Epidemiological study of bovine leukemia virus infections in eastern Hokkaido, Japan and northern Vietnam

2018

DAO Duy Tung

Doctoral Program in Animal and Food Hygiene

Graduate School of Animal Husbandry

Obihiro University of Agriculture and Veterinary Medicine

日本の北海道東部とベトナムのハノイにおける
牛白血病感染の疫学的研究

平成 30 年

(2018)

帯広畜産大学大学院畜産学研究科

博士後期課程　畜産衛生学専攻

ダオ　ズイ　トウング

Contents

Chapter 2

The first study on bovine leukemia virus infection in dairy and beef cattle of small and medium holding farms in northern Vietnam

Chapter 3

Usefulness of a SureSelect target enrichment method for next-generation sequencing and its application to obtain the complete genome of Vietnamese bovine leukemia virus, it's genetic variability, phylogenetic and phylogeographic analysis

Abbreviations

A	aa	Amino acid
B	BLV:	Bovine leukemia virus
D	DNA:	Deoxyribonucleic acid
E	EDTA:	Ethylene-diamine-tetraacetic acid
	ELISA:	Enzyme-linked immunosorbent assay
	env:	Envelope (surface unit (SU) + transmembrane (TM))
G	gp51:	Glycoprotein 51
	gp30:	Glycoprotein 30
I	IPTG	Isopropylthio-β-galactoside
K	kDa:	Kilodalton
L	*LTRs*:	Long Terminal Repeat sequences
	LB:	Luria-Bertani
N	NGS:	Next Generation Sequencing
	nt	nucleotide
O	orf:	Open reading frame
	OD:	optical density

P	PCR:	Polymerase chain reaction
	p12:	Nucleocapsid
	p15:	Matrix
	p24:	Capsid
	pro:	Protease
	pol:	Reverse Transcriptase Polymerase
	PBMC	Peripheral blood mononuclear cells
R	RNA:	Ribonucleic acid
T	TAE	Tris-acetate-EDTA
X	X-Gal	5-Bromo-4-Chloro-3-Indolyl- D-Galactopyranoside

Unit abbreviation

bp:	Base pair
^{o}C:	Degree Celsius
min:	Minute
sec:	Second
ml:	Milliliter
µl:	Microliter
ng:	Nano gram
µM	Micromole

General introduction

I. Background

Bovine leukemia virus (BLV) is an etiological agent that causes Enzootic Bovine Leucosis (EBL) [11]. It is an oncogenic B-lymphocytotropic retrovirus belonging to the genus *Deltaretrovirus* in the *Retroviridae* family [29, 60]. Cattle and water buffaloes are natural hots for BLV. However, several animal species have been experimentally infected with BLV [66]. BLV infections have been reported in all continents and imposed a huge economic impact on the dairy cattle industry [6, 18, 66]. The disease process can be divided into three states including asymptomatic, persistent lymphocytosis (PL) and lymphoma [3]. Most of BLV infected animals are asymptomatic, approximately one-third of them develop PL, and 1-5% of PL animals proceed to lymphoma [11, 23]. The infection can be transmitted by infected blood or by free viral particles produced in cell culture. The viremia only finds in blood during the first two weeks of infection, and infected animals develop a serological response to viral capsid and envelop proteins 2-8 weeks after infection [66].

II. Biology

The genome structure, sequence, and regulation of expression of BLV are close to those of human T-cell leukemia virus type 1 (HTLV-1). When BLV infects host B-cell, DNA provirus is generated after reverse transcription of the viral genome and integrates randomly into the DNA of the host B cell genomes. The complete sequence of the integrated BLV is 8,714 nucleotides long, and the 5' and 3' extremities of the integrated BLV present the same nucleotide sequence called Long Terminal Repeat (LTRs) [64]. BLV's genome structure can be divided into two main parts including the essential part which consist of four open reading frames (orfs); encoding for the capsid proteins *gag*, the viral protease *pro*, the reverse transcriptase polymerase *pol*, the envelope proteins *env*, and the regulation part encodes for at least four proteins including the *R3*, *G4, rex*, and *tax* proteins [23, 66] (Fig. a).

The long terminal repeat (*LTR*)

The 5' *LTR* composes three consecutive regions named *U3*, *R,* and *U5*. These regions are believed to regulate the viral transcription. The *U3* region contains a triplicate copy of a 21 bp sequence called TxRE enhancer regions which are a major regulator of viral expression [23].

The viral essential proteins

The proteins encoded by *gag* gene are involved in the formation of the viral capsid which consists of the main capsid protein *p24*, the nucleocapsid *p12* and the matrix protein *p15*. The post-translational of *gag* gene is carried out by the viral protease *p14* which is encoded by a region located between *gag* and *pol* genes. The *pol* gene encodes an 80 kDa reverse transcriptase and has an integrase activity [23, 66]. The envelope gene encodes a 72 kDa glycosylated precursor *env*, and it is cleaved into two subunits: the extracellular gp51 (SU) and the transmembrane gp30 (TM) proteins. SU and TM associate through disulfide bonds, and help membrane fusion during infection of the host cell [10, 66]. The envelope glycoprotein gp51 plays an important role in the virus life cycle. It is also required for cell entry and the target of neutralizing antibodies [12, 26, 38, 57]. The three conformational epitopes F, G, and H, which are located in the N-terminal half of the gp51 protein, play a critical role in viral infection and syncytium formation [10, 57]. The linear epitopes A, B, D, and E are located within the C-terminal half, which interacts with the TM hydrophobic region, likely via a disulfide bond [26]. In addition, the TM protein gp30 is poorly immunogenic while the SU gp51 protein induces a massive expression of specific antibodies in infected animals [37]. Therefore, it is normally used for both diagnostic and genotyping purposes [23]. Recent phylogenetic analyses of BLV *env* gp51 gene sequences from viral strains isolated worldwide revealed that BLV has been classified into 10 genotypes (G1–G10) [5, 13, 32, 39, 40, 46, 51-54, 59, 61, 74].

The regulation of BLV expression

The regulation part of BLV genome encodes for at least four proteins, the accessory proteins R3, G4 and the regulatory protein rex and tax. The R3 and G4 proteins are involved in post-transcriptional regulations of viral expression [23]. R3 protein contributes to the maintenance of infectivity, and it is located in the nucleus and in cellular membranes [22, 73]. G4 protein is likely implicated in cell transformation, and it is particularly relevant to leukemogenesis [33]. BLV tax protein is a transcriptional activator of viral expression, and it is a target of the host immune response with T and B epitopes [47, 65]. Beside of that, the tax protein is believed to play a critical role in BLV induced leukemogenesis [3], and it is essential for viral infectivity *in vivo* [2, 70]. The rex protein has a role in the nuclear export of viral RNA, and it requires for the accumulation of genomic and singly-spliced *env* RNAs [17, 19]. In addition, BLV rex protein is interchangeable for purpose of post-transcriptional regulation [63].

III. Pathogenicity

BLV induces a persistent and chronic infection in B lymphocyte cells [21]. The infected animal will be positive for proviral and serological test for whole life. The lifelong antibodies will develop 2–8 weeks after infection [66]. The BLV proviral load correlates with BLV infection capacity and lymphoma progression [3, 25]. BLV is not very pathogenic, when animals are infected with BLV, in the years following the infection, around 70% of infected animals develop B cell lymphocytes. This state is defined as an asymptomatic state. Among these animals, approximately one-third develop PL which corresponds to a stable increase in the number of circulating lymphocytes. The lymphoma might occur typically in animals over 3 years of age in 1–5% of the PL animals [11, 23]. The symptoms/signs of lymphoma caused by BLV infection depend on the occurrence site of the tumors (Fig. b) but include reduced milk production, loss of appetite, weight loss, weakness, digestive disturbance or general debility, and various neurological manifestations (*Chapter 2.4.11, OIE Terrestrial Manual 2012*).

IV. Transmission and Economic impact

BLV is transmitted by infected cells (B lymphocytes, monocytes/macrophages, etc.), probably via blood or milk due to the high number of BLV-infected lymphocytes. The horizontal transmission can occur during animal manipulation procedures, such as ear tattooing, dehorning, rectal palpation, injection or by hematophagous insect bites. The vertical transmission possibly via colostrum, milk or the placenta [7, 48, 66]. (Fig. c). In natural conditions, only domestic cattle, water buffaloes, and capybaras have been found to be infected with BLV. However, several species including sheep, goats, pigs, rats, rabbits and non-human primates have been experimentally infected [4, 36, 49, 50, 66]. Recent BLV studies have reported that the BLV infections widely spread to all continents, and its prevalence varies from country to country [13, 30, 32, 39, 40, 46, 51-54, 61, 71, 72, 74]. BLV infection has broad impacts on the dairy industry, not only because of the neoplastic feature but also due to the harmful influences of reduced fertility and decreased milk production in dairy cattle. The average production of PL animals is 3–10 % lower than that of the herd and it can, therefore, lead to early culling and international trade restrictions. In detail, the economic loss due to BLV infection was estimated to 86 million dollars in 1992 in the United States [6, 14, 18, 56, 58].

V. Diagnosis

BLV can be detected by non-malignant polyclonal B-cells expansion in *in-vitro* cultivation of peripheral blood mononuclear cells (PBMC) of infected animal. The virus is present in PBMC and in tumor cells as provirus integrating into the DNA of the infected cells. The virus can also be found in the cellular fraction of various body fluids including nasal and bronchial fluids, saliva, colostrum and milk (*Chapter 2.4.11, OIE Terrestrial Manual 2012*). The BLV provirus can be detected by PCR or nested PCR methods followed by gel electrophoresis. Recently, quantitative real-time PCR and direct blood-based PCR assays have been developed and applied worldwide to BLV detection [55]. Antibodies can first be detected from 3-16 weeks after infection. Maternally derived antibodies may take up to 6-7 months to disappear. There is no way of distinguishing passively transferred antibodies

from those resulting from active infection. However, the active infection can be confirmed by the detection of BLV provirus. Agar gel immunodiffusion, passive hemagglutination assay, radioimmunoassay and enzyme-linked immunosorbent assay (ELISA) are conventional methods for detecting antibodies in bovine serum, milk or the supernatants of BLV infected cell cultures. Among that ELISA is a highly sensitive and easily implemented procedure which can be used to analyze both serum and milk samples [55].

VI. Objectives of the present study

BLV infection has a worldwide distribution and it has significant impacts on the agricultural economy and international trade of the cattle industry. Some European countries have been successfully eradicated BLV infection by the national control program [1, 44]. However, BLV infection has increased rapidly in Japan and it is a notifiable disease which has been subjected to passive surveillance since 1998 [41]. In Japan, the highest number of cattle is raised in Hokkaido, especially in eastern Hokkaido. Accordingly, many EBL cases have been reported in Hokkaido. However, the current status of BLV infection among cattle in eastern Hokkaido has not been monitored in detail.

In contrast, BLV infection has not been paid much attention in Vietnam. Although BLV infection has been known to have significant impacts on the agricultural economy as mentioned above, up to date, no study on BLV infection and its adverse economic effects has been conducted in Vietnam. Therefore, an actual status of BLV infection among cattle in Vietnam should be investigated in detail, by which BLV control strategy available in Vietnam should be established, based on the understanding distribution of current BLV strains and identifying the routes of BLV introduction into cattle herds.

The study was aimed at the following:

1. To develop and evaluate an absolute quantitative real-time PCR assay for detection of BLV proviral copy numbers.

2. To examine the prevalence of BLV infection, and molecular characterization of current BLV strains among cattle in eastern Hokkaido, Japan.
3. To examine the prevalence of BLV infection among cattle in northern Vietnam, and to molecularly characterize BLVs prevalent in the country.
4. To evaluate the efficiency and usefulness of a SureSelect target enrichment method for deep sequencing of the whole BLV genome using Next generation sequencing (NGS), and to sequence the complete genomes of some Vietnamese BLVs to obtain understanding of the genetic variability, phylogeny, and phylogeography of Vietnamese BLVs in comparison with the known BLV strains.

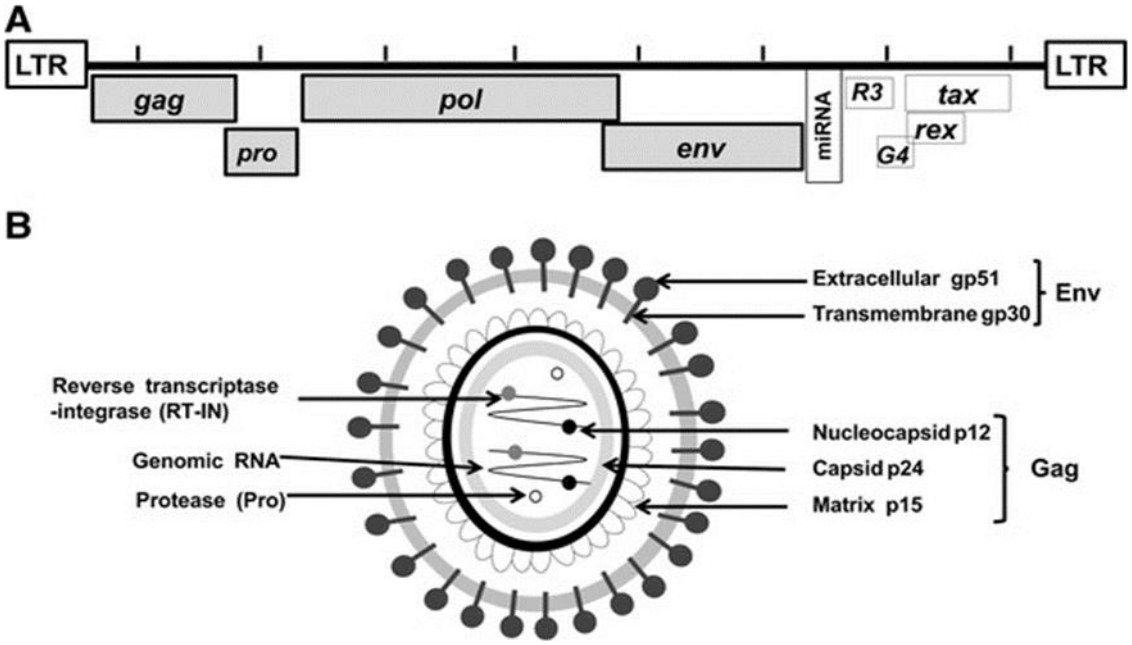


Fig. a Schematic representation of the BLV genome structure (**A**) and viral particle (**B**). The structural and enzymatic genes, *gag*, *pro*, *pol*, and *env*; regulatory genes, *tax,* and *rex*; accessory genes *R3* and *G4*; and microRNA (miRNA) are indicated in (**A**). Proteins encoded by structural and enzymatic genes, including the Env glycoproteins (gp51 and gp30) encoded by the *env* gene, the Gag proteins (p12, p24, and p15) encoded by the *gag* gene, reverse transcriptase and integrase (RT-IN) encoded by the *pol* gene, and protease (Pro) encoded by the *pro* gene are indicated in (**B**) [55]

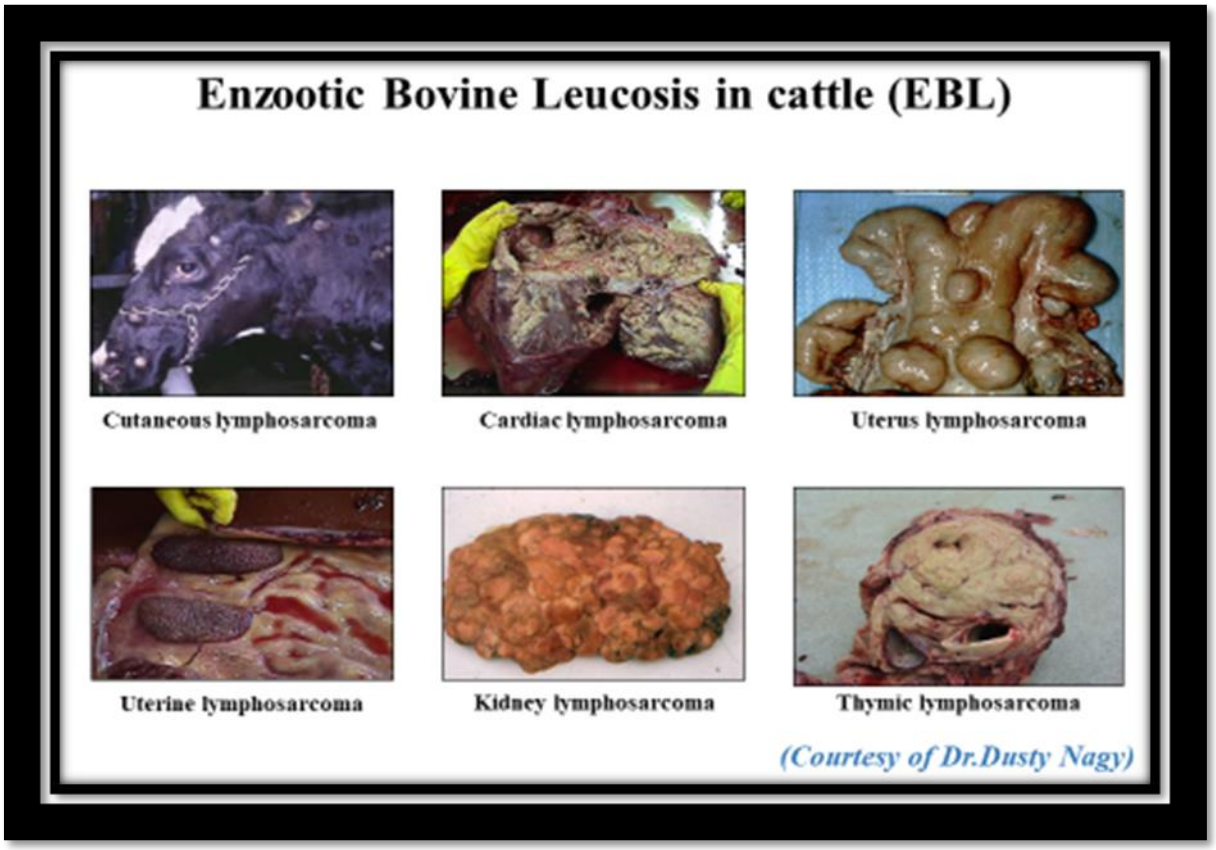


Fig. b Occurrence sites of the tumors found in cattle with EBL

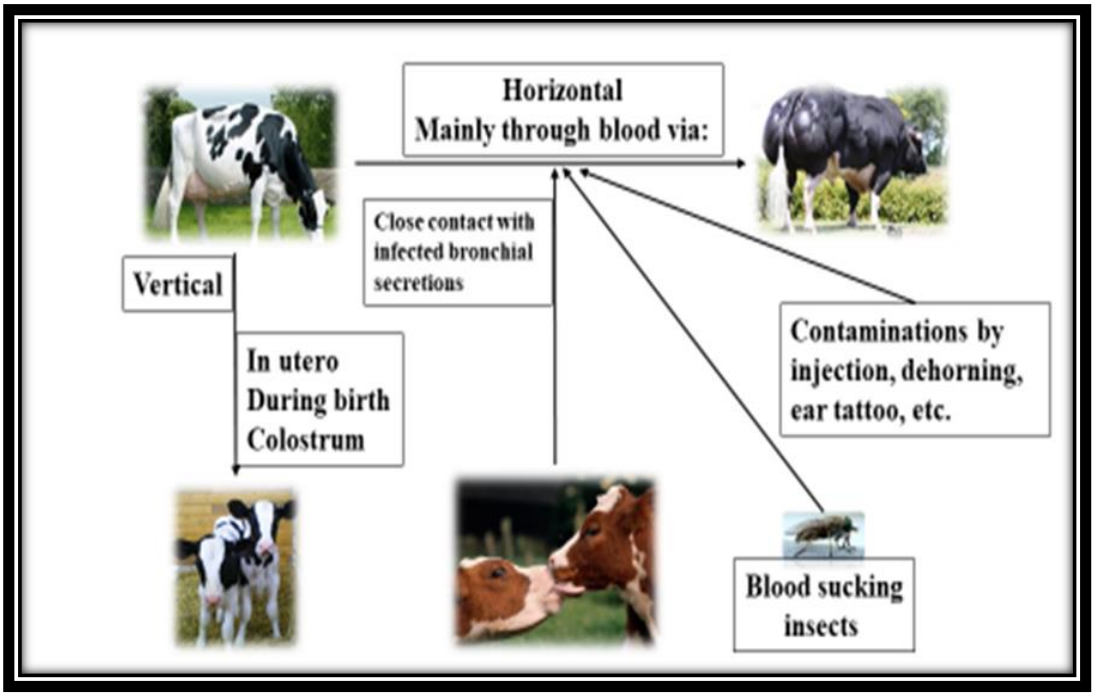


Fig. c Schematic main routes of BLV transmission

Chapter 1

Surveillance of BLV infection in cattle in eastern Hokkaido, Japan

1.1 Introduction

BLV infection has a worldwide distribution, and it impacts on the dairy cattle husbandry due to reduced milk production [43] , and reproductive inefficiency [6]. The progress of enzootic bovine leucosis (EBL) can be divided into three states including asymptomatic, persistent lymphocytosis (PL) and lymphoma [3]. Each state of the disease might be differentiated by the increasing of BLV proviral copy numbers in the blood and lymph nodes of infected animal [3, 67]. Some European countries have successfully eradicated BLV infections by the national control program [1, 44]. In the recent years, BLV infection have been increased rapidly in Japan, according to the nationwide survey of BLV infection among dairy and beef breeding cattle in Japan from 2009-2011 [42]. The intra-herd BLV seroprevalence in dairy and beef breeding cattle was 40.9% and 28.7%, respectively. The seroprevalence of dairy and beef breeding cattle were approximately 10- and 4-fold higher than that previously reported in nationwide survey from 1980-1982 in Japan. It is unclear why or how BLV has become so widespread across Japan.

