLV infected cattle to a country free from disease and also introduction of infected cattle from farm to farm contribute significantly to widespread of BLV prevalence. I observed cattle introduction in all of the studied area. Animal breeding strategies might also cause viral distribution. Most of the studied area, to increase milk and beef productivity, cattle were crossed with exotic breeds which might increase the possibility of infection through physical contact. Cattle management system, such as free housing, and iatrogenic pathways, namely common use of blood containing device (for dehorning and vaccination) also associated with BLV prevalence. Interestingly, I detected the transmission of viral infection through milk during the milking process only in the Philippines but not other countries.

Phylogenetic analysis based on gp51 sequences showed that two different BLV genotypes, G-1 and -6, are present in the Philippines. Genotyping analysis of South America revealed that Peruvian and Paraguayan BLV strains are of G-1, -2, and -6. Of interest, current findings indicate that Bolivian BLV strains are clustered into G-1, -2, -6, and a novel G-9. Surprisingly, Myanmar BLV strains were neither classified into any of eight known genotypes nor novel G-9, but separately grouped into another novel genotype, namely G-10. The BLV classification were

further confirmed by nucleotide sequences similarity, substitution patterns of nucleotide and amino acid sequences of *env-gp51* via comparing BLV strains isolated in this study with previously known BLV strains. A variety of BLV genotypes detected in each country can be explained by cattle introduction, modern cattle trading, and geographical affinity with other neighboring countries. Philippine BLV strains showed similarity to strains from Japan, Brazil and Argentine which might indicate cattle introduction from same origin among these countries. Previous studies indicate the presence of BLV genotype-1, -2 and -6 in Brazil and Argentina to where Peru, Paraguay and Bolivia were geographically close. The classification of the same BLV genotypes (G-1, -2 and -6) in South American countries might be the consequences of virus spread throughout animal exchange and also the historical background of human colonization by which viral host, cattle, were distributed. The presence of two or more than two genotypes circulating together in a certain geographical region can be the outcomes of separate infection of different viral origins and further transmission of the virus through cattle exchange.

The most important finding of current study is the confirmation of novel genotypes G-9 in South America and G-10 in Asia. Novel genotype-9 was only detected in Bolivia, but neither other South American countries in this study, including Peru, Paraguay, Bolivia, nor previously BLV genotypes determined countries, namely Argentine, Brazil, Chile, Uruguay. As far as the breeds were concerned, genotype-9 was detected in Holstein, Nelore and Yacumeño Creole cattle in Bolivia, and it might be very reasonable that Transmission of BLV genotype-9 was occurred among the above mentioned cattle breeds through crossbreeding or other unknown routes. Even though it is too early to decide the origin of BLV genotype-9 only based on current study, it is clear that this genotype is limited in Bolivia, and not spread out to other neighboring countries. Since phylogenetic analysis of the BLV strains in most of Asian countries are unknown, except from Japan, the Philippines, Korea, Mongolia and Turkey, the origin of genotype-10 in Myanmar remains to be still unknown. Most cases, the viral spread has been linked with trade of cattle or import of animals for breeding, determining novel genotype-10 requires a comprehensive study which involves broad range of sampling within Myanmar and in neighboring countries especially samples from Thailand, Laos, Bangladesh, Bhutan, Eastern India, and in Southern China.

Even though BLV worldwide were classified into a variety of genotypes, different genotypes neither associated with unique serotypes, nor lead to specific clinical manifestations (Rodriguez

et al., 2009). The biological significance of these novel genotypes, G-9 and -10, are not clear. However, unique amino acid mutation, especially those found in functional domains, is supposed to have potential effects on fusion and infectivity of BLV. For example, genotype-9 specific amino acid substitution A133V in the 2nd ND, genotype-10 specific mutation V106A in 1st ND, S137F and Q143R in the 2nd ND, and P177S in linear E-epitope in SU gp51 protein might affect virus infectivity, syncytium formation and immune responses. Another one mutation in Myanmar strains, D181N, is locate in leucine-rich domain of regulatory protein Tax which might involved in heterologous protein interactions (Willems *et al.*, 1991).

Studying the prevalence and genetic diversity of a virus could help to estimate the infection level in a certain area, and to gain better understanding of the correlation between variation in genotype and difference in infectivity and possible impact of viral variability on its functions. Current study provides important information on BLV infection levels and will enable the implementation of appropriate cattle-management policies in addition to providing supplementary information for the development of more-effective methods of BLV eradication in South America and Asia. Sequences variability of different genotypes might be of importance in vaccine development, especially those functional domains conserved in all genotypes.

AKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor Professor Yoko Aida of Laboratory of Viral Infectious Diseases, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, and Viral Infectious Diseases Unit, RIKEN, for the continuous support of my Ph.D study and related research, for her patience, motivation, valuable advice, discussions, encouragement, and for her immense knowledge throughout my study. I was touched by my supervisor's attitude of decisiveness and persistence to her work, and got encouragement from her self-discipline and hard working with passion.

I am thankful to Dr. Shin-nosuke Takeshima of Laboratory of Viral Infectious Diseases, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, and Viral Infectious Diseases Unit, RIKEN, for his technical support, suggestions and valuable contributions to my graduate study, and for being a member of defense committee. I appreciate Prof. Kaoru Uchimaru, Prof. Jun-Ichiro Inoue, Prof. Hitoshi Satoh and Prof. Matsumoto Yasu, The University of Tokyo, for being my evaluator and defense committee member, and for reviewing my doctoral thesis.

I am also expressing my gratitude to Dr. Kazuyoshi Hosomichi of Kanazawa University for his help in nest generation sequencing.

I thank to members of Viral Infectious Diseases Unit, RIKEN, namely: Dr. Hiroyuki Otsuki, Dr.Hirotaka Sato, Dr. Nopporn Chutiwitoonchai, Dr. LanLan Bai, Dr. Lowela Siarot and Dr.Ayumu Ohno, for their experience, stimulating discussions, assistance and their technical support during my study. I am grateful to Kanako Hamada, Ayuko Ishizuka, and Yoko Hashimoto, Viral Infectious Diseases Unit, RIKEN, for arrangements in settling my life here, reservation of lecture rooms and their kind help in language assistance. I appreciate Saeki Negishi, Toshie Wakamatsu, Sonoko Watanuki, Yoshiko Sakuma, Viral Infectious Diseases Unit, RIKEN, for the happy moments we shared during my study.

I would like to express my sincere gratitude to the program for Junior Scientists of RIKEN for granting me an International Program Associate position at the laboratory of Viral Infectious Diseases Unit, RIKEN, for supporting me financially in the entirety of my doctor course study.

I am truly and deeply grateful to my family: my parents Polat Pazil and Raziyem Ala; my sisters Pazilat and Dilraba for their unconditional love, encouragement, prayers, and constant support in taking care of my daughter and me. I also appreciate my husband Sidikjan Gupur for caring, understanding and being with me. I especially thank my 4-years-old daughter, Subinuer, for making me happy and relaxed by her smile, and for making me trust the value of persuading further study as a mother.

I also thank all the veterinarians and collaborators in my research. To Drs. T. Miyasaka, M.A. Villanueva, A.J. Salces, Libertado C. Cruz and Mr. M. Ohashi, Dr. Claro Niegos Mingala, the management and staff of the Philippine Carabao Center, Dr. Carlos Javier Panei, Dr. Ester Teresa Gonzalez, Dr. Guillermo Giovambattista of Universidad national De La PLATA, Dr. Hla Hla Moe and Kyaw Kyaw Moe of University of Veterinary Science, Yezin, Myanmar, and to Dr. Takeshi Shimogiri of Kagoshima University for their help in obtaining the experiment samples.

Finally, I am thankful to the journal of "Archives of Virology" in which chapter 1 of this Ph.D thesis was already published.

References

Acaite, J., Tamosiunas, V., Lukauskas, K., Milius, J. & Pieskus, J. (2007). The eradication experience of enzootic bovine leukosis from Lithuania. *Prev Vet Med* 82, 83-89.

Aida, Y., Miyasaka, M., Okada, K., Onuma, M., Kogure, S., Suzuki, M., Minoprio, P., Levy, D. & Ikawa, Y. (1989). Further phenotypic characterization of target cells for bovine leukemia virus experimental infection in sheep. *Am J Vet Res* 50, 1946-1951.

Aida, Y., Murakami, H., Takahashi, M. & Takeshima, S. N. (2013). Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia virus. *Front Microbiol* 4, 328.

Aida, Y., Okada, K. & Amanuma, H. (1993). Phenotype and ontogeny of cells carrying a tumor-associated antigen that is expressed on bovine leukemia virus-induced lymphosarcoma. *Cancer Res* 53, 429-437.

Alexandersen, S., Carpenter, S., Christensen, J., Storgaard, T., Viuff, B., Wannemuehler, Y., Belousov, J. & Roth, J. A. (1993). Identification of alternatively spliced mRNAs encoding potential new regulatory proteins in cattle infected with bovine leukemia virus. *J Virol* 67, 39-52.

Alfonso, R., Almansa, J. E. & Barrera, J. C. (1998). Serological prevalence and evaluation of the risk factors of bovine enzootic leukosis in the Bogota savannah and the Ubate and Chiquinquira Valleys, Colombia. *Rev Sci Tech* 17, 723-732.

Altanerova, V., Ban, J., Kettmann, R. & Altaner, C. (1990). Induction of leukemia in chicken by bovine leukemia virus due to insertional mutagenesis. *Arch Geschwulstforsch* 60, 89-96.

Altanerova, V., Portetelle, D., Kettmann, R. & Altaner, C. (1989). Infection of rats with bovine leukaemia virus: establishment of a virus-producing rat cell line. *J Gen Virol* 70 (Pt 7), 1929-1932.

Arainga, M., Takeda, E. & Aida, Y. (2012). Identification of bovine leukemia virus tax function associated with host cell transcription, signaling, stress response and immune response pathway by microarray-based gene expression analysis. *BMC Genomics* 13, 121.

Asfaw, Y., Tsuduku, S., Konishi, M., Murakami, K., Tsuboi, T., Wu, D. & Sentsui, H. (2005). Distribution and superinfection of bovine leukemia virus genotypes in Japan. *Arch Virol* 150, 493-505.

Bai, L., Takeshima, S. N., Isogai, E., Kohara, J. & Aida, Y. (2015). Novel CD8(+) cytotoxic T cell epitopes in bovine leukemia virus with cattle. *Vaccine* 33, 7194-7202.

Balic, D., Lojkic, I., Periskic, M., Bedekovic, T., Jungic, A., Lemo, N., Roic, B., Cac, Z., Barbic, L. & Madic, J. (2012). Identification of a new genotype of bovine leukemia virus. *Arch Virol* 157, 1281-1290.

Ban, J., Czene, S., Altaner, C., Callebaut, I., Krchnak, V., Merza, M., Burny, A., Kettmann, R. & Portetelle, D. (1992). 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* 73 (Pt 9), 2457-2461.

Bech-Nielsen, S., Piper, C. E. & Ferrer, J. F. (1978). Natural mode of transmission of the bovine leukemia virus: role of bloodsucking insects. *Am J Vet Res* 39, 1089-1092.

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.

Bondoc, O. L. (2013). Genetic Diversity and Relationship of Domestic Buffalo and Cattle Breeds and Crossbreeds (Subfamily Bovinae) in the Philippines Based on the Cytochrome C Oxidase I (COI) Gene Sequence. *Philipp Agric Sci* 96, 93-102.

Bruck, C., Mathot, S., Portetelle, D., Berte, C., Franssen, J. D., Herion, P. & Burny, A. (1982a). Monoclonal antibodies define eight independent antigenic regions on the bovine leukemia virus (BLV) envelope glycoprotein gp51. *Virology* 122, 342-352.

Bruck, C., Portetelle, D., Burny, A. & Zavada, J. (1982b). Topographical analysis by monoclonal antibodies of BLV-gp51 epitopes involved in viral functions. *Virology* 122, 353-362.

Bruck, C., Portetelle, D., Mammerickx, M., Mathot, S. & Burny, A. (1984). Epitopes of bovine leukemia virus glycoprotein gp51 recognized by sera of infected cattle and sheep. *Leuk Res* 8, 315-321.

Brym, P., Rusc, A. & Kaminski, S. (2013). Evaluation of reference genes for qRT-PCR gene expression studies in whole blood samples from healthy and leukemia-virus infected cattle. *Vet Immunol Immunopathol* 153, 302-307.

Buehring, G. C., Shen, H. M., Jensen, H. M., Choi, K. Y., Sun, D. & Nuovo, G. (2014). Bovine leukemia virus DNA in human breast tissue. *Emerg Infect Dis* 20, 772-782.

Buehring, G. C., Shen, H. M., Jensen, H. M., Jin, D. L., Hudes, M. & Block, G. (2015). Exposure to Bovine Leukemia Virus Is Associated with Breast Cancer: A Case-Control Study. *PLoS One* 10, e0134304.

Bundza, A., Greig, A. S., Chander, S. & Dukes, T. W. (1980). Sporadic bovine leukosis: a description of eight calves received at Animal Diseases Research Institute from 1974-1980. *Can Vet J* 21, 280-283.

Burny, A., Bruck, C., Cleuter, Y., Couez, D., Deschamps, J., Ghysdael, J., Gregoire, D., Kettmann, R., Mammerickx, M., Marbaix, G. & et al. (1985). Bovine leukemia virus, a versatile agent with various pathogenic effects in various animal species. *Cancer Res* 45, 4578s-4582s.

Burny, A., Cleuter, Y., Kettmann, R., Mammerickx, M., Marbaix, G., Portetelle, D., Vandenbroeke, A., Willems, L. & Thomas, R. (1988). Bovine Leukemia - Facts and Hypotheses Derived from the Study of an Infectious Cancer. *Vet Microbiol* 17, 197-218.

Burridge, M. J., Thurmond, M. C., Miller, J. M., Schmerr, M. J. & Van Der Maaten, M. J. (1982). Fall in antibody titer to bovine leukemia virus in the periparturient period. *Can J Comp Med* 46, 270-271.

Callebaut, I., Voneche, V., Mager, A., Fumiere, O., Krchnak, V., Merza, M., Zavada, J., Mammerickx, M., Burny, A. & Portetelle, D. (1993). Mapping of B-neutralizing and T-helper cell epitopes on the bovine leukemia virus external glycoprotein gp51. *J Virol* 67, 5321-5327.

Camargos, M. F., Pereda, A., Stancek, D., Rocha, M. A., dos Reis, J. K., Greiser-Wilke, I. & Leite, R. C. (2007). Molecular characterization of the env gene from Brazilian field isolates of Bovine leukemia virus. *Virus Genes* 34, 343-350.

Camargos, M. F., Stancek, D., Rocha, M. A., Lessa, L. M., Reis, J. K. & Leite, R. C. (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.

Castillo, A. C. BACKYARD BEEF PRODUCTION IN THE PHILIPPINES.

Castro, R. S., Leite, R. C., Abreu, J. J., Lage, A. P., Ferraz, I. B., Lobato, Z. I. & Balsamao, S. L. (1992). Prevalence of antibodies to selected viruses in bovine embryo donors and recipients from Brazil, and its implications in international embryo trade. *Trop Anim Health Prod* 24, 173-176.

Ch, A. H. (1983). Bovine leukaemia virus infection in Peru. *Trop Anim Health Prod* 15, 61.

Coulston, J., Naif, H., Brandon, R., Kumar, S., Khan, S., Daniel, R. C. & Lavin, M. F. (1990). Molecular cloning and sequencing of an Australian isolate of proviral bovine leukaemia virus DNA: comparison with other isolates. *J Gen Virol* 71 (Pt 8), 1737-1746.

D'Angelino, J. L., Garcia, M. & Birgel, E. H. (1998). Epidemiological study of enzootic bovine leukosis in Brazil. *Trop Anim Health Prod* 30, 13-15.

de Brogniez, A., Bouzar, A. B., Jacques, J. R., Cosse, J. P., Gillet, N., Callebaut, I., Reichert, M. & Willems, L. (2015). Mutation of a Single Envelope N-Linked Glycosylation Site Enhances the Pathogenicity of Bovine Leukemia Virus. *J Virol* 89, 8945-8956.

Derse, D., Diniak, A. J., Casey, J. W. & Deininger, P. L. (1985). Nucleotide sequence and structure of integrated bovine leukemia virus long terminal repeats. *Virology* 141, 162-166.

DiGiacomo, R. F., Darlington, R. L. & Evermann, J. F. (1985). Natural transmission of bovine leukemia virus in dairy calves by dehorning. *Can J Comp Med* 49, 340-342.

Djilali, S. & Parodi, A. L. (1989). The BLV-induced leukemia--lymphosarcoma complex in sheep. *Vet Immunol Immunopathol* 22, 233-244.

Djilali, S., Parodi, A. L., Levy, D. & Cockerell, G. L. (1987). Development of leukemia and lymphosarcoma induced by bovine leukemia virus in sheep: a hematopathological study. *Leukemia* 1, 777-781.

Dube, S., Abbott, L., Dube, D. K., Dolcini, G., Gutierrez, S., Ceriani, C., Juliarena, M., Ferrer, J., Perzova, R. & Poiesz, B. J. (2009). The complete genomic sequence of an in vivo low replicating BLV strain. *Virol J* 6, 120.

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.

Fechner, H., Blankenstein, P., Looman, A. C., 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.

Felber, B. K., Derse, D., Athanassopoulos, A., Campbell, M. & Pavlakis, G. N. (1989). Cross-activation of the Rex proteins of HTLV-I and BLV and of the Rev protein of HIV-1 and nonreciprocal interactions with their RNA responsive elements. *New Biol* 1, 318-328.

Felmer, R., Munoz, G., Zuniga, J. & Recabal, M. (2005). Molecular analysis of a 444 bp fragment of the bovine leukaemia virus gp51 env gene reveals a high frequency of non-silent point mutations and suggests the presence of two subgroups of BLV in Chile. *Vet Microbiol* 108, 39-47.

Ferrer, J. F. & Piper, C. E. (1978). An evaluation of the role of milk in the natural transmission of BLV. *Ann Rech Vet* 9, 803-807.

Ferrer, J. F. & Piper, C. E. (1981). Role of colostrum and milk in the natural transmission of the bovine leukemia virus. *Cancer Res* 41, 4906-4909.

Florins, A., Boxus, M., Vandermeers, F., Verlaeten, O., Bouzar, A. B., Defoiche, J., Hubaux, R., Burny, A., Kettmann, R. & Willems, L. (2008). Emphasis on cell turnover in two hosts infected by bovine leukemia virus: a rationale for host susceptibility to disease. *Vet Immunol Immunopathol* 125, 1-7.

Florins, A., Gillet, N., Boxus, M., Kerkhofs, P., Kettmann, R. & Willems, L. (2007). Even attenuated bovine leukemia virus proviruses can be pathogenic in sheep. *J Virol* 81, 10195-10200.

Fukai, K., Sato, M., Kawara, M., Hoshi, Z., Ueno, S., Chyou, N. & Akashi, H. (1999). A case of an embryo transfer calf infected with bovine leukemia virus from the recipient cow. *Zentralbl Veterinarmed B* 46, 511-515.

Gatot, J. S., 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 differently affect viral fusion and infectivity in vivo. *J Virol* 76, 7956-7967.

Gillet, N., Florins, A., Boxus, M., Burteau, C., Nigro, A., Vandermeers, F., Balon, H., Bouzar, A. B., Defoiche, J., Burny, A., Reichert, M., Kettmann, R. & Willems, L. (2007). Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in human. *Retrovirology* 4, 18.

Gillet, N. A., Hamaidia, M., de Brogniez, A., Gutierrez, G., Renotte, N., Reichert, M., Trono, K. & Willems, L. (2016). Bovine Leukemia Virus Small Noncoding RNAs Are Functional Elements That Regulate Replication and Contribute to Oncogenesis In Vivo. *PLoS Pathog* 12, e1005588.

Graves, D. C. & Ferrer, J. F. (1976). In vitro transmission and propagation of the bovine leukemia virus in monolayer cell cultures. *Cancer Res* 36, 4152-4159.

Gutierrez, G., Alvarez, I., Politzki, R., Lomonaco, M., Dus Santos, M. J., Rondelli, F., Fondevila, N. & Trono, K. (2011). Natural progression of Bovine Leukemia Virus infection in Argentinean dairy cattle. *Vet Microbiol* 151, 255-263.

