Single nucleotide polymorphisms in the bovine MHC region of Japanese Black cattle are associated with bovine leukemia virus proviral load

Shin-nosuke Takeshima¹, Shinji Sasaki², Polat Meripet¹, Yoshikazu Sugimoto² and Yoko Aida¹*

Abstract

Bovine leukemia virus (BLV) is the causative agent of enzootic bovine leukosis, a malignant B cell lymphoma that has spread worldwide and causes serious problems for the cattle industry. The BLV proviral load, which represents the BLV genome integrated into host genome, is a useful index for estimating disease progression and transmission risk. Here, we conducted a genome-wide association study to identify single nucleotide polymorphisms (SNPs) associated with BLV proviral load in Japanese Black cattle. The study examined 93 cattle with a high proviral load and 266 with a low proviral load. Three SNPs showed a significant association with proviral load. One SNP was detected in the CNTN3 gene on chromosome 22, and two (which were not in linkage disequilibrium) were detected in the bovine major histocompatibility complex region on chromosome 23. These results suggest that polymorphisms in the major histocompatibility complex region affect proviral load. This is the first report to detect SNPs associated with BLV proviral load in Japanese Black cattle using whole genome association study, and understanding host factors may provide important clues for controlling the spread of BLV in Japanese Black cattle.

Keywords: Bovine leukemia virus, Whole genome association study, Major histocompatibility complex

Main text

Bovine leukemia virus (BLV), which infects cattle worldwide [1–6], belongs to the family Retroviridae (genus Deltaretrovirus), together with human T cell leukemia virus types 1 and 2 (HTLV-1 and -2) [7]. Historically, the economic losses caused by BLV infection were thought to be related only to the onset of bovine leukosis, which occurs in only 1–5% of BLV-infected cows within 5 years post-infection [7]. However, recent reports show that BLV infection also reduces milk production [6, 8–11] and causes a high incidence of infectious disease [12] and reproductive inefficiency, resulting in high culling rates [13]; thus BLV eradication is of utmost importance.

Previous studies show that the proviral load is an important index for estimating the stage of BLV infection because it is associated with disease progression [14–16], lymphocyte count [17], viral biokinetics [18], and virus shedding into saliva and nasal secretions [19]. Indeed, one study shows that cattle with a low proviral load are not a source of BLV transmission [20]. Therefore, determining host factors associated with an increased proviral load is important if we are to develop eradication programs for BLV.

Studies of BLV-associated host factors identified polymorphisms within the bovine major histocompatibility complex (MHC) (BoLA) [21–29]. Recently, Miyasaka et al. revealed that polymorphisms within BoLA class II haplotypes were strongly associated with BLV proviral load in Japanese Black cattle, the main breed of beef cattle in Japan, but less so with that in European breeds [22]. However, no group has undertaken a genome-wide association study (GWAS) to identify such host factors.
Therefore, to identify proviral load-associated polymorphisms, we performed a GWAS using DNA samples from 676 Japanese Black cattle [30]. Genomic DNA was isolated from peripheral blood, and the BLV proviral load was measured using the BLV-CoCoMo-qPCR-2 method [31]. BLV provirus was detected in samples from 444 animals (range, 1 copy/10^5 cells to 132,230 copies/10^5 cells; median value, 5498 copies/10^5 cells) (Fig. 1a). We then compared the proviral load in animals used for the GWAS with that in Japanese Black cattle selected randomly from whole areas of Japan. We found no significant difference in the proviral load between animals used for GWAS and the randomly selected group (Fig. 1b). In most cases, the animals in both groups showed a proviral load of <10,000 copies/10^5 cells. A proviral load >100,000 copies/10^5 cells was rare.