I was involved in a BLV surveillance study in dairy cattle in eastern Hokkaido, Japan. Therefore, a useful and effective diagnostic tool to determine BLV infection was required. Although the BLV provirus can be detected by a PCR or nested PCR method followed by gel electrophoresis, those assays cannot determine the proviral load in the infected cattle. Therefore, the objective of this study was to develop and evaluate an in-house absolute quantitative real-time PCR assay for the detection of BLV proviral copy numbers, and to examine the prevalence of BLV infection, and molecular characteristics of current BLV strains among cattle in eastern Hokkaido, Japan.

1.2 Materials and methods

1.2.1 Sampling

Two cattle farms in eastern Hokkaido were selected for the study. Blood of all the cattle in two farms, named K and U, were sampled. In detail, a total of 211 whole blood samples in K farm and 170 samples in U farm were collected in 2016-2017. All the blood samples were used for molecular detection purpose while serum samples were used for serological examination.

1.2.2 Genomic DNA extraction

Genomic DNA was extracted from 200 µl of whole blood using High Pure PCR Template Preparation Kit (Roche) according to the manufacturer's recommendation. DNA samples were stored at 4ºC for molecular tests until use.

1.2.3 Detection of BLV DNA provirus by a real-time PCR targeting the BLV *pol* gene

Genomic DNA samples were examined using an absolute quantitative real-time PCR targeting the BLV *pol* gene (in-house *pol*- rt-PCR), which was modified based on the TaqMan Probe BLV *pol* real-time PCR previously described by Heenemann et al. (2012) [24]. A control plasmid was constructed for the in-house *pol*- rt-PCR to verify the assay sensitivity and to quantify the proviral copy number in the testing samples. The 525 bp fragment of the BLV *pol* gene from a positive BLV DNA provirus sample was amplified by a primer pair (BLV 525 F, BLV 525 R) (Table. 1.1). The electrophoresed PCR product was purified by Qiagen Gel Purification Kit, and ligated to the pGEM-T Easy vector ligation system (Promega, USA) with T4 ligase according to the instructions supplied by the manufacturer. After that, approximately 2 *µl* of the ligation reaction was transformed into 20 *µl* *Escherichia coli (E.coli)* DH5∝ competent cells. The transformed *E. coli* were cultured on LB agar plates containing 100 *µg/ml* ampicillin, 20 *µg/ml* X-Gal, and 100 *µM* IPTG then incubated at 37ºC overnight. Individually white colonies were selected then cultured in tubes containing 2 *ml* LB with 100 *µg/ml* ampicillin. For plasmid DNA isolation, the Qiaprep Spin Miniprep kit (Qiagen, Germany) was used according to the instructions from the manufacturer. The inserted gene was confirmed by

digestion using *Eco*RI restriction enzyme and by DNA sequencing to check the correct inserted size and correct inserted sequence, respectively. The plasmid contains 525 bp fragment that covers all binding region for primers (BLV-pol-F, BLV-pol-R) and probe (BLV-pol-probe) (Table. 1.1) which were used for detection of BLV *pol* gene in BLV DNA provirus samples. The concentration of the plasmid was determined by NanoDrop spectrophotometer (Thermo SCIENTIFIC), and the exact number of plasmid molecules was calculated using the formula: **[X (g/ µl DNA) / (length in nucleotides x 660)] x 6.022 x 10^{23} = number of molecules / µl**.

An absolute quantitative real-time PCR assay was set up to determine BLV *pol* gene provirus copy numbers following the protocol: Master mix components contained 12.5 μl Eagle Taq Master Mix (Roche), 0.6 μl primers (10 μM each), 0.3 μl probe (5 μM), 6 μl distilled water. Five microliters of DNA was added, total 25 μl volume of the real-time PCR reaction mixture was determined by a LightCycler Nano real-time PCR system (Roche) by the following thermal cycle conditions: initial denaturation at 50^{o}C 2 *min* and 95^{o}C at 10 *min*, followed by 45 cycles of 95^{o}C at 15 *sec* and 58^{o}C at 1 *min*. Fluorescence data were collected during the annealing step.

1.2.4 Detection of BLV provirus by a real-time PCR detection commercial kit targeting the BLV *tax* gene

All DNA samples were tested using the commercial BLV detection kit targeting *tax* gene (commercial *tax*-rt-PCR) (TAKARA Bio Inc., Japan), using the running conditions provided in the manufacturer's instructions. Standard plasmid and the positive control were also provided with the kit. An Eagle Taq Master Mix was used to amplify BLV *tax* gene and the gene was detected by a LightCycler Nano real-time PCR system (Roche).

1.2.5 Detection of anti-BLV antibody by ELISA

All serum samples were examined to detect antibodies against the BLV gp51 using a commercially available BLV ELISA kit (JNC Corporation, Tokyo, Japan) according to the

manufacturer's instructions. Color development was measured by reading the optical density (OD) at 450 *nm* with a spectrophotometer (TECAN Genios Pro; Tecan Group Ltd, Männedorf, Switzerland). The antibody-positive sample was determined by calculating the sample-to-positive (S/P) ratio from the OD values obtained for the tested serum, and the positive and negative control sera provided with the kit. The sample was regarded as positive if the value of the S/P ratio was greater than 0.3 and the positive control was greater than 0.6.

1.2.6 Nested PCR assay for amplifying complete BLV *env* gp51 gene

All DNA samples collected from cattle which were identified as ELISA positive but real-time PCR assay negative were confirmed by a nested PCR assay. The nested PCR assay was designed for the study to amplify the complete BLV *env* gp51 gene of BLV DNA provirus, with two pairs of primers (BLV Outer F, BLV Outer R) and (BLV Inner F, BLV Inner R) (Table. 1.1). For both the first and second runs of nested PCR, same PCR reaction mixture was used. The mixture contained 12.5 μl 2XGoTaq Green Master Mix (Promega, USA), 1 μl primers (each 10 μM), 5.5 μl distilled water. Five μl of DNA samples were used for the first run, then five μl PCR product from the first run was added to the second run of the following conditions: initial denaturation 95^{o}C for 5 *min*, followed by 25 cycles consisting of denaturation at 95^{o}C for 30 *sec*, annealing at 50^{o}C for 30 *sec* and extension at 72^{o}C for 3 *min*. The last extension was run at 72^{o}C for 10 *min*; and condition: initial denaturation 95^{o}C for 5 *min*, followed by 40 cycles consisting of denaturation at 95^{o}C for 30 *sec*, annealing at 55^{o}C for 30 *sec* and extension at 72^{o}C for 2 *min*. The last extension was run at 72^{o}C for 10 *min*. PCR products were analyzed in 1.5% agarose gel stained with ethidium bromide (1$\mu g/ml$) in 1XTAE buffer. The target bands (1175 bp) were cut from gel then purified using QIAGEN Gel extraction kit (Qiagen, Germany) according to the manufacturer's instructions. Finally, the DNA was eluted with 50 μl distilled water and measured the concentration by NanoDrop spectrophotometer.

1.2.7 DNA sequencing

Approximately 250 *ng/µl* DNA of each sample which was amplified and purified from the nested PCR were used for sequencing the complete BLV *env* gp51 gene encoding for the gp51 surface glycoprotein (903 bp). Two each of PCR products for each DNA sample were sequenced in both directions using the same nested PCR primers (BLV Inner F, BLV Inner R) (Table. 1.1): primer BLV Inner F for amplifying positive strand and primer BLV Inner R for amplifying negative strand by ABI Prism Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosciences, Foster City, CA, USA). The resultant products were then purified using the Big Dye XTerminator Purification Kit (Applied Biosciences). The sequences were analyzed by 3500 Genetic Analyzer (Life Technologies, Carlsbad, CA, USA). All raw sequencing data were obtained from both strands of each PCR product for verification, and multiply aligned by sequence analysis program GENETYX Ver.13.0.3 (GENETYX Corp., Tokyo, Japan). The consensus sequences were generated from two bidirectional repeats of each sample.

1.2.8 Phylogenetic analysis

The BLV *env gp51* gene sequences obtained in this study were aligned using CLUSTALW multiple alignment in BioEdit version 7.2.5 (http://www.mbio.ncsu.edu/bioedit/bioedit) with worldwide BLV reference strains illustrating all ten genotypes (G1-G10) which were identified in previous studies [5, 32, 46, 51-54, 59, 74]. The complete BLV *env gp51* nucleotide sequences (903 bp) of all BLVs were aligned with 96 full-length BLV *env gp51* reference sequences available on GenBank. Phylogenetic analyses were conducted by Maximum Likelihood (ML) method by MEGA6 software [69] with parameter settings of 1,000 bootstrap replicates, and the best fit model K2 + G. The phylogenetic trees were then rooted at genotype 5 (G5) to easily compared with other previous studies [5, 32].

1.3 Results

1.3.1 BLV infection prevalence in two farms in eastern Hokkaido, Japan

The BLV infection prevalence which was determined by ELISA in two farms, were mostly higher than those determined by real-time PCR assays. In detail, the positive rate in U farm detected by in-house *pol*-rt-PCR and commercial *tax*-rt-PCR was the same at 14.12% (24 positives/170 tested samples), while the positive rate detected by ELISA was 17.65% (30 positives/170 tested samples). The positive rates detected by ELISA and both real-time PCR assays in K farm were at 46.44% (98 positives/211 tested samples) and 38.86% (82 positives/211 tested samples), respectively (Table. 1.2). A total of 22 samples which were negative by real-time PCR assays but positive by ELISA in two farms (6 samples in U farm and 16 samples in K farm), were confirmed by the nested PCR assay. The results of nested PCR indicated that there were more 11 positive samples over total 22 tested samples (including 1 positive in U farm and 10 positives in K farm) (Table. 1.3).

1.3.2 Evaluation of in-house *pol*-real-time PCR assay

The mean of BLV proviral copy number which was determined by in-house *pol*-rt-PCR and by commercial *tax*-rt-PCR in U farm was 8,615 copies/μl and 7,485 copies/μl, respectively. The average of Ct value of both assays was also similar each other (around 30) (Table. 1.4). The Ct value correlation of in-house *pol*-rt-PCR and commercial *tax*-rt-PCR demonstrated a significant correlation supporting by R square value at 0.7278 (Fig. 1.1).

1.3.3 The positive rate within age groups in two farms

The highest positive rates were detected in the oldest groups, while the lowest positive rates were observed in the youngest groups in both farms. In detail, the positive proportion of the oldest group in K (9-13 years old) and U (4-9 years old) farms was 93.33% and 31.43%, respectively. The positive rate of the youngest groups with < 1-year-old was 15.9% in K farm and 5.26% in U farm (Table. 1.5, Table. 1.6). Interestingly, the highest mean of copy number (per μl) which was calculated by in-house *pol*-rt-PCR were found in the group with age from 1- to 3-year-old in both farms at 6,914 copy in K farm and 9,767 copy in U farm (Table. 1.5, Table. 1.6).

1.3.4 DNA sequencing

A total of 11 DNA samples, which were negative by real-time PCR assays but positive by Nested PCR (Table.1.3), were forwarded to sequencing the complete BLV *env gp51* gene. Additional 5 positive samples (1 in K farm and 4 in U farm) were randomly selected to sequence the complete BLV *env gp51* gene. A total of 16 BLV DNA provirus were successfully sequenced.

1.3.5 Phylogenetic analysis

In the phylogenetic analysis using the complete BLV *env* gp51 gene sequences (903 bp) comparing the ten genotypes, all 16 BLV *env gp51* sequences in this study were clearly clustered into genotype 1 (G1) with a high supporting value at 97 in ML analysis (Fig. 1.2). In detail, the eleven BLVs in K farm were separated into two subgroups. Seven BLVs in K farm together with five BLVS in U farm built one subgroup and four other BLVs in K farm were clustered in another subgroup.

1.3.6 Possible transmission route of infected calves with age < 1-year-old

Both vertical and horizontal transmissions were identified in infected calves with age smaller than 1-year-old in both farms. In detail, 6 cases in U and 1 cases in K farms were found vertically transmitted. And horizontal transmission was observed 2 cases in U and 1 cases in K farms (Table. 1.7).

1.4 Discussion

BLV infection has increased dramatically and spread widely across Japan in recent years. In this study, all cattle in two farms in eastern Hokkaido were screened for BLV infection by serological and molecular detection methods. The number of positive samples that were determined by ELISA was higher than by the real-time PCR assays in both farms. Indeed, total 6 positive samples in U farm, and 16 positive samples in K farm were ELISA positive but in-house *pol*-rt-PCR negative (Table. 1.2). One possible explanation for this phenomenon is that the ELISA positives of the cattle with <1-year-old are more likely to be related to passive maternal antibodies. However, the nested PCR assay

confirmed more 11 positive samples which could not be detected by in-house *pol*-rt-PCR (1 sample in U farm and 10 samples in K farm) (Table. 1.3). These results suggested that the nested PCR assay was more sensitive than in-house *pol*-rt-PCR assays. Therefore, it is necessary to improve the sensitivity of in-house *pol*-rt-PCR assay.

The accordant results of the in-house *pol*-rt-PCR and the commercial *tax*-rt-PCR implied that the in-house *pol*-rt-PCR assay can be used for BLV DNA proviral detection. The highest positive rate within age groups was found in the oldest groups in both farms. Interestingly, however, the highest mean copy number which was calculated by the in-house *pol*-rt-PCR was found in the age group of 1–3 year-old in both farms. In additional, there is a possibility to distinguish EBL cattle from BLV infected cattle by observing the BLV proviral copy number in lymph node samples. Difference in the BLV proviral copy number in lymph node samples between EBL cattle and BLV infected cattle was significantly different [67]. Therefore, to identify EBL from the BLV infected cattle, it is required to collect lymph node samples of BLV infected cattle by using core needle biopsy or fine-needle aspirate. Moreover, the positive rate and the BLV copy number in the young cattle age group of 1–3 year-old may be used as an indicator to assess the BLV infection status in herds.

The phylogenetic analysis presented that the genetic distribution of BLV among cattle in two farms in eastern Hokkaido was simple; only G1 group was found among the infected cattle. However, the 11 BLVs in K farm were divided into 2 subgroups of G1, suggesting that there were two sources of BLV infection that were distributed to cattle population in K farm. This situation might be a result of changes in transmission modes or other influencing factors, such as cattle import and movement. Furthermore, according to the possible transmission route of infected calves with age less than 1-year-old in two farms, BLV might be vertically and horizontally transmissions in both farms; however, definitive evidence was not obtained. Therefore, continuing surveillance and identifying the key risk factor of BLV transmission to control new infection should be given a priority. In addition, the infected

cattle are a reservoir of BLV leading to transmission among herd. Accordingly, in order to control and prevent new infections, step by step replace the infected cattle is critical, and feeding pool colostrum and bulk milk without the treatment to newborn calves not be recommended.

1.5 Summary

This study revealed that the seroprevalence of BLV infection among cattle in eastern Hokkaido were variety (high at 46.44% in K farm and low at 17.65% in U farm). Molecular characterization demonstrated that BLVs belonging to G1 are circulating among cattle in eastern Hokkaido. In-house absolute quantitative real-time PCR for detecting the BLV *pol* gene was evaluated. Basically, this assay can be applied to detection of BLV proviral DNA in infected cattle. The results also indicated that continuing surveillance and finding the route of horizontal transmission in Japanese farming system, as well as eliminate and control that risk factors in order to prevent new infections are critical.

Table 1.1 Primers and probes used for in-house *pol*-rt-PCR, and the nested PCR targeting the *env* gp51 gene

Name	**Sequence (5' – 3')**	**Position***
BLV-pol-F	CCCTGGCCTACTTCCAGACC	3304–3323
BLV-pol-R	CTTGGCATACGAGCTTAAGGCC	3395–3416
BLV-pol-probe	FAM- TTGACTGACAACCAAGCCTCACCT-BHQ1	3327–3350
BLV 525 F	GACCCTAGGGCCATCATCC	3126–3144
BLV 525 R	AGCCCCGTCACTAAAGAGG	3632–3650
BLV Outer F	TTCCCTACAACCCCACAAGT	4444–4463
BLV Outer R	GGTCAAGCATTTTATCAGG	6190–6208
BLV Inner F	TCAGAGGGCGGAGAAACAC	4650–4668
BLV Inner R	AGGTGAGTCTCTGATGGCTAAG	5804–5825

*Nucleotide positions according to the GenBank accession No. LC164086.

Table 1.2 BLV prevalence detected by different detection assays in 2 farms

Assays	U farm	K farm
	Positives/total samples tested (%)	**Positives/total samples tested (%)**
In-house *pol*-rt-PCR	24/170 (14.12%)	82/211 (38.86%)
Commercial *tax*-rt-PCR	24/170 (14.12%)	ND*
ELISA	30/170 (17.65%)	98/211 (46.44%)

*: ND: Not done

Table 1.3 Nested PCR results of samples determined to be negative by in-house *pol*-rt-PCR but positive by ELISA

Farm	No of negative samples tested by in-house *pol*-rt-PCR	No of positive/total samples tested by Nested PCR
U	6	1/6
K	16	10/16
Total	22	11/22

Table 1.4 Comparison of BLV copy number detected by different absolute quantitative real-time PCR assays

Assays	U farm		K farm	
	Mean copy number per *µl*	**Mean Ct value**	**Mean copy number per *µl***	**Mean Ct value**
In-house *pol*-rt-PCR	8,615	30.01	5,854	30,35
Commercial *tax*-rt-PCR	7,485	29.75	ND*	ND

*: ND: not done

Table 1.5 The positive rate of age groups cattle determined by in-house *pol*-rt-PCR in U farms

Age group	% Positives	Mean copy number per ***µl***
0	5.26	1,329
1-3	13.83	**9,767 ***
4-9	31.43	7,015

*: The highest mean copy number per ***µl***

Table 1.6 The positive rate of age groups cattle determined by in-house *pol*-rt-PCR in K farms

Age group	% Positives	Mean copy number per ***µl***
0	15.9	3,095
1-3	56.25	**6,914** *
4-8	58.49	3,449
9-13	93.33	3,016

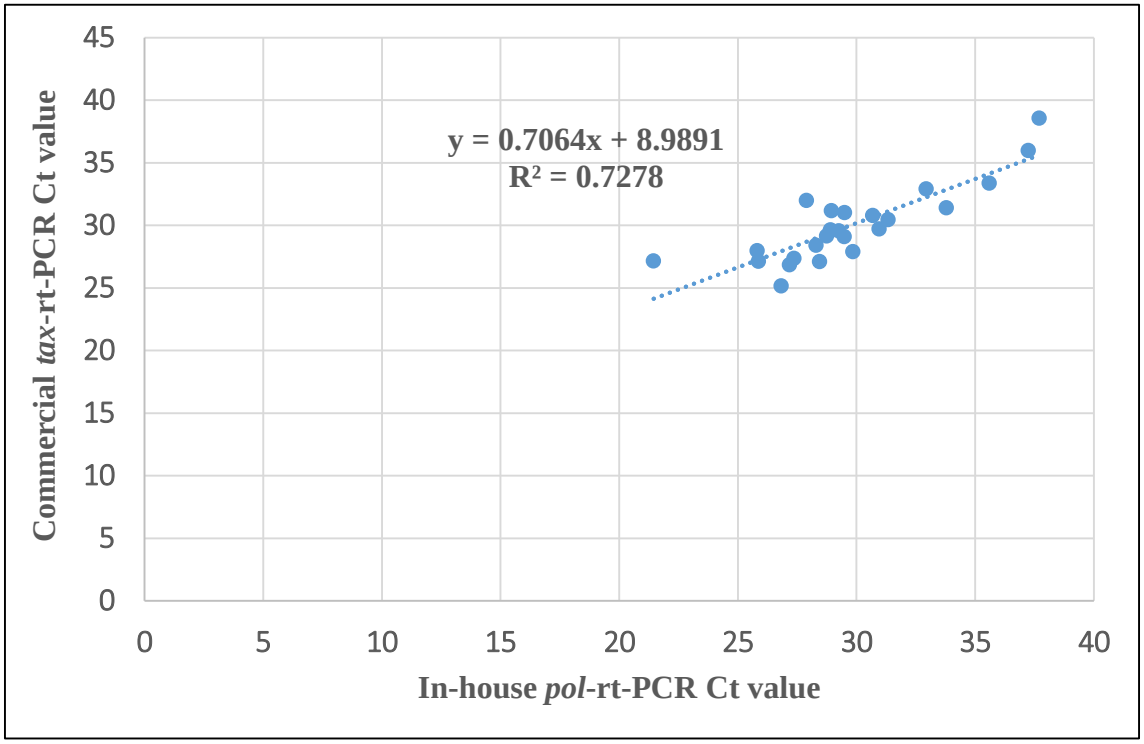
*: The highest mean copy number per ***µl***

Table 1.7 The possible transmission route of infected calves (< 1-year-old)

	Farm	Calf with BLV infected	Dam with BLV infected
Vertical transmission*	U	6	6
	K	1	1
Horizontal transmission**	U	2	0
	K	1	0

*: BLV infected calves were transmitted from their mother.