Gutierrez, G., Rodriguez, S. M., de Brogniez, A., Gillet, N., Golime, R., Burny, A., Jaworski, J. P., Alvarez, I., Vagnoni, L., Trono, K. & Willems, L. (2014). Vaccination against delta-retroviruses: the bovine leukemia virus paradigm. *Viruses* 6, 2416-2427.

Hamard-Peron, E. & Muriaux, D. (2011). Retroviral matrix and lipids, the intimate interaction. *Retrovirology* 8, 15.

Hemmatzadeh, F. (2007). Sequencing and phylogenetic analysis of gp51 gene of bovine leukaemia virus in Iranian isolates. *Vet Res Commun* 31, 783-789.

Hernández-Herrera, D., Posso-Terranova, A., Benavides, J., Muñoz-Flórez, J., Giovambattista, G. & Álvarez-Franco, L. (2011). Bovine leukosis virus detection in Creole Colombian breeds using nested-PCR. *ACTA AGRONÓMICA* 60, 311-317.

Hopkins, S. G., Evermann, J. F., DiGiacomo, R. F., Parish, S. M., Ferrer, J. F., Smith, S. & Bangert, R. L. (1988). Experimental transmission of bovine leukosis virus by simulated rectal palpation. *Vet Rec* 122, 389-391.

Inabe, K., Ikuta, K. & Aida, Y. (1998). Transmission and propagation in cell culture of virus produced by cells transfected with an infectious molecular clone of bovine leukemia virus. *Virology* 245, 53-64.

Inabe, K., Nishizawa, M., Tajima, S., Ikuta, K. & Aida, Y. (1999). The YXXL sequences of a transmembrane protein of bovine leukemia virus are required for viral entry and incorporation of viral envelope protein into virions. *J Virol* 73, 1293-1301.

Jewell, N. A. & Mansky, L. M. (2000). In the beginning: genome recognition, RNA encapsidation and the initiation of complex retrovirus assembly. *J Gen Virol* 81, 1889-1899.

Jimba, M., Takeshima, S. N., Matoba, K., Endoh, D. & Aida, Y. (2010). BLV-CoCoMo-qPCR: Quantitation of bovine leukemia virus proviral load using the CoCoMo algorithm. *Retrovirology* 7, 91.

Jimba, M., Takeshima, S. N., Murakami, H., Kohara, J., Kobayashi, N., Matsuhashi, T., Ohmori, T., Nunoya, T. & Aida, Y. (2012). BLV-CoCoMo-qPCR: a useful tool for evaluating bovine leukemia virus infection status. *BMC Vet Res* 8, 167.

Johnson, R., Gibson, C. D. & Kaneene, J. B. (1985). Bovine Leukemia-Virus - a Herd-Based Control Strategy. *Prev Vet Med* 3, 339-349.

Johnson, R. & Kaneene, J. B. (1992). bovine leukemia virus and enzootic bovine leukesis. *Veterinary Bullitin* 62, 287-312.

Johnston, E. R., Albritton, L. M. & 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.

Johnston, E. R. & Radke, K. (2000). The SU and TM envelope protein subunits of bovine leukemia virus are linked by disulfide bonds, both in cells and in virions. *J Virol* 74, 2930-2935.

Jukes, T. H. & Canter, C. R. (1969). *Evolution of protein molecules*. New York.

Juliarena, M. A., Lendez, P. A., Gutierrez, S. E., Forletti, A., Rensetti, D. E. & Ceriani, M. C. (2013). Partial molecular characterization of different proviral strains of bovine leukemia virus. *Arch Virol* 158, 63-70.

Kettmann, R., Burny, A., Callebaut, I., Droogmans, L., Mammerickx, M., Willems, L. & Portelle, D. (1994). Bovine leukemia virus. 3, 39-81.

Kettmann, R., Meunier-Rotival, M., Cortadas, J., Cuny, G., Ghysdael, J., Mammerickx, M., Burny, A. & Bernardi, G. (1979). Integration of bovine leukemia virus DNA in the bovine genome. *Proc Natl Acad Sci U S A* 76, 4822-4826.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16, 111-120.

Kincaid, R. P., Burke, J. M. & Sullivan, C. S. (2012). RNA virus microRNA that mimics a B-cell oncomiR. *Proc Natl Acad Sci U S A* 109, 3077-3082.

Kincaid, R. P. & Sullivan, C. S. (2012). Virus-encoded microRNAs: an overview and a look to the future. *PLoS Pathog* 8, e1003018.

Kirkland, P. D. & Rodwell, B. J. (2005). Enzootic Bovine Leukosis. *Australia and New Zealand Standard Diagnostic Procedures*, 1-14.

Kobayashi, S., Hidano, A., Tsutsui, T., Yamamoto, T., Hayama, Y., Nishida, T., Muroga, N., Konishi, M., Kameyama, K. & Murakami, K. (2014). Analysis of risk factors associated with bovine leukemia virus seropositivity within dairy and beef breeding farms in Japan: A nationwide survey. *Res Vet Sci* 96, 47-53.

Kono, Y., Sentsui, H., Arai, K., Ishida, H. & Irishio, W. (1983). Contact transmission of bovine leukemia virus under insect-free conditions. *Nihon Juigaku Zasshi* 45, 799-802.

Kurdi, A., Blankenstein, P., Marquardt, O. & Ebner, D. (1999). [Serologic and virologic investigations on the presence of BLV infection in a dairy herd in Syria]. *Berl Munch Tierarztl Wochenschr* 112, 18-23.

LanLan, B., Otsuki, H., Sato, H., Kohara, J., Isogai, E., Takeshima, S.-n. & Aida, Y. (2015). Identification and characterization of common B cell epitope in bovine leukemia virus via high-throughput peptide screening system in infected cattle. *Retrovirology*.

Lassauzet, M. L., Thurmond, M. C., Johnson, W. O., Stevens, F. & Picanso, J. P. (1990). Effect of brucellosis vaccination and dehorning on transmission of bovine leukemia virus in heifers on a California dairy. *Can J Vet Res* 54, 184-189.

Lee, E., Kim, E. J., Joung, H. K., Kim, B. H., Song, J. Y., Cho, I. S., Lee, K. K. & Shin, Y. K. (2015). Sequencing and phylogenetic analysis of the gp51 gene from Korean bovine leukemia virus isolates. *Virol J* 12, 64.

Leisering, A. (1871). Hypertrophy der Malpigischen Korperchen der Milz. *Bericht uber das Veterinarwesen im Konigreich Sachsen* 16, 15-16.

Levy, D., Deshayes, L., Parodi, A. L., Levy, J. P., Stephenson, J. R., Devare, S. G. & Gilden, R. V. (1977). Bovine leukemia virus specific antibodies among French cattle. II. Radioimmunoassay with the major structural protein (BLV p24). *Int J Cancer* 20, 543-550.

Lew, A. E., Bock, R. E., Miles, J., Cuttell, L. B., Steer, P. & Nadin-Davis, S. A. (2004). Sensitive and specific detection of bovine immunodeficiency virus and bovine syncytial virus by 5 ' Taq nuclease assays with fluorescent 3 ' minor groove binder-DNA probes. *J Virol Methods* 116, 1-9.

Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.

Licursi, M., Inoshima, Y., Wu, D., Yokoyama, T., Gonzalez, E. T. & Sentsui, H. (2002). Genetic heterogeneity among bovine leukemia virus genotypes and its relation to humoral responses in hosts. *Virus Res* 86, 101-110.

Licursi, M., Inoshima, Y., Wu, D., Yokoyama, T., Gonzalez, E. T. & Sentsui, H. (2003). Provirus variants of bovine leukemia virus in naturally infected cattle from Argentina and Japan. *Vet Microbiol* 96, 17-23.

Lim, S. I., Jeong, W., Tark, D. S., Yang, D. K. & Kweon, C. H. (2009). Agar gel immunodiffusion analysis using baculovirus-expressed recombinant bovine leukemia virus envelope glycoprotein (gp51/gp30(T-)). *J Vet Sci* 10, 331-336.

Liron, J. P., Peral-Garcia, P. & Giovambattista, G. (2006). Genetic characterization of Argentine and Bolivian Creole cattle breeds assessed through microsatellites. *J Hered* 97, 331-339.

Mamoun, R. Z., Astier, T., Guillemain, B. & Duplan, J. F. (1983). Bovine lymphosarcoma: expression of BLV-related proteins in cultured cells. *J Gen Virol* 64 (Pt 9), 1895-1905.

Mamoun, R. Z., 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.

Matsumura, K., Inoue, E., Osawa, Y. & Okazaki, K. (2011). Molecular epidemiology of bovine leukemia virus associated with enzootic bovine leukosis in Japan. *Virus Res* 155, 343-348.

Merezak, C., Pierreux, C., Adam, E., Lemaigre, F., Rousseau, G. G., Calomme, C., Van Lint, C., Christophe, D., Kerkhofs, P., Burny, A., Kettmann, R. & Willems, L. (2001). Suboptimal enhancer sequences are required for efficient bovine leukemia virus propagation in vivo: implications for viral latency. *J Virol* 75, 6977-6988.

Miller, J. M., Miller, L. D., Olson, C. & Gillette, K. G. (1969). Virus-like particles in phytohemagglutinin-stimulated lymphocyte cultures with reference to bovine lymphosarcoma. *J Natl Cancer Inst* 43, 1297-1305.

Mingala, C. N., Konnai, S., Cruz, L. C., Onuma, M. & Ohashi, K. (2009). Comparative moleculo-immunological analysis of swamp- and riverine-type water buffaloes responses. *Cytokine* 46, 273-282.

Mirsky, M. L., Olmstead, C. A., Da, Y. & Lewin, H. A. (1996). The prevalence of proviral bovine leukemia virus in peripheral blood mononuclear cells at two subclinical stages of infection. *J Virol* 70, 2178-2183.

Molnár, É., Molnár, L., Dias, H., Silva, A. & Vale, W. (1999). Occurrence of enzootic bovine leukosis in the State of Pará, Brazil. *Pesquisa Veterinária Brasileira* 19, 7-11.

Molteni, E., Agresti, A., Meneveri, R., Marozzi, A., Malcovati, M., Bonizzi, L., Poli, G. & Ginelli, E. (1996). Molecular characterization of a variant of proviral bovine leukaemia virus (BLV). *Zentralbl Veterinarmed B* 43, 201-211.

Monke, D. R. (1986). Noninfectivity of semen from bulls infected with bovine leukosis virus. *J Am Vet Med Assoc* 188, 823-826.

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

Monti, G. E., Frankena, K., Engel, B., Buist, W., Tarabla, H. D. & de Jong, M. C. (2005b). Evaluation of a new antibody-based enzyme-linked immunosorbent assay for the detection of bovine leukemia virus infection in dairy cattle. *J Vet Diagn Invest* 17, 451-457.

Moratorio, G., Iriarte, A., Moreno, P., Musto, H. & Cristina, J. (2013). A detailed comparative analysis on the overall codon usage patterns in West Nile virus. *Infect Genet Evol* 14, 396-400.

Moratorio, G., Obal, G., Dubra, A., Correa, A., Bianchi, S., Buschiazzo, A., Cristina, J. & Pritsch, O. (2010). Phylogenetic analysis of bovine leukemia viruses isolated in South America reveals diversification in seven distinct genotypes. *Arch Virol* 155, 481-489.

Murakami, K., Aida, Y., Kageyama, R., Numakunai, S., Ohshima, K., Okada, K. & Ikawa, Y. (1994a). Immunopathologic study and characterization of the phenotype of transformed cells in sheep with bovine leukemia virus-induced lymphosarcoma. *Am J Vet Res* 55, 72-80.

Murakami, K., Kobayashi, S., Konishi, M., Kameyama, K., Yamamoto, T. & Tsutsui, T. (2011). The recent prevalence of bovine leukemia virus (BLV) infection among Japanese cattle. *Vet Microbiol* 148, 84-88.

Murakami, K., Okada, K., Ikawa, Y. & Aida, Y. (1994b). Bovine leukemia virus induces CD5- B cell lymphoma in sheep despite temporarily increasing CD5+ B cells in asymptomatic stage. *Virology* 202, 458-465.

Naif, H. M., Brandon, R. B., Daniel, R. C. W. & Lavin, M. F. (1990). Bovine Leukemia Proviral DNA Detection in Cattle Using the Polymerase Chain-Reaction. *Vet Microbiol* 25, 117-129.

National Consultative Committee, M. (2013). National Report on Animal genetic Resources, The Union of Myanmar, 2013 edn. Myanmar: National Consultative committe.

Ochirkhuu, N., Konnai, S., Odbileg, R., Nishimori, A., Okagawa, T., Murata, S. & Ohashi, K. (2016). Detection of bovine leukemia virus and identification of its genotype in Mongolian cattle. *Arch Virol* 161, 985-991.

Ohno, A., Takeshima, S. N., Matsumoto, Y. & Aida, Y. (2015). Risk factors associated with increased bovine leukemia virus proviral load in infected cattle in Japan from 2012 to 2014. *Virus Res* 210, 283-290.

Ohshima, K. I., Okada, K., Numakunai, S., Yoneyama, Y., Sato, S. & Takahashi, K. (1981). Evidence on Horizontal Transmission of Bovine Leukemia-Virus Due to Blood-Sucking Tabanid Flies. *Jpn J Vet Sci* 43, 79-81.

OIE (2009). World Animal Health infromation database-version: 1.4. Paris, France: World organisation for animal Health.

OIE (2012). Chapter 2.4.11 Enzootic Bovine Leukosis, Seventh Edition edn. Edited by W. O. F. A. Health. France.

Olson, C., Kettmann, R., Burny, A. & Kaja, R. (1981). Goat lymphosarcoma from bovine leukemia virus. *J Natl Cancer Inst* 67, 671-675.

Onuma, M., Wada, M., Yasutomi, Y., Yamamoto, M., Okada, H. M. & Kawakami, Y. (1990). Suppression of immunological responses in rabbits experimentally infected with bovine leukemia virus. *Vet Microbiol* 25, 131-141.

Panei, C. J., Takeshima, S. N., Omori, T., Nunoya, T., Davis, W. C., Ishizaki, H., Matoba, K. & Aida, Y. (2013). Estimation of bovine leukemia virus (BLV) proviral load harbored by lymphocyte subpopulations in BLV-infected cattle at the subclinical stage of enzootic bovine leucosis using BLV-CoCoMo-qPCR. *BMC Vet Res* 9, 95.

Perzova, R. N., Loughran, T. P., Dube, S., Ferrer, J., Esteban, E. & Poiesz, B. J. (2000). Lack of BLV and PTLV DNA sequences in the majority of patients with large granular lymphocyte leukaemia. *Br J Haematol* 109, 64-70.

Polat, M., Ohno, A., Takeshima, S. N., Kim, J., Kikuya, M., Matsumoto, Y., Mingala, C. N., Onuma, M. & Aida, Y. (2015). Detection and molecular characterization of bovine leukemia virus in Philippine cattle. *Arch Virol* 160, 285-296.

Polat, M., Takeshima, S. N., Hosomichi, K., Kim, J., Miyasaka, T., Yamada, K., Arainga, M., Murakami, T., Matsumoto, Y., de la Barra Diaz, V., Panei, C. J., Gonzalez, E. T., Kanemaki, M., Onuma, M., Giovambattista, G. & Aida, Y. (2016). A new genotype of bovine leukemia virus in South America identified by NGS-based whole genome sequencing and molecular evolutionary genetic analysis. *Retrovirology* 13, 4.

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 NH2 part of the envelope glycoprotein gp51. *Virology* 169, 27-33.

Qualley, D. F. & Boleratz, B. L. (2014). Expression, purification, and characterization of full-length bovine leukemia virus Gag protein from bacterial culture. *Protein Expr Purif* 93, 32-37.

Rabbani, B., Mahdieh, N., Hosomichi, K., Nakaoka, H. & Inoue, I. (2012). Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders. *J Hum Genet* 57, 621-632.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol* 29, 24-26.

Rodriguez, S. M., Florins, A., Gillet, N., de Brogniez, A., Sanchez-Alcaraz, M. T., Boxus, M., Boulanger, F., Gutierrez, G., Trono, K., Alvarez, I., Vagnoni, L. & Willems, L. (2011). Preventive and therapeutic strategies for bovine leukemia virus: lessons for HTLV. *Viruses* 3, 1210-1248.

Rodriguez, S. M., Golemba, M. D., Campos, R. H., Trono, K. & Jones, L. R. (2009). Bovine leukemia virus can be classified into seven genotypes: evidence for the existence of two novel clades. *J Gen Virol* 90, 2788-2797.

Rola-Luszczak, M., Pluta, A., Olech, M., Donnik, I., Petropavlovskiy, M., Gerilovych, A., Vinogradova, I., Choudhury, B. & Kuzmak, J. (2013). The molecular characterization of bovine leukaemia virus isolates from Eastern Europe and Siberia and its impact on phylogeny. *PLoS One* 8, e58705.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M. A. & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61, 539-542.

Rosewick, N., Momont, M., Durkin, K., Takeda, H., Caiment, F., Cleuter, Y., Vernin, C., Mortreux, F., Wattel, E., Burny, A., Georges, M. & Van den Broeke, A. (2013). Deep sequencing reveals abundant noncanonical retroviral microRNAs in B-cell leukemia/lymphoma. *Proc Natl Acad Sci U S A* 110, 2306-2311.

Rovnak, J., Boyd, A. L., Casey, J. W., Gonda, M. A., Jensen, W. A. & Cockerell, G. L. (1993). Pathogenicity of molecularly cloned bovine leukemia virus. *J Virol* 67, 7096-7105.

Sagata, N., Yasunaga, T., Ohishi, K., Tsuzuku-Kawamura, J., Onuma, M. & Ikawa, Y. (1984). 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* 3, 3231-3237.

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 U S A* 82, 677-681.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4, 406-425.

Samara, S. I., Lima, E. G. & Nascimento, A. A. (1997). Monitoring of enzootic bovine leukosis in dairy cattle from the Pitangueiras region in São Paulo, Brazil. *Brazilian Journal of Veterinary Research and Animal Science* 34, 349-351.

Stanton, E. S. (2010). The Philippines Dairy Cattle Farming Sector - A Briefing for Canadian Livestock Genetics Suppliers. Edited by A. a. A.-f. C. T. E. f. C. i. t. Philippines.

Tajima, S. & Aida, Y. (2000a). 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* 74, 10939-10949.

Tajima, S. & Aida, Y. (2000b). 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* 74, 10939-10949.

Tajima, S. & Aida, Y. (2002). Mutant tax protein from bovine leukemia virus with enhanced ability to activate the expression of c-fos. *J Virol* 76, 2557-2562.

Tajima, S., Ikawa, Y. & Aida, Y. (1998). Complete bovine leukemia virus (BLV) provirus is conserved in BLV-infected cattle throughout the course of B-cell lymphosarcoma development. *J Virol* 72, 7569-7576.

Tajima, S., Takahashi, M., Takeshima, S. N., Konnai, S., Yin, S. A., Watarai, S., Tanaka, Y., Onuma, M., Okada, K. & Aida, Y. (2003). A mutant form of the tax protein of bovine leukemia virus (BLV), with enhanced transactivation activity, increases expression and propagation of BLV in vitro but not in vivo. *J Virol* 77, 1894-1903.

Takahashi, M., Tajima, S., Okada, K., Davis, W. C. & Aida, Y. (2005). Involvement of bovine leukemia virus in induction and inhibition of apoptosis. *Microbes Infect* 7, 19-28.

Takahashi, M., Tajima, S., Takeshima, S. N., Konnai, S., Yin, S. A., Okada, K., Davis, W. C. & Aida, Y. (2004). Ex vivo survival of peripheral blood mononuclear cells in sheep induced by bovine leukemia virus (BLV) mainly occurs in CD5- B cells that express BLV. *Microbes Infect* 6, 584-595.

Takeshima, S. N., Kitamura-Muramatsu, Y., Yuan, Y., Polat, M., Saito, S. & Aida, Y. (2015). BLV-CoCoMo-qPCR-2: improvements to the BLV-CoCoMo-qPCR assay for bovine leukemia virus by reducing primer degeneracy and constructing an optimal standard curve. *Arch Virol* 160, 1325-1332.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10, 512-526.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28, 2731-2739.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30, 2725-2729.

Tawfeeq, M. M., Horiuchi, N., Kobayashi, Y., Furuoka, H. & Inokuma, H. (2013). Evaluation of Gene Expression in Peripheral Blood Cells as a Potential Biomarker for Enzootic Bovine Leukosis. *J Vet Med Sci* 75, 1213-1217.

Thompson, K. G., Johnstone, A. C. & Hilbink, F. (1993). Enzootic bovine leukosis in New Zealand--a case report and update. *N Z Vet J* 41, 190-194.