We categorized the 444 BLV-infected cows into four groups according to proviral load: Low (0 < provirus load ≤ 13,819, 266 heads), Medium (14,237 < provirus load ≤ 40,698, 85 heads), High (42,605 < provirus load ≤ 73,145, 60 heads), and Very High (76,397 < provirus load ≤ 132,230, 33 heads). We then performed a GWAS using these traits as a binary variable, as is done in 93 case (High + Very High group)−266 control (Low group) studies. The 359 animals were genotyped using a SNP50 K BeadChip comprising probes targeting 54,001 single nucleotide polymorphisms (SNPs). In all, 32,919 autosomal SNPs met the quality control criteria (call rate >99%; minor allele frequency >0.01; Hardy–Weinberg equilibrium, \( p > 0.001 \)). Analyses were then performed using GEMMA software [32], which uses a linear-mixed model approach based on a genetic-relationship matrix estimated from SNP genotypes to model correlations between the phenotypes of sample subjects. The genomic-inflation factor (\( \lambda_{GC} \)) for this analysis was 1.021, indicating that a sample was appropriate for inclusion in an association study. The quantile–quantile (Q–Q) plot showed that three SNPs showed a significant deviation from the null hypothesis (Fig. 2b, Bonferroni-corrected threshold for genome-wide significance (\( p < 1.5 \times 10^{-6} \) add threshold line in A). Three significant genome-wide associations were detected: rs29026690 (\( p = 1.91 \times 10^{-7} \), odds ratio = 2.745) and rs17872126 (\( p = 1.91 \times 10^{-7} \), odds ratio = 0.414) on bovine chromosome 23 (BTA23) and rs110616206 (\( p = 5.37 \times 10^{-7} \), odds ratio = 6.589) on BTA22 (Fig. 2b; Table 1). The two SNPs on BTA23 were found within an 800 Kb window located at 27,421,348–28,223,274 bp; these two SNPs did not show
Odds ratio is the effective value for estimating how strongly the SNPs associated to the proviral load, using following formula:

\[ \text{Odds ratio} = \frac{ \frac{d}{e} \times \frac{f}{c} }{ \frac{b}{a} } \]

- **a** SNP ID assigned by Illumina, Inc
- **b** Reference SNP (refSNP) ID assigned in the single nucleotide polymorphism database (dbSNP)
- **c** Positions are based on the bovine genome, assembled in UMD3.1
- **d** Minor allele is minor frequency allele determined in this study
- **e** Major allele is major frequency allele determined in this study
- **f** Odds ratio is the effective value for estimating how strongly the SNPs associated to the proviral load, using following formula

| Chromosome | Ilumina_ID | Reference cluster ID | Position | p   | Minor allele | Minor allele | Minor allele | Major allele | Odds ratio |
|------------|------------|----------------------|----------|-----|--------------|--------------|--------------|--------------|-----------|
| 23         | Hapmap57616-rs29026690 | rs29026690 | 27421348 | 1.91 × 10⁻⁷ | A | 0.2903 | 0.1297 | G | 2.745 |
| 23         | ARS-BFGL-NGS-113235 | rs17872126 | 28223274 | 1.91 × 10⁻⁷ | G | 0.2849 | 0.4906 | A | 0.414 |
| 22         | Hapmap33580-BTA-136506 | rs110616206 | 27280154 | 5.37 × 10⁻⁷ | A | 0.0914 | 0.01504 | G | 6.589 |

*Table 1* **SNPs showing a significant association with BLV proviral load**

linkage disequilibrium (LD) \((r^2 = 0.117)\), indicating that BTA23 harbored two independent quantitative trait loci (QTL)s (Figs. 2b, 3; Table 1).

Genes within or near these regions were then analyzed using the UMD3.1 genome assembly tool. Hapmap57616-rs29026690 (27,421,348 bp on BTA23) was located between ENSBTAG00000000580 and ABHD16A (ENSBTAG00000000578) (Additional file 1: Table S1), whereas ARS-BFGL-NGS-113235 (28,223,274 bp on BTA23) was located between the 4th and 5th exons of PRR3 (ENSBTAG000000005914). These SNPs reside within the BoLA class III and class I regions, respectively (Figs. 2b, 3; Additional file 1: Table S1). Therefore, the gene density was much higher than that in other areas of the genome, and a number of candidate genes that could be used to estimate proviral load were present around the detected SNPs [33]. Hapmap33580-BTA-136506 was located on the centromeric side of BTA22, at a distance of 6.5 kb from the CONTACTIN3 (CNTN3) gene (Table 1; Additional file 1: Table S1, Additional file 2: Fig. S1).

