論 文 の 内 容 の 要 旨

論文題目　**Study on molecular genetic analysis of bovine leukemia virus in South America and Asia**

（南米とアジアにおける牛白血病ウイルスの分子遺伝学的解析）

氏　　名　　**买热帕提　帕拉提**

Background

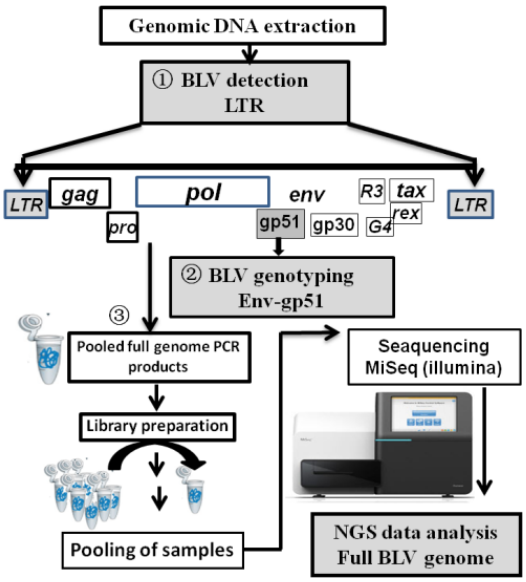
Bovine leukemia virus (BLV) is an oncogenic member of *retroviridae* family belonging to the genes *deltaretrovirus*, and is the etiological agent of enzootic bovine leukosis (EBL), the most common neoplastic disease of cattle. BLV infects cattle worldwide, some of BLV-infected cattle suffer from lymphomas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

BLV complete genome is constituted of 8714 nucleotides, including structural and enzymatic *gag*, *pro*, *pol*, and *env* genes which is indispensible in the synthesis of viral particles (Fig. 1). BLV genome also contains a pX region which encodes regulatory proteins Tax and Rex, and accessory proteins R3 and G4. The structural genes and pX region are surrounded by two identical long terminal repeats (LTR). Among the structural genes, *env* gene encodes ENV polyprotein precursor (gPr72) which further cleaved into Env gp51 and gp30 glycoprotein. The Env gp51 glycoprotein plays an essential role in the viral life cycle, and is required for cell entry and the target of neutralizing antibodies. The N-terminal half of BLV gp51 contains conformational epitopes -F,-G and -H, and plays an important role in viral infectivity and syncytium formation, while the C-terminal half of BLV gp51 contains the linear epitopes A, B, D, and E. Therefore, the gp51 region has been widely used for BLV genotyping studies and recent phylogenetic studies of this region from viral strains isolated worldwide demonstrate that BLV can be classified into at least eight genotypes.

Objective:

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

Fig. 1



Materials and Methods:

Genomic DNA was extracted from blood samples obtained from a total of 3386 cattle which contained 2204 samples in five South American countries (328 from Peru, 139 from Paraguay, 507 from Bolivia, 420 from Argentina, and 810 from Chile) and 1116 cattle from the Philippines and the rest 66 cattle from Myanmar. BLV infection was detected by amplification of BLV long terminal repeats (LTRs) using BLV-CoCoMo-qPCR-2 and/ or nested PCR (Fig. 1 ①). BLV positive samples were used for amplification of BLV-*env* gp51 by nested PCR (Fig. 1 ②) and sequenced for phylogenetic analyses. Phylogenetic trees were constructed either by using the neighbor-joining (NJ) method with the Tamura-Nei model of nucleotide substitution for the Philippine BLV strains or Bayesian Inference (BI) and/ or maximum-likelihood (ML) method with the Tamura-Nei model of nucleotide substitution for South

American BLV strains or Kimura-2 parameter model with gamma distribution (K2+G) for Myanmar BLV strains. Full genome sequences of new genotypes were obtained by NGS-based whole genome sequencing for G-9 (Fig. 1 ③) or clone-sequencing for G-10, respectively.