Thompson, P. D. & Miller, R. H. (1974). Retrograde Flow of Milk within Machine-Milked Teats. *J Dairy Sci* 57, 1489-1496.

Trono, K. G., Perez-Filgueira, D. M., Duffy, S., Borca, M. V. & 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.

Uera, J. A., Lazaro, J. V. & Mingala, C. N. (2012). Detection of Enzootic Bovine Leukosis in Cattle using Nested Polymerase Chain Reaction Assay. *Thai J Vet Med* 42, 319-324.

USDA (2008). Bovine Leukosis Virus (BLV) on U.S.Dairy Operations, 2007.

Van der Maaten, M. J., Miller, J. M. & Schmerr, M. J. (1981). In utero transmission of bovine leukemia virus. *Am J Vet Res* 42, 1052-1054.

Voneche, V., Portelle, D., Kettmann, R., Willems, L., Limbach, K., Paoletti, E., Ruysschaert, J. M., Burny, A. & Brasseur, R. (1992). Fusogenic Segments of Bovine Leukemia-Virus and Simian Immunodeficiency Virus Are Interchangeable and Mediate Fusion by Means of Oblique Insertion in the Lipid Bilayer of Their Target-Cells. *Proc Natl Acad Sci U S A* 89, 3810-3814.

Wang, C. T. (1991). Bovine leukemia virus infection in Taiwan: epidemiological study. *J Vet Med Sci* 53, 395-398.

Wang, H., Norris, K. M. & Mansky, L. M. (2003). Involvement of the matrix and nucleocapsid domains of the bovine leukemia virus Gag polyprotein precursor in viral RNA packaging. *J Virol* 77, 9431-9438.

Watanabe, T., Masangkay, J. S., Wakana, S., Saitou, N. & Tomita, T. (1989). Mitochondrial DNA polymorphism in native Philippine cattle based on restriction endonuclease cleavage patterns. *Biochem Genet* 27, 431-438.

Willems, L., Burny, A., Collete, D., Dangoisse, O., Dequiedt, F., Gatot, J. S., Kerkhofs, P., Lefebvre, L., Merezak, C., Peremans, T., Portetelle, D., Twizere, J. C. & Kettmann, R.

(2000). Genetic determinants of bovine leukemia virus pathogenesis. *AIDS Res Hum Retroviruses* 16, 1787-1795.

Willems, L., Gatot, J. S., Mammerickx, M., Portetelle, D., Burny, A., Kerkhofs, P. & Kettmann, R. (1995). The Yxxl Signaling Motifs of the Bovine Leukemia-Virus Transmembrane Protein Are Required for in-Vivo Infection and Maintenance of High Viral Loads. *J Virol* 69, 4137-4141.

Willems, L., Grimonpont, C., Heremans, H., Rebeyrotte, N., Chen, G., Portetelle, D., Burny, A. & Kettmann, R. (1992). Mutations in the Bovine Leukemia-Virus Tax Protein Can Abrogate the Long Terminal Repeat-Directed Transactivating Activity without Concomitant Loss of Transforming Potential. *Proc Natl Acad Sci U S A* 89, 3957-3961.

Willems, L., Grimonpont, C., Kerkhofs, P., Capiau, C., Gheysen, D., Conrath, K., Roussef, R., Mamoun, R., Portetelle, D., Burny, A., Adam, E., Lefebvre, L., Twizere, J. C., Heremans, H. & Kettmann, R. (1998). Phosphorylation of bovine leukemia virus Tax protein is required for in vitro transformation but not for transactivation. *Oncogene* 16, 2165-2176.

Willems, L., Heremans, H., Chen, G., Portetelle, D., Billiau, A., Burny, A. & Kettmann, R. (1990). Cooperation between Bovine Leukemia-Virus Transactivator Protein and Ha-Ras Oncogene Product in Cellular-Transformation. *EMBO J* 9, 1577-1581.

Willems, L., Kerkhofs, P., Dequiedt, F., Portetelle, D., Mammerickx, M., Burny, A. & Kettmann, R. (1994). Attenuation of bovine leukemia virus by deletion of R3 and G4 open reading frames. *Proc Natl Acad Sci U S A* 91, 11532-11536.

Willems, L., Kettmann, R. & Burny, A. (1991). The amino acid (157-197) peptide segment of bovine leukemia virus p34tax encompass a leucine-rich globally neutral activation domain. *Oncogene* 6, 159-163.

Wyatt, C. R., Wingett, D., White, J. S., Buck, C. D., Knowles, D., Reeves, R. & Magnuson, N. S. (1989). Persistent infection of rabbits with bovine leukemia virus associated with development of immune dysfunction. *J Virol* 63, 4498-4506.

Zarkik, S., Decroly, E., Wattiez, R., Seidah, N. G., Burny, A. & Ruysschaert, J. M. (1997). Comparative processing of bovine leukemia virus envelope glycoprotein gp72 by subtilisin/kexin-like mammalian convertases. *FEBS Lett* 406, 205-210.

Zhao, X. & Buehring, G. C. (2007). Natural genetic variations in bovine leukemia virus envelope gene: possible effects of selection and escape. *Virology* 366, 150-165.

Tables

Table 1 Summary of common diagnostic techniques of BLV prevalence

Diagnostic Assay		Sample	Target	Advantages	Disadvantages
type	**Assay**				
Serological test	AGID	Serum	Antibodies (p24, p51)	Specific, simple and easy to perform Large scale screening Economically less expensive Time saving	Less sensitive and inconclusive Cannot evaluate disease states of infected cattle
	ELISA	Serum Milk Bulk milk	Antibodies (p24, p51)	Specific and sensitive Large scale screening Time saving	False negative-cattle in early infection phase False positive-materially derived antibodies Cannot evaluate disease states of infected cattle
	PHA	Virus Virus particle	BLV glycoprotein	Sensitive Specific in detecting BLV Large scare titration Economically less expensive Time saving	Affecting by PH and temperature Hemagglutination activity reduced by trypsin, potassium periodate and neuraminidase
Proviral DNA detection	PCR (single PCR; semi-nested PCR; nested PCR)	Blood PBMC Tumor sample Buffy coat Milk somatic cells Semen Saliva Nasal secretions	Provirus	Direct: Fast: sensitive A variety of samples can be used BLV detection in the early phase of infection or in the presence of colostral antibodies	Cannot detect if proviral load is too low Cross contamination easily Requires specific primers Require equipment (PCR machine)
	RT-PCR (RT-PCR; CoCoMo-qPCR; CoCoMo-qPCR-2; Takara RT-PCR)	Blood PBMC Tumor sample Buffy coat Milk somatic cells Semen Saliva Nasal secretions	Provirus	Direct, Fast, sensitive Low risk of contamination A variety of samples can be used Distinguishing EBL from SBL BLV detection in the early phase of infection or in the presence of colostral antibodies Quantitatively measure proviral load	Requires specific primers and probe Require equipment (real time PCR machine) Economically expensive Complicated sample preparation procedure

Table 2 Worldwide geographical distribution of eight known BLV genotypes based on *env-gp51* sequences

Geographic divisions	Country	Genotype								References
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
Asia	Korea	1		3						Lim *et al*. (Lim *et al.*, 2009); Lee *et al*. (Lee *et al.*, 2015)
	Japan	1	2	3						Sagata *et al*. (Sagata *et al.*, 1985); Licursi *et al*. (Licursi *et al.*, 2003); Zhao & Buehring (Zhao & Buehring, 2007)
	Mongolia	1			4			7		Ochirkhuu *et al*. (Ochirkhuu *et al.*, 2016)
North America	USA	1		3	4					Derse *et al*. (Derse *et al.*, 1985); Mamoun *et al*. (Mamoun *et al.*, 1990); Zhao & Buehring (Zhao & Buehring, 2007)
Central America	Costa Rica	1				5				Zhao & Buehring (Zhao & Buehring, 2007)
South America	Argentina	1	2		4		6			Dube *et al*. (Dube *et al.*, 2000); Licursi *et al*. (Licursi *et al.*, 2003); Monti *et al*. (Monti *et al.*, 2005a); Dube *et al*. (Dube *et al.*, 2009); Rodriguez *et al*. (Rodriguez *et al.*, 2009)
	Brazil	1	2			5	6	7		Camargos *et al*. (Camargos *et al.*, 2002); Camargos *et al*. (Camargos *et al.*, 2007); Moratorio *et al*. (Moratorio *et al.*, 2010)
	Chile				4			7		Felmer *et al*. (Felmer *et al.*, 2005)
	Uruguay	1								Moratorio *et al*. (Moratorio *et al.*, 2010)
Australia	Australia	1								Coulston *et al*. (Coulston *et al.*, 1990)
Eastern Europe	Belarus				4					Rola-Luszczak *et al*. (Rola-Luszczak *et al.*, 2013)
	Russia				4			7	8	Rola-Luszczak *et al*. (Rola-Luszczak *et al.*, 2013)
	Urkaine				4			7	8	Rola-Luszczak *et al*. (Rola-Luszczak *et al.*, 2013)
Central Europe	Croatia								8	Balic *et al*. (Balic *et al.*, 2012)
	Poland				4			7		Rola-Luszczak *et al*. (Rola-Luszczak *et al.*, 2013)
Western Europe	Belgium				4					Mamoun *et al*.(Mamoun *et al.*, 1990); Zhao & Buehring (Zhao & Buehring, 2007)
	France			3	4					Mamoun *et al*. (Mamoun *et al.*, 1990)
	Germany	1			4					Fechner *et al*. (Fechner *et al.*, 1997)
	Italy							7		Molteni *et al*. (Molteni *et al.*, 1996)

Table 3 The BLV detection results for Bohol Island as determined by BLV-CoCoMo-qPCR-2 and nested PCR

Island	City / Municipality	Farm	Breed	Sample Number	Number of BLV positive samples		
					CoCoMo-qPCR	Nested-PCR	Both
Bohol	I	A	Native	22	0	0	0
			Native x Brahman	3	0	0	0
		B	Native	24	0	0	0
		C	Native x Brahman	2	0	0	0
		D	Native x Holstein	1	0	0	0
	II	E	Native	20	0	0	0
Total Number of Samples				72	0	0	0

Table 4 The BLV detection results for Cebu Island as determined by BLV-CoCoMo-qPCR-2 and nested PCR

Island	City / Municipality	Farm	Breed	Sample Number	Number of BLV positive samples		
					CoCoMo - qPCR	Nested-PCR	Both
Cebu	I	A	Holstein x Sahiwal	4	0	0	0
		B	Holstein x Sahiwal	9	0	0	0
		C	Holstein x Sahiwal	7	0	0	0
		D	Holstein x Sahiwal	10	0	0	0
		E	Holstein x Sahiwal	14	0	0	0
	II	F	Holstein x Sahiwal	8	0	0	0
			Holstein x Brahman x Sahiwa	3	0	0	0
		G	Holstein x Sahiwal	6	0	0	0
		H	Holstein x Jersey	1	0	0	0
			Holstein x Sahiwal	40	0	0	0
Total Number of Samples				102	0	0	0

Table 5 The BLV detection results for Leyte Island as determined by BLV-CoCoMo-qPCR-2 and nested PCR

Island	City / Municipality	Farm	Breed	Sample Number	Number of BLV positive samples		
					CoCoMo-qPCR	Nested-PCR	Both
Leyte	I	A	Brahman [e]	1	0	0	0
			Native [e]	10	0	0	0
			Native x Brahman [e]	1	0	0	0
		B	Native x Holstein [d]	1	0	0	0
			Native [e]	9	0	0	0
		C	Native x Brahman [e]	2	0	0	0
			Native [e]	29	0	0	0
		D	Native [e]	20	0	0	0
		E	Native [e]	6	0	0	0
			Native x Brahman [e]	8	0	0	0
		F	Native [e]	12	0	0	0
		G	Native [e]	12	0	0	0
		H	Native x Brahman [e]	2	0	0	0
			Native [e]	16	0	0	0
		I	Native x Brahman [e]	3	0	0	0
			Native [e]	16	0	0	0
		J	Native x Brahman [e]	8	0	0	0
			Native [e]	13	0	0	0
	II	K	Native x Brahman [e]	1	0	0	0
			Native [e]	11	0	0	0
		L	Native x Brahman [e]	2	0	0	0
			Native [e]	3	0	0	0
		M	Native x Brahman [e]	1	0	0	0
			Native [e]	10	1	0	0
		N	Native x Brahman [e]	1	0	0	0
			Native [e]	4	0	0	0
		O	Native x Brahman [e]	3	0	0	0
			Native [e]	24	1	0	0
		P	Native [e]	7	0	0	0
		Q	Native x Brahman [e]	3	0	0	0
			Native [e]	38	0	0	0
		R	Native [e]	1	1	0	0
		S	Native x Brahman [e]	8	0	0	0
			Native [e]	21	1	0	0
		T	Native x Brahman [e]	4	0	0	0
			Native [e]	2	0	0	0
		U	Native x Holstein [d]	1	0	0	0
			Native [e]	3	0	0	0
	III	V	Brahman [e]	51	0	0	0
			Sahiwal x Brahman [d]	4	0	0	0
Total Number of Samples				372	4	0	0

[d] Dairy cattle ; [e] Beef cattle;

Table 6 The BLV detection results for Iloilo Island as determined by BLV-CoCoMo-qPCR-2 and nested PCR

Island	City / Municipality	Farm	Breed	Sample Number	Number of BLV positive samples		
					CoCoMo - qPCR	Nested-PCR	Both
Iloilo	I	A	Native [e]	19	3	0	0
			Native x Holstein [d]	20	0	0	0
			Native x Holstein x Sahiwal [d]	2	0	0	0
		B	Native [e]	46	1	1	1
			Native x Brahman [e]	3	0	0	0
		C	Native [e]	15	0	0	0
			Native x Brahman [e]	3	0	0	0
		D	Native [e]	17	1	0	0
			Native x Brahman [e]	7	0	0	0
			Native x Holstein [d]	1	0	0	0
			Native x Holstein x Brahman [d]	1	0	0	0
		E	Native [e]	5	0	0	0
			Native x Holstein [d]	2	1	0	0
		F	Native [e]	18	3	2	2
Total Number of Samples				159	9	3	3

[d] Dairy cattle
[e] Beef cattle

Table 7 The BLV detection results for Luzon Island as determined by BLV-CoCoMo-qPCR-2 and nested PCR

Island	City / Municipality	Farm	Breed	Sample Number	Number of BLV positive samples		
					CoCoMo-qPCR	Nested-PCR	Both
Luzon	I	A	Native x Brahman [e]	8	0	0	0
		B	Native x Brahman [e]	23	0	0	0
		C	Native x Brahman [e]	13	0	0	0
	II	D	Holstein x Sahiwal [d]	17	6	6	6[a]
	III	E	Native x Brahman [e]	4	0	0	0
	IV	F	Holstein x Sahiwal [d]	31	5	3	3[b]
	V	G	Holstein x Sahiwal [d]	105	50	37	37[c]
	VI	H	Brahman[e]	39	7	0	0
			Native [e]	11	3	0	0
		I	Brahman[e]	31	5	3	3[b]
			Native [e]	1	1	1	1
		J	Brahma[e]	73	14	1	1
			Native [e]	13	2	0	0
		K	Brahman[e]	38	2	0	0
			Native [e]	4	0	0	0
Total Number of Samples				411	95	51	51

[a] Indicates two samples were sequenced
[b] Indicates three samples were sequenced
[c] Indicates 35 samples were sequenced
[d] Dairy cattle
[e] Beef cattle

Table 8 Summary of BLV detection by nested PCR for five South American countries

Country	City/Municipality	Farm	Breed	Positive % (+/All)[1]
Peru	Lima	A	Holstein	58.6 (61/104)
	Pucallpa	B	Brahman x Nelore	40.0 (78/195)
	Huancayo	C	Holstein	0.0 (0/29)
Paraguay	Asuncion	A	Holstein	71.4 (25/35)
		B	Holstein	60.0 (18/30)
		C	Holstein	61.3 (19/31)
		D	Holstein	33.3 (1/3)
		E	Holstein	28.6 (8/28)
			Brown Swiss	36.4 (4/11)
			Holstein x Brown Swiss	100.0 (1/1)
Bolivia	Yapacani	A	Yacumeño	27.7 (10/36)
		B	Yacumeño	19.0 (4/21)
		C	Yacumeño	4.5 (1/22)
			Brown Swiss	0.0 (0/1)
		D	Yacumeño	24.3 (7/29)
	Montero	E	Holstein	38.0 (19/50)
		F	Gir	18.5 (20/108)
			Nelore	20.0 (21/105)
	Portachuelo	G	Holstein	64.2 (70/109)
	Santacruz	H	Brahman	0.0 (0/3)
			Gir	0.0 (0/2)
			Holstein	100.0 (1/1)
			Montana	0.0 (0/2)
			Nelore	9.0 (1/11)
			Brahman x Nelore	28.6 (2/7)
Argentina	Buenos Aires	D	Holstein	90.9 (60/66)
		E	Holstein	79.4 (152/194)
		F	Holstein	70.6 (113/160)
Chile	Valdivia	A	Overo Colorado	12.0 (3/25)
		B	Chilean Wagyu	50.0 (1/2)
			Simmental	0.0 (0/1)
			Jersey	0.0 (0/1)
			Overo Colorado	0.0 (0/5)
			Overo Negro	0.0 (0/2)
			Red Angus	0.0 (0/2)
		C	Overo Colorado	100.0 (30/30)
	Mafil	D	Holstein	0.0 (0/4)
			Overo Colorado	3.1 (1/32)
		E	Holstein	30.0 (9/30)
		F	Hereford x Overo Colorado	30.0 (6/20)
	San Jose de la	G	Hereford x Jersey	91.7 (66/72)
	Mariquina	H	Holstein	50.0 (10/20)
	Panguipulli	I	Overo Colorado	7.7 (6/13)

		Overo Negro	30.0 (10/33)
	J	Overo Colorado	16.6 (1/6)
		Overo Negro	38.5 (20/52)
Rio Bueno	K	Holstein	12.1 (4/33)
	L	Holstein	39.3 (11/28)
Lago Ranco	M	Black Angus	0.0 (0/17)
		Red Angus	0.0 (0/7)
	N	Overo Colorado	10.5 (2/19)
La Union	O	Overo Negro	5.0 (1/20)
Futrono	P	Black Angus	41.7 (15/36)
		Red Angus	19.7 (12/61)
	Q	Chilean Wagyu	2.5 (2/79)
		Hereford	0.0 (0/1)
		Red Angus	12.5 (2/16)
	R	Hereford x Angus	14.7 (14/95)
Llefen	S	Hereford	20.8 (10/48)

[1](+/All) indicates the ratio of the number of BLV positive samples to that of all samples.

Table 9 Average sequencing depths for 25 aligned BAM files generated from 25 samples sequenced by MiSeq sequencer.

BAM file	Sample name	Average of depth
Portachuelo108-new	Portachuelo 108	317
Portachello-14b	Portachuelo 14	953
Portachello-2b	Portachuelo 2	1270
Portachello-20b	Portachuelo 20	1190
Portachello-25b	Portachuelo 25	1075
Portachello-28b	Portachuelo 28	1190
Portachello-46b	Portachuelo 46	858
Portachello-57b	Portachuelo 57	1321
Portachello-71b	Portachuelo 71	1094
Portachello-84b	Portachuelo 84	952
Portachello-87b	Portachuelo 87	1558
Portachello-93b	Portachuelo 93	1349
Monetro-41b	Montero41	895
Monetro-1b	Montero1	1248
Monetro-22b	Montero22	669
Monetro-28b	Montero28	818
Monetro-17b	Montero17	1013
Lima40G2	Lima40	396
Paraguay17G1	Paraguay17	595
Asuncion1-1G1	Asuncion1	718
Asuncion5-1G1	Asuncion5	425
Paraguay7G1	Paraguay7	495
Paraguay62-2G6	Paraguay62	364
Paraguay89-1G6	Paraguay89	1492
Paraguay91-1G6	Paraguay91	143

Table 10 Nucleotide and amino acid substitution rate for 25 novel BLV complete provirus and eight known BLV whole genome.

	LTR	p15	p24	p12	pro	pol	gp51	gp30	tax	rex	R3	G4	All
Nucleotide*1	0.013	0.023	0.030	0.022	0.025	0.023	0.023	0.027	0.023	0.016	0.027	0.013	0.023
Amino-acid*2	-	0.020	0.006	0.012	0.021	0.014	0.018	0.018	0.053	0.030	0.057	0.028	-

*1 Average nucleotide substitution per site calculated using the Jukes-Canter model.