**: BLV infected calves were transmitted from other routes, not from their mother.



Pearson r	**0.8531**
95% confidence interval	**0.6857 to 0.9348**
P value (two-tailed)	**< 0.0001**
Is the correlation significant? (alpha=0.05)	**Yes**
R square	**0.7278**

Fig. 1.1 Ct value correlation of in-house *pol*-rt- PCR with commercial *tax*-rt-PCR assays

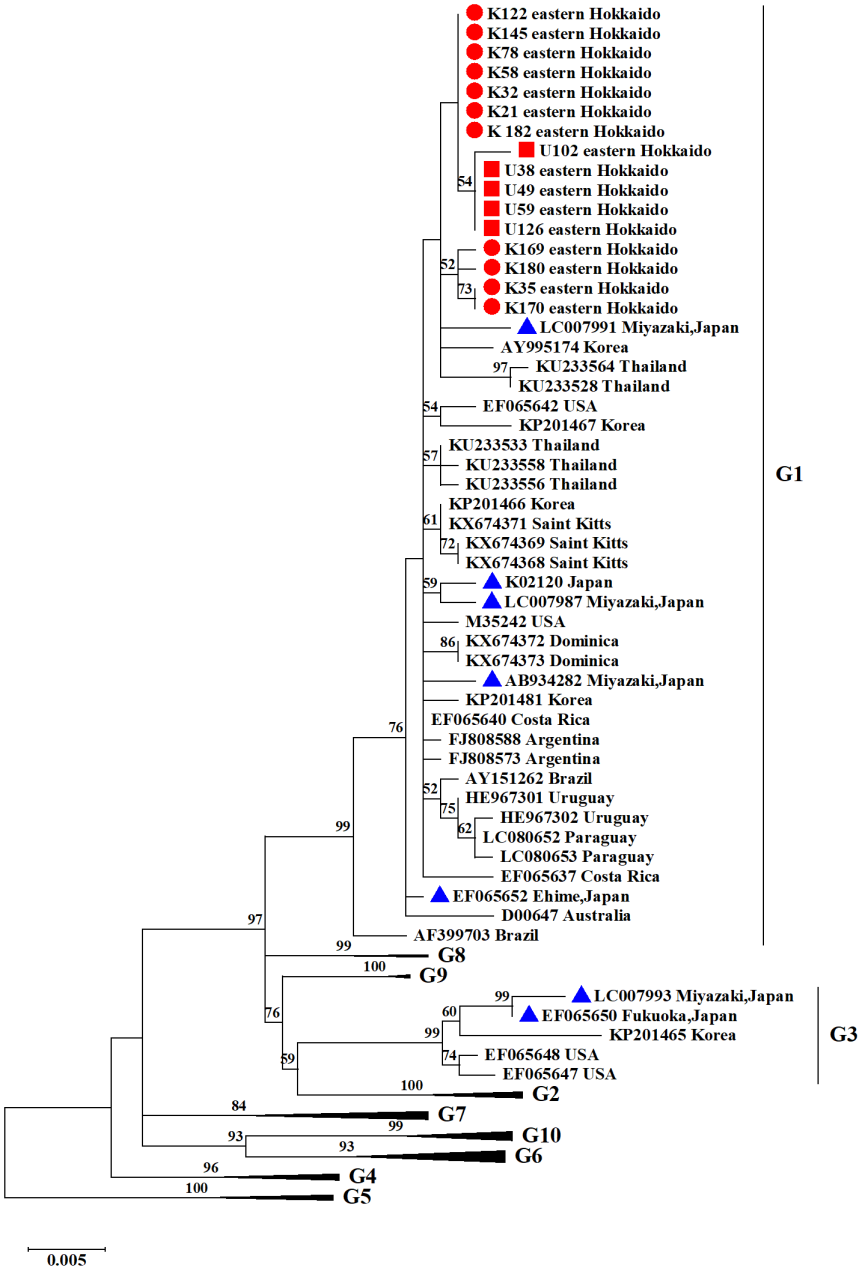


Fig. 1.2 Maximum-likelihood phylogenetic tree of the complete *env gp51* nucleotide sequences (903bp) of ten strains of BLV (n=96). Hokkaido BLV strains (●), (■) and other Japanese strains (▲)

Chapter 2

The first study on bovine leukemia virus infection in dairy and beef cattle of small and medium holding farms in northern Vietnam

2.1 Introduction

Vietnam is a developing country, located in Southeast Asia. Its economy depends mainly on agricultural industries. Animal husbandry plays a significant role in the development of the economy in country area, and cattle husbandry is one of the most important occupations for farmers in Vietnam. In recent years Vietnamese cattle population and ruminant productions (meat, milk and other milk production) have increased dramatically. According to statistics from Vietnam's Husbandry (http://www.channuoivietnam.com) and the General Statistics Office of Vietnam (http://www.gso.gov.vn) in 2016, a total cattle population was approximately 6.1 million including 3.1 million beef cattle, 2.5 million water buffaloes, and 463,000 dairy cattle. Every year Vietnamese cattle husbandry produces approximately 308,000 tons of beef and 795,000 tons of milk. The cattle populations are prominently distributed to northern Vietnam, including the Hanoi region. To develop the cattle husbandry in Vietnam, the spread of cattle diseases has been strictly controlled by the Vietnamese government, with a primary focus on foot-and-mouth disease, bovine anthrax, bovine pasteurellosis, and so on.

BLV is the causative agent of EBL, the most common neoplastic disease of cattle [11, 66]. BLV has been spread worldwide and had a high impact on the economy of cattle husbandry. To date, several studies on BLV infection prevalence and BLV genetic identification have been reported in many countries [5, 13, 32, 39, 40, 46, 51-54, 59, 61, 74]. However, up to date, no study on BLV infection and its adverse economic effects has been conducted in Vietnam. Therefore, an actual status of BLV infection among cattle in Vietnam should be investigated in detail.

2.2 Materials and methods

2.2.1 Samples and localities

Total of 93 whole blood and serum samples (45 dairy cattle and 48 beef cattle) were randomly collected in 464 individuals from 22 farms in northern Vietnam, including Vinh Tuong district and 3 districts of Hanoi (viz., Ba Vi, Chuong My, and My Duc) (Fig. 2.1). Most of dairy cattle were 'Lai HF' (mix with Vietnam native cattle x Holstein Friesian) and the beef cattle were Vietnam native cattle (Lai Sind) or cattle F1 BBB (mix with native cattle x Blanc Bleu Belge). The size of farms was differentiated into two groups according to the number of animals present on the farm. Those holding 20 or fewer cattle were considered as small farms, whereas those holding 20–100 cattle were considered as medium farms. In this study, the blood samples were obtained from 17 small and 5 medium holding farms and accounted for approximately 20% of the total cattle population on each farm. All the cattle involved in this study appeared healthy and their ages ranged from 12 to 36 months.

2.2.2 Genomic DNA extraction

The same methods were described in the subsection of 1.2.2 in Chapter 1.

2.2.3 Detection of BLV DNA provirus by in-house *pol-rt*-PCR

The same methods were described in the subsection of 1.2.3 in Chapter 1.

2.2.4 Detection of BLV DNA provirus by commercial *tax*-rt-PCR

The same methods were described in the subsection of 1.2.4 in Chapter 1.

2.2.5 Detection of anti-BLV antibody by ELISA

The same methods were described in the subsection of 1.2.5 in Chapter 1.

2.2.6 Nested PCR for amplification of the complete BLV *env gp51* gene

The same methods were described in the subsection of 1.2.6 in Chapter 1.

2.2.7 DNA sequencing

The methods were described in the subsection of 1.2.7 in Chapter 1. All the positive samples were used for sequencing the complete BLV *env gp51* gene (903 bp).

2.2.8 Phylogenetic analysis

The BLV *env gp51* nucleotide sequences obtained in this study were aligned with sequences from worldwide BLV reference strains of all 10 genotypes (G1–G10), using the CLUSTALW multiple alignment tools in BioEdit v.7.2.5 (http://www.mbio.ncsu.edu/bioedit/bioedit). First, the complete BLV *env gp51* gene sequences (903 bp) of all Vietnam strains were aligned with those of 56 BLV strains available in GenBank. Second, a 444 bp fragment of the partial *env gp51* gene at the 277–720 nt position, which was previously identified by restriction fragment length polymorphism and has often been used for genotyping of BLV [20, 35], was applied for the alignment of the 33 Vietnam strains with 66 worldwide strains. Phylogenetic analyses based on the two datasets (903 bp and 444 bp of the BLV *env gp51* gene) were conducted using two methods: maximum-likelihood (ML) analysis in MEGA 6 [69], applying 1,000 bootstrap replicates and the best-fit model K2+G; and Bayesian inference (BI) analysis in MrBayes v.3.2.6 [62], applying the best evolutionary model (l set nst = 2, rates = gamma [corresponding to model HKY+G]), as estimated by MrModeltest2 in PAUP4 (http://paup.sc.fsu.edu/index.html). In the BI analysis, two runs and four Markov chains were used for 10,000,000 generations, and the trees were sampled at every 1,000 generations. The first 25% of the BI trees were discarded as "burn-in." A consensus tree was constructed from the output file using FigTree v.1.4.3 (http://tree.bio.ed.ac.uk/). The phylogenetic trees were then rooted at G5 for ease of comparison [5, 32].

2.2.9 Mean nucleotide and amino acid distances and amino acid substitutions in BLVs

The average percentage of nucleotide and amino acid distances for partial 444 bp BLV *env gp51* within (intra-sub genotype) and between (inter-sub genotype) six sub-genotype 6 groups which

were named in this study, were estimated using the Kimura-2-parameter model with the rate variation among sites was modeled with a gamma distribution, and Poisson model in MEGA 6 [69], respectively.

Total of 301 amino acids (aa) of the complete gp51 *env* protein of 56 BLVs belonging to ten genotype groups, and all Vietnamese BLVs were observed throughout the functional regions of the BLV *env gp51* protein sequences that were described in the previous studies [53, 75].

2.3 Results

2.3.1 Prevalence of BLV infection in Hanoi, Vietnam

The antibodies against BLV were detected by ELISA among cattle in 3 districts VinhTuong, BaVi, and ChuongMy, and the positive prevalence were 14.29 %, 58.54 %, and 30 %, respectively. Similarly, the percentage of positives determined by in-house *pol*-rt-PCR in the same 3 districts were 14.29%, 58.54%, and 30%, respectively, while the positive rates were determined by commercial *tax*-rt-PCR in those districts were 9.52%, 53.66%, and 30%, respectively (Table. 2.1). All cattle tested in 4 small beef cattle farms in My Duc district showed negative results in both serological and molecular detection methods.

Totally, the prevalence of BLV infection in northern Vietnam was 35.48% (33/93 samples tested) determined by ELISA and by in-house *pol*-rt-PCR. The mean of DNA provirus copies number of in-house *pol*-rt-PCR was 19,966 copies/μl with a mean of Ct value at 28.43. On the other hand, the positive samples which were determined by commercial *tax*-rt-PCR were only 32.26% (30/93 samples tested) (Table. 2.1), with the mean of DNA provirus copies was 1,504 copies/μl and mean of Ct value at 30.56 (Table. 2.2). Overall, the BLV infections were detected in 5 out of the 9 dairy cattle farms, while the BLV infections were found in only 3 out of the 13 beef cattle farms. Notably, the highest prevalence of infected BLV was observed in BaVi district (58.54%) (Table. 2.1).

2.3.2 Phylogenetic analysis and genotyping of the Vietnam BLV strains

The complete BLV *env gp51* nt sequence (903 bp) was successfully obtained for all the 33 Vietnamese samples. The nt sequences of nine representative strains were deposited to GenBank under the accession numbers MF817716–MF817724. The phylogenetic relationship of the complete BLV *env gp51* nt sequences (903 bp) among ten genotypes showed consistent topologies in both approaches of ML and BI with high supporting value, indicating that 33 Vietnamese BLVs and 56 BLVs in other countries were clearly classified into ten genotype groups (G1-G10) (Fig. 2.2, Fig. 2.3). In detail, 33 Vietnam BLVs could be clustered into two distinguished genotypes, the 30 strains belonging to genotype 1 (G1) and the 3 remain strains belong to genotype 6 (G6).

The phylogenetic analysis of 444 bp partial *env gp51* gene constructed by two methods ML and BI illustrated that Vietnamese BLVs were grouped into two clusters. Similarly, most of Vietnamese BLVs (30 strains) could be classified into G1 with high bootstrap value (75) in ML and high posterior probabilities value (1) in BI (Fig. 2.4, Fig. 2.5). However, other BLVs genotype 6 (G6), genotype 10 (G10) and 3 Vietnamese BLVs genotype 6 (G6VN) were grouped into the same big cluster that included some polytomies with supporting value low in ML analysis (<50) but high in BI analysis (>0.7). The existence of G6 subgroups which was obtained in 444 bp BI phylogenetic tree (Fig. 2.5) supporting by high posterior probabilities value (0.99) suggested that the existing topology of G6 was clearly separated into at least six sub-genotypes.

Indeed, one of three Vietnamese BLVs G6 (VinhTuong_Vietnam S9F16), which was isolated from the dairy cattle farm in Vinh Tuong district, was grouped into one clade, together with six Thai BLVs, while the two remaining Vietnamese BLVs G6 (BaVi_Vietnam S38F21 and BaVi_Vietnam S39F21) were separated into a distinguished group, both strains existed in the same herd from a medium milking farm in Ba Vi district.

Taken together I designated that BLVs G6 could be classified into at least six sub-genotypes, with five known sub-genotypes G6a, G6b, G6c, G6d, G6e [52] and one additional new subtype G6f (Fig. 2.2, Fig. 2.3, Fig. 2.4 and Fig. 2.5).

Overall, 33 Vietnamese BLVs were identified into 2 genotypes. The most dominant genotype was G1 (30/33 strains) distributing in Vinh Tuong, Vinh Phuc, and Ba Vi, while only 3 Vietnamese BLVs were grouped into G6 which were found in Vinh Tuong and Ba Vi districts.

2.3.3 Mean nucleotide and amino acid distances and amino acid substitutions in Vietnamese BLV strains

The average percentage of nucleotide (nt) and amino acid (aa) distances based on the partial 444 bp BLV *env gp51* sequences within (intra-sub genotype) and between (inter-sub genotype) of six exotic sub-genotype 6 groups of BLV indicated that new Vietnamese sub-genotype (G6f) strains were distantly from other exotic sub-genotype 6 group strains. In detail, the percentage of mean nucleotide distances showed that G6f sub-genotype group were most divergent to G6d sub-genotype group at 4.1%, and most closely to G6c sub-genotype group at 2.31 %. On the other hand, the percentage of mean amino acid distances presented that G6f was close to G6a at 0.72% and distant to G6c at 1.45% (Table. 2.3).

All aa substitution found in the complete BLV *env* gp51 sequences of the Vietnamese and 10 known BLV genotypes strains are summarized in Table. 2.4. A total of nine aa substitutions were found in the Vietnamese sequences. One unique aa substitution, P179L (proline to leucine), located in the overlapping region of the CD8+ T-cell epitope and epitope E, was observed in a Vietnamese G1 strain. Two other unique substitutions, Q51R (glutamine to arginine) at a “not-determined” region, and S189N (serine to asparagine) in the antigenic region E or E′ linear epitope, were observed in Vietnamese G6 strains. Three other aa mutations were found on antigenic determinant regions of gp51 in G6 strains. Two of them were located at overlapping functional regions between the secondary

neutralizing domain (ND2) and the zinc-binding peptide: one was H142R (histidine to arginine), which has also been found in BLV G7 strains in Poland and G9 strains in Bolivia; and the other was I144T (isoleucine to threonine), which has been demonstrated in BLV strains from multiple countries (G6 strains in Thailand, Brazil, Argentina, the Philippines, and Paraguay; G4 strains in Moldova, Russia, France, and Mongolia; and G10 strains in Thailand and Myanmar). Another aa substitution of the Vietnamese G6 strains, R267K (arginine to lysine), was located in the epitope D and D′ domain; this substitution was also found in G6 strains from Thailand, Brazil, Paraguay, and Argentina strains, and in G8 strains from Russia and Ukraine. Several other aa substitutions were found in the leader peptide region: C28S (cysteine to serine) in G1 strains, and P12E (proline to glutamic acid) and I31T (isoleucine to threonine) in the G6 strains of Vietnam and other countries.

2.4 Discussion

This study provided the first report of the prevalence of BLV infection in beef and dairy cattle in northern Vietnam, along with the genetic characterization of the strains isolated. Among 93 cattle tested, 33 (35.48%) were determined to be BLV-infected by ELISA and in-house *pol*-rt-PCR. One remarkable finding was that the commercial *tax*-rt-PCR failed to detect three of the positive samples (Table. 2.2). Interestingly, phylogenetic analysis based on the complete BLV *env* gp51 gene clearly indicated that the three Vietnamese strains not detected by the commercial *tax*-rt-PCR belonged to the genotype G6 (Fig. 2.2, Fig. 2.3). In the previous study of BLV infections in Myanmar, the commercial *tax*-rt-PCR also failed to detect 3 out of 6 BLV positives which were determined by a real-time BLV-CoCoMo-qPCR2 targeting the BLV LTRs, and all 6 positive samples were then identified as genotype 10 (G10) [54]. Unfortunately, the target region of the commercial *tax*-rt-PCR are not currently available. Therefore, I hypothesize that possibly some specific mutations in Vietnamese BLVs G6 *tax* genes that occurred in the target regions of the commercial *tax*-rt-PCR might cause the failure. Further studies are needed to prove this hypothesis.

It is noteworthy that the BLV infections were found in both beef and dairy cattle farms in Ba Vi district with the highest prevalence (58.54%). Whereas, in other districts, BLV infections were found in either a beef cattle farm or a dairy cattle farm (Table. 2.1). Ba Vi district is one of the main livestock districts and holds 65% of the total cattle population in Hanoi, according to a report in November 2016 of the statistic of Vietnam's Husbandry and General Statistics Office of Vietnam (https://www.channuoivietnam.com; https://www.gso.gov.vn). Some farms in the Ba Vi district raise both beef and dairy cattle together, and a husbandry practice that may increase the risk of horizontal BLV transmission between the two types of cattle. Both of the two beef cattle farms and three out of four dairy cattle farms tested were BLV-positive in this district. According to cattle husbandry recommendations of the Vietnam Department of Livestock Production (http://www.cucchannuoi.gov.vn), beef cattle are usually kept for two years and then slaughtered, whereas dairy cattle are kept for five to seven years for producing milk. Therefore, the monitoring of BLV infection in dairy cattle may be important to control EBL caused by the viral infection. Factors affecting the introduction of BLV into farms include the import, transport, and crossbreeding of cattle in Vietnam, which may be causing the genetic diversification of this virus in this country.