Results and discussion

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

To investigate the spread of BLV infection, 3386 samples were screened for BLV infection by amplifying BLV LTR region. Among the 2204 cattle tested in the South America, all South American countries showed relatively high level of infection (Table 1): 139 cattle out of 328 (42.3 %) were positive for the BLV provirus in Peru, prevalence rates were up to 58.6 % at the farm level. In Paraguay, 76 cattle samples out of 139 (54.5 %) were BLV positive. Of 507 samples collected from Bolivia, 156 (30.7 %) were positive for BLV provirus and the highest BLV prevalence reached to 100 % at the individual level. Argentine samples (n = 420) demonstrated extremely high levels of BLV prevalence (77.4 %), with up to 90.9 % at the herd level. In Chile, 810 samples were collected from relatively wide geographical area. Of the samples screened, 236 (29.1 % prevalence) were BLV provirus positive. In contrast to the BLV prevalence level in South America, Asia showed remarkable low level of BLV infection. In the Philippines, a total of 9.7 % of tested samples (108/1116) were determined as BLV positive, while only 5 out of 66 cattle from Myanmar showed the prevalence of BLV with an infection level of 7.6 %.

Table 1. Summary of BLV detection

Territory	Country	tested sample	BLV-positive result	
			Number	%
Asia (1182)	Philippines	1116	108	9.7
	Myanmar	66	5	7.6
South America (2204)	Peru	328	139	42.3
	Paraguay	139	76	54.5
	Bolivia	507	156	30.7
	Argentina	420	325	77.4
	Chile	810	236	29.1

2. Phylogenetic analysis of partial *env* gp51 sequences

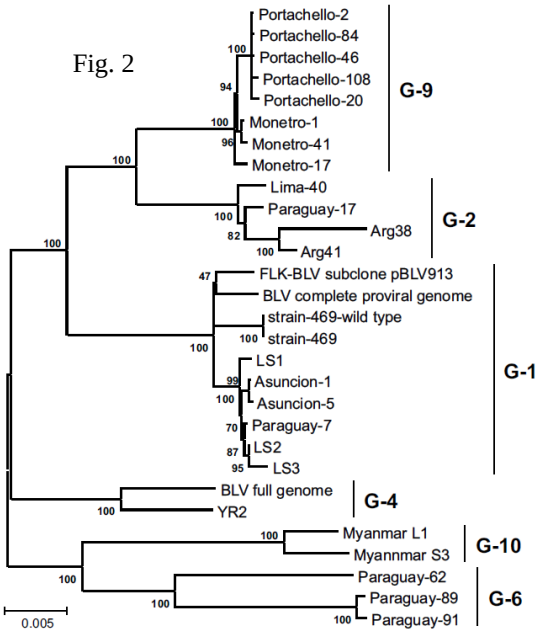
To gain insight into the degree of genetic variability of BLV strains in South America and Asia, especially Peru, Paraguay, Bolivia, the Philippines and Myanmar where no molecular genetic analysis have been done, *env* gp51 from BLV positive samples were amplified and used for phylogenetic analysis to determine the genotypes of BLV strains circulating in each country. Total of 32 philippine BLV strains were assigned to G-1, while the rest 11 strains were grouped into G-6. In South America, Peru BLV strains were assigned to G-1, -2, and -6. Likewise, the majority of Paraguayan BLV strains clustered into G-1 and -6, with a small number in G-2. Interestingly, the BLV strains collected from Bolivia clustered not only into G-1, -2, and -6, together with our Peruvian and Paraguayan strains, but also into a unique clade, which was distinct from the eight previously known BLV genotypes as novel G-9. Surprisingly, Myanmar strains were not clustered within any known genotypes, but separately located in a different branch indicating that Myanmar strains might also be a new genotype, termed as G-10.

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

To get better insight into the amino acid changes and variability, deduced amino acid sequence of each genotype BLV strains circulating in South America, in the Philippines and Myanmar were aligned. Comparison of predicted amino acid sequences of Philippine BLV strains revealed highly conservation. Philippine strains assigned to G-6 showed G-6-specific common amino acid substitution of isoleucine by threonine at residue 144 of second neutralizing domain (2nd ND). In addition, seven single amino acid substitutions were observed. Substitutions Y108C at $CD4^{+}$ epitope, A119P at 2nd ND and L202F were detected in G-1 Philippine BLV strains, while amino acid mutations K175E at $CD8^{+}$epitope, T231S and S234N in the B-epitope were found in G-6 Philippine BLV strains. South American BLV strains were strictly conserved. G-2 and -6 South American BLV strains all shared G-2 and -6 specific amino acid substitution of N141D and I144T at 2nd ND, which are common among all G-2 and -6 BLV strains worldwide. Surprisingly, all South American new G-9 BLV strains showed common amino acid substitution of alanine by valine at residue 133 (A133V), which is firstly detected in our study. Myanmar BLV strains showed a variety of unique amino acid substitutions, which were not detected in other strains. One G-6-specific amino acid substitution from isoleucine to threonine at residue 144 (I144T) which is common to all G-6 BLV strains worldwide is also detected in 2nd ND of