*2 Average amino acid substitution per site calculated using the p-distance model.

Table 11 Primers used for sequencing of full-length gp51 of Myanmar strains

gp51 full-length sequence primers		
Primers	Binding position	Sequences
pBLV-F0	4611-4631	5`-AGATGGGAGCTACACCATTCA-3`
gp51-R-5639	5613-5634	5`-AWCAACAACCTCTGGGAAGGGT-3`
pBLV-F1	5812-5830	5`-TCAGAGACTCACCTCCCTG-3`
pBLV-0R	6055-6076	5`-GTCTGTAGAGACTCTTTGCGAG-3`

Note: binding position refers to pBLV-IF sequences

Table 12 Primers used for sequencing the complete genome sequences of Myanmar BLV strains

Myanmar BLV complete genome sequence primers		
Primers	Binding position	Sequences
pBLV-R6	1294-1315	5`-GGCCTGAAGCCAGAGGTTTTGA-3`
pBLV-R5	1970-1989	5`-GGTGTCCACAAGCATGAGGG-3`
pBLV-R4.5	2499-2516	5`-CGGAATGGGCTTTGTAAG-3`
pBLV-R3.5	2996-3013	5`-ATTGGGGATGAGATCTGC-3`
pBLV-R2.5	3560-3577	5`-GCCCTGGTGATTAAGGTC-3`
pBLV-R1.5	4191-4208	5`-ATGGGTTATATCGGCCTG-3`
pBLV-R1	4811-4828	5`-CATTTGAAGGCTTTCAGC-3`
pBLV-1st LTR-R	1099-1115	5`-GGTCAGCCGGAGTAGGG-3`
pBLV-F0	4611-4631	5`-AGATGGGAGCTACACCATTCA-3`
gp51-R-5639	5613-5634	5`-AWCAACAACCTCTGGGAAGGGT-3`
pBLV-F1	5812-5830	5`-TCAGAGACTCACCTCCCTG-3`
pBLV-0R	6055-6076	5`-GTCTGTAGAGACTCTTTGCGAG-3`
pBLV-1.5F	6314-6331	5`-ATCTACTCTCACCTCTCC-3`
pBLV-2.5F	6801-6818	5`-TTTACGCCCTGTTGCACA-3`
pBLV-3F	7410-7429	5`-ATCAACTGGACCGCCGATGG-`3
T7 promoter*		5`-TAATACGACTCACTATAGG-3`

Note: primer binding position refers to pBLV-IF sequence

* indicates T7 promoter sequence of pGEM-T easy vector

Table 13 BLV-detection result as determined by Nested PCR and CoCoMo-qPCR-2 targeting LTR, and Nested PCR targeting *gp51*,and real time PCR for *tax* in Myanmar

Sampling place	LTR		*tax*	*gp51*	Final result
	Nested PCR	CoCoMo-qPCR-2	Real time-PCR	Nested PCR	
Lewe Township	2/35[a]	2/35[a]	1/2[a]	2/2[b]	2/35[b] (5.7)[c]
Zeyathiri township	0/7	0/7	0/0	0/0	0/7 (0)
Tatkone township	3/4	3/4	1/3	2/3	3/4 (75.0)
Pyinmana township	0/20	1/20	1/1	1/1	1/20 (5.0)

[a]BLV-positive animal number/tested animal number
[b]BLV-positive sample number /animal number detected as positive by nested PCR and CoCoMo-qPCR-2
[c]BLV-positive sample number detected by at least two method/total sample number
[d](%) shows the percentages of positive animals

Figures and Figure legends

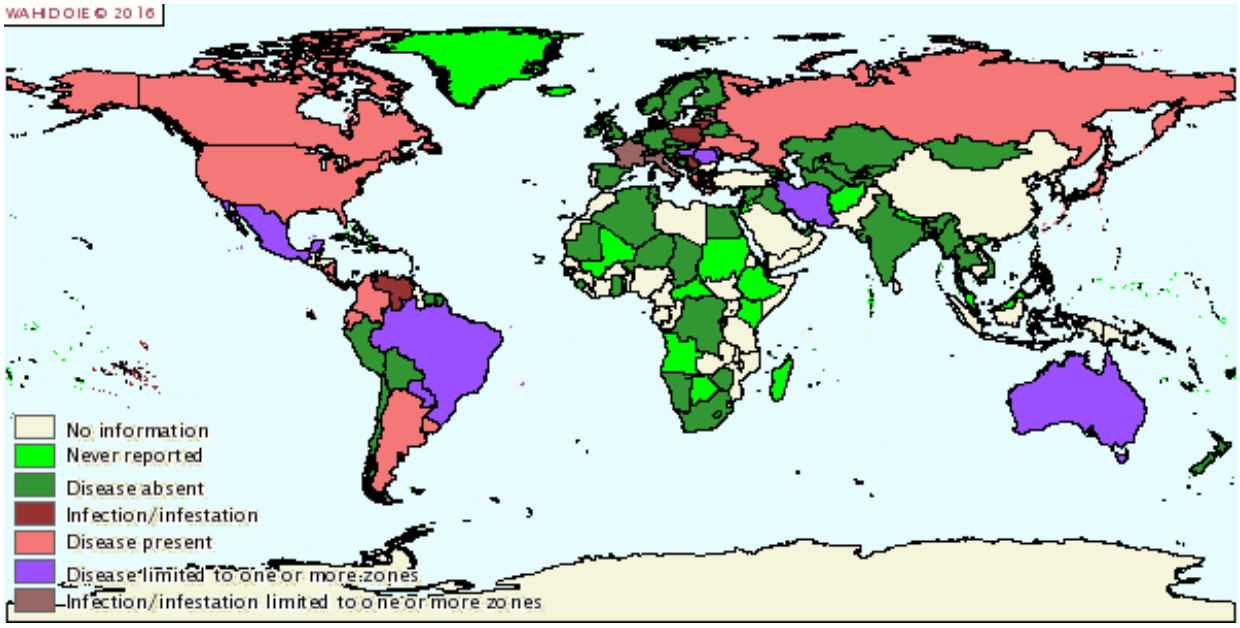


Figure 1 The worldwide prevalence of enzootic bovine leukosis.

This figure is obtained from World Animal Health Information system (WAHIS) interface 2015. (http://www.oie.int/wahis_2/public/wahid.php/Diseaseinformation/Diseasedistributionmap?disease_type_hidden=&disease_id_hidden=&selected_disease_name_hidden=&disease_type=0&disease_id_terrestrial=35&species_t=0&disease_id_aquatic=999&species_a=0&sta_method=semesterly&selected_start_year=2015&selected_report_period=1&selected_start_month=1&date_submit=OK) (Opened on June 1st)

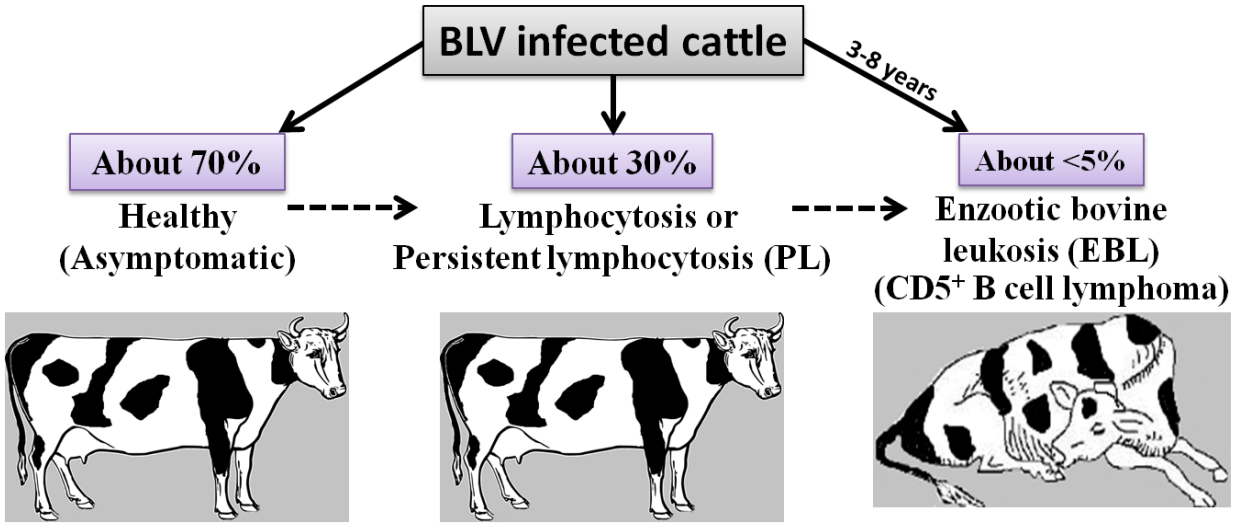


Figure 2 Schematic representation of disease progression in BLV-infected cattle .

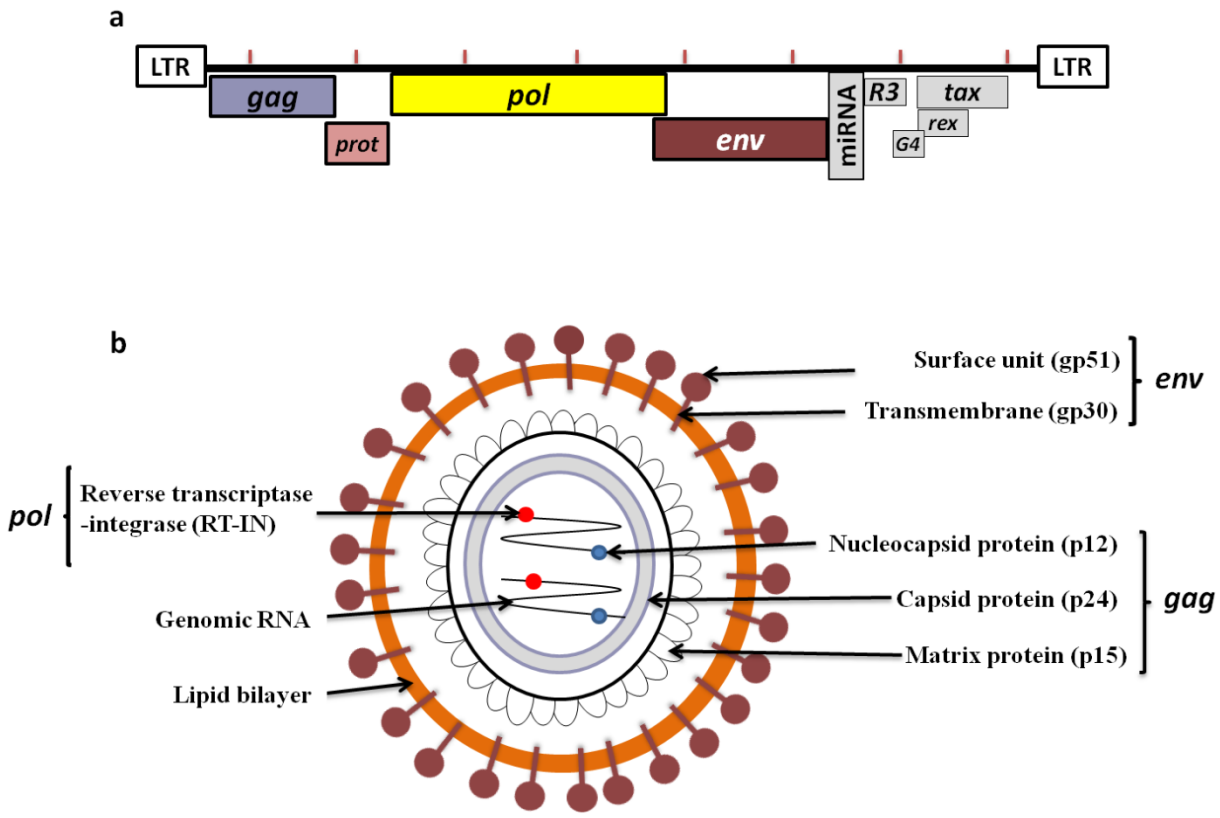


Figure 3 Schematic representations of BLV genome structure (a) and viral particle (b).

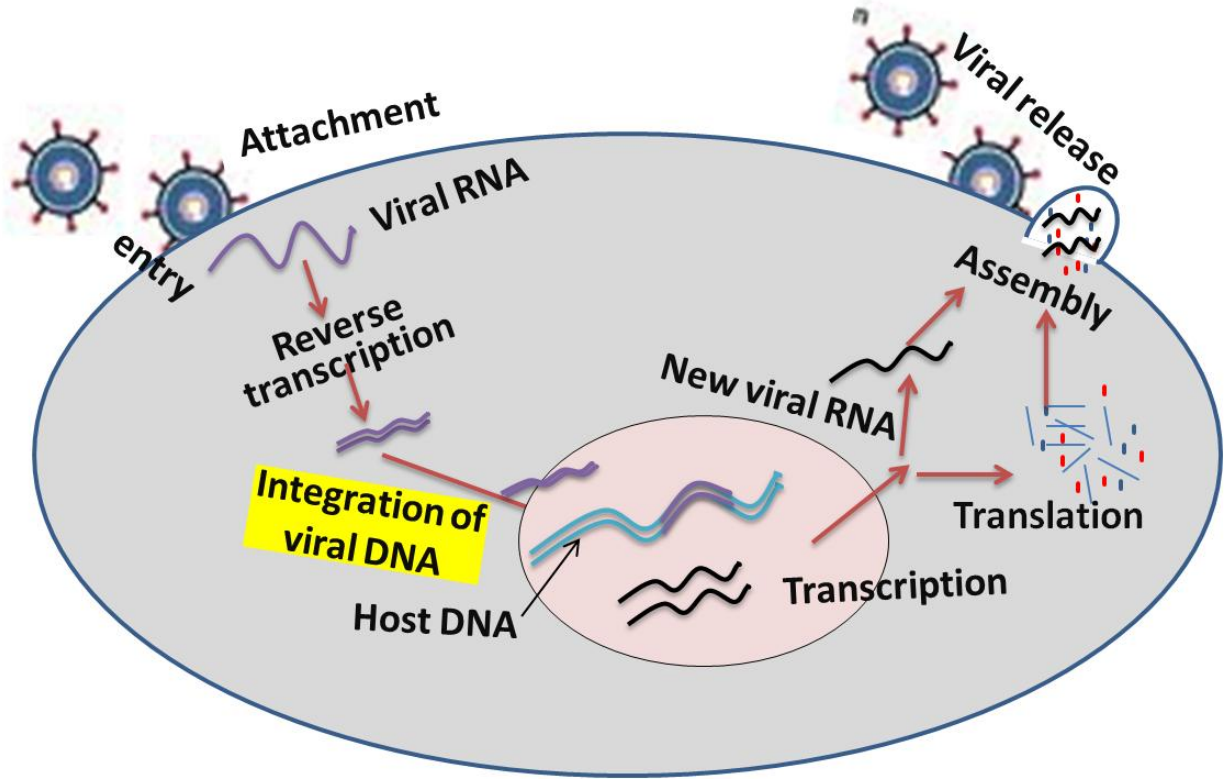


Figure 4 Lifecycle of BLV

The brief description of BLV lifecycle is: attachment of the virus to its target cell; fusion of virus and cellular membrane; viral uncoating and reverse transcription; proviral integration; viral transcripts; viral assembly and viral release; maturation into infectious viral particles

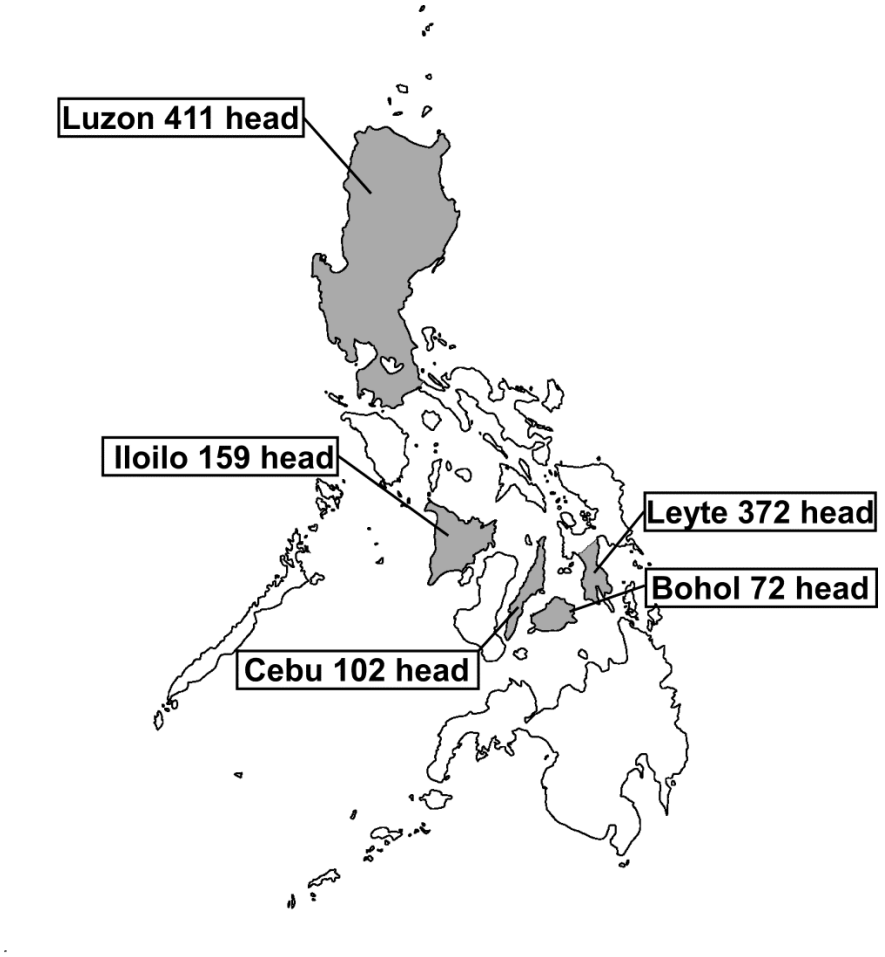


Figure 5 Map of the Philippine islands showing the number of cattle from islands included in the study. The five islands where sampling was performed are indicated and colored with gray in the figure.

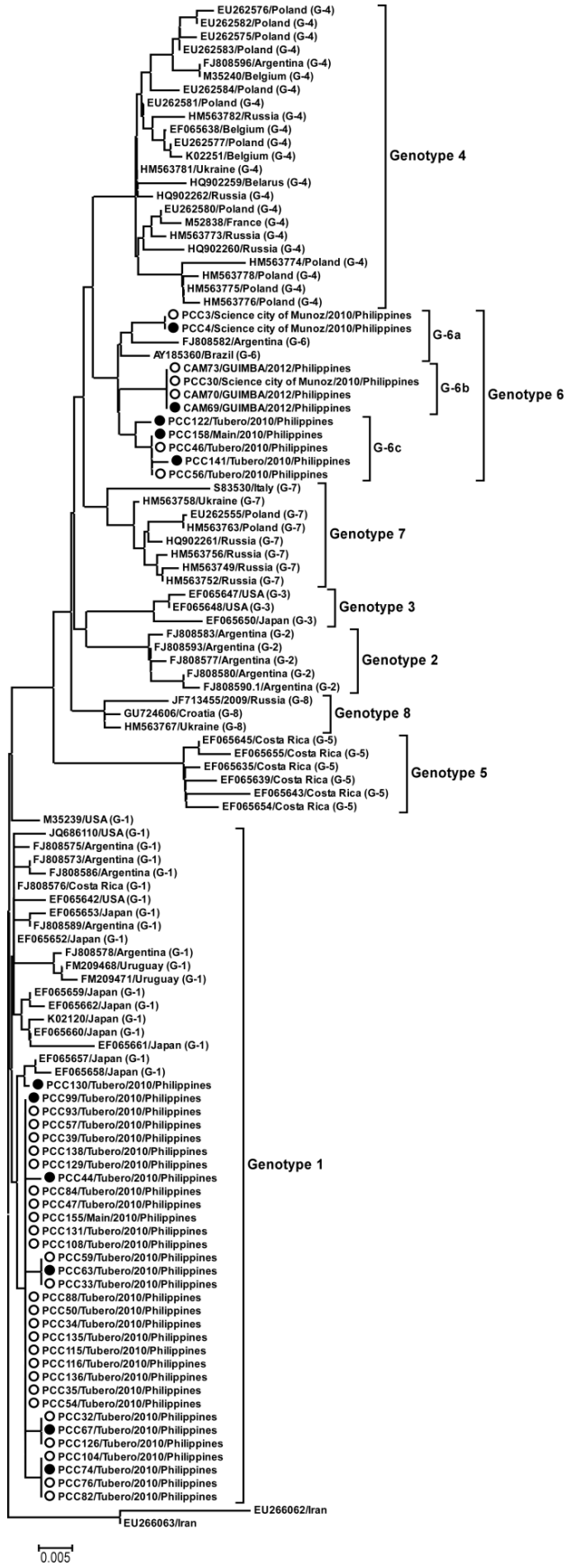


Figure 6 Neighbor-joining phylogenetic tree based on 423-bp nucleotide sequences of *env* genes from 43 BLV strains in the Philippines and BLV strains isolated elsewhere. Philippine BLV strains are indicated by the sample ID together with the collection time and country name. The remaining isolates in the tree are indicated by accession number and country of origin. Philippine BLV strains are marked by open (○) and filled (●) circles. Typical BLV isolates from identical samples were indicated by filled circles (●). Genotypes are indicated by the numbers on the right of the figure. The bar at the bottom of the figure denotes distance. Genotypes are indicated by the numbers on the right of the figure. The bar at the bottom of the figure denotes distance.