The phylogenetic analysis based on the complete BLV *env gp51* gene (903 bp) illustrated that 33 Vietnamese BLVs were clearly clustered into two genotypes, G1 and G6 (Fig. 2.2, Fig. 2.3). G6 has five known subgenotypes, designated G6a through G6e. I found two Vietnamese G6 strains that were distinguishable from the other G6 subgenotypes and classified into a new subgenotype, G6f (Fig. 2.4 and Fig. 2.5). One interesting point is that two Vietnamese BLVs G1 (VinhTuong_Vietnam S5F15, VinhTuong_Vietnam S6F15) and one BLVs G6 (VinhTuong_Vietnam S9F16), which were isolated from VinhTuong district, could be clustered together with Thai BLVs G1 and G6 groups, respectively. It was suggested that Vietnamese BLVs isolated from Vinh Tuong district had the same origin as Thai BLVs. This might be as a consequence of trading between two countries. Another point

to be discussed in this study is that the phylogenetic trees on the 444 bp fragment of the BLV *env gp51* gene which has been used commonly for BLV genotyping, had a limitation in classifying BLVs G6 and G10 in both methods ML and BI (Fig. 2.4, Fig. 2.5). Therefore, to differentiate BLVs G6 and G10, the complete BLV *env gp51* gene sequence are recommend for phylogenetic analysis.

Analysis of the complete BLV *env gp51* protein sequences of 33 Vietnam BLVs denoted that the sequences contained three aa substitutions (P12E, C28S, and I31T) in the leader peptide region, which may be biologically significant, as well as one unique mutation (Q51R) in a "not-determined" region. The H142R and I144T aa changes were observed in the overlapping region between the ND2, CD8+ T-cell epitope (N11), and zinc-binding peptide of the G6 strains (Table. 2.4). The region includes surface protein aa residues at 137–156, which affects the fusion and infectivity of BLV *in vivo* [26, 27]. Therefore, the mutations observed in the Vietnamese G6 strains possibly affect the fusion and infectivity of the virus. Two unique mutations that have not been described previously were found in the functional region of the Vietnam BLV gp51 protein; namely, P179L in the overlapping region between the CD8+ T cell and E epitope of the G1 strain, and S189N in the E epitope domain of the G6 strain. These two unique substitutions, along with the R267K substitution found in the Vietnam G6 strain, are located in the C-terminal half of the gp51 protein. The region contains the linear epitopes A, B, D, and E, and interacts with the TM hydrophobic domain via a disulfide bond [10, 37]. Therefore, a mutation occurring in this region may affect the interaction between the virus and the host cell membrane.

2.5 Summary

In conclusion, this is the first study on BLV infection conducted in Vietnam. The BLV infection was detected in 8 out of the 22 tested cattle farms in 4 districts in northern Vietnam, with the high positive proportion (35.48 %) as determined by both methods, ELISA and in-house *pol*-rt-PCR. One remarkable finding is that the commercial *tax*-rt-PCR failed to detect Vietnamese BLV G6. The

33 Vietnamese strains isolated were genotypes G1 and G6, one of which was a new subgenotype, G6f. The findings obtained herein will be useful for further studies on BLV infection not only in Vietnam but also in other countries that require better cattle husbandry practices. Further studies are needed to gain more understanding about the BLV infection in Vietnam.

Table 2.1 Positive rates of BLV infection in cattle farms in Hanoi, Vietnam, as determined by ELISA and real-time PCR targeting *pol* and *tax* genes

District	% ELISA-positive	% *pol* gene-positive	% *tax* gene-positive	No.positives/no.of farm tested by farm size**		No.positives/no.of farm tested by farm type	
				Small	Medium	Beef	Dairy
VinhTuong	14.29 (3/21) *	14.29 (3/21)	9.52 (2/21)	2/7	None	0/2	2/5
BaVi	58.54 (24/41)	58.54 (24/41)	53.66 (22/41)	1/2	4/4	2/2	3/4
ChuongMy	30.00 (6/20)	30.00 (6/20)	30.00 (6/20)	0/4	1/1	1/5	None
MyDuc	0.00 (0/11)	0.00 (0/11)	0.00 (0/11)	0/4	None	0/4	None
Total	35.48 (33/93)	35.48 (33/93)	32.26 (30/93)	3/17	5/5	3/13	5/9

*: No. of positives/no. of cattle tested

**: The farm holding the BLV-infected cattle.

Small: small cattle farm holding 20 or fewer cattle.

Medium: medium cattle farm holding 20-100 cattle.

Table 2.2 Molecular detection of BLV provirus DNA in cattle blood samples collected in Hanoi, Vietnam by absolute quantitative real-time PCR targeting *pol* and *tax* genes

Assays	Tested number	Positives for G1	Positives for G6	Mean copy number per *µl*	Mean Ct value
In-house *pol* rt-PCR	93	30	3	19,966	28.43
Commercial *tax*-rt-PCR	93	30	0	1,504	30.56

The mean copy number and mean Ct value was calculated from the values obtained for the positive samples.

Table 2.3 Average percentages of nucleotide and amino acid distances in the 444-bp partial BLV *env* gp51 gene sequences within (intra-subgenotype) and between (inter-subgenotype) six G6 subgenotypes

Subgenotype	G6a	G6b	G6c	G6d	G6e	G6f
G6a	**0.72 / 0.00***	0.36	0.72	0.36	0.31	**0.72[b]**
G6b	1.40	**1.21 / 0.72**	1.09	0.72	0.67	1.08
G6c	1.18	1.42	**0.48 / 0.96**	1.09	1.04	**1.45[a]**
G6d	2.63	3.14	2.85	**0.32 / 0.48**	0.67	1.08
G6e	1.20	1.36	1.19	2.39	**0.27 / 0.62**	1.03
G6f	2.35	2.61	**2.31[b]**	**4.10[a]**	2.38	**0.00 / 0.00**

*: The intra-subgenotype and inter-subgenotype distances were calculated using the Kimura 2-parameter model and Poisson model, respectively.

Left lower matrix: mean nucleotide distances. Right upper matrix: mean amino acid distances.

The values on the left in bold along the diagonal are the intra-subgenotype nucleotide distances.

The values on the right in bold along the diagonal are the intra-subgenotype amino acid distances.

Numbers in bold marked by [a], [b] indicate the maximum and minimum nucleotide and amino acid distances of G6f from the other G6 subgenotypes, respectively.

Table 2.4 Amino acid substitutions in the Vietnamese BLV G1 and G6 strains compared with those in BLV G1–G10 strains in other countries

Functions[a]	Positions[b]	Dominant [c]	G1	G2	G4	G6	G7	G8	G9	G10
Leader Peptide (1-33)[d]	12	**P**	E (AUS)			**E (VN**, THA, BRA, PRY)				
	28	**C**	**S (VN**, THA, KOR)							
	31	**I**	T (ARG, KOR)	V (PER)		**T (VN**, THA)				
Not determined	51	**Q**				**R (VN)***	L (UKR)			H (THA)
ND2 (132-149) Zinc-binding peptide (137-156)	142	**H**				**R (VN)**	R (POL)		R (BOV)	
	144	**I**			T (MDV, RUS, FRA, MNG)	**T (VN**, THA, BRA, ARG, PHL, PRY)				T (THA, MAM)
CD8+ T-cell epitope (154-183)	179	**P**	**L (VN)*** R (JP)							
Epitope E (175-194), E'(189-199)	189	**S**				**N (VN)***				
Epitope D and D' (251-270)	267	**R**				**K (VN**, THA, BRA, PRY, ARG)		K (RUS, UKR)		

Vietnamese BLV results are indicated in bold, and unique amino acid changes found only in the Vietnamese strains are denoted by asterisks.

[a]: Functions and [b]: positions of BLV *env* gp51 amino acids.

[c]: The most common amino acid observed in the 10 genotypes of worldwide BLV strains.

[d]: The position of function domain on BLV *env* gp51 amino acid sequence.

Abbreviations: ND2 (secondary neutralizing domain), VN (Vietnam), MAM (Myanmar) THA (Thailand), ARG (Argentina), JP (Japan), KOR (Korea), MNG (Mongolia), RUS (Russia), UKR (Ukraine), PER (Peru), BRA (Brazil), PHL (The Philippines), PRY (Paraguay), FRA (France), MDV (Moldova), POL (Poland), BOV (Bolivia), AUS (Australia).

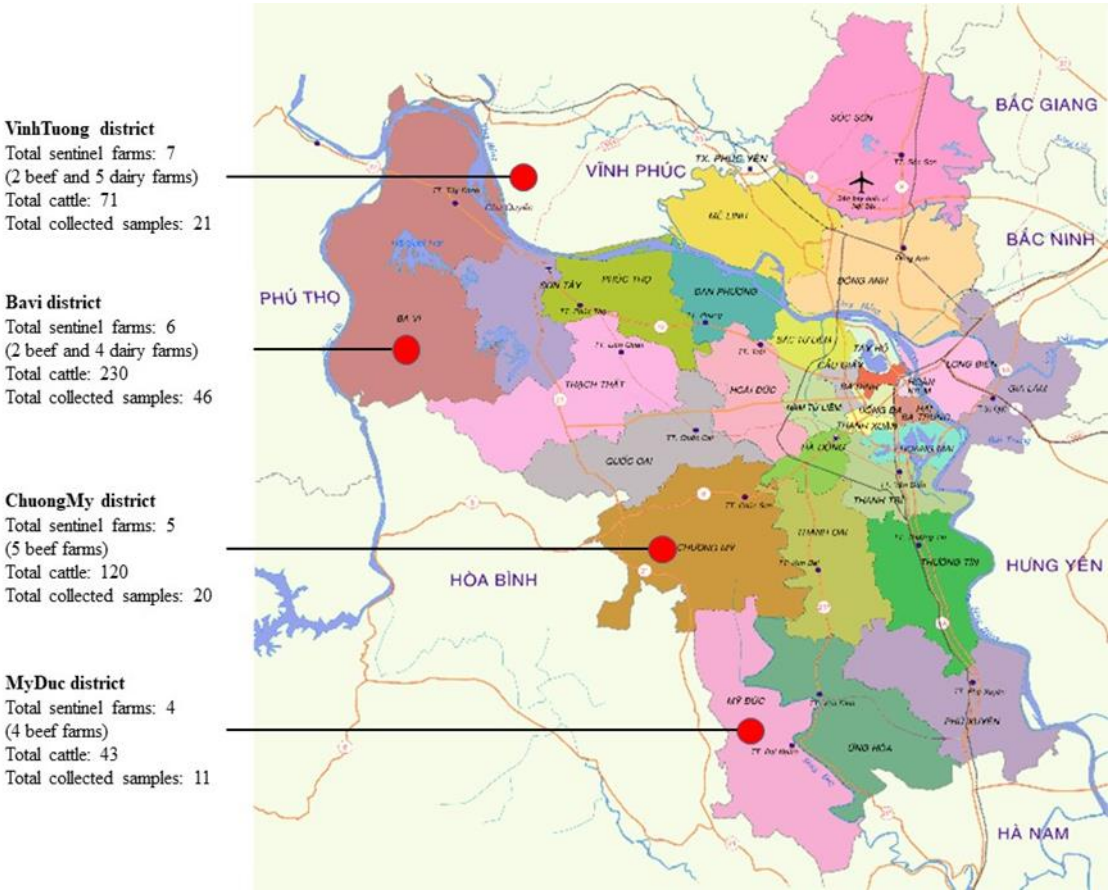


Fig. 2.1 Map of northern part of Vietnam, showing the distribution of cattle farms and the numbers of cattle sampled in the four districts tested in this study.

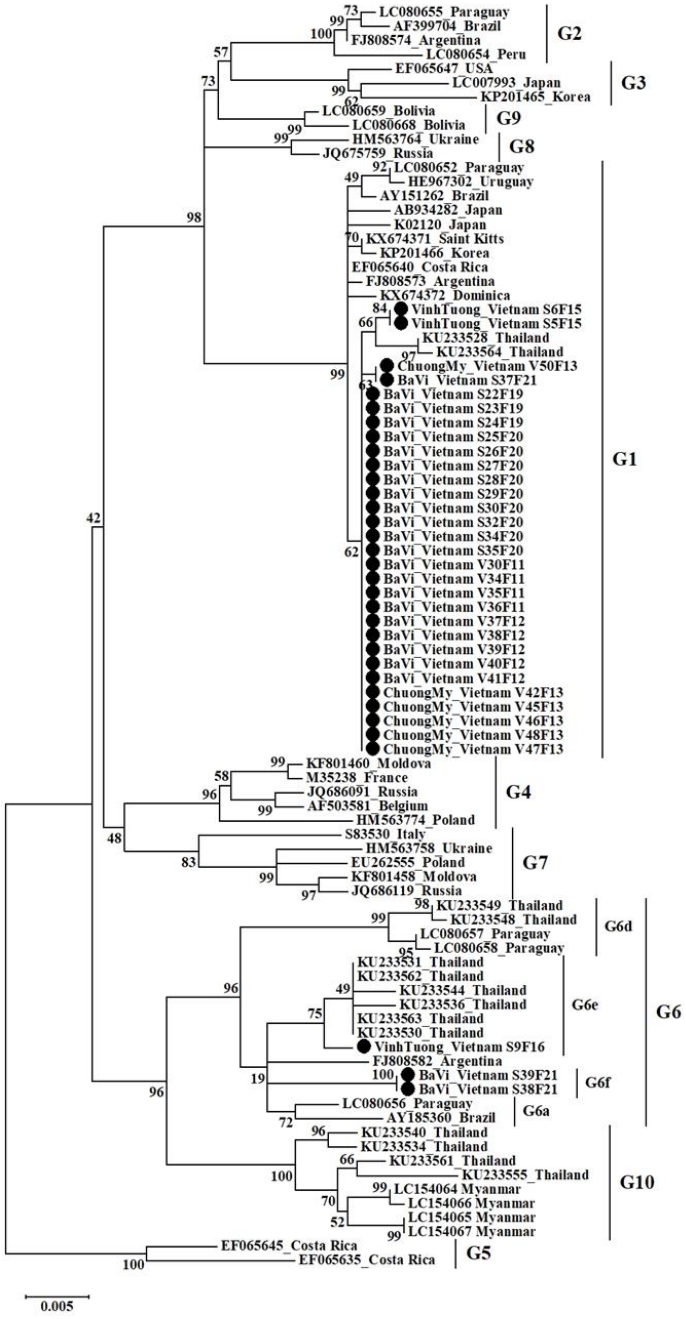


Fig. 2.2 Maximum-likelihood phylogenetic tree based on the complete nucleotide sequences of the BLV *env* gp51 gene (903 bp) from 33 Vietnam strains (marked as ●) and 56 worldwide BLV strains of 10 genotypes. Numbers at the nodes indicate the bootstrap value. Genotypes (G1–G10) are indicated by the numbers on the right of the figure. The bar at the bottom of the figure denotes the distance. The tree was rooted on G5.

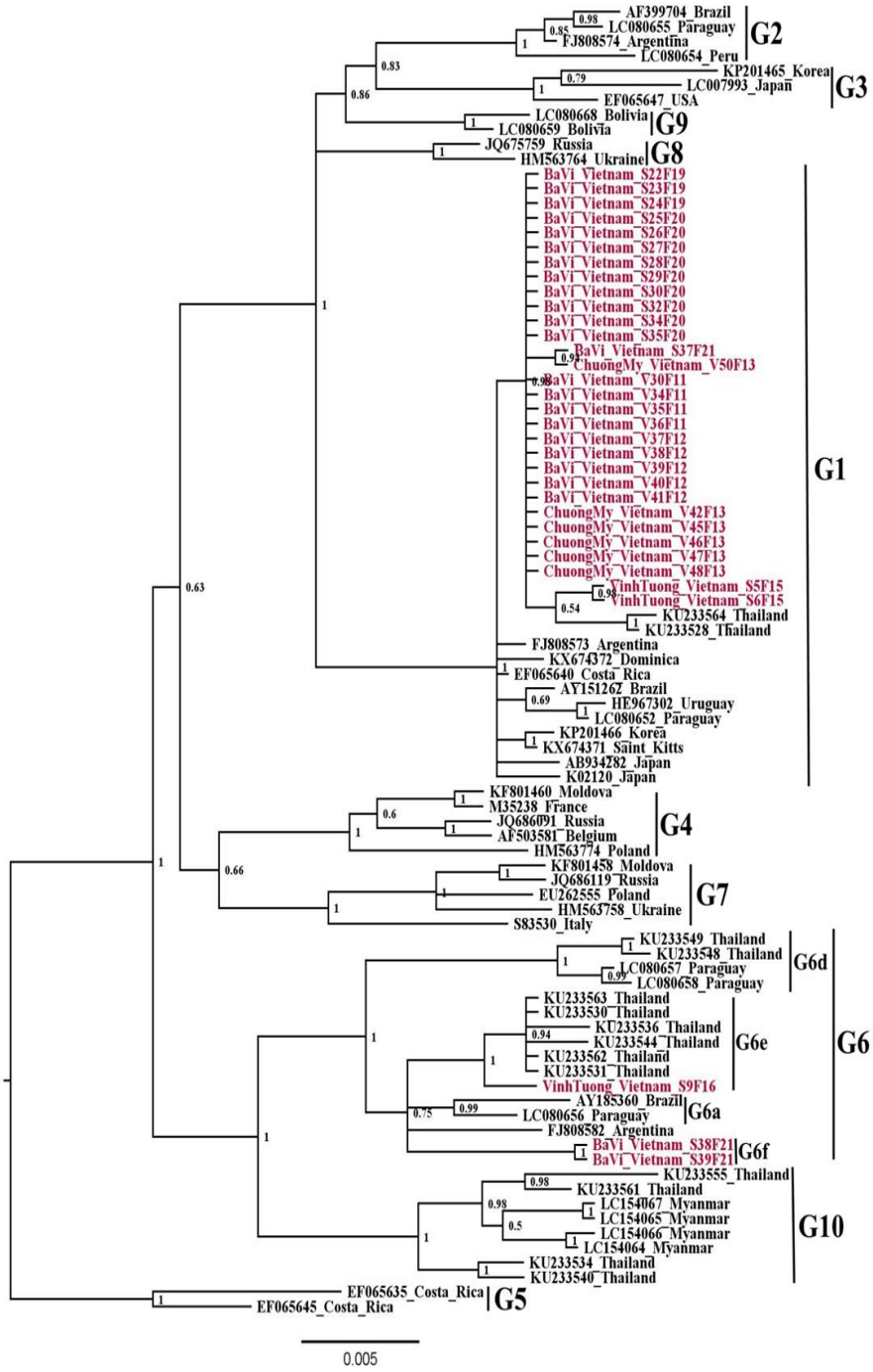


Fig. 2.3 Phylogenetic relationship of the complete nucleotide sequences of the BLV *env* gp51 gene (903 bp) from 33 Vietnam strains (showed in red) and 56 worldwide BLV strains of 10 genotypes, constructed by Bayesian inference. Numbers at the nodes indicate the posterior probabilities of sampling the node among 15,002 trees. Genotypes (G1–G10) are indicated by the numbers on the right of the figure. The bar at the bottom of the figure denotes the distance. The tree was rooted on G5.