Myanmar BLV strains. Besides that, four unique amino acid mutations were observed only in Myanmar strains as follows: amino acid substitution of valine replaced by alanine at residue 106 (V104A) which is the common amino acid residue of both 1st ND and CD4+-epitope region. Two amino acid substitutions of serine to phenylalaine at residue 137 (S137F) and of glutamine to arginine at residue 143 (Q143R) were located at 2nd ND. The last common amino acid substitutions from proline to serine at residue 177 (P177S) were located in CD8+- and E-epitope.

Fig. 2



4. BLV complete genome sequencing and phylogenetic analysis

To get full insight into comprehensive phylogenetic characteristic of complete genome sequence, and to confirm the new G-9 and -10, complete BLV genome sequences of 16 South American strains including eight of G-9 strains, and two Myanmar G-10 strains were obtained, then full ML tree was constructed. The full genome ML tree clearly demonstrates stratification of BLV genotypes, including G-1, -2, -4, and -6, and the novel G-9, into separate clades (Bootstrap values 100 % for every clade). Thus, this analysis provides evidence for the existence of a novel G-9 as shown in Fig. 2. Two Myanmar strains were located in a different branch separated from other clusters (Bootstrap values 100 % for every clade), indicating that Myanmar strains is also a new genotype, G-10 (Fig. 2). The ML tree further confirm the existence of new G-9 and -10 firstly in our study.

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

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

Genotype	Sample	Gag	Pro	Pol	Env
		3333	111	111222233344444445666666677888888	1111111112244447 8 0 5
		3466891466	145893441	949460036579023488604888909012333	1247803334447562778 0 0
		8813980805 6	392114891	0757659180099377016244475123460 9	2984263471346747309 0 0 4
G-1	BLV-FLK	THGTRGAVPAT	IVSIVVVR	APASEVIAQTTIIDTDPVSRHKVVVNVDALH	QRAKSVADSNQIFPSRNQLTHV
	LS1	.....MS..	.....I.	.S.....A.....PK.....	.....N.....
	LS2	.....MS..	.....I.	.S.....A.....PK.....	.....N....S.....
	LS3	.....MS..	.....I.	.S.....A.....PK.....	.....N....S.....
	Japan	.....MS..	.....	.F.....K.....	.....
	Asun1	.....MS..	.....I.	.S.....A.....PK.....	.....N....S.....
	Asun5	.....MS..	.....I.	.S.....A.....PK.....	.....N....S.....
	Para7	.....MS..	.....I.	.S.....A.....PK.....	.....N....S.....
G-2	Arg38	.YS...TI.T.	T..TIA.P	TST.....E...L.K.RA.I.I..P.	.Q..F...D....L..E.ART
	Arg41	.YS...TI.T.	T..TI..P	TST.....E...L.K.RA.I.I..P.	.Q..F...D....L..EFART
	Lima40	..S...TI.T.	T..TIA.P	TSTP.....E...L.K.RA.I.I..P.	.QV.F...D....L..EFART
	Para17	..N...TI.T.	T..TIA.P	TST.....E...L.K.RA.I.I..P.	.Q..F...D....L..EFART
G-4	USA	...AKE...T.	V.....	TS.....E...L.K...I...N.P.	.QTRF.....L...FP.T
G-6	Par62	.YVKE.T.T.	...T...S.	S.P..MTHA.V.EA..L.K...I...N.PQ	EQTRF.....T..LK.F..T
	Par89	.YVKE.T.T.	.....IS.	S.P..MTHA.V.EA..L.K..AI...N.PQ	EQTRF.....T..LKD.F..T
	Par91	.YVKE.T.T.	...T...IS.	S.P..MTHA.V.EA..L.K..AI...N.PQ	EQTRF.....T..LKD.F..T
G-10	Myan-L1	AYSVKET...A	.VF.I...	.....L...A..VEA.SL.K...I...NVP.	.QTRFA..F.RT.SL...F..T
	Myan-S3	AYSVKET...A	..F.....	.....L...A..VE..SL.K...I...NVP.	.QTRFA..F.RT.SL...F..T
G-9	Mon1	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y.....TP.	.Q..F.V.....L.DEFART
	Mon17	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y.....TP.	.Q..F.V.....L.DEFART
	Mon41	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y.....TP.	.Q..F.V.....L.DEFART
	Port2	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y....S...TP.	.Q..L.V.....L.DEFART
	Port46	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y....S...TP.	.Q..L.V.....L.DEFART
	Port84	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y....S...TP.	.Q..L.V.....L.DEFART
	Port20	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y....S...TP.	.Q..L.V.....L.DEFART
	Port108	..S...TI.T.	T..TIA..	TS..D.....E.G.L..Y....S...TP.	.Q..L.V.....L.DEFART