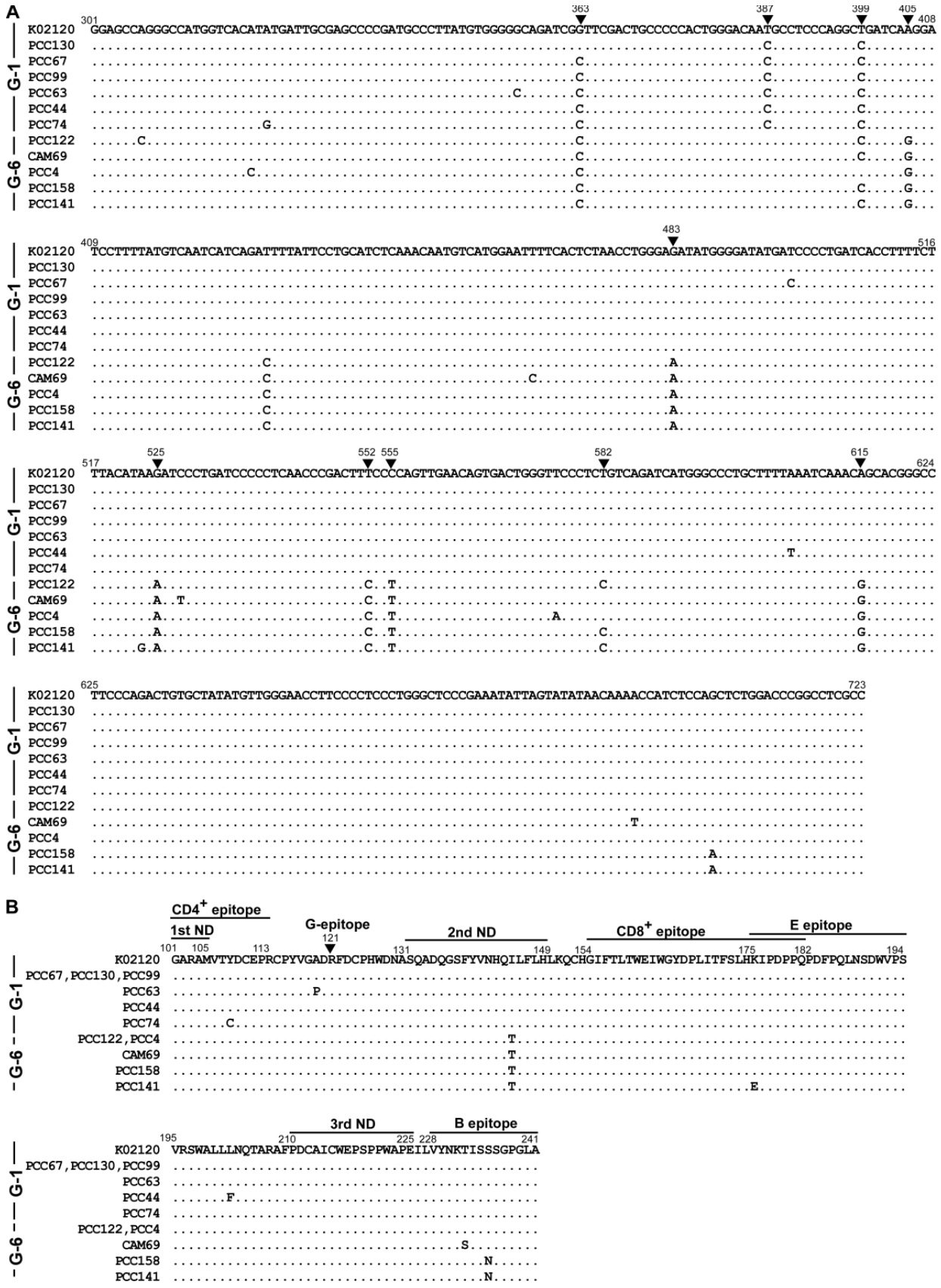

A
G-1
G-6
K02120
PCC130
PCC67
PCC99
PCC63
PCC44
PCC74
PCC122
CAM69
PCC4
PCC158
PCC141
B
CD4+ epitope
1st ND
G-epitope
2nd ND
CD8+ epitope
E epitope
3rd ND
B epitope
PCC67,PCC130,PCC99
PCC63
PCC44
PCC74
PCC122,PCC4
CAM69
PCC158
PCC141

Figure 7 Alignment of a partial nucleotide sequence (A) and deduced amino acid sequence (B) of the *env* gene of BLV strains in the Philippines. One typical sequence from identical sequences of BLV strains was submitted to GenBank (accession numbers: KJ668809 (PCC130), KJ668810 (PCC67), KJ668811 (PCC99), KJ668812 (PCC63), KJ6688113 (PCC44), KJ668814 (PCC74), KJ668815 (PCC122), KJ668816 (CAM69), KJ668817 (PCC4), KJ668818 (PCC158), and KJ668819 (PCC141). Alignments were performed for 11 typical nucleotide sequences (A) and eight unique deduced amino acid sequences (B) of the *env* gene from 43 Philippine BLV strains. Philippine BLV strains are shown by the sample ID. Nucleotide silent substitutions were showed by numbers and marked by dark rectangular mark (▾). The first, second and third neutralizing domain (ND), and other epitopes, are shown on the top of the alignment in B. Numbers above the sequences are amino acid residue numbers that indicate the start and end of each domain. Genotypes (G-1 or G-6) are indicated by the black bars in the far left of the figure. Dots indicate identity with K02120 (accession number: AY151262), used as a reference in this study.

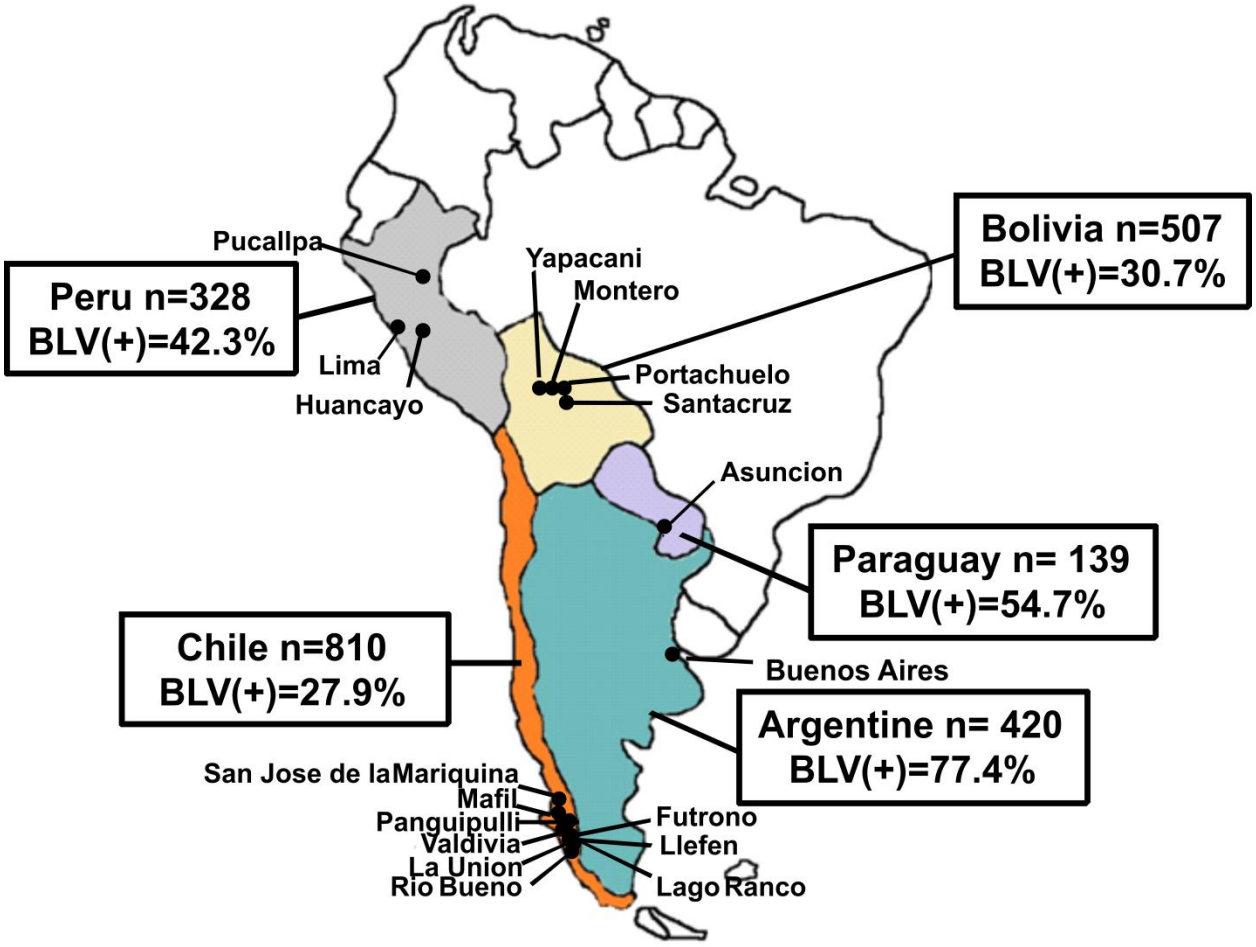
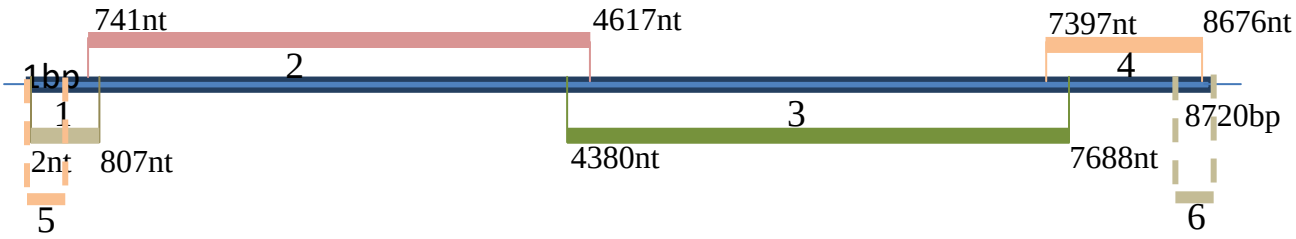


Figure 8 Map of South America showing the number of cattle (head) sampled from countries included in the study. The five countries and cities/municipalities where sampling was performed are indicated in the figure. n= indicates the total number of samples in each country and BLV (+) =% indicates the BLV prevalence rate.



Amplified position		Primers (Binding position)	Sequence	Length of amplicon
1		LRT-F (2-23nt; 8201-8222nt)	5`-GTATGAAAGATCATGCCGACCT-3`	806 bp
		807-R (781-807n t)	5`-CCAGAAGTGAAAGTCCATGGTT-3`	
2	1st PCR (External)	94F (94-113nt)	5`-GGCTAGAATCCCCGTACCTC-3`	5760 bp
		5853R (5834-5853nt)	5`-TGATCTTGCTCCAGAACGTG-3`	
	2nd PCR (External)	741F (741-765nt)	5`-CCGATTTAAAGAATTACATCCATTG -3`	3877 bp
		4617R (4601-4617nt)	5`-CCCATCTGGTCTTTAGAATTGG-3`	
3		4380F (4380-4403nt)	5`-AACTACACCTCCAAAACCTTTGTC-3`	3309 bp
		4380-G2-F (4380-4403nt)*	5`- AACTACACCTCTAGAACCTTTGCC-3`	
		7688R (7665-7688nt)	5`-GAGGACAGGATGCGTTACTAAGTT-3`	
4		7397F (7397-7416nt)	5`-CGAGACCCACCGTATCAACT-3`	1280 bp
		LRT-R (455-477nt; 8654-8676 nt)	5`- GCGAGAAACAGAAAGTAAGACAGG-3`	

Figure 9 Strategy and primers used for the amplification of full-length BLV genomes of South American strains. A scheme on the top shows the strategy for amplifying the BLV genome in four PCRs (1-4) using appropriate primers shown in the table. The full length of BLV complete genomes were shown as numbers in the beginning and ending of the scheme. The binding positions of the primers were also shown in the table. Binding position refers to Bovine leukemia virus cell-line FLK-BLV sub-clone pBLV913 complete genome (Accession number: EF600696). * indicates the forward primer used specifically for the amplification of the fragment from 4380 nt to 7688 nt of South American BLV genotype-2 strains. Numbers with nt indicates the PCR amplification region shown as the starting and ending nucleotide position. Square dot box 5 and 6 indicate the sequences of the region were obtained from NGS.

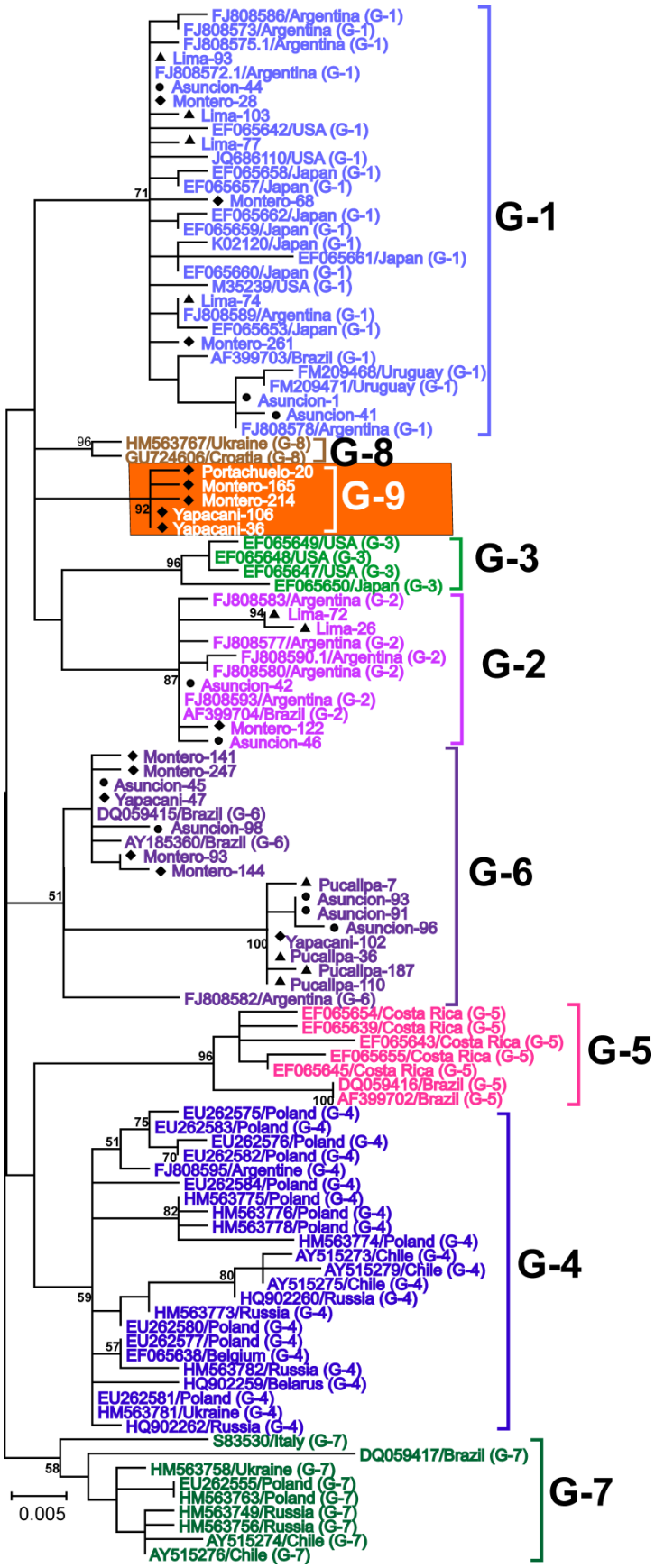


Figure 10 Maximum likelihood phylogenetic tree based on partial sequences of *env* gene of BLV strains from different geographical locations worldwide. A Maximum Likelihood phylogenetic tree was constructed from 35 distinct 475 bp BLV *env* sequences (submitted to the GenBank nucleotide sequence database and assigned accession numbers LC075543-LC075577) and with 74 sequences from known BLV strains, representing eight different BLV genotypes from other South American countries, including Argentina, Brazil, Chile, and Uruguay, and other parts of the world. These 74 nucleotide sequences were obtained from the GenBank nucleotide sequence database. The South American BLV strains in this study are indicated by sample ID and location of collection. Other isolates are shown in the tree by accession number and country of origin. South American BLV strains in this study are marked according to the country of origin as follows: Peru ▲, Paraguay ●, and Bolivia ◆. Genotypes are indicated by numbers to the right of the figure. Genotypes-1, -2, -3, -4, -5, -6, -7, -8 and -9 are shown in pastel blue, violet, sea green, blue, pink, twilight violet, avocado green, brown and orange, respectively. The bar at the bottom of the figure denotes distance.

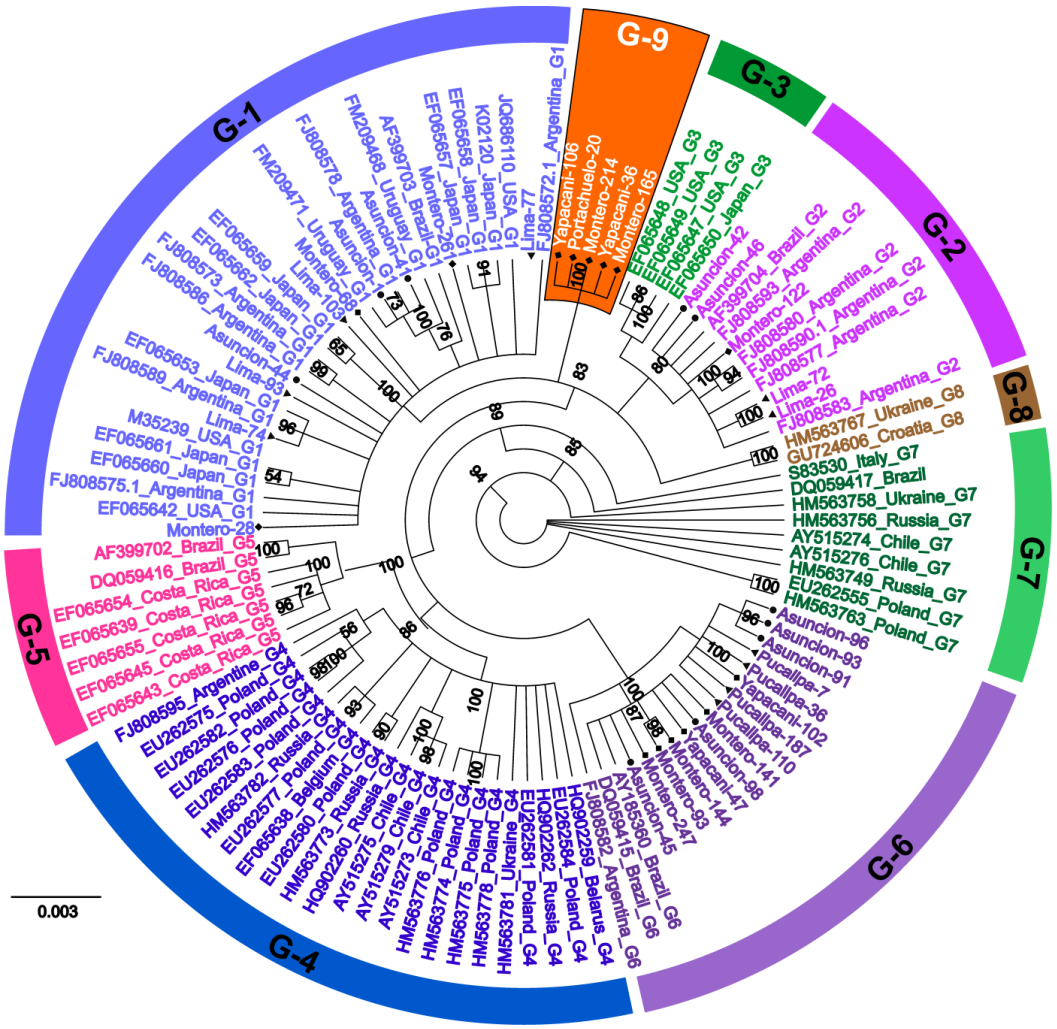


Figure 11 Bayesian phylogenetic tree of BLV *env* partial sequences from different geographical locations worldwide. A Bayesian phylogenetic tree was constructed from 35 distinct 475 bp BLV *env* sequences generated in this study, together with 74 sequences from known BLV strains, representing eight different BLV genotypes from other South American countries, including Argentina, Brazil, Chile and Uruguay, and other parts of the world. Strains in the tree are indicated by accession number and country of origin. South American BLV strains from this study are marked by country of origin as follows: Peru ▲, Paraguay ●, and Bolivia ◆. Genotypes are indicated by numbers around the circumference of the figure. Genotypes-1, -2, -3, -4, -5, -6, -7, -8, and -9 are shown in pastel blue, violet, sea green, blue, pink, twilight violet, avocado green, brown and orange, respectively. Numbers at nodes indicate posterior probabilities. The bar at the bottom of the figure denotes distance.