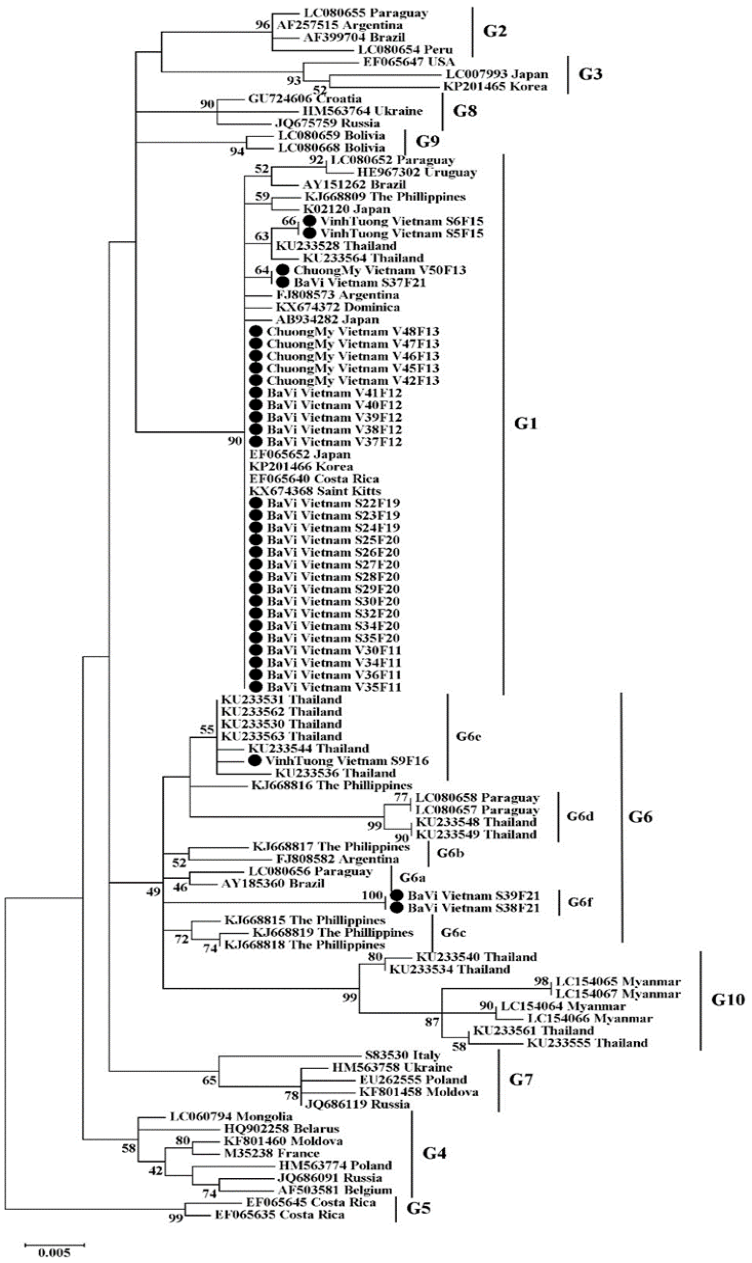


Fig. 2.4 Maximum-likelihood phylogenetic tree based on the 444-bp partial nucleotide sequences of the BLV *env* gp51 gene from 33 Vietnam strains (marked as ●) and 66 worldwide BLV strains of 10 genotypes. Numbers at the nodes indicate the bootstrap value. Genotypes (G1–G10) are indicated by the numbers on the right of the figure. The bar at the bottom of the figure denotes the distance. The tree was rooted on G5.

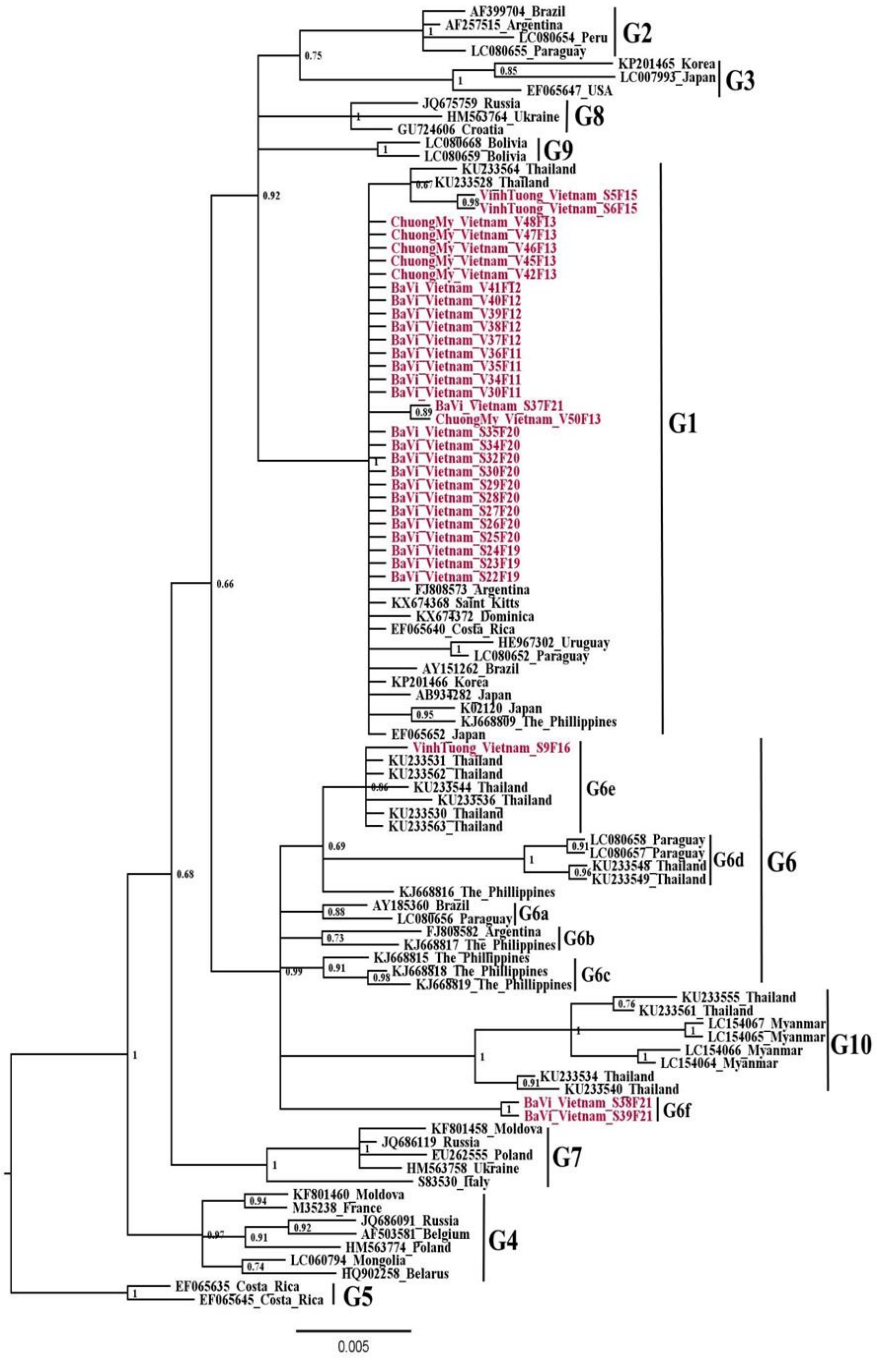


Fig. 2.5 Phylogenetic relationship of the 444-bp partial nucleotide sequences of the BLV *env* gp51 gene from 33 Vietnam strains (showed in red) and 66 worldwide BLV strains of 10 genotypes, constructed by Bayesian inference. Numbers at the nodes indicate the posterior probabilities of sampling the node among 15,002 trees. Genotypes (G1–G10) are indicated by the numbers on the right of the figure. The bar at the bottom of the figure denotes the distance. The tree was rooted on G5.

Chapter 3

Usefulness of a SureSelect target enrichment method for next-generation sequencing and its application to obtain the complete genome of Vietnamese bovine leukemia virus, it's genetic variability, phylogenetic and phylogeographic analysis

3.1 Introduction

I was engaged in surveillance and epidemiology studies on BLV infections in Vietnamese cattle. I am interested in sequencing the complete genome some Vietnamese BLV strains to gain understanding the genetic variability of Vietnamese BLV strains, especially, by comparing the genetic variability of *pol* and *tax* gene of Vietnamese BLV G1 strains (G1VN) with Vietnamese BLV G6 strains (G6VN). This may give insight to better understand why G6VN was not to detected by the commercial *tax*-rt-PCR (Takara Bio Inc.), which was successfully determined to be positive by in-house *pol*-rt-PCR as reported in Chapter 2. However, BLV is an exogenous virus which integrates its genome to host B-cell genome and coexists with the host B-cell as provirus [23]. In addition, genomic viral DNA extracts of the BLV infected samples contains a large proportion of the host cell genome. Therefore, sequencing the complete BLV genome by Sanger sequencing or conventional next-generation sequencing (NGS) methods required very sensitive PCR techniques and plenty of time to process sequencing [53, 54, 72].

A SureSelect target enrichment method has been developed and applied successfully to the whole genome sequencing of multiple viruses [8, 9, 15]. This method selectively captures the nucleic acids of the virus by using custom-designed streptavidin bead-bound biotinylated RNAs that are complementary to all the genome viral. The captured viral nucleic acids are amplified by PCR, and then proceed to NGS of the complete viral genomics.

Agilent Technologies (Tokyo, Japan) has developed a custom DNA capture-solution enrichment method for NGS called Agilent SureSelect^{QXT}. Recently, we have collaborated with Research and Education Center for Prevention of Global Infectious Diseases of Animals, Tokyo University of Agriculture and Technology (TUAT) for evaluating a new comprehensive method for detecting livestock-related pathogenic viruses using the Agilent SureSelect^{QXT} system [45]. Here I applied the custom-made livestock-related pathogenic virus detection system to the whole genome sequencing of the Vietnamese BLVs. The aims of this study were to evaluate the efficiency and usefulness of SureSelect target enrichment method for deep sequencing the whole BLV genome, and to sequence the complete some Vietnamese BLVs to gain understanding the genetic variability, phylogeny, and phylogeography of Vietnamese BLVs.

3.2 Materials and Methods

3.2.1 Samples

Four extracted genomic DNA samples which were isolated from the whole blood of BLV infected cattle in three districts (Vinh Tuong, Chuong My, and Ba Vi) in northern Vietnam, were used to evaluate the efficiency and usefulness of Agilent SureSelect^{QXT} system. Four selected samples including 'VinhTuong Vietnam S6F15, VinhTuong Vietnam S9F15, ChuongMy Vietnam V50F13 and BaVi Vietnam S38F21' belong to two Vietnamese BLV genotypes G1 and G6 as reported in Chapter 2.

3.2.2 Custom capture library

The SureSelect XT custom capture library, which was designed by Research and Education Center for Prevention of Global Infectious Diseases of Animals, TUAT [45] to capture various viral pathogens of domestic animals including BLV, was used for the study.

3.2.3 Sample preparation for library construction

BLV DNA provirus samples were quantified by NanoDrop spectrophotometric, and the DNA libraries were generated by randomly fragmenting genomic DNA samples using the SureSelect QXT Reagent kit (Agilent Technologies). Then, the DNA samples were purified and amplified by Agencourt AMPure XP (Beckman Coulter, Inc., Brea, CA, USA), and the SureSelect QXT Primer Mix, respectively. After that, the DNA samples were purified using Agencourt AMPure XP. The adapter-attached DNA libraries were hybridized to the SureSelect XT custom capture library, and the captured DNA was purified by using Dynabeads MyOne Streptavidin T1 beads (Thermo Fisher Scientific). Finally, the nucleic acids of the target viruses were enriched with index primers and purified by Agencourt AMPure XP. The obtained purified product was subjected to NGS analysis. All conditions for each step were conducted according to the recommendation of the SureSelect QXT Target Enrichment kit.

3.2.4 Deep sequencing and sequencing data analysis

Deep sequencing was performed on a MiSeq benchtop sequencer (Illumina, San Diego, CA, USA). The constructed library was analyzed as a read of a 76-bp paired-end by means of the MiSeq Reagent Kit v3 (150 cycles) (Illumina). The sequence of each obtained read was output in FASTAQ format, using a MiSeq reporter, and analyzed in CLC Genomic Workbench 6.5.1 (CLC bio, Aarhus, Denmark). Each read was processed by a quality trim command to trim low-quality sequences, and contigs were obtained via the *de novo* assembly command. A BLAST search was conducted on all contigs using a virus database obtained from NCBI.

3.2.5 Phylogenetic analysis

The complete BLV nucleotide sequences of four Vietnamese strains were aligned with those of 28 worldwide BLV reference strains of genotypes 1, 2, 4, 6, 9 and 10, which are available on GenBank, using the CLUSTALW multiple alignment tool in BioEdit v.7.2.5 (http://www.mbio.ncsu.edu/bioedit/bioedit). A maximum-likelihood (ML) tree for full BLV genome

sequences was constructed using MEGA 6 [69] applying 1,000 bootstrap replicates with the best-fit model TN93+G.

3.2.6 Phylogeographic analysis

To get insight into a temporal history of spatial molecular epidemiology of BLV infection in Vietnam, I modeled a phylogeographic analysis for BLV infection. Nine Vietnamese BLV *env gp51* gene nucleotide sequences obtained in Chapter 2 (GenBank accession numbers MF817716–MF817724) were aligned using CLUSTALW multiple alignment in BioEdit v7.2.5 with 105 worldwide BLVs which were isolated from 24 countries and presented for all ten genotypes (G1-G10) as described in the previous studies [5, 32, 46, 51-54, 59, 72, 74]. The decimal date of each sequence was collected based on its released date on GenBank. The Bayesian maximum clade credibility (MMC) location discrete traits tree was constructed by using software BEAST v1.8.4 (http://tree.bio.ed.ac.uk/software/beast/). The molecular clock rate was estimated by software TempEst v1.5.1 (http://tree.bio.ed.ac.uk/software/tempest/). The uncorrelated lognormal relaxed molecular clock and SRD06 model of the normal distribution of nucleotide substitution were used, with a constant population size coalescent tree prior, and the location discrete traits was applied gamma distribution rate variation. Time to most recent common ancestor (TMRCA) was set up for Vietnamese BLV group to understand when nine Vietnamese BLVs met their common ancestor. The Bayesian Markov Chain Monte Carlo (MCMC) was run two individual times at 50,000,000 generations and sampled at every 5,000 generations. The effective sample size (ESSs) of the analysis was checked by software Tracer v1.6 (http://tree.bio.ed.ac.uk/software/tracer/). The MMC location discrete traits output tree was reconstructed based on the location of sampled sequences by using software FigTree v.1.4.3 (http://tree.bio.ed.ac.uk/). Finally, the MMC location discrete traits tree was plotted on software SPREAD v1.0.6 (https://www.kuleuven.be/aidslab/phylogeography/SPREAD.html) to get an overview of the temporal history of spatial molecular epidemiology of worldwide BLV infections, and

for generating KMC file which was used for visualizing MMC tree on Google Earth software (http://earth.google.com).

3.2.7 Genetic analysis

Nucleotide sequences of the four Vietnamese BLV strains were aligned with worldwide BLVs and the FLK-BLV sub-clone pBLV913 sequence (NCBI, EF600696) was used as a consensus sequence. All amino acids (aa) sequences of worldwide BLVs and four Vietnamese BLVs were compared with the reference strains throughout the essential part (including *gag*, *pro*, *pol,* and *env*) and the regulator part (including *tax*, *rex*, *R3*, and *G4)* of BLV genome.

To compare the mean diversity of 4 Vietnamese BLVs with 28 worldwide BLVs, I calculated the average substitution rate of nucleotides and amino-acids for whole genome and each of the structural genes (including *gag* (*p15*, *p12* and *p21*), *pro*, *pol* and *env* (*gp51* and *gp30*), and non-structural genes (including *tax*, *rex*, *R3*, *G4* and *LTRs*) for total 32 BLVs using Jukes-Cantor and p-distance model in MEGA 6 [28, 68], respectively.

To compare genetic variability of *pol* and *tax* gene within and between BLVs G1VN and BLVs G6VN, the average percentage of nucleotide distances of BLV *pol* and *tax* gene sequences within (intra-genotype) and between (inter-genotype) BLVs G1VN and G6VN were calculated by Jukes-Cantor model in MEGA 6, respectively.

Unfortunately, the detection site of the commercial *tax*-rt-PCR is not currently available. Therefore, to identify the specific mutations which possibly abrogate the *tax* gene detection of the commercial *tax*-rt-PCR kit, which failed to detect BLVs G6VN and G10 Myanmar BLV (described in Chapter 2). Four full *tax* gene sequence of Vietnamese BLVs obtained in this study, were aligned with those of the G10 Myanmar strains and worldwide BLV strains to find out the exclusive mutations on the both G6VN and G10 Myanmar *tax* gene.

3.3 Results

3.3.1 Evaluation of SureSelect target enrichment method for deep sequencing the complete BLV genome.

The complete four BLVs genomics of Vietnamese BLV DNA provirus samples were sequenced successful by using the SureSelect target enrichment method. The whole process was implemented in less than one week but satisfied with high read depth values and genome coverage of 100% across all sequenced samples. Obtained sequences were deposited to GenBank under accession numbers (MH170027-MH170030).

3.3.2 Phylogenetic analysis

The phylogenetic analysis based on the complete BLV genome nt sequences indicated that the four Vietnamese BLV strains clearly clustered into genotype 1 and genotype 6 (Fig. 3.1) supporting by 100 bootstrap value for every clade.

3.3.3 Phylogeographic analysis

The MMC location discrete traits tree based on the BLV *env gp51* gene analysis demonstrated that Vietnamese BLVs were clearly distinguished into two groups. One group had the same origin as Japanese and Korean strains, while another group had the same origin as the strains isolated in Brazil, Argentina, Paraguay, Thailand, and China (Fig. 3.2). In addition, TMRCA among Vietnamese BLVs was determined in 1909.497 (1909/6/30) (data not show) that was close to the TMRCA of the whole tree in 1902.73 (1902/9/23). Taken together it is suggested that the current BLV infection in Vietnam was transmitted from at least two origins. Moreover, the plotted MMC location discrete traits tree on SPREAD illustrated that Vietnamese BLVs were related to Thailand and Japan (Fig. 3.3). Visualizing the result on Google Earth software demonstrated that, BLV was transmitted to Thailand from the USA in 1905, and 45 years later Thai BLVs began spreading to Vietnamese cattle. On the other hand, BVL was imported to Japan from the USA in 1982, then it started spreading to

Vietnamese cattle in 2004 (Fig. 3.4). Overall, BLV might circulate among Vietnamese cattle in 2011 as a consequence of the transmission from Thailand and Japan.

3.3.4 Genetic analysis

Aligning Vietnamese BLVs LTR nucleotide sequences with worldwide BLVs reference sequences, 2 deletions (at positions 324 and 373), and 4 insertions (at position 402) of nucleotides in Vietnamese G6VN sequences were observed (Fig. 3.5). Comparing throughout the aa sequences of the essential part and the regulator part of Vietnamese BLVs and worldwide BLVs genomes, total 33 unique aa substitutions were observed in four Vietnamese BLVs (10 in BLVs G1VN and 23 in BLVs G6VN) as summarized in Fig. 3.6. In detailed, 6 substitutions located in the essential part (including Y103H, V318M, P379S in *gag*; Q275R, A476T in *pol*; and P179L in *env*), 4 substitutions located in the regulator part (including R85C, P93R, G148D in *rex* and C25G in *R3*) of BLVs G1VN. The 23 unique aa substitutions in BLVs G6VN contained 13 substitutions in the essential part including V318A, V318T in *gag*; V91I, V91L, R149C in *pro*; I106V, Q295R, N374S, G438R, K6277R in *pol*, Q51R, S189N, L327F in *env*, and 10 substitutions in the regulator part including: N90S, Q99R, K149R, R271H, R287K in *tax*; P93Q, S105A, P116G in *rex*; L9P, I80V in G4.

Analysis of the mean diversity of 4 Vietnamese BLVs with 28 worldwide BLVs demonstrated that the average nucleotide substitution rate of the whole BLV genome was 0.036 per site. Comparison between each gene revealed that *p15* gene had the highest average rate of nucleotide substitution at 0.048 while *LTRs* gene was significant conserved at 0.019. Comparison of amino acid substitution rate between BLV proteins indicated that *R3* was the most polymorphic at 0.098, while *p24* was very conservative at 0.009 (Table. 3.1).

The comparison of mean nucleotide distances of BLV *pol* and *tax* gene sequences within (intra-genotype) and between (inter-genotype) BLVs G1VN and BLVs G6VN revealed that BLVs G6VN *tax* gene was more divergent than its *pol* gene. Indeed, the BLVs G6VN is distant from BLVs G1VN at

4.6 % for *pol* gene and at 5.8 % for *tax* gene. The divergence within BLVs G1VN *pol* gene was 0.3 %, and *tax* gene was 0.7 %. On the other hand, the divergence within BLVs G6VN *pol* gene was 1.4 % and *tax* gene was 1.6%. (Table. 3.2).

The alignment of four full *tax* gene sequences of Vietnamese BLVs with those of the G10 Myanmar and worldwide BLVs illustrated a total 7 specific nucleotide mutations that were found exclusively in the BLV *tax* gene sequences of the G6 BLV strains including the Vietnamese strains and G10 Myanmar strains (Fig. 3.7).