Genotype	Sample	Tax	Rex	R3	G4
		11111111222	11111111		2234567
		2456600134578	1345557801224445	1222334	2267556
		4046490800022419379	2903682033560266	2267560	7267550
G-1	BLV-FLK	VQRNCTPFRINVISDFLRI	QMSSNKTASASGLTLT	QVNKHILF	CFPHFFL R
	LS1	.....	.....	.....	.....
	LS2	.....	.....	.....	.....
	LS3	.....	.....	L	.....
	Japan	.....	.....	.....	.....
	Asun1	.....	.....	.....	.....
	Asun5	.....	.....	.....	.....
	Para7	.....	.....	.....	.....
G-2	Arg38	...TFA...V...P....L	.T....PS.....S.	.I...V.L	...RSSL..
	Arg41	...TFA...V...P....L	.T....PS.....S.	.I...V.L	...RSSL..
	Lima40	...TFA...V...P....L	.T....PS.....S.	.I...V.L	.S.RSSL..
	Para17	...TFA...V...P....L	.T....PS.....S.	.I...V.L	.S.RSSL..
G-4	USA	...TF...Q....P...GL	...N..PL...S..S.	.I.....L	.L..S....
G-6	Par62	DRHTFA..Q.....	E.TNDRPL..NS.MS.	EI.EY.SL	RSL.S....
	Par89	D.HTF...Q....P....L	E.TN.RPS..NS.MS.	EI.EY.SL	R.L.S.S..
	Par91	D.HTF...Q....P....L	E.TN.RPS..NS.MS.	EI.EY.SL	R.L.S.S..
G-10	Myan-L1	...TF..LQ.KET.NDF.L	...N.RPSF..SP..N	.I.N..SL	.S..S.I P R
	Myan-S3	...TF..LQ.kET.NDF.L	...N..PSF.NSP..N	.I.N..SL	.S..S.. P Q
G-9	Mon1	.RTF..L.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Mon17	.RTF..L.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Mon41	.RTF..L.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Port2	.RTF.SL.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Port46	.RTF.SL.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Port84	.RTF.SL.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Port20	.RTF.SL.....P...GL	.T..D.PS.E....S.	.IH....L	....S.L..
	Port108	.RTF.SL.....P...GL	.T..D.PS.E....S.	.IH....L	....S L..

Fig. 3

Conclusion

The present study provides the first evidence of the prevalence of BLV infection among cattle in Paraguay, Peru and Bolivia and Myanmar, and also confirms the widespread distribution of BLV infection in Argentina, Chile and in the Philippines. All of five South American countries showed higher level of BLV prevalence, and also indicated tendency of dramatic increase of BLV infection in South America. By contrast, low to medium level of BLV prevalence were observed in the Philippines and Myanmar in Asia. Our phylogenetic analysis demonstrated that two different BLV G-1 and -6 are present in the Philippines. BLV genotyping analysis of South America showed that Peruvian and Paraguayan BLV strains are of G-1, -2, and -6. Of interest, our findings indicate that Bolivian BLV strains are clustered into G-1, -2, -6, and a new G-9. In addition, Myanmar BLV strains were classified into another novel G-10. Amino acid comparison studies showed that ten unique amino acid substitutions in South American G-9 and 22 unique amino acid substitutions in Myanmar G-10 BLV strains were observed. These unique amino acid substitutions clearly distinguish novel G-9 and -10 from previously known genotypes, further indicating the existence of new genotypes. Thus, the detection of G-9 and -10 is a novel finding of this study. These results provide important information on BLV infection levels and will enable the implementation of appropriate cattle-management policies in addition to providing supplementary information for the development of more-effective methods of BLV eradication in South America and Asia.