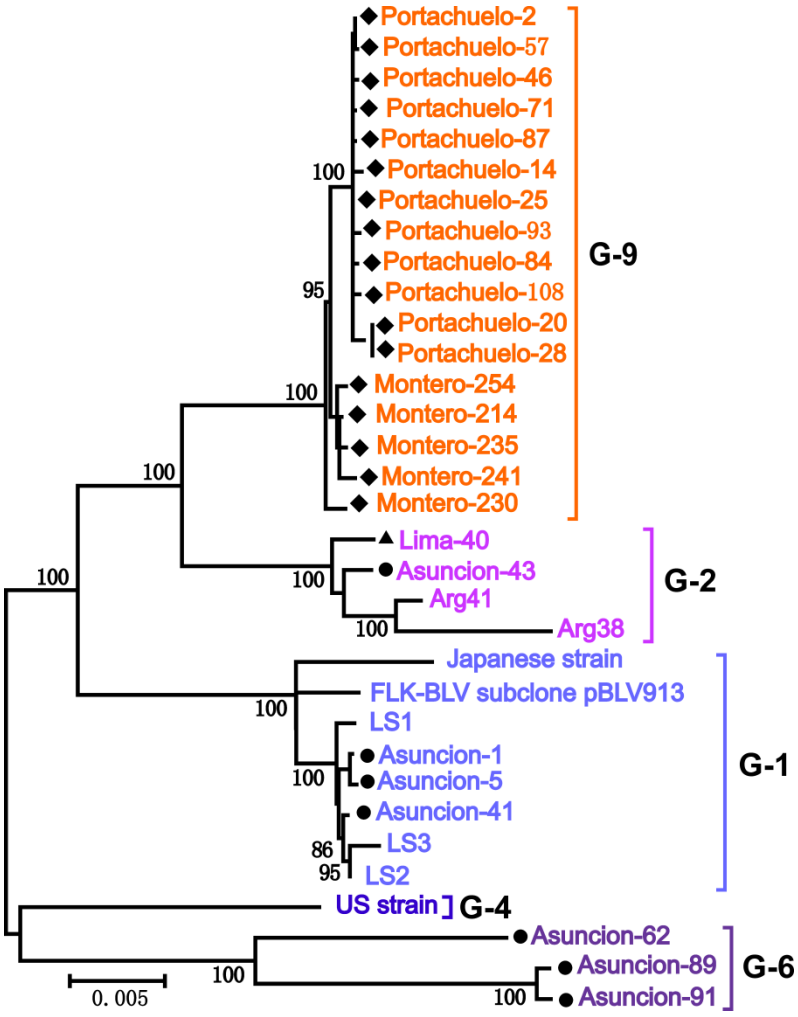


Figure 12 Maximum likelihood phylogenetic tree constructed using complete BLV genomic sequences. A Maximum likelihood phylogenetic tree was constructed from complete genomic sequences of BLV strains from a total of 25 samples (submitted to GenBank nucleotide sequence database and assigned accession numbers LC080664, LC080670, LC080669, LC080671, LC080673, LC080665, LC080667, LC080674, LC080672, LC080675, LC080666, LC080668, LC080663, LC080659, LC080661, LC080662, LC080660, LC080654, LC080655, LC080651, LC080652, LC080653, LC080656, LC080657, LC080658) including one sample from Peru, seven from Paraguay and 17 from Bolivia, together with eight reference sequences obtained from the GenBank nucleotide sequence database (whole genomes of FLK-BLV pBLV913, LS1, LS2, LS3, Arg41, Arg38, Japan, and USA with accession numbers EF600696, HE967301, HE967302, HE967303, FJ914764, AF257515, K02120, and AF033818, respectively). Gaps/missing data were treated as partial deletions and, finally, 8,374 of 8,728 sites were used for the calculation.

One thousand replications were performed for calculation of bootstrap values (indicated on the tree). Strains from this study are indicated in the tree by the sample ID together with the location of collection. Reference sequences in the tree are indicated by strain name. Genotypes are indicated by numbers to the right of the figure. South American BLV strains in this study are marked according to country of isolation as follows: Peru ▲, Paraguay ●, and Bolivia ◆. Genotypes-1, -2, -4, -6, and -9 are shown in pastel blue, violet, blue, twilight violet and orange, respectively. The bar at the bottom of the figure denotes distance.

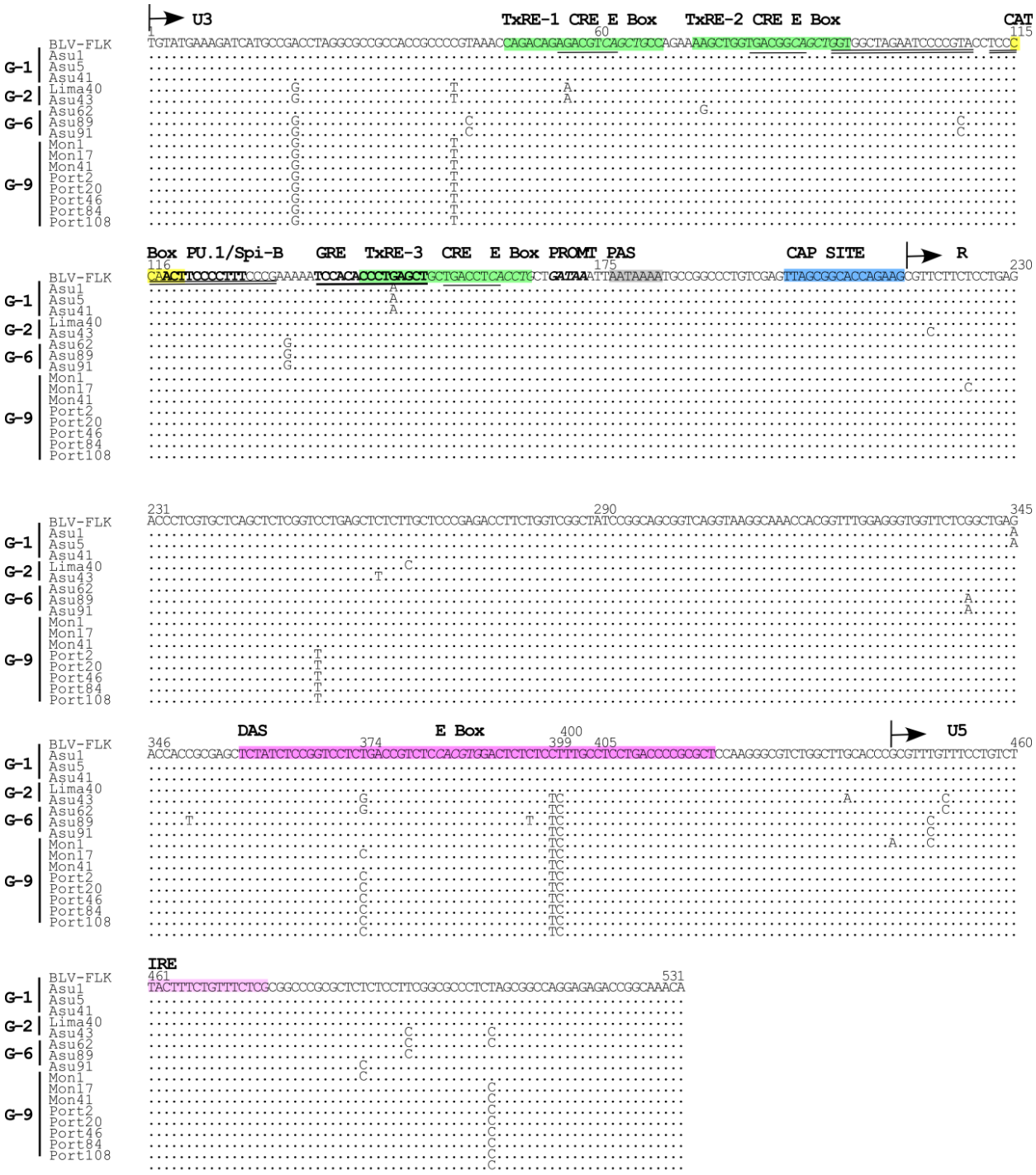

U3
TxRE-1 CRE E Box
TxRE-2 CRE E Box
CAT
Box PU.1/Spi-B
GRE
TxRE-3
CRE
E Box
PROMT
PAS
CAP SITE
R
DAS
E Box
U5
IRE
BLV-FLK
Asu1
Asu5
Asu41
Lima40
Asu43
Asu62
Asu89
Asu91
Mon1
Mon17
Mon41
Port2
Port20
Port46
Port84
Port108
G-1
G-2
G-6
G-9

**Figure 13 Alignment of 5`LTR nucleotide sequences of BLV strains from South America.** Alignment of 5`LTR nucleotide sequences from a total of 16 South American BLV strains, including one, seven, and eight samples from Peru, Paraguay, and Bolivia, respectively, together with that of the FLK-BLV strain pBLV913. Strains from South America are shown by location of sample collection and sample ID. Identity with the FLK-BLV strain is indicated by dots. The U3, R, and U5 regions are indicated above 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 and underlined. Nuclear factor κB (NF-κB) 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 grey and the CAP site is shown in blue. The downstream activator sequence (DAS) and the interferon regulatory factor (IRF) binding sites are shown in pink and light pink respectively.

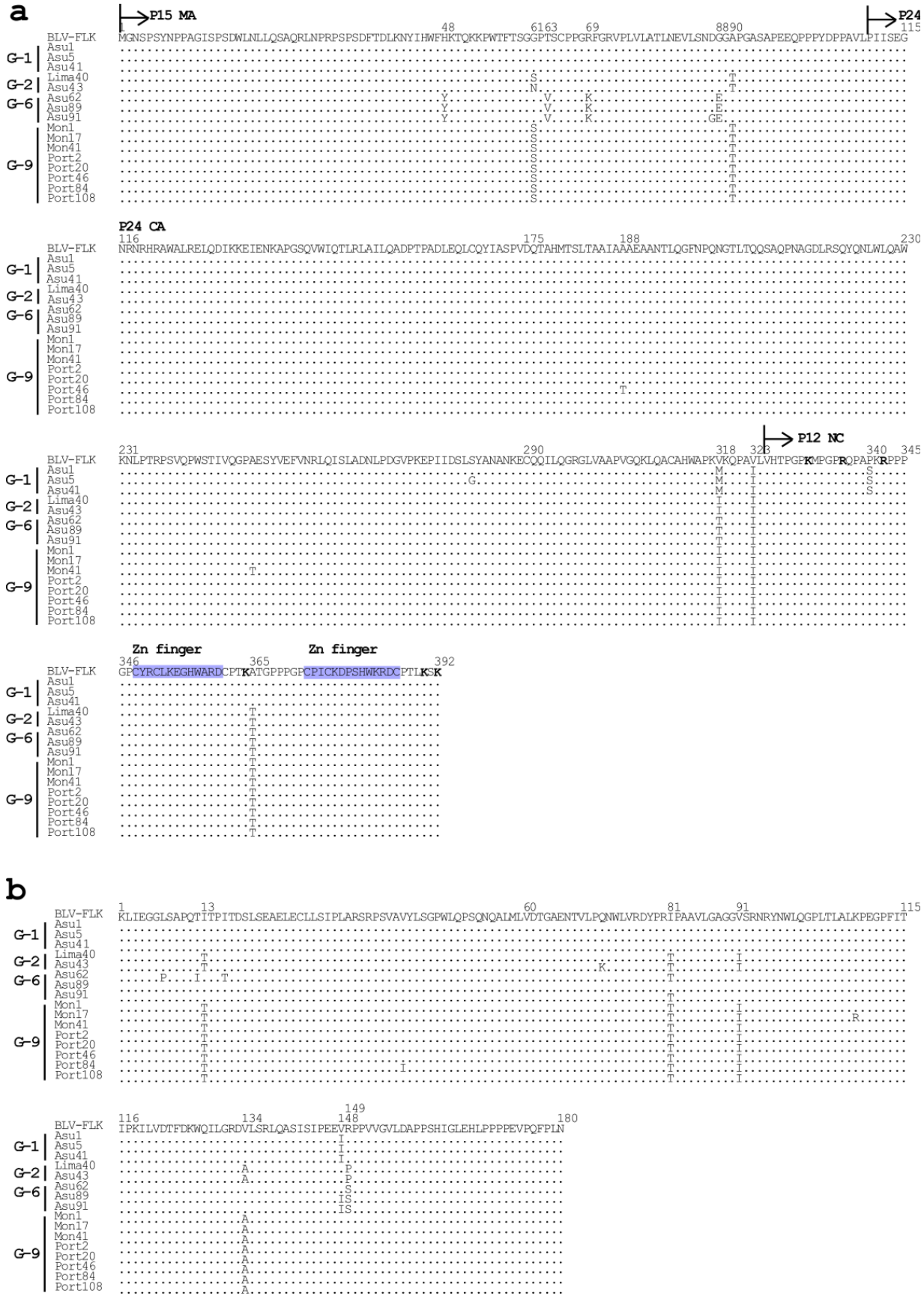

a
P15 MA
P24
P24 CA
P12 NC
Zn finger
Zn finger
b
BLV-FLK
Asu1
Asu5
Asu41
Lima40
Asu43
Asu62
Asu89
Asu91
Mon1
Mon17
Mon41
Port2
Port20
Port46
Port84
Port108
G-1
G-2
G-6
G-9

Figure 14 Alignment of deduced amino acid sequences of BLV *Gag and Protease (Pro)*. Alignment of amino acid residues corresponding to Gag polyprotein (a) and Pro protein (b) from a total of 16 South American BLV strains including one sample from Peru, seven from Paraguay, and eight from Bolivia, together with that of FLK-BLV strain pBLV913. Arrows designate the putative Matrix (p15MA), Capsid (p24CA) and Nucleocapsid (p12NC) protein sequences. Zinc finger domains are shown in purple. NC residues in boldface indicate basic residues. Identity with the FLK-BLV strain is indicated by dots.

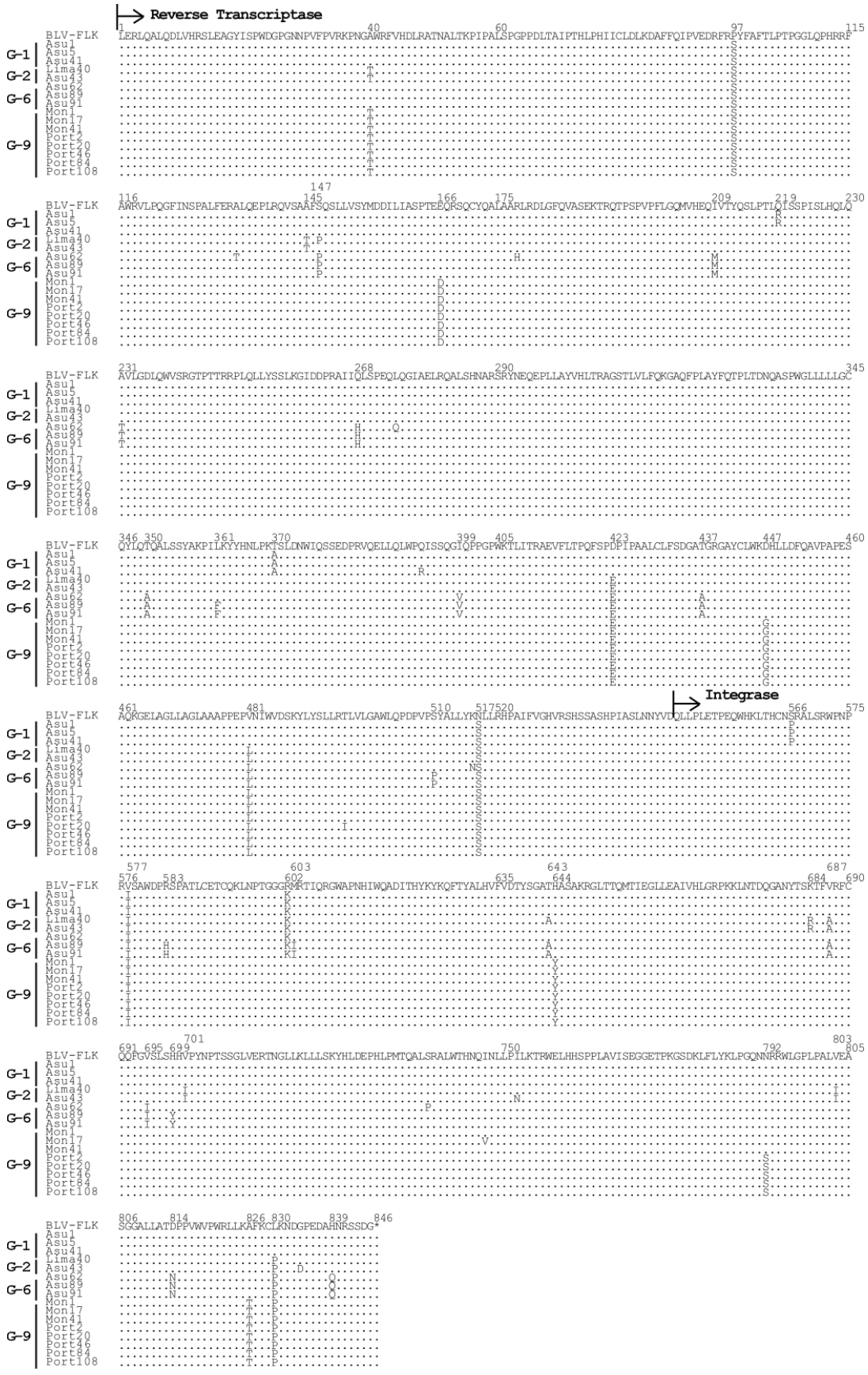

Reverse Transcriptase
Integrase
BLV-FLK
Asu1
Asu5
Asu41
Lima40
Asu43
Asu62
Asu89
Asu91
Mon1
Mon17
Mon41
Port2
Port20
Port46
Port84
Port108
G-1
G-2
G-6
G-9

Figure 15 Alignment of deduced amino acid sequences of BLV *Polymerase* (*Pol*) fragment. Alignment of amino acid residues corresponding to Pol proteins of a total of 16 South American BLV Strains, including one sample from Peru, seven from Paraguay and eight from Bolivia, together with that of the FLK-BLV strain pBLV913. Arrows designate the putative reverse transcriptase (RT) and integrase (IN) regions. Identity with the FLK-BLV strain is indicated by dots.