3.4 Discussion

The SureSelect target enrichment method is a highly effective method for deep sequencing the complete BLV genome from DNA provirus samples across genotypes, with high read depth values and genome coverage of 100% across all sequenced samples. Recent studies have reported BLV whole genome sequencing using overlapping PCR amplicons and Sanger sequencing [72], cloning of PCR product and Sanger sequencing [54] or using combination of overlapping PCR amplicons with SureSelect QXT Library prep for Illumina Multiplexed Sequencing [53]. Those methods required very sensitive PCR assays and plenty of time to process. In contrast, the new comprehensive method for detecting livestock-related pathogenic viruses using the Agilent SureSelect^{QXT} system that was developed by Oba et al. (2018) was applied successfully to sequence the whole genome of Vietnamese BLVs from DNA provirus samples. This method is usefulness, efficiency, faster turn-around times and easy to implement.

The phylogeographic analysis revealed that BLV might circulate among Vietnamese cattle in 2011 as a consequence of the BLV transmission from Thailand and Japan. This might be the consequence of cattle trading between these countries. The complete references BLV *gp51* nt sequences available were limited on GenBank. Therefore, I could not collect enough representative references of BLV strains isolated elsewhere, and this might affect the phylogeographic analysis

result due to miss some middle transmission locations. However, it could not decline that the phylogeographic analysis in this study gave an overview to gain better understanding the temporal history of spatial molecular epidemiology of BLV infections in 25 worldwide countries (Fig. 3.2, Fig.3.3 and Fig. 3.4).

The genetic analysis results showed that one nucleotide deletion and 4 nucleotide insertions in LTRs nt sequences of BLV G6VN 'BaVi Vietnam S38F21' were located in the downstream activator sequence (DAS) [16] (Fig. 3.5). Those mutations may influence the expression of BLVs G6VN. The comparison of aa substitutions clearly showed that Vietnamese BLVs were different from the others worldwide BLVs. The occurrence of unique aa substitutions on Vietnamese BLVs demonstrated that each genotype harbored genotype-specific mutations that distinguished it from the others.

In addition, the higher divergence of BLVs G6VN *tax* gene and the seven specific nucleotide mutations were found exclusively in the BLV *tax* gene sequences of the G6 BLV strains including the Vietnamese strains and G10 Myanmar strains. Those mutations possibly abrogate the gene detection of the commercial *tax*-rt-PCR kit. This is a plausible explanation why BLV G6VN were not detected by the commercial *tax*-rt-PCR kit. In addition, the mean diversity analysis of total 32 BLVs presented the significant conservation of LTRs gene that can be used as a targeting region, to design the primers and probe for developing a new real-time PCR detection assay that possibly improve the sensitivity and efficiency of BLV real-time PCR detection assays.

3.5 Summary

This study provided the information of the complete Vietnamese BLV genome sequences at the first time. The phylogeographic analysis gave an overview the temporal history of spatial molecular epidemiology of BLV infection in Vietnam, that were imported from Thailand and Japan as consequence of trading between countries. However, the study on BLV infection is very limited in Vietnam. Therefore, it is necessary to do a nationwide survey in Vietnam to understand better the

dynamic and evolution history of BLV infection in Vietnam, in order to prevent and control new infections. Additionally, the analysis of the sequence data obtained on this study gave new insight into the genetic diversity of BLV *tax* genes, which are used in the commercial *tax*-rt-PCR detection kits. The results obtained may contribute to improving the reliability of such detections systems. Moreover, the study evaluated that the SureSelect target enrichment system is an efficient and useful method for deep sequencing the whole BLV genome from DNA provirus samples.

Table 3.1 The average substitution rate per site for each gene and the complete genome of the four Vietnamese BLV strains in comparison with those of the 28 BLV strains obtained worldwide.

		Structural genes							Non-structural genes				
	LTR	**p15**	**p24**	**p12**	**pro**	**pol**	**gp51**	**gp30**	**tax**	**rex**	**R3**	**G4**	**All**
Nucleotide	**0.019**	**0.048**	0.044	0.041	0.039	0.035	0.035	0.033	0.041	0.028	0.042	0.024	0.036
Amino acid	-	0.047	**0.009**	0.031	0.027	0.021	0.03	0.025	0.04	0.062	**0.098**	0.05	-

The average nucleotide substitution rate per site for each gene and complete gene were calculated by Jukes-Cantor model.

The average amino acid substitution rate per site for each gene and complete gene were calculated by p-distance model.

Values in bold indicate the minimum and maximum of average substitution rate per site of nucleotide and amino acid.

Table 3.2 The average percentage of nucleotide distances of BLV *pol* and *tax* gene sequences within (intra-genotype) and between (inter-genotype) G1VN-BLV and G6VN-BLV

Genotype	**BLV *pol* gene**		**BLV *tax* gene**	
	G1VN-BLV	G6VN-BLV	G1VN-BLV	G6VN-BLV
G1VN-BLV	**0.3**[a]		**0.7**[a]	
G6VN-BLV	**4.6**[b]	**1.4**[a]	**5.8**[b]	**1.6**[a]

The values for intra-genotype and inter-genotype were calculated by Jukes-Cantor model in MEGA 6, respectively.

[a]: the intra-genotype nucleotide distances

[b]: the inter-genotype nucleotide distance

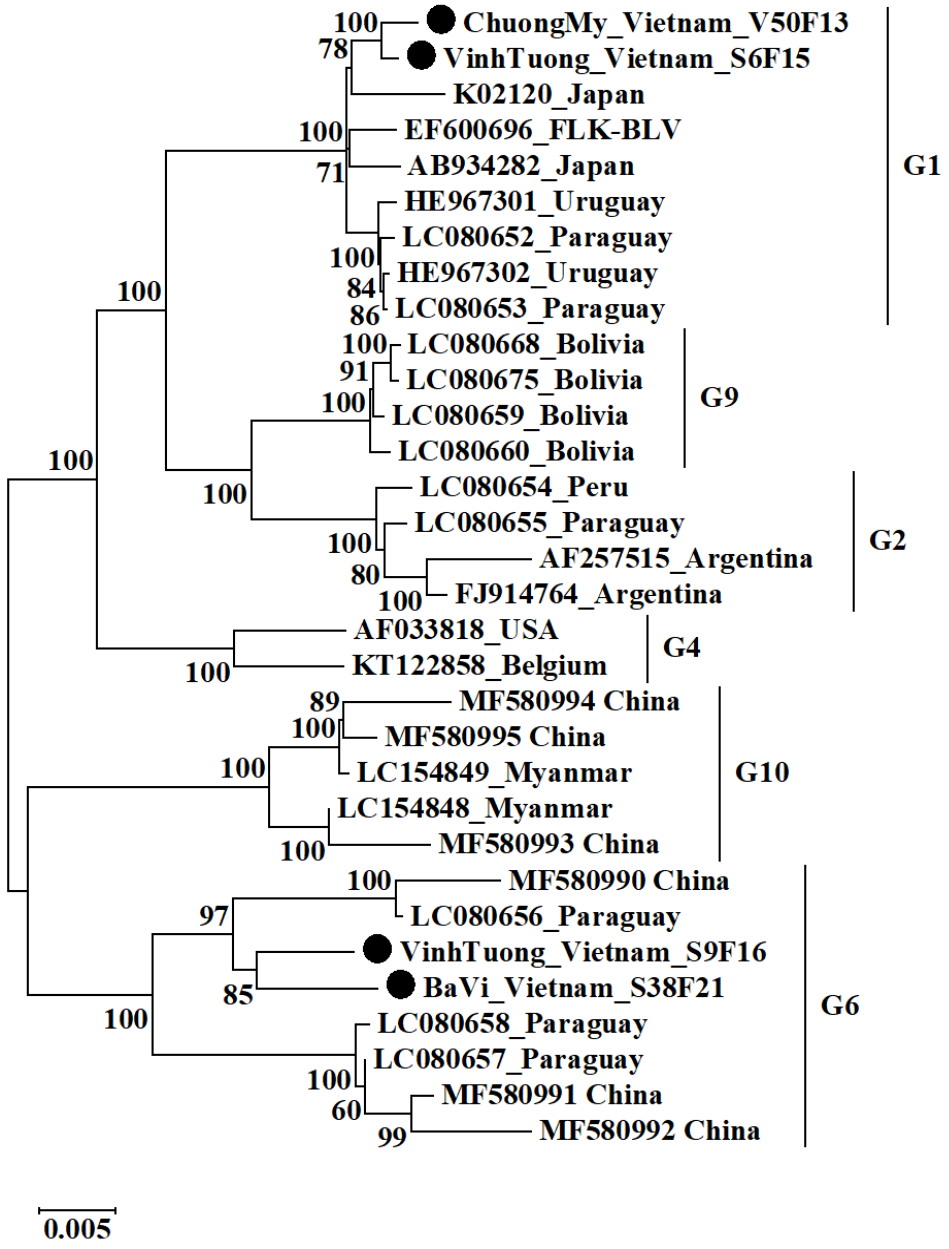
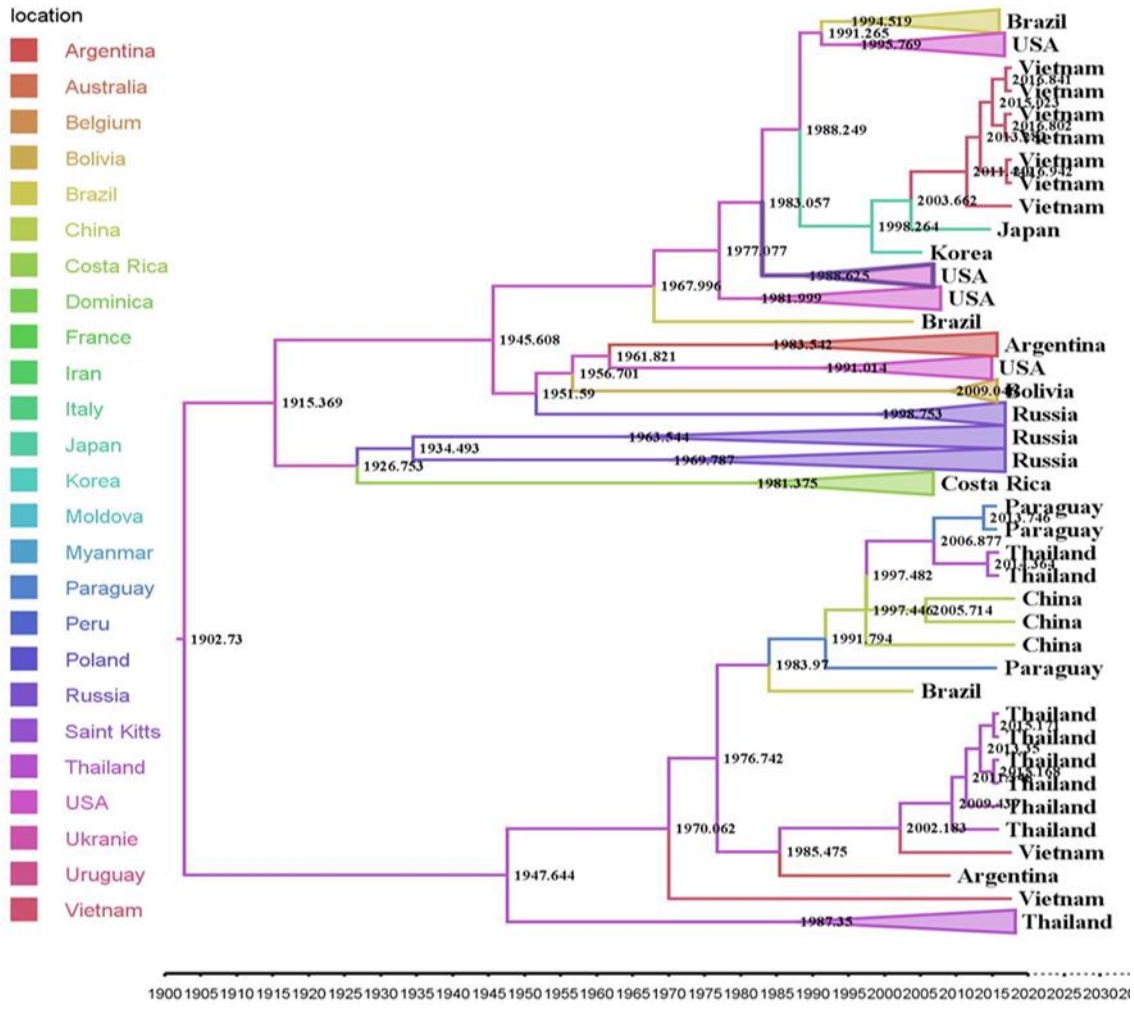


Fig. 3.1 ML phylogenetic tree constructed from the complete 4 Vietnamese BLV nucleotide sequences (marked as ●) and 28 complete BLV reference sequences obtained from the GenBank nucleotide sequence database. Genotypes (G1-G10) are indicated by numbers on the right of the figure. The bar at the bottom of the figure denotes evolutionary distance.



Time to Most Recent Common Ancestor (TMRCA) ←

Fig. 3.2 MMC location discrete traits tree of total 114 worldwide BLV gp51 gene sequences. Bar scale indicates the divergent time of the tree.

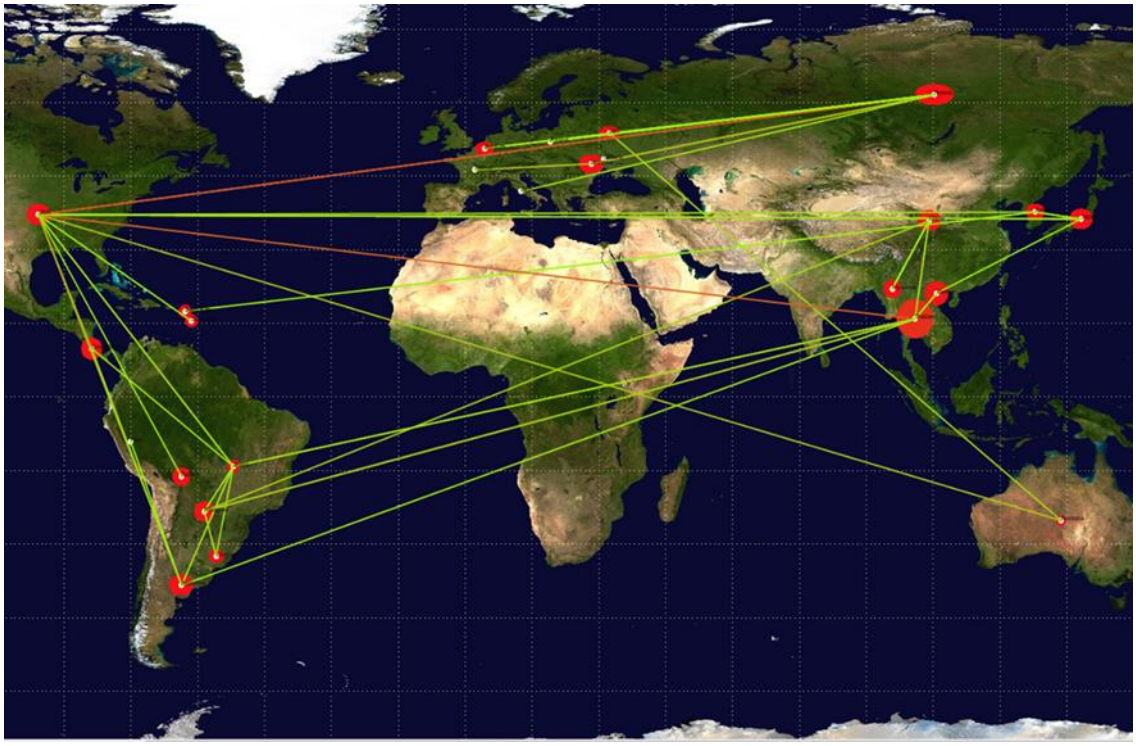


Fig. 3.3 Temporal linkage of spatial molecular epidemiology of worldwide BLV infections. Red circles indicate countries where BLV infection were reported, green and red lines indicate correlative transmission of BLV infection between two countries, (red lines indicate strong correlation)

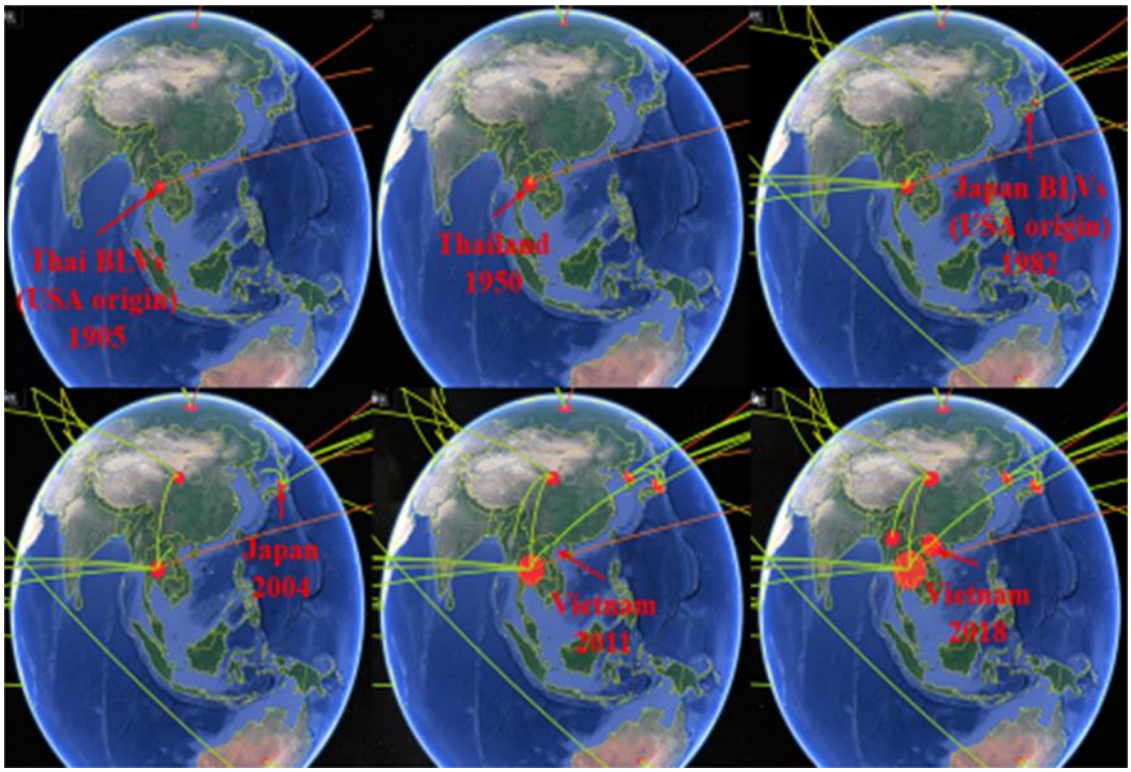


Fig. 3.4 Temporal history of spatial molecular epidemiology of BLV infection in Vietnam. Red circles indicate countries where BLV infection were reported, green and red lines indicate correlative transmission of BLV infection between two countries, (red lines indicate strong correlation)

