

1 **Diagnosis and early prediction of lymphoma using**
2 **high-throughput clonality analysis of bovine leukemia**
3 **virus-infected cells**

4

5 Tomohiro Okagawa^{1, #}, Honami Shimakura^{2, #}, Satoru Konnai^{1,2,*}, Masumichi Saito^{3,4},
6 Takahiro Matsudaira⁵, Naganori Nao^{6,7}, Shinji Yamada^{8,9}, Kenji Murakami^{8,9}, Naoya
7 Maekawa², Shiro Murata^{1,2}, Kazuhiko Ohashi^{1,2}

8

9 ¹Department of Advanced Pharmaceutics, Faculty of Veterinary Medicine, Hokkaido
10 University, Sapporo, 060-0818, Japan

11 ²Department of Disease Control, Faculty of Veterinary Medicine, Hokkaido University,
12 Sapporo, 060-0818, Japan

13 ³Department of Virology II, National Institute of Infectious Diseases, Tokyo, 162-8640,
14 Japan

15 ⁴Center for Emergency Preparedness and Response, National Institute of Infectious
16 Diseases, 162-8640 Tokyo, Japan

17 ⁵Biotechnological Research Support Division, FASMAC Co., Ltd., Atsugi, 243-0021,
18 Japan

19 ⁶Division of International Research Promotion, International Institute for Zoonosis
20 Control, Hokkaido University, Sapporo, 001-0020, Japan

21 ⁷One Health Research Center, Hokkaido University, Sapporo, 060-0818, Japan

22 ⁸Cooperative Department of Veterinary Medicine, Faculty of Agriculture, Iwate
23 University, Morioka, 020-8550, Japan

24 ⁹Graduate School of Veterinary Sciences, Iwate University, Morioka, 