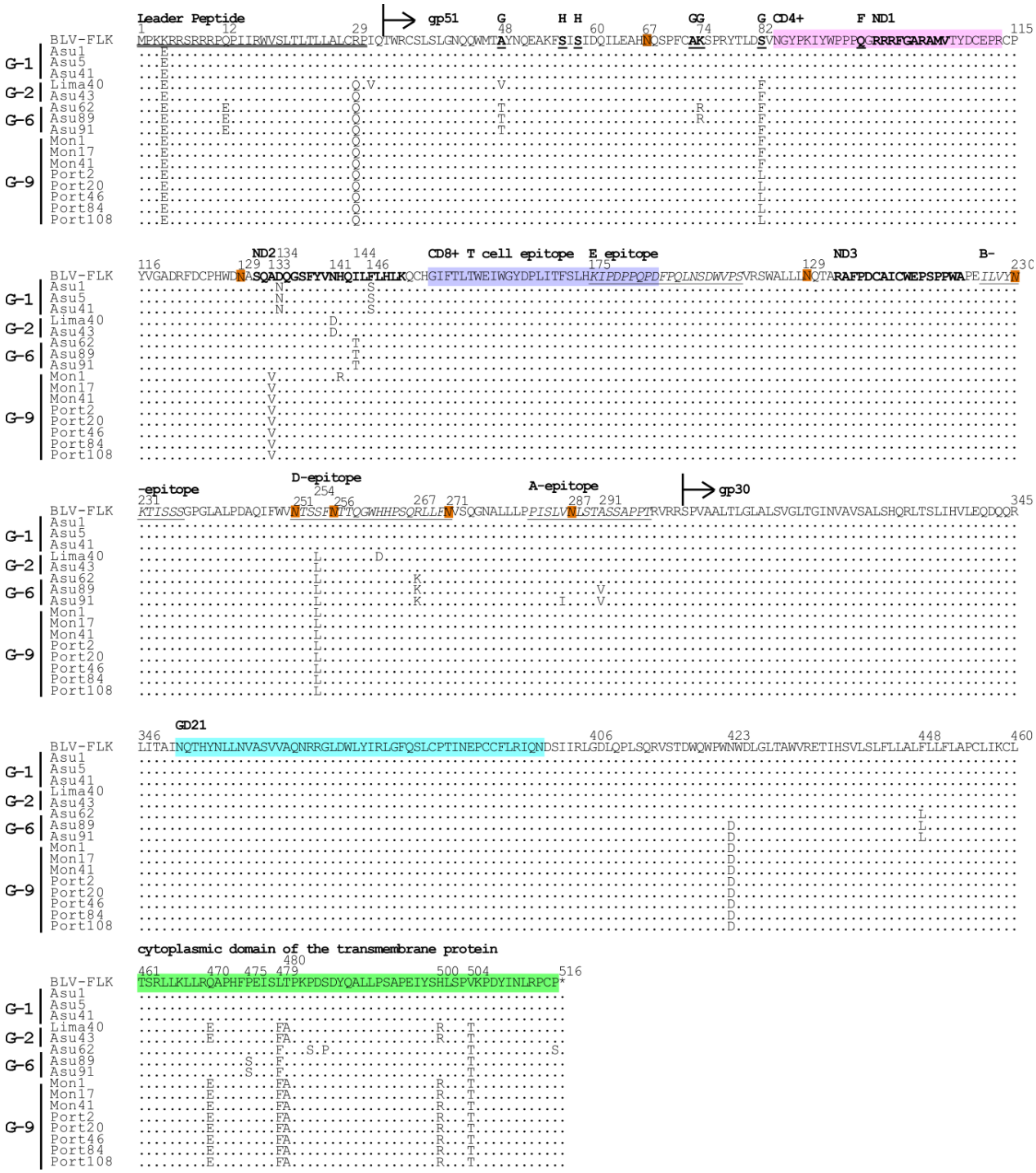

Leader Peptide
gp51
G
H H
GG
G CD4+
F ND1
BLV-FLK
Asu1
Asu5
Asu41
Lima40
Asu43
Asu62
Asu89
Asu91
Mon1
Mon17
Mon41
Port2
Port20
Port46
Port84
Port108
G-1
G-2
G-6
G-9
ND2
CD8+ T cell epitope
E epitope
ND3
B-
-epitope
D-epitope
A-epitope
gp30
GD21
cytoplasmic domain of the transmembrane protein

Figure 16 Alignment of deduced amino acid sequences of BLV *env* region. Alignment of amino acid residues corresponding to Env proteins from a total of 16 South American BLV strains, including one sample from Peru, seven from Paraguay and eight from Bolivia, together with that of the FLK-BLV strain pBLV913. Arrows designate the putative gp51 (SU) and gp30 (TM) proteins. The leader peptide is double underlined. The conformational epitopes, F, G and H are indicated in bold and underlined. Linear epitopes, A, B, D and E are shown in italics and underlined. Amino acids involved in neutralization domains are shown in bold. The $CD4^+$ T and $CD8^+$ T epitopes are shown in pink and purple, respectively. GD21 is shown in blue. The cytoplasmic domain is indicated in green. Identity with the FLK-BLV strain is indicated by dots. The eight N-glycosylation sites are shown in orange box.

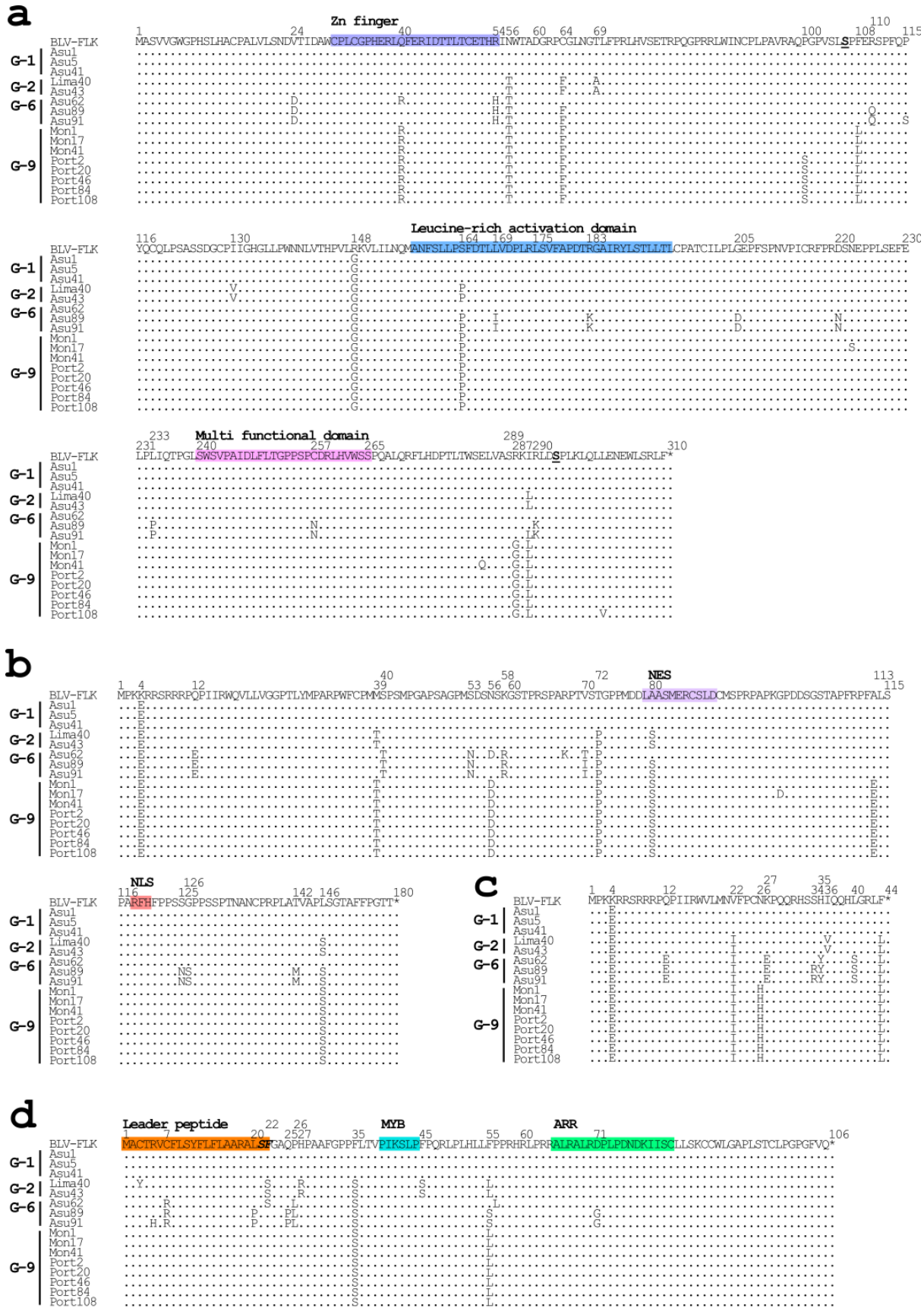

a
Zn finger
Leucine-rich activation domain
Multi functional domain
b
NES
NLS
c
d
Leader peptide
MYB
ARR
BLV-FLK
G-1
G-2
G-6
G-9
Asu1
Asu5
Asu41
Lima40
Asu43
Asu62
Asu89
Asu91
Mon1
Mon17
Mon41
Port2
Port20
Port46
Port84
Port108

Figure 17 Alignment of amino acid sequences of the BLV regulatory and accessory proteins. Alignment of amino acid residues corresponding to the regulatory proteins Tax (a) and Rex (b), and the accessory proteins R3 (c) and G4 (d) from a total of 16 South American BLV strains including one sample from Peru, seven from Paraguay and eight from Bolivia, together with that of the FLK-BLV strain pBLV913. Identity with the FLK-BLV strain is indicated by dots. (a) A putative zinc finger, leucine-rich activation, and multiple functional domains in the Tax protein are shown in purple, blue, and pink, respectively. Sites of phosphorylation in the Tax protein are indicated in bold, italics, and underlined. (b) The nuclear export signal (NES) and the nuclear localization signal (NLS) in the Rex protein are shown in light purple and light red, respectively. (d) The leader peptide, the myb-like motif (MYB), and the arginine-rich nucleus targeting RNA-binding region (ARR) are shown in orange, blue, and green, respectively. Two putative cellular protease cleavage sites are in italics and bold. Termination codons are shown by an asterisk.

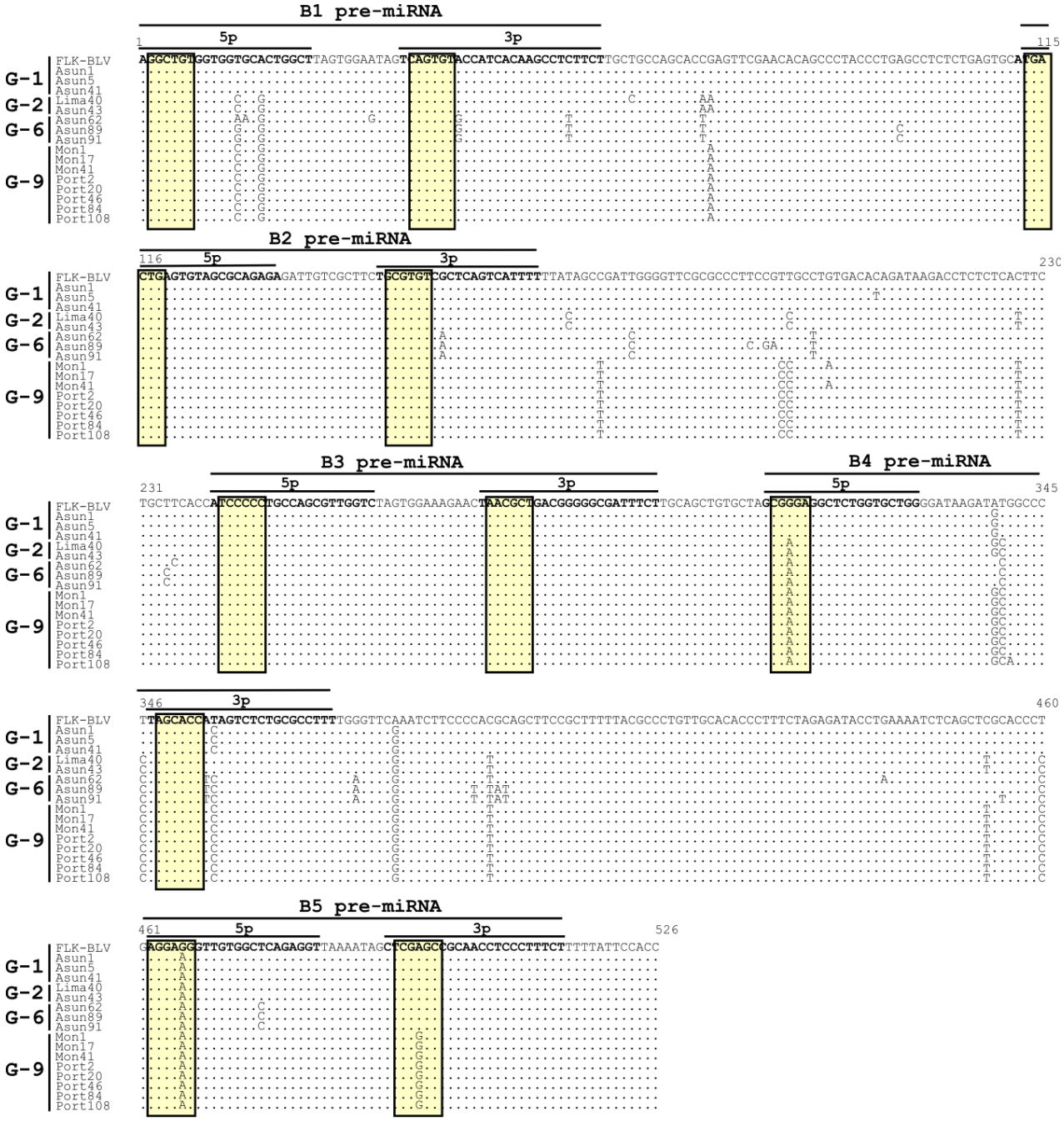


Figure 18 Alignment of nucleotide sequences of BLV miRNAs. Alignments of the nucleotide sequences of miRNAs encoded by BLV strains from 16 distinctive South American samples, including one from Peru, seven from Paraguay, and 8 from Bolivia, together with eight reference sequences from GenBank. miRNA seed regions are shown in yellow.

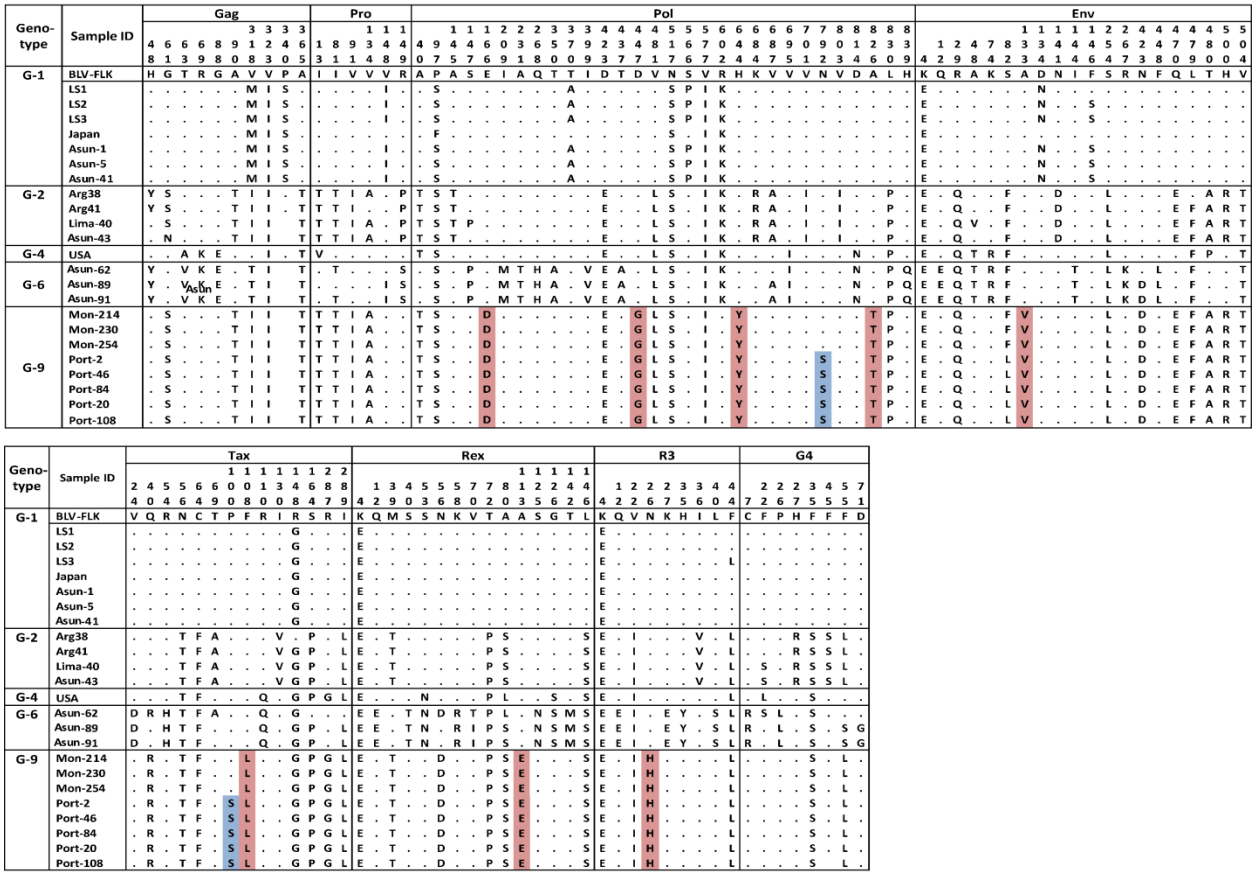
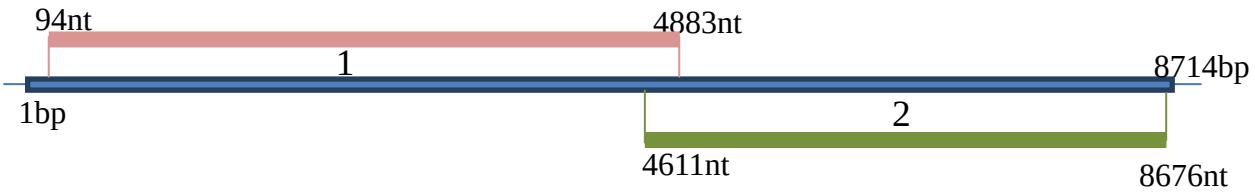


Figure 19 Summary of amino acid substitutions in sixteen novel and eight known BLV whole genome sequences. Amino acid substitutions identified in at least three samples are shown. Genotypes are indicated by numbers in the figure. Light red color indicates genotype-9-specific mutations detected in all samples, and light blue color indicates genotype-9-specific mutations detected only samples collected from Portachuelo in Bolivia.



Amplified position		Primers (Binding position)	Sequence	Length of amplicon
1	1st PCR (External)	BLV-LTR-F2 (2-23nt)	5`-GTATGAAAGATCATGCCGGCCT-3`	5830 bp
		5853R (5834-5853nt)	5`-TGATCTTGCTCCAGAACGTG-3`	
	2nd PCR (External)	BLV-LTR-F94 (94-113nt)	5`-GGCTAGAATCCCCGCACCTC -3`	4789 bp
		BLV-R4883 (4860-4883nt)	5`-CGAGGAGAGTGAGAGTGAGACTTA -3`	
2	*1*st PCR	pBLV-F0 (4611-4631nt)	5`-AGATGGGAGCTACACCATTCA-3`	4065 bp
		LTR-R (8654-8676nt)	5`-GCGAGAAACAGAAAGTAAGACAGG-3`	

Figure 20 Strategy and primers used for the amplification of full-length BLV genomes of Myanmar strains. A scheme on the top shows the strategy for amplifying the BLV genome in two separate PCRs (1-2) using appropriate primers shown in the table. The full length of BLV complete genome was shown as numbers in the beginning and ending of the scheme. The binding positions of the primers were also shown in the table. Binding position refers to Bovine leukemia virus cell-line FLK-BLV sub-clone pBLV913 complete genome (Accession number: EF600696). Numbers with nt indicates the PCR amplification region shown as the starting and ending nucleotide position.