To the best of our knowledge, this is the first report to detect SNPs associated with BLV proviral load in Japanese Black cattle using GWAS. Two of the identified SNPs were located in the BoLA region. We found it interesting that these two SNPs were located within the class III and class I regions because a previous study reported involvement of only class II genes [22]. The genome reference sequences for the BoLA region have many gaps, mainly because class I genes were difficult to genotype, making associations with class I genes difficult to determine. Target resequencing of high density SNPs across the MHC region using a next generation sequencer should be undertaken to confirm which genes are truly responsible for regulating the proviral load. Our result showed that the MHC polymorphism is important factor for proviral load. The reason why MHC polymorphisms were associated with proviral load is the polymorphism of classical MHC directly associate with antigen presentation and the difference of antigen presentation in each allele leads to the immunological difference in each host.

Taken together, the results described herein show that MHC genotyping of class III and class I alleles can identify cows with a low proviral load. In the farm with high infection rate, eliminating high proviral load cow is an effective way for eradicating BLV because proviral load is major risk factor for transmitting BLV to other host [20]. Therefore, farmer should frequently check the proviral load because the proviral load is variable, although it is not cost-effective. Taken together with the information of our finding 3 SNPs and our previously report about resistant BoLA class II allele [22], we can identify...
Fig. 3 Regional Manhattan plot of the locus on chromosome 23 that harbors SNPs associated with BLV proviral load. The imputed SNPs are shown by arrows. Genes (Chr23:27,116,737 to Chr23:28,311,070) are listed, and the positions of SNPs associated with the BLV proviral load are indicated by arrows. The horizontal blue lines represent the Bonferroni-corrected thresholds for genome-wide significance ($-\log_{10}(p) = 5.82$). The indicated positions are based on the bovine genome (assembled in UMD3.1).
the BLV resistant cow. It will be helpful to develop a low cost method of eradicating BLV from farms because we can reduce the frequency of measurement of proviral load.

**Additional files**

**Additional file 1: Table S1.** Genes around rs110616206 SNP (Tab “CNTNS”) and rs29026690 and rs17872126 SNPs (Tab “BTA23”).

**Additional file 2: Figure S1.** Regional Manhattan plot showing the association of 33,006 SNPs (BovineSNP BeadChip) with BLV proviral load in 359 Japanese Black cattle. Regional plot of the locus on chromosome 22 that harbors SNPs associated with BLV proviral load. The imputed SNPs are indicated by arrows. The horizontal blue lines represent the Bonferroni-corrected thresholds for genome-wide significance (−log10(5) = 5.82). The indicated positions are based on the bovine genome (assembled in UMD3.1).

**Authors’ contributions**

Study conception and design: YA. Data acquisition, analysis, and interpretation: SS, MP, and ST. Contribution of reagents/materials/analysis tools: YA and YS. Drafting and revising the manuscript: ST, YA, and SS. Approval of the final version: ST, SS, MP, YS, and YA. All authors read and approved the final manuscript.

**Author details**

1. Viral Infectious Diseases Unit, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. 2. Shirakawa Institute of Animal Genetics, Japan Livestock Technology Association, Odakura, Nishigo, Fukushima 961-8061, Japan.

**Acknowledgements**

We thank the Support Unit at the Biomaterial Analysis, RIKEN BSI Research Resources Center, for help with sequence analysis.

**Competing interests**

The authors declare that they have no competing interests.

**Availability of data and materials**

The datasets analyzed herein are available from the corresponding author upon request.

**Ethics approval and consent to participate**

All protocols for the collection of blood samples were reviewed and approved by the Shirakawa Institute of Animal Genetics Committee on Animal Research (H21-2).

**Funding**

The study was supported by Grants-in-Aid for Scientific Research (A and C) from the Japan Society for the Promotion of Science (JSPS) (Grant Nos. 18255013, 25450405), by a grant from the Integration Research for Agriculture and Interdisciplinary Fields in Japan (Grant No. 14538311). This research was also supported by grants from the Project of the NARO Bio-oriented Technology Research Advancement Institution (Grant No. 16B17983) (the special scheme project on regional developing strategy).