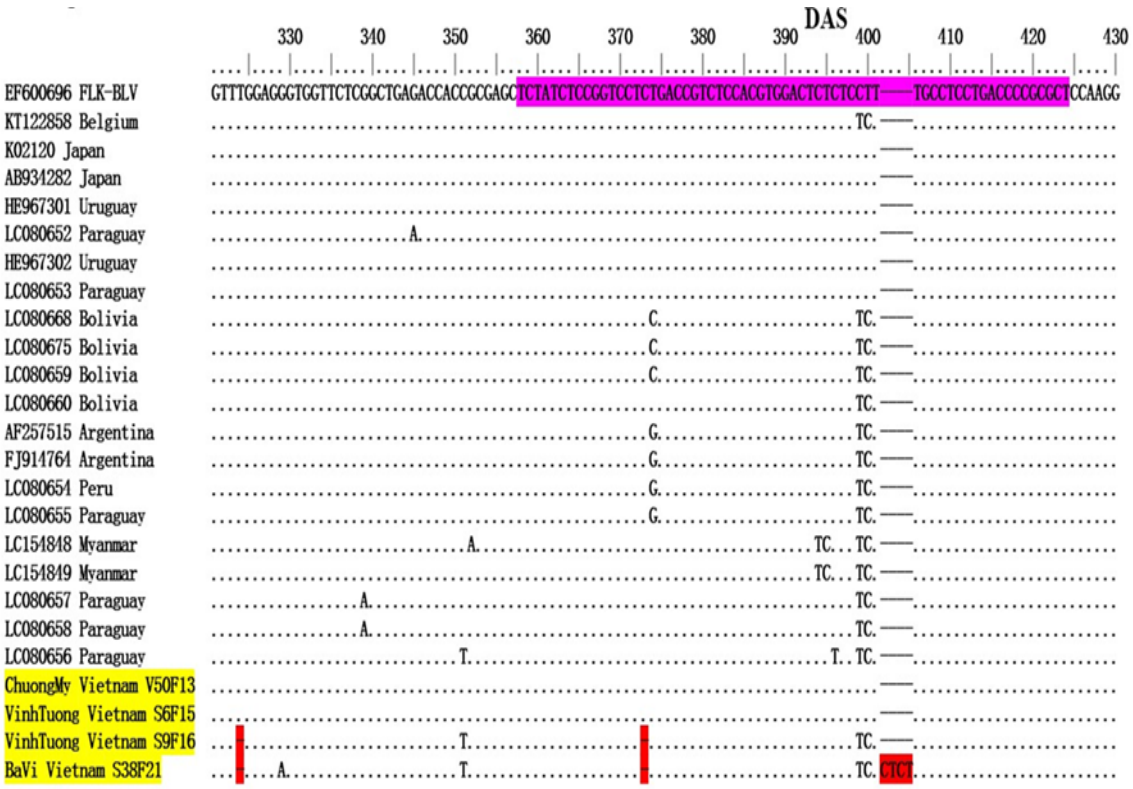


Fig. 3.5 Alignment of 5' LTR nucleotide sequences (nt 320-430) of four Vietnamese BLV strains (showed in yellow) with FLK-BLV strain and worldwide BLVs. Identity with the FLK-BLV strain is indicated by dots. The nucleotide deletions and insertions in Vietnamese BLVs are indicated by red. The downstream activator sequence (DAS) is indicated in purple

Genotype	Accession number	Gag 103	Gag 318	Gag 379	Pro 91	Pro 149	Pol 160	Pol 275	Pol 295	Pol 374	Pol 438	Pol 476	Pol 672	Env 51	Env 179	Env 189	Env 327	Tax 90	Tax 99	Tax 149	Tax 271	Tax 287	Rex 85	Rex 93	Rex 105	Rex 116	Rex 148	R3 25	G4 9	G4 80
G1	EF600696_FLK	Y	V	P	V	R	I	Q	Q	N	G	A	K	Q	P	S	L	N	Q	K	R	R	R	P	S	P	G	C	L	I
	K02120_Japan	.	M	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	AB934282_Japan	.	M	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	HE967301_Uruguay	.	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LC080652_Paraguay	.	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	HE967302_Uruguay	.	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LC080653_Paraguay	.	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ChuongMy_Vietnam_V50F13	H	M	S	.	.	.	R	.	.	.	T	.	.	L	.	.	.	.	.	.	.	C	.	.	.	D	G	.	.
	VinhTuong_Vietnam_S6F15	.	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.
G2	AF257515_Argentina	.	I	L	I	P	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	FJ914764_Argentina	.	I	L	I	P	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LC080654_Peru	.	I	.	I	P	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LC080655_Paraguay	.	I	.	I	P	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
G4	AF033818_USA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	H	.	.	.	.	.	.	.
	KT122858_Belgium	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
G6	LC080656_Paraguay	.	T	.	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.
	LC080657_Paraguay	.	T	.	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LC080658_Paraguay	.	T	.	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	VinhTuong_Vietnam_S9F16	.	A	.	I	C	V	.	.	S	.	.	R	R	.	.	F	.	.	.	H	.	.	Q	.	.	.	.	.	.
	BaVi_Vietnam_S38F21	.	T	.	L	C	.	.	R	.	R	.	.	.	.	N	.	S	R	R	.	K	.	Q	A	G	.	.	P	V
G9	LC080668_Bolivia	.	I	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
	LC080675_Bolivia	.	I	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
	LC080659_Bolivia	.	I	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
	LC080660_Bolivia	.	I	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
G10	LC154848_Myanmar	.	.	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LC154849_Myanmar	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.

Fig. 3.6 Summary of unique amino acid substitutions in the four complete genome sequence of Vietnamese BLVs compared with worldwide BLVs

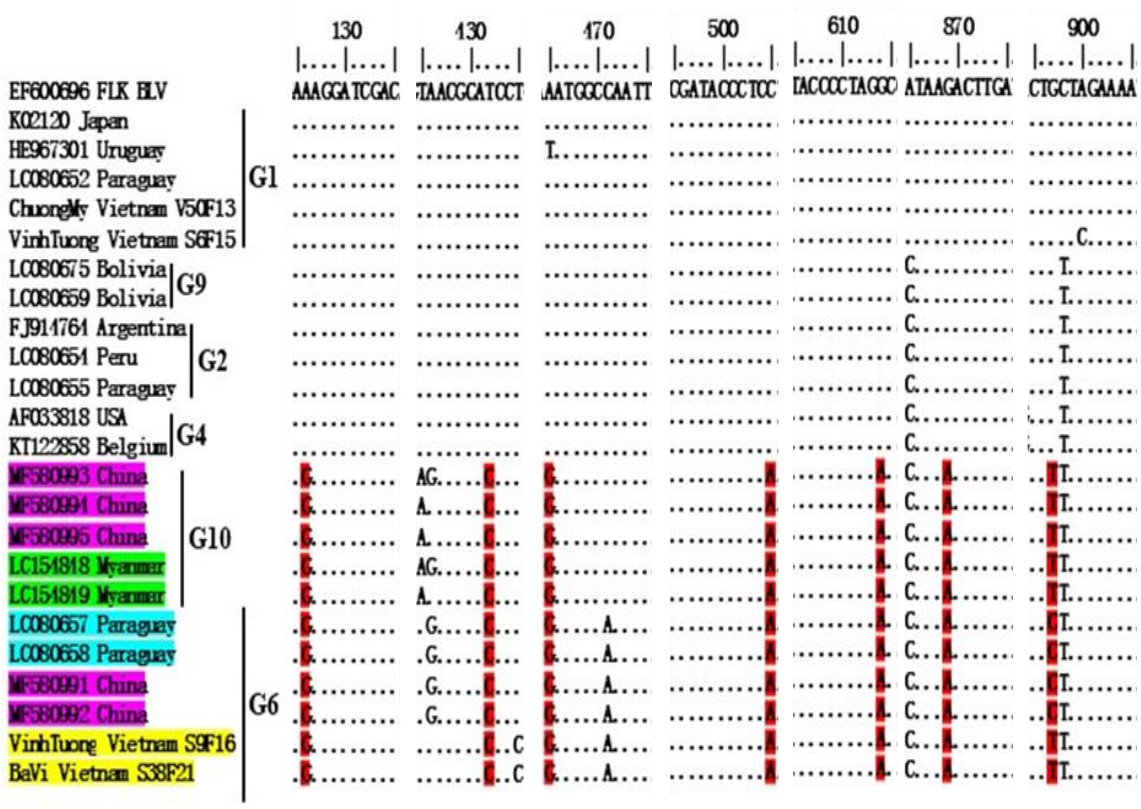


Fig. 3.7 Identifying exclusive mutations on BLV tax gene sequences for G6 and G10

Identify with the consensus FLK-BLV strain is indicated by dots.

Yellow: Vietnamese BLVs G6

Blue: Paraguay BLVs G6

Green: Myanmar BLVs G10

Purple: Chinese BLVs G6 & G10

Red: Specific mutations

General discussion

BLV is the contagious agent that causes EBL, the most common neoplastic disease of cattle [11, 66]. BLV infection has been reported all continents and caused a severe economic impact on the dairy cattle industry [56]. The virus induces a persistent and chronic infection that essentially affects the B lymphocyte population. The progress of EBL can be divided into three states including asymptomatic, persistent lymphocytosis (PL) and lymphoma [3]. Each state of the disease can be differentiated by the increasing of BLV proviral copy numbers in the lymph nodes of an infected animal [3, 67]. According to the Somura et al. (2014) the slight difference of BLV proviral copy number in whole blood samples of EBL and BLV infected cattle led to the difficulty to diagnosis EBL cattle from infected cattle. However, the BLV proviral copy number between EBL cattle and BLV infected cattle in lymph node samples was significantly different. Therefore, it is possible to distinguish EBL cattle from BLV infected cattle by observing the BLV proviral copy number in lymph node samples. In this study, the whole blood samples of cattle were collected for molecular and serology detection of BLV infection. The in-house *pol*-rt-PCR can be used for detection BLV DNA proviral in blood or lymph node samples. Therefore, for differential diagnosis of EBL from the BLV infected cattle, it is required to collect lymph node sample of BLV infected cattle by using core needle biopsy or fine-needle aspirate.

Phylogenetic analysis in this study implied that using the 444 bp fragment of the partial *env gp51* gene at the 277–720 nt position, which was previously identified by restriction fragment length polymorphism and has often been used for BLV genotyping [20, 35, 46, 52], had a limitation in classifying Vietnam BLVs G6 and Myanmar G10 in both Maximum Likelihood (ML) and Bayesian Inference (BI) methods (Fig. 2.4, Fig. 2.5). Therefore, the complete BLV *env gp51*sequence are recommend to use for phylogenetic analysis. Moreover, the genetic analysis of the complete genomes

of Vietnamese and foreign BLVs indicated that each genotype encodes specific amino acid substitutions in both structural and non-structural gene regions (Fig. 3.6). In addition, the mean diversity analysis of total 32 worldwide BLV complete genome sequences presented the significant conservation of BLV LTRs gene (Table. 3.1), suggesting the possibility of using this target region to design the primers and probe that may improve sensitivity and efficiency of BLV real-time PCR detection assays.

Recently, the phylogenetic inference from molecular sequences is a popular tool to trace the patterns of pathogen dispersal. To date, phylogeographic analysis becomes a common approach to visualize the spatial and temporal information in molecular epidemiology of pathogens, by connecting historical processes in evolution with spatial distributions [31]. To infer biogeography from genetic data for many organisms, a phylogeographic analysis framework was developed to draw conclusions about the epidemic origin and epidemiological linkage between locations [34]. Therefore, to understand the temporal history of spatial molecular epidemiology of BLV infections, I modeled this concept of the phylogeographic framework using the complete BLV *env gp51* sequences represented for 25 worldwide BLVs that were available on GenBank. Ignoring the possibility of missing middle transmission locations, the result of the phylogeographic analysis in this study illustrated the correlation of BLV infection among 25 worldwide countries. Moreover, it provided an overview to gain better understanding the temporal history of spatial molecular epidemiology of worldwide BLV infections that has not been conducted before (Fig. 3.2, Fig. 3.3, Fig. 3.4).

Finally, although BLVs is present worldwide, BLV genotyping studies are limited to many countries. Therefore, the accumulation of the complete BLV genome sequences is desirable. This study introduced and evaluated a new comprehensive method based on the SureSelect target enrichment system for deep sequencing the complete BLV genome form DNA provirus sample. This method is usefulness, efficacy, faster turn-around times and easy to implement.

General summary

BLV is an oncogenic virus belong to the genus *Deltaretrovirus* in the *Retroviridae* family [29, 60]. BLV is prevalent and infects cattle worldwide, imposing a severe economic impact especially on the dairy industry. Phylogenetic analysis using the complete BLV *gp51 env* gene sequence in current BLV studies demonstrated that BLV can be classified into ten genotypes (G1- G10). In the recent years, BLV infection has been rapidly increased in Japan and it is a notifiable disease which has been subjected to passive surveillance since 1998 [41]. I was involved in a BLV surveillance study in dairy cattle in eastern Hokkaido, Japan. In contrast, BLV infection has not been paid much attention in Vietnam, and the understanding of the disease is very limited. So far, no study on the BLV infection has been conducted in Vietnam. Therefore, I conducted an epidemiological study of BLV infections in eastern Hokkaido, Japan, and northern Vietnam to understand a current status of BLV infection among cattle.

The study in Chapter 1 evaluated an absolute quantitative in-house *pol*-rt-PCR, and monitored the current status of BLV infections among cattle in eastern Hokkaido from 2016－2017. All cattle in two farms (named K and U) were screened for BLV infections by serological and molecular detection methods. The antibody positive rates determined by ELISA were higher than the BLV gene detection rates by the real-time PCR in both farms. Overall, the seroprevalence of BLV in K farm was as high as 46.44%, while as low as 17.65% in U farm. The results of the in-house *pol*-rt-PCR mostly accorded with the commercial absolute quantitative *tax*-rt-PCR suggested that the in-house *pol*-rt-PCR assay can be applied to the detection of BLV genes in the infected cattle. Phylogenetic analysis presented that BLV G1 is circulating among infected cattle. Moreover, 11 BLVs in K farm were divided into 2 subgroups, suggesting that there were two sources of BLV infection that were distributed to cattle population in K farm. This situation might be a result of changes in transmission modes or other

influencing factors, such as cattle import, or movement. Furthermore, according to the correlation between calves and their dams in two farms, BLV might be vertically and horizontally transmissions in both farms; however, definitive evidence was not obtained. Therefore, continuing surveillance and identifying the key risk factor of BLV transmission to control a new infection should be given a priority.

The study in Chapter 2 provided the information on BLV infection in Vietnam for the first time. The results presented that BLV infections were detected among cattle in 8 out of the 22 tested farms in 4 districts in northern Vietnam in 2017. The highest prevalence of BLV infections (58.54%) was found in the BaVi district, where is one of the main livestock districts and holds 65% of the total cattle population in Hanoi. Overall, the BLV infected prevalence among cattle in northern Vietnam was 35.48 % as determined by both ELISA and the in-house *pol*-rt-PCR. One remarkable finding is that the commercial *tax*-rt-PCR failed to detect Vietnamese BLVs genotype 6 (BLV G6VN), similar to the previous report on Myanmar BLV genotype 10 (G10). Phylogenetic analysis indicated that at least two genotypes groups, G1 and G6, were circulated among cattle population in northern Vietnam. Among them, a new sub-genotype 6, named G6f, was found. In addition, the phylogenetic analysis showed that Vietnamese BLVs could be clustered with Thai BLVs in both genotype groups (G1 and G6). It was suggested that Vietnamese BLV may have the same origin as Thai BLVs as the consequence of cattle trading between two countries. Considering the management of Vietnamese farming system, cattle replacement from different herds of farms, crossbreeding between native cattle with imported cattle, and extended calving interval may be the factors that caused the genetic diversification of this virus in this country. Moreover, the phylogenetic analysis reported that to differentiate BLVs G6 and G10, the complete *env gp51* (903 bp) sequences are recommend to use instead of the partial *env gp51* sequences (444 bp), which have been normally used to genotype BLVs.

The study in Chapter 3 evaluated the SureSelect target enrichment method, which is an efficient and useful support method for next-generation sequencing (NGS) of the whole BLV genome

from DNA provirus samples. Four Vietnamese complete BLV genomes were successfully sequenced, which is the first report in Vietnam. Analyzing the genetic diversity of BLV *pol* and *tax* genes of Vietnamese BLVs revealed that BLV G6VN *tax* gene is more diverse than its *pol* gene. In addition, thoroughly comparing the BLV G6VN *tax* gene sequences with the BLV reference *tax* gene sequences, total of 7 specific nucleotide mutations were exclusively found in the BLV *tax* gene sequences of the BLV G6 strains including the BLV G6VN and G10 strains. It is assumed that any of the 7 nucleotide mutations in BLV *tax* gene possibly abrogate the gene detection by the commercial *tax*-rt-PCR. The results obtained may contribute to improving the reliability of such detection systems. The phylogeographic analysis illustrated that Vietnamese BLVs were clearly distinguished into two groups. One group had the same origin as Japanese and Korean strains, while another group had the same origin as the strains isolated in Brazil, Argentina, Paraguay, Thailand, and China. Moreover, visualizing the result on Google Earth software demonstrated that BLV was transmitted to Thailand from the USA in 1905, and 45 years later Thai BLVs began spreading to Vietnam cattle. On the other hand, BVL was imported to Japan from the USA in 1982, then it started spreading to Vietnamese cattle in 2004.

In conclusion, this study provides the informative data of current BLV infection status among cattle population in eastern Hokkaido, Japan, and northern Vietnam. The results supported that BLV is more divergent than previously thought. Evidently, a new sub-genotype G6f was found in Vietnam in this study. In addition, the study showed the SureSelect target enrichment method is an efficient and useful method for NGS to sequence the whole BLV genome from DNA provirus samples. The results can be obtained in less than one week but satisfied the requirement for high read depth values and a genome coverage of 100% across all sequenced samples. The phylogeographic analysis in this study gave an overview to better understanding the temporal history of spatial molecular epidemiology of BLV infections not only in Vietnam but also in worldwide countries. The genetic

diversity analysis of the BLV *tax* genes may contribute to improving the reliability of the commercial *tax*-rt-PCR. The findings obtained herein would be useful for further epidemiological studies on BLV infection not only in Japan and Vietnam but also in other countries that require better cattle husbandry practices.

Acknowledgments

The research studies were partially supported by JSPS KAKENHI Grant Number 15H05260 from the Japan Society for the Promotion of Science, Japan.

I am deeply indebted to Prof. IMAI Kunitoshi, for giving me an opportunity to be a student of the doctoral program in Animal and Food Hygiene in Obihiro University of Agriculture and Veterinary Medicine. Your guidance supports, and comments were unmeasurable.

I owe my deepest to Prof. OGAWA Haruko, for her valuable suggestions, heartwarming encouragements, tireless supports and continuing guidance.

I would like to express my special appreciation and thank Prof. IGARASHI Ikuo for giving me valuable comments and suggestions during my research.

It is great pleasure to thank Prof. XUAN Xuenan and Prof. YOKOYAMA Naoaki for their beneficial comments and special reviews of my thesis.

I am very grateful to Prof. MIZUTANI Tetsuya, Associate Prof. OMATSU Tsutomu, Ms. KATAYAMA Yukie in Research and Education Center for Prevention of Global Infectious Disease of Animal, Tokyo University of Agriculture and Technology, for giving me an opportunity to implement my internship in their laboratory. Without their support, it would not be possible to conduct this research deeply.

I am obliged to Ms. MATSUDA Sachiko and students of the Diagnostic Center for Animal Health and Food Safety for their help and support.

My sincere thanks to Dr. Bui Nghia Vuong and all staff member of National Institute of Veterinary Research, Vietnam for their support and consultant.

I would like to thank Graduate School of Animal Husbandry, Obihiro University of Agriculture and Veterinary Medicine to Provide me the opportunity to study here.

I am very appreciating Onishi san, your generosity and helps make me and my family had a cozy life when we lived in Japan.

Finally, I warmly thank and appreciate my mother, my wife and my children for their material and spiritual support in all aspects of my life.