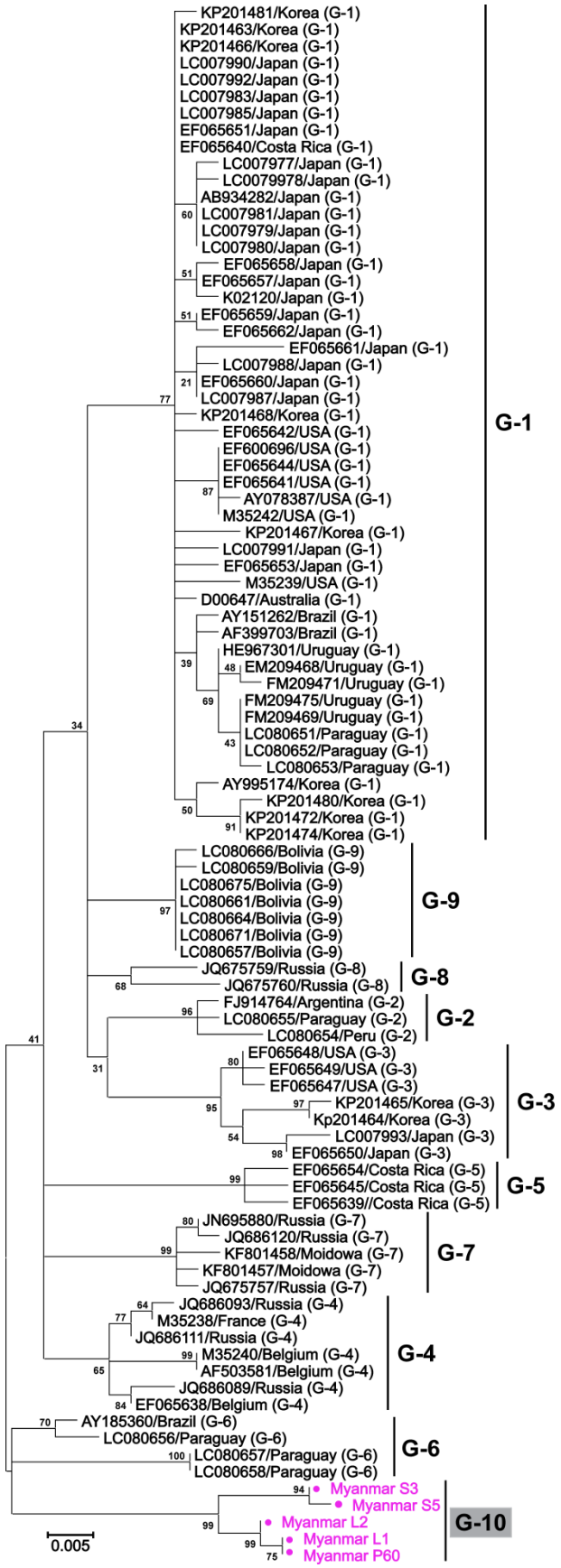


Figure 21 Maximum likelihood phylogenetic tree based on 475 bp of BLV *env-gp51* partial sequences from different geographical locations worldwide. A Maximum Likelihood phylogenetic tree was constructed from 475 bp BLV *env-gp51* partial sequences of five Myanmar strains and with 94 sequences from known BLV strains, representing eight different BLV genotypes from different locations of the world and new genotype-9 from Bolivia. These 94 nucleotide sequences were obtained from the GenBank nucleotide sequence database. The Myanmar BLV strains in this study are indicated by sample ID and country name of Myanmar. Other isolates are shown in the tree by accession number and country of origin. Myanmar BLV strains in this study are marked as ●. Genotypes are indicated by numbers to the right of the figure. Genotype 10 strains were highlighted by pink color and G-10 was highlighted by grey color. The bar at the bottom of the figure denotes distance.

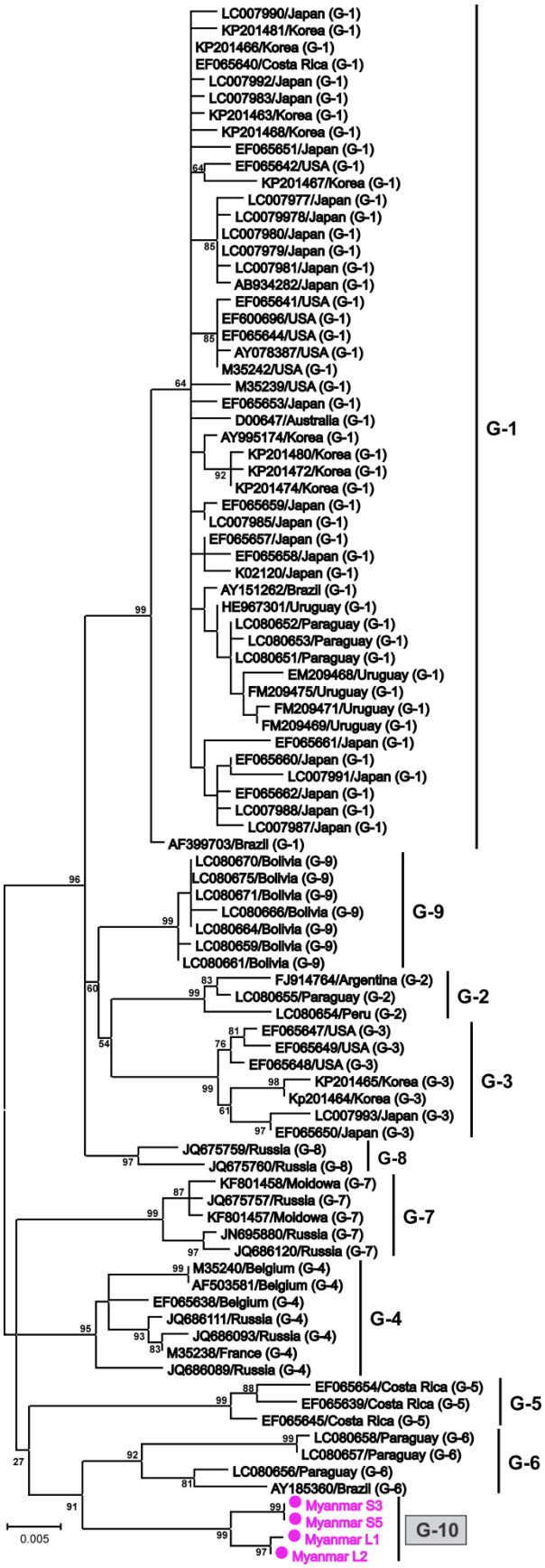


Figure 22 Maximum likelihood phylogenetic tree of BLV *env-gp51* full sequences from different geographical locations worldwide. A Maximum Likelihood phylogenetic tree was constructed from 807 bp BLV *env-gp51* full sequences of Myanmar strains (submitted to the GenBank nucleotide sequence database and assigned accession numbers LC154064~LC154067) and with 88 sequences from known BLV strains, representing eight different BLV genotypes from different locations of the world and new genotype-9 from Bolivia. These 88 nucleotide sequences were obtained from the GenBank nucleotide sequence database. The Myanmar BLV strains in this study are indicated by sample ID and country name of Myanmar. Other isolates are shown in the tree by accession number and country of origin. Myanmar BLV strains in this study are marked as ●. Genotypes are indicated by numbers to the right of the figure. Genotype 10 strains were highlighted by pink color and G-10 was highlighted by grey color. The bar at the bottom of the figure denotes distance.

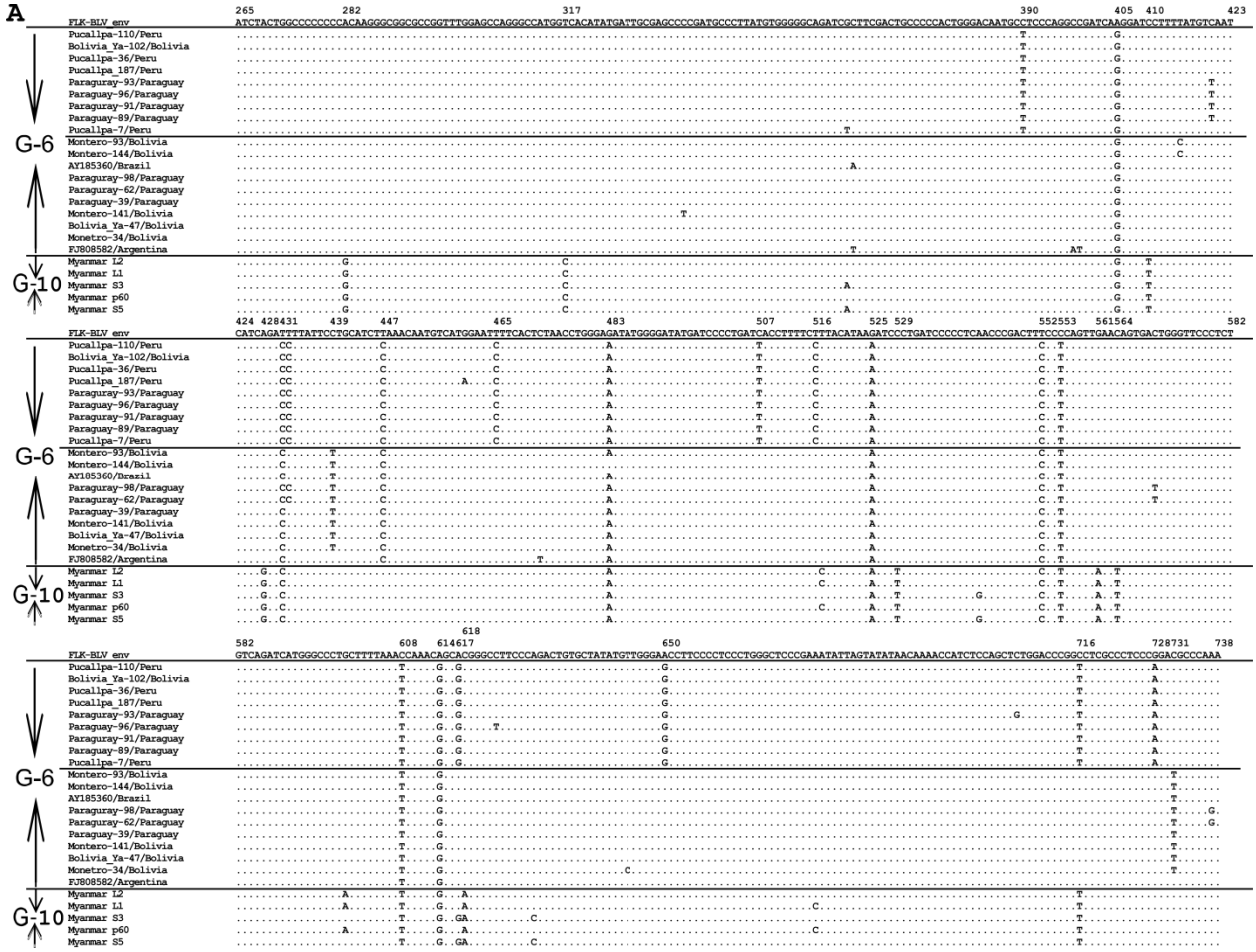

A
FLK-BLV env
Pucallpa-110/Peru
Bolivia_Ya-102/Bolivia
Pucallpa-36/Peru
Pucallpa_187/Peru
Paraguray-93/Paraguay
Paraguay-96/Paraguay
Paraguray-91/Paraguay
Paraguay-89/Paraguay
Pucallpa-7/Peru
Montero-93/Bolivia
Montero-144/Bolivia
AY185360/Brazil
Paraguray-98/Paraguay
Paraguray-62/Paraguay
Paraguay-39/Paraguay
Montero-141/Bolivia
Bolivia_Ya-47/Bolivia
Monetro-34/Bolivia
FJ808582/Argentina
Myanmar L2
Myanmar L1
Myanmar S3
Myanmar p60
Myanmar S5
G-6
G-10
B
CD4+Epitope
1st ND
F Epitope
G Eitope
2nd ND
CD8+ Epitope
E Epitope
3rd ND
B Epitope

Figure 23 Alignment of partial nucleotide sequences (A) and deduced amino acid sequence (B) of the BLV *env* gene from strains in Myanmar. Nucleotide (A) and deduced amino acid sequences (B) of five Myanmar BLV strains were aligned with 19 sequences of other genotype-6 strains from worldwide. The Myanmar BLV strains in this study are located lower part of alignment. Genotype is indicated by the black bars at the far left of the figure. Numbers on the top of sequences indicates the first and last nucleotide positions, and the position of nucleotide substitutions. Numbers above the deduced amino acid sequence are amino acid residue numbers that indicate the start and end of each domain. The first, second and third neutralizing domains (ND) and other epitopes are shown at the top of alignment. Dots indicate identity with FLK-BLV (accession number: EF600696), which was used as a reference in this study.

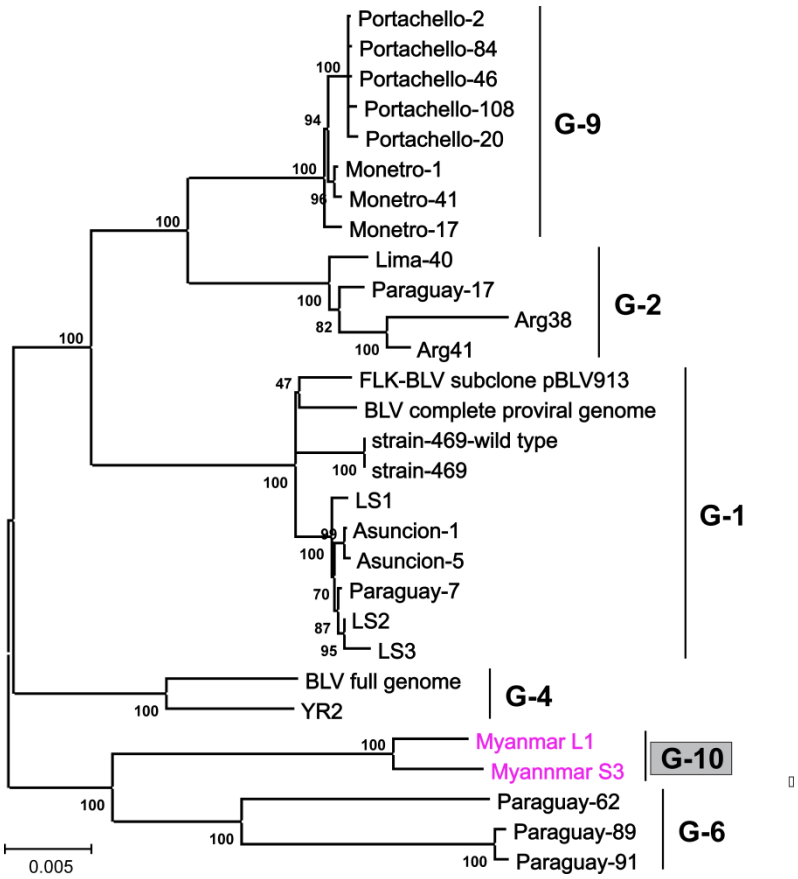
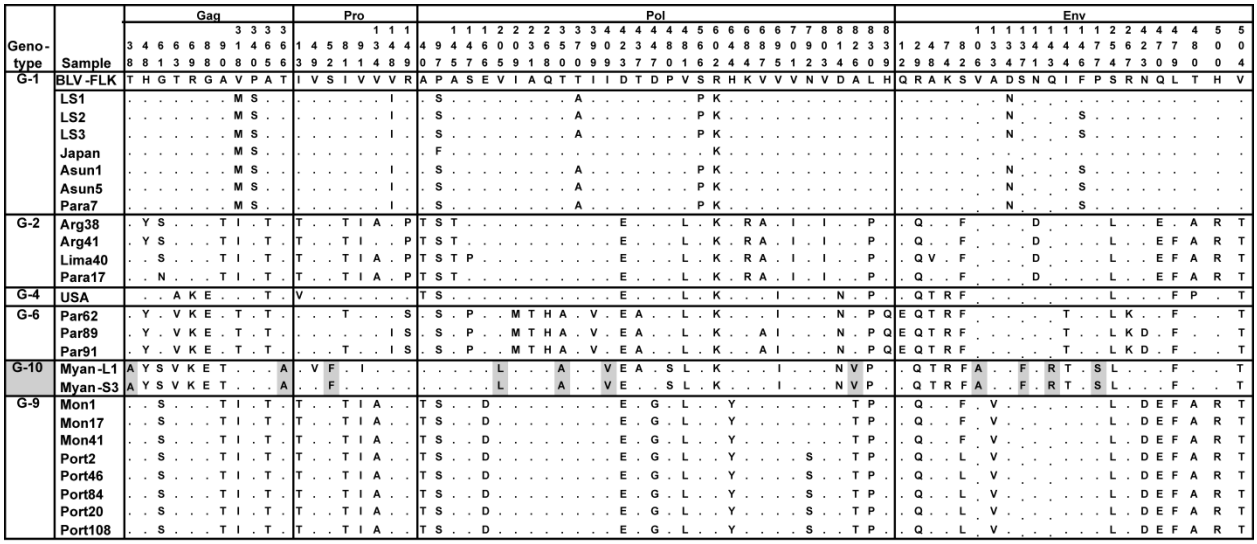


Figure 24 Maximum likelihood phylogenetic tree constructed using complete BLV genomic sequences. A maximum likelihood phylogenetic tree was constructed from complete genomic sequences of two BLV strains from Myanmar (submitted to GenBank nucleotide sequence database and assigned accession numbers LC154848-LC154849), together with 27 reference sequences obtained from the GenBank nucleotide sequence database (whole genomes of FLK-BLV subclone pBLV913, LS1, LS2, LS3, Arg41, Arg38, Japan, and USA with accession numbers EF600696, HE967301, HE967302, HE967303, FJ914764, AF257515, K02120, and AF033818, respectively) and from DDBJ database (accession numbers: LC080664, LC080670, LC080669, LC080671, LC080673, LC080665, LC080667, LC080674, LC080672, LC080675, LC080666, LC080668, LC080663, LC080659, LC080661, LC080662, LC080660, LC080654, LC080655, LC080651, LC080652, LC080653, LC080656, LC080657, LC080658). Gaps/missing data were treated as partial deletions. One thousand replications were performed for calculation of bootstrap values (indicated on the tree). Strains from this study are indicated in the tree by the sample ID together with the location of collection. Reference sequences in the tree are indicated by strain name. Genotypes are indicated by *numbers* to the *right* of the figure. Genotype 10 strains were highlighted by pink color, and Genotype 10 was highlighted by grey color. The *bar* at the *bottom* of the figure denotes distance.



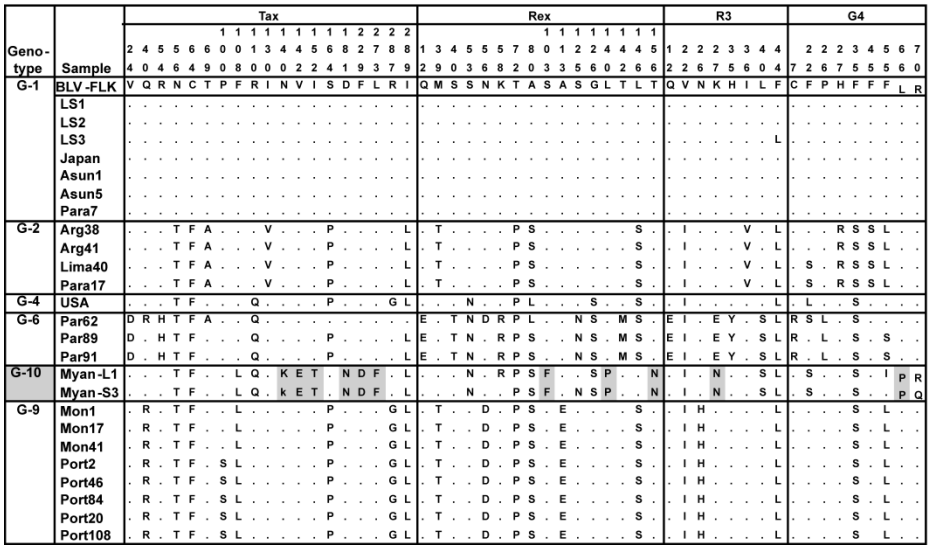


Figure 25 Summary of amino acid substitutions in two BLV strains and 27 known BLV whole genome sequences. Amino acid substitutions identified in at least three samples are shown. Genotypes are indicated by numbers to the far left of the figure. Each structural and regulatory gene encoded glycoprotein was shown on the top of the figure. Numbers above the deduced amino acid sequence are amino acid residue numbers that indicate the position of amino acid substitutions. Genotype 10 and genotype 10 specific mutations were highlighted by grey color. Dots indicate identity with FLK-BLV (accession number: EF600696), which was used as a reference in this study.