**Received:** 10 January 2017  **Accepted:** 21 March 2017 **Published online:** 04 April 2017

**References**

1. Erskine RJ, Bartlett PC, Byrem TM, Render CL, Fębaw C, Houseman JT. Using a herd profile to determine age-specific prevalence of bovine leukemia virus in Michigan dairy herds. Vet Med Int. 2012;2012:350374.

2. Polat M, Ohno A, Takeshima SN, Kim J, Kikuya M, Matsumoto Y, Mingala CN, Ohmura M, Aida Y. Detection and molecular characterization of bovine leukemia virus in Philippine cattle. Arch Virol. 2015;160:285–96.

3. Polat M, Takeshima SN, Hosomioki K, Kim J, Miyasaka T, Yamada K, Arainga M, Murakami T, Matsumoto Y, de la Barra Diaz V, et al. A new genotype of bovine leukemia virus in South America identified by NGS-based whole genome sequencing and molecular evolutionary genetic analysis. Retrovirology. 2016;13:4.

4. Polat M, Moe HH, Shimogiri T, Moe KK, Takeshima SN, Aida Y. The molecular epidemiological study of bovine leukemia virus infection in Malaysian cattle. Arch Virol. 2017;162:425–37.

5. Zhang R, Jiang J, Sun W, Zhang J, Huang K, Gu X, Yang Y, Xu X, Shi Y, Wang C. Lack of association between bovine leukemia virus and breast cancer in Chinese patients. Breast Cancer Res. 2016;18:101.

6. Yang Y, Fan W, Mao Y, Yang Z, Liu G, Zhang R, Zhang H, Saito C, Wang C. Bovine leukemia virus infection in cattle of China: association with reduced milk production and increased somatic cell score. J Dairy Sci. 2016;99:3688–97.

7. Aida Y, Murakami H, Takeshima M, Takeshima SN. Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia. Front Microbiol. 2013;4:528.

8. Emanuelsen U, Schering L, Pettersson H. Relationships between herd bovine leukemia-virus infection status and reproduction, disease incidence, and productivity in Swedish dairy herds. Prev Vet Med. 1992;12:121–31.

9. Da Y, Shahs RD, Stewart JA, Lewin HA. Milk and fat yields decline in bovine leukemia virus-infected Holstein cattle with persistent lymphocytosis. Proc Natl Acad Sci USA. 1993;90:6538–41.

10. Norby B, Bartlett PC, Byrem TM, Erskine RJ. Effect of infection with bovine leukemia virus on milk production in Michigan dairy cows. J Dairy Sci. 2016;99:2043–52.

11. Nekouei O, VanLeeuwen J, Stryhn H, Kelton D, Keefe G. Lifetime effects of infection with bovine leukemia virus on longevity and milk production of dairy cows. Prev Vet Med. 2016;133:1–9.

12. Sandev N, Koleva M, Binew R, Jleva D. Influence of enzootic bovine leukosis virus upon the incidence of subclinical mastitis in cows at a different stage of infection. Vet Arh. 2004;74:411–6.

13. Bartlett PC, Norby B, Byrem TM, Parmeelee A, Ledergerber JT, Erskine RJ. Bovine leukemia virus and cow longevity in Michigan dairy herds. J Dairy Sci. 2013;96:1591–7.

14. Jimba M, Takeshima SN, Matoba K, Endoh D, Aida Y. BLV-CoCoMo-qPCR quantitation of bovine leukemia virus proviral load using the CoCoMo algorithm. Retrovirology. 2010;7:91.

15. Jimba M, Takeshima SN, Murakami H, Kohara J, Kobayashi N, Matsumo Y, Ohmori T, Unno N, Aida Y. BLV-CoCoMo-qPCR: a useful tool for evaluating bovine leukemia virus infection status. BMC Vet Res. 2012;8:167.

16. Somura Y, Sugiyama E, Fujikawa H, Murakami K. 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. 2014;159:2693–7.

17. Takahashi II, Takeshima SN, Aida Y. Detection of bovine leukemia virus proviral load in infected cattle in Japan from 2012 to 2014. Virus Res. 2015;210:283–90.

18. Panei CJ, Takeshima SN, Omori T, Nunoa T, Davis WC, Ishizaki H, Matoba K, Aida Y. Estimation of bovine leukemia virus (BLV) proviral load harbored by lymphocyte subpopulations in BLV-infected cattle at the subclinical stage of enzootic bovine leucosis using BLV-CoCoMo-qPCR. BMC Vet Res. 2013;9:96.

19. Yuan Y, Kita Murumatsu Y, Saito S, Ishizaki H, Nakano M, Haga S, Matoba K, Ohno A, Murakami H, Takeshima SN, Aida Y. Detection of the BLV provirus from nasal secretion and saliva samples using BLV-CoCoMo-qPCR: comparison with blood samples from the same cattle. Virus Res. 2015;210:248–54.

20. Jularema MA, Barrios CN, Ceriani MC, Esteban EN. Hot topic: bovine leukemia virus infection with low proviral load are not a source of infection for BLV-free cattle. J Dairy Sci. 2016;99:4526–9.

21. Nikbakht Brujeni G, Ghobanpour R, Esmailnejad A. Association of BoLA-II haplotypes associated with variations in bovine leukemia virus proviral load in Iranian Holstein cattle. Biochem Genet. 2016;54:194–207.

22. Miyasaka T, Takeshima SN, Jimba M, Matsumoto Y, Kobayashi N, Matsumo Y, de la Barra Diaz V, et al. Retrovirology. 2017;14:24.
23. Juliarena MA, Poli M, Sala L, Ceriani C, Gutierrez S, Dolcini G, Rodriguez EM, Marino B, Rodriguez-Dubra C, Esteban EN. Association of BLV infection profiles with alleles of the BoLA-DRB3.2 gene. Anim Genet. 2008;39:432–8.

24. Aida Y, Nishino Y, Amanuma H, Murakami K, Okada K, Ikawa Y. The role of tumor-associated antigen in bovine leukemia virus-induced lymphosarcoma. Leukemia. 1997;11(Suppl 3):216–8.

25. Zanotti M, Poli G, Ponti W, Polli M, Rocchi M, Bolzani E, Longeri M, Russo S, Lewin HA, van Eijk MJ. Association of BoLA class II haplotypes with subclinical progression of bovine leukemia virus infection in Holstein–Friesian cattle. Anim Genet. 1996;27:337–41.

26. Xu A, van Eijk MJ, Park C, Lewin HA. Polymorphism in BoLA-DRB3 exon 2 correlates with resistance to persistent lymphocytosis caused by bovine leukemia virus. J Immunol. 1993;151:6977–85.

27. Stear MJ, Dimmock CK, Newman MJ, Nicholas FW. BoLA antigens are associated with increased frequency of persistent lymphocytosis in bovine leukemia virus infected cattle and with increased incidence of antibodies to bovine leukemia virus. Anim Genet. 1988;19:151–8.

28. Lewin HA, Wu MC, Stewart JA, Nolan TJ. Association between BoLA and subclinical bovine leukemia virus infection in a herd of Holstein–Friesian cows. Immunogenetics. 1988;27:338–44.

29. Lewin HA, Bernoco D. Evidence for BoLA-linked resistance and susceptibility to subclinical progression of bovine leukaemia virus infection. Anim Genet. 1986;17:197–207.

30. Sasaki S, Ibi T, Watanabe T, Matsuhashi T, Ikeda S, Sugimoto Y. Variants in the 3’ UTR of general transcription factor IIIF, polypeptide 2 affect female calving efficiency in Japanese Black cattle. BMC Genet. 2013;14:41.

31. Takeshima SN, Kitamura-Muramatsu Y, Yuan Y, Polat M, Saito S, Aida Y. BLV-CoCoMo-qPCR-2: improvements to the BLV-CoCoMo-qPCR assay for bovine leukemia virus by reducing primer degeneracy and constructing an optimal standard curve. Arch Virol. 2015;160:1325–32.

32. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44:821–U136.

33. Consortium TMs. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. Nature. 1999;401:921–3.