References

1. 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
2. Adam E, Kerkhofs P, Mammerickx M, Kettmann R, Burny A, Droogmans L, Willems L (1994) Involvement of the cyclic AMP-responsive element binding protein in bovine leukemia virus expression in vivo. J Virol 68:5845-5853
3. Aida Y, Murakami H, Takahashi M, Takeshima SN (2013) Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia virus. Front Microbiol 4:328
4. 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
5. 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
6. Bartlett PC, Norby B, Byrem TM, Parmelee A, Ledergerber JT, Erskine RJ (2013) Bovine leukemia virus and cow longevity in Michigan dairy herds. J Dairy Sci 96:1591-1597
7. Bech-Nielsen S, Piper CE, Ferrer JF (1978) Natural mode of transmission of the bovine leukemia virus: role of bloodsucking insects. Am J Vet Res 39:1089-1092
8. Bodi K, Perera AG, Adams PS, Bintzler D, Dewar K, Grove DS, Kieleczawa J, Lyons RH, Neubert TA, Noll AC, Singh S, Steen R, Zianni M (2013) Comparison of commercially available target enrichment methods for next-generation sequencing. J Biomol Tech 24:73-86
9. Brown JR, Roy S, Ruis C, Yara Romero E, Shah D, Williams R, Breuer J (2016) Norovirus whole-genome sequencing by sureSelect target enrichment: a robust and sensitive method. J Clin Microbiol 54:2530-2537
10. Bruck C, Mathot S, Portetelle D, Berte C, Franssen JD, Herion P, Burny A (1982) Monoclonal antibodies define eight independent antigenic regions on the bovine leukemia virus (BLV) envelope glycoprotein gp51. Virology 122:342-352
11. Burny A, Bruck C, Cleuter Y, Couez D, Deschamps J, Gregoire D, Ghysdael J, Kettmann R, Mammerickx M, Marbaix G, et al. (1985) Bovine leukaemia virus and enzootic bovine leukosis. Onderstepoort J Vet Res 52:133-144
12. 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

13. Coulston J, Naif H, Brandon R, Kumar S, Khan S, Daniel RC, Lavin MF (1990) Molecular cloning and sequencing of an Australian isolate of proviral bovine leukaemia virus DNA: comparison with other isolates. J Gen Virol 71 (Pt 8):1737-1746
14. Da Y, Shanks RD, Stewart JA, Lewin HA (1993) Milk and fat yields decline in bovine leukemia virus-infected Holstein cattle with persistent lymphocytosis. Proc Natl Acad Sci U S A 90:6538-6541
15. Depledge DP, Palser AL, Watson SJ, Lai IY, Gray ER, Grant P, Kanda RK, Leproust E, Kellam P, Breuer J (2011) Specific capture and whole-genome sequencing of viruses from clinical samples. PLoS ONE 6:e27805
16. Derse D, Casey JW (1986) Two elements in the bovine leukemia virus long terminal repeat that regulate gene expression. Science 231:1437-1440
17. Deshayes L, Levy D, Parodi AL, Levy JP (1977) Proteins of bovine leukemia virus. I. Characterization and reactions with natural antibodies. J Virol 21:1056-1060
18. Erskine RJ, Bartlett PC, Byrem TM, Render CL, Febvay C, Houseman JT (2012) Association between bovine leukemia virus, production, and population age in Michigan dairy herds. J Dairy Sci 95:727-734
19. Felber BK, Derse D, Athanassopoulos A, Campbell M, Pavlakis GN (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
20. 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
21. Ferrer JF, Marshak RR, Abt DA, Kenyon SJ (1978) Persistent lymphocytosis in cattle: its cause, nature and relation to lymphosarcoma. Ann Rech Vet 9:851-857
22. 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
23. Gillet N, Florins A, Boxus M, Burteau C, Nigro A, Vandermeers F, Balon H, Bouzar AB, Defoiche J, Burny A, Reichert M, Kettmann R, Willems L (2007) Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in human. Retrovirology 4:18
24. Heenemann K, Lapp S, Teifke JP, Fichtner D, Mettenleiter TC, Vahlenkamp TW (2012) Development of a Bovine leukemia virus polymerase gene-based real-time polymerase chain reaction and comparison with an envelope gene-based assay. J Vet Diagn Invest 24:649-655

25. Jimba M, Takeshima S, 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
26. Johnston ER, 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
27. Johnston ER, Albritton LM, Radke K (2002) Envelope proteins containing single amino acid substitutions support a structural model of the receptor-binding domain of bovine leukemia virus surface protein. J Virol 76:10861-10872
28. Jukes TH, Cantor CR (1969) CHAPTER 24 - Evolution of Protein Molecules A2 - MUNRO, H.N. Mammalian Protein Metabolism. Academic Press, pp 21-132
29. Kettmann R, Portetelle D, Mammerickx M, Cleuter Y, Dekegel D, Galoux M, Ghysdael J, Burny A, Chantrenne H (1976) Bovine leukemia virus: an exogenous RNA oncogenic virus. Proc Natl Acad Sci U S A 73:1014-1018
30. Khudhair YI, Hasso SA, Yaseen NY, Al-Shammari AM (2016) Serological and molecular detection of bovine leukemia virus in cattle in Iraq. Emerg Microbes Infect 5:e56
31. Knowles LL, Maddison WP (2002) Statistical phylogeography. Mol Ecol 11:2623-2635
32. Lee E, Kim EJ, Ratthanophart J, Vitoonpong R, Kim BH, Cho IS, Song JY, Lee KK, Shin YK (2016) Molecular epidemiological and serological studies of bovine leukemia virus (BLV) infection in Thailand cattle. Infect Genet Evol 41:245-254
33. Lefebvre L, Vanderplasschen A, Ciminale V, Heremans H, Dangoisse O, Jauniaux JC, Toussaint JF, Zelnik V, Burny A, Kettmann R, Willems L (2002) Oncoviral bovine leukemia virus G4 and human T-cell leukemia virus type 1 p13(II) accessory proteins interact with farnesyl pyrophosphate synthetase. J Virol 76:1400-1414
34. Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian Phylogeography Finds Its Roots. PLoS ONE 5:e1000520
35. Licursi M, Inoshima Y, Wu D, Yokoyama T, Gonzalez ET, Sentsui H (2002) Genetic heterogeneity among bovine leukemia virus genotypes and its relation to humoral responses in hosts. Virus Res 86:101-110
36. Mammerickx M, Portetelle D, de Clercq K, Burny A (1987) Experimental transmission of enzootic bovine leukosis to cattle, sheep and goats: infectious doses of blood and incubation period of the disease. Leuk Res 11:353-358
37. Mamoun RZ, Astier T, Guillemain B, Duplan JF (1983) Bovine lymphosarcoma: expression of BLV-related proteins in cultured cells. J Gen Virol 64 (Pt 9):1895-1905
38. Mamoun RZ, Morisson M, Rebeyrotte N, Busetta B, Couez D, Kettmann R, Hospital M, Guillemain B (1990) Sequence variability of bovine leukemia virus env gene and its relevance to the structure and antigenicity of the glycoproteins. J Virol 64:4180-4188

39. 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
40. Monti G, Schrijver R, Beier D (2005) Genetic diversity and spread of Bovine leukaemia virus isolates in Argentine dairy cattle. Arch Virol 150:443-458
41. 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
42. Murakami K, Kobayashi S, Konishi M, Kameyama K, Tsutsui T (2013) Nationwide survey of bovine leukemia virus infection among dairy and beef breeding cattle in Japan from 2009-2011. J Vet Med Sci 75:1123-1126
43. Nekouei O, VanLeeuwen J, Stryhn H, Kelton D, Keefe G (2016) Lifetime effects of infection with bovine leukemia virus on longevity and milk production of dairy cows. Prev Vet Med 133:1-9
44. Nuotio L, Rusanen H, Sihvonen L, Neuvonen E (2003) Eradication of enzootic bovine leukosis from Finland. Prev Vet Med 59:43-49
45. Oba M, Tsuchiaka S, Omatsu T, Katayama Y, Otomaru K, Hirata T, Aoki H, Murata Y, Makino S, Nagai M, Mizutani T (2018) A new comprehensive method for detection of livestock-related pathogenic viruses using a target enrichment system. Biochem Biophys Res Commun 495:1871-1877
46. 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
47. Ogawa Y, Sagata N, Tsuzuku-Kawamura J, Koyama H, Onuma M, Izawa H, Ikawa Y (1987) Structure of a defective provirus of bovine leukemia virus. Microbiol Immunol 31:1009-1015
48. Ohshima K, 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. Nihon Juigaku Zasshi 43:79-81
49. Olson C, Kettmann R, Burny A, Kaja R (1981) Goat lymphosarcoma from bovine leukemia virus. J Natl Cancer Inst 67:671-675
50. Onuma M, Wada M, Yasutomi Y, Yamamoto M, Okada HM, Kawakami Y (1990) Suppression of immunological responses in rabbits experimentally infected with bovine leukemia virus. Vet Microbiol 25:131-141
51. Pluta A, Rola-Luszczak M, Kubis P, Balov S, Moskalik R, Choudhury B, Kuzmak J (2017) Molecular characterization of bovine leukemia virus from Moldovan dairy cattle. Arch Virol 162:1563-1576

52. Polat M, Ohno A, Takeshima S-n, Kim J, Kikuya M, Matsumoto Y, Mingala CN, Onuma M, Aida Y (2015) Detection and molecular characterization of bovine leukemia virus in Philippine cattle. Arch Virol 160:285-296
53. Polat M, Takeshima SN, Hosomichi K, Kim J, Miyasaka T, Yamada K, Arainga M, Murakami T, Matsumoto Y, de la Barra Diaz V, Panei CJ, Gonzalez ET, 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
54. Polat M, Moe HH, Shimogiri T, Moe KK, Takeshima SN, Aida Y (2017) The molecular epidemiological study of bovine leukemia virus infection in Myanmar cattle. Arch Virol 162:425-437
55. Polat M, Takeshima S, Aida Y (2017) Epidemiology and genetic diversity of bovine leukemia virus. Virol J 14
56. Pollari FL, Wangsuphachart VL, DiGiacomo RF, Evermann JF (1992) Effects of bovine leukemia virus infection on production and reproduction in dairy cattle. Can J Vet Res 56:289-295
57. 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
58. Rhodes JK, Pelzer KD, Johnson YJ (2003) Economic implications of bovine leukemia virus infection in mid-Atlantic dairy herds. J Am Vet Med Assoc 223:346-352
59. Rodriguez SM, Golemba MD, Campos RH, Trono K, Jones LR (2009) Bovine leukemia virus can be classified into seven genotypes: evidence for the existence of two novel clades. J Gen Virol 90:2788-2797
60. Rodriguez SM, Florins A, Gillet N, de Brogniez A, Sanchez-Alcaraz MT, 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
61. 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
62. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539-542

63. Sagata N, Tsuzuku-Kawamura J, Nagayoshi-Aida M, Shimizu F, Imagawa K, Ikawa Y (1985) Identification and some biochemical properties of the major XBL gene product of bovine leukemia virus. Proc Natl Acad Sci U S A 82:7879-7883
64. 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
65. Sakakibara N, Kabeya H, Ohashi K, Sugimoto C, Onuma M (1998) Epitope mapping of bovine leukemia virus transactivator protein Tax. J Vet Med Sci 60:599-605
66. Schwartz I, Levy D (1994) Pathobiology of bovine leukemia virus. Vet Res 25:521-536
67. Somura Y, Sugiyama E, Fujikawa H, Murakami K (2014) Comparison of the copy numbers of bovine leukemia virus in the lymph nodes of cattle with enzootic bovine leukosis and cattle with latent infection. Arch Virol 159:2693-2697
68. 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
69. 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
70. Twizere JC, Kruys V, Lefebvre L, Vanderplasschen A, Collete D, Debacq C, Lai WS, Jauniaux JC, Bernstein LR, Semmes OJ, Burny A, Blackshear PJ, Kettmann R, Willems L (2003) Interaction of retroviral Tax oncoproteins with tristetraprolin and regulation of tumor necrosis factor-alpha expression. J Natl Cancer Inst 95:1846-1859
71. Walrand F, Fumoux F, Roelants G, Parodi AL, Levy D (1986) Incidence of bovine leukemia virus-specific antibodies in West African cattle. Int J Cancer 37:619-621
72. Wang M, Wang Y, Baloch AR, Pan Y, Xu F, Tian L, Zeng Q (2018) Molecular epidemiology and characterization of bovine leukemia virus in domestic yaks (Bos grunniens) on the Qinghai-Tibet Plateau, China. Arch Virol 163:659-670
73. 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
74. Yang Y, Kelly PJ, Bai J, Zhang R, Wang C (2016) First Molecular Characterization of Bovine Leukemia Virus Infections in the Caribbean. PLoS ONE 11:e0168379
75. Zhao X, Buehring GC (2007) Natural genetic variations in bovine leukemia virus envelope gene: possible effects of selection and escape. Virology 366:150-165

学位論文要旨

牛白血病ウイルス (BLV) は、レトロウイルス科デルタウイルス属に分類される腫瘍ウイルスである。BLV は地方病性牛白血病の原因ウイルスであり、宿主の B 細胞に持続感染を引き起こす。外因性の BLV が宿主の B 細胞に感染した場合には、ウイルス RNA の逆転写により産生された DNA が宿主 B 細胞のゲノムに組み込まれ、プロウイルスとして宿主 B 細胞ゲノムと共存するウイルス感染様式をとる。

BLV は世界的に多くの牛群に感染し、重度の経済的損失、取り分け酪農産業に大きな影響を与えている。持続性 B リンパ球増多症を示す牛の乳生産量は牛群の平均的な牛乳生産量に比べて 3〜10%低い値を示す。その結果、感染牛は早期の淘汰の対象であり、国際的な貿易において取引制限を受ける。

BLV gp51 *env* 遺伝子の塩基配列を用いた系統樹解析の結果から、BLV は 10 の遺伝子型 (G1– G10) に分類されている。ヨーロッパのいくつかの国では、国家的コントロールプログラムによって BLV が牛群から排除されている。一方、日本を含む多くの国では、現在でも BLV 感染を排除することに成功していない。2009 年から 2011 年にかけて行われた乳牛および肉牛群における BLV 感染の全国調査により、日本では急激に BLV 感染が増加していることが明らかになったが、乳牛および肉牛群における BLV 抗体陽性率は、それぞれ 40.9 %と 28.7%であった。これらの乳牛および肉牛群における抗体陽性率は、1980 年から 1982 年にかけて行われた全国調査より、それぞれ 10 倍および 4 倍高かった。

学位申請者は、北海道東部の搾乳牛群における BLV サーベイランスに参加する機会を得た。一方、学位申請者の母国であるベトナムでは、BLV 感染は注目されてこなかったために本疾病に関する理解が非常に限られており、これまでに BLV 研究は行われてこなかった。従って本論文においては、北海道東部とベトナム国ハノイ地域の牛群における BLV 感染の現在の状況を明らかにする目的で BLV 感染に関する疫学的研究を実施した。

第 1 章では、BLV の逆転写酵素とインテグラーゼをコードする *pol* 遺伝子を標的とする in-house pol-rt-PCR の有用性を評価すると共に、2016 年から 2017 にかけて北海道東部の牛群における BLV 感染状況について調査した結果が記載されている。調査対象の K 農場と U 農場の全頭の牛について、血清学的および遺伝学的検査法により BLV 感染状況を調べた。

両農場の ELISA による抗体陽性率は、real-time PCR の遺伝子検出率より高かった。K 農場の抗体陽性率は 46.4%であったが、U 農場の抗体陽性率は 17.6%であった。in-house pol-rt-PCR の結果は、BLV の転写活性化遺伝子である *tax* 遺伝子を標的とする市販の tax-rt-PCR キットの結果とほぼ一致した。この結果は、in-house pol-rt-PCR は、感染牛における BLV 遺伝子の検出に応用可能であることを示している。系統樹解析の結果から、遺伝子型 1 (G1) の BLV が牛群の中で感染伝播していることが示唆された。さらに、K 農場で検出された BLV 遺伝子の解析結果から、11 株の BLV は 2 つのサブグループに分類された。この結果から、K 農場に導入されて感染を引き起こしている BLV には、2 つの異なる感染源があることが示され、新たな伝播経路、牛の導入や移動、あるいは交配といった他の要因によってもた

らされた可能性が考えられた。さらに、両農場の子牛とその母牛の間で水平感染と垂直感染が起こった可能性が示唆されたが、明確な証拠は得られなかった。以上の結果は、新たな BLV 感染を防ぐためには継続的なサーベイランスと BLV 伝播に関与する主要なリスクファクターの同定が最優先されるべきであることを示している。

第 2 章では、ベトナムにおける最初の BLV 感染状況に関する情報が記載されている。2017 年にハノイ地域の 4 地区において調べられた 22 の農場のうち 8 農場において BLV 感染が認められた。最も高い BLV 感染率 (58.54%) が、BaVi 地区において認められた。本地区は、ハノイ地域における主要な畜産地域の 1 つであり、ハノイ地域の飼育頭数の 65%を占めている。ELISA と BLV *pol* 遺伝子を検出する in-house pol-rt-PCR の検査結果から、ハノイ地域全体の BLV 陽性率は 35.48 %であった。予想に反して市販の tax-rt-PCR キットは、遺伝子型 6 (G6VN) の遺伝子を検出できなかった。本キットはミャンマーの牛群に感染を引き起こしている遺伝子型 10 (G10) を検出できなかったと報告されている。

系統樹解析の結果から、ハノイ地域の牛群では少なくとも 2 つの遺伝子型 （G1, G6）が存在していることが示された。さらに G6 の新しい亜遺伝子型である G6f の存在が初めて明らかにされた。加えて、系統樹解析の結果から、ベトナムの BLV はタイの BLV とともに、G1 と G6 に分類された。このことから、ベトナムとタイの BLV は、同じ起源をもつ可能性が示唆された。ベトナムにおける飼育形態を考慮すると、農場内の異なる牛群間の牛の移動、土着の牛と輸入牛との交配、長期間にわたる牛の飼育がこの国の BLV の遺伝的多様性

を引き起こした要因であるかもしれない。さらに、これまで BLV の遺伝子型の決定に用いられてきた 444 bp の *env* gp51 塩基配列に代わりに、完全長の *env* gp51 (903 bp) 遺伝子の塩基配列を用いる方がより正確に遺伝子型を決定することができることが示唆された。

第 3 章では、BLV プロウイルスサンプルから完全長の BLV ゲノムの次世代シークエンシング (NGS) をする際の効率的で有用な NGS 用サンプル調整法でとして知られる SureSelect Target Enrichment 法の有用性を評価した。通常、完全長の BLV ゲノムの塩基配列を決定するには、サンガー法あるいは従来法の NGS が用いられてきたが、BLV 感染牛から抽出されるゲノム DNA サンプル中に多量の宿主細胞由来のゲノムが含まれていることから、これらの方法におけるシークエンシングの過程において、高感度の PCR 法の使用と多くの時間を必要とする問題点があった。

SureSelect Target Enrichment 法は、標的分子 (DNA) と相補的なビオチン化 RNA 結合ストレプトアビジン粒子を使って、選択的に標的 DNA を捕捉するために開発された。捕捉された標的 DNA は、PCR 法により特異的に増幅後に NGS を用いて塩基配列が決定される。

本法と NGS を用いて、ベトナムの BLV 4 株の完全長のゲノムの塩基配列の決定に初めて成功した。BLV *pol* 遺伝子と *tax* 遺伝子の多様性解析から、BLV G6VN の *tax* 遺伝子が *pol* 遺伝子よりもより多様性を示すことがわかった。BLV G6VN *tax* 遺伝子の塩基配列を BLV 参照株の *tax* 遺伝子の塩基配列と比較したところ、計 7 箇所の変異が、BLV G6VN を含む G6 株および G10 株の *tax* 遺伝子の配列にのみ検出された。恐らくこれら 7 カ所の塩基配列の変

異のいずれかが市販の tax-rt-PCR キットの *tax* 遺伝子検出に影響を与えたものと推察された。得られた結果は、遺伝子検出システムの信頼性を高めることに役立つ情報であると考えられた。

生物系統地理学的解析から、ベトナムの BLV は明らかに 2 つのグループに分類された。一つのグループは、日本と韓国由来の BLV 株と同一の由来であったが、もう一つのグループは、ブラジル、アルゼンチン、パラグアイ、タイおよび中国の株と同じ由来を持っていた。さらに、Google Earth software を用いて、本研究で得られた結果を可視化して解析したところ、1905 年に BLV がアメリカ合衆国からタイへ伝播し、45 年後にタイの BLV がベトナムの牛群に拡散していったことが推察された。一方、1982 年にアメリカ合衆国から日本に導入された BVL は、その後 2004 年にベトナムの牛群に拡がり始めたと考えられた。

以上、本研究は、北海道東部とベトナムのハノイ地域の牛群における BLV 感染の現状に関する有益な情報をもたらした。本研究により、BLV にはこれまでに考えられていた以上に多様性があることが判明した。実際、新しい亜遺伝子型（G6f）がベトナムの牛群に存在することが明らかになった。さらに、BLV プロウイルスサンプルから完全長の BLV ゲノム配列の決定に使用される NGS 用のサンプル調整法として、SureSelect Target Enrichment 法が効率的で有用な方法であることが示された。本法の使用により、全塩基配列の決定が１週間以内に可能であり、しかも正確に塩基配列を読むことができた。生物系統地理学的

解析は、時空間的にどのような歴史的過程を辿り、ベトナムばかりではなく世界的な BLV 感染状況がどのように成立してきたかを包括的に理解するのに役立った。

BLV *tax* 遺伝子の遺伝学的多様性に関する解析は、市販の tax-rt-PCR キットの信頼性向上に貢献するものと期待される。

本研究において得られた知見は、日本やベトナムばかりではなく、牛群の管理システムの改善が求められている他の国々における BLV 感染に関する疫学的研究の遂行に役立つものと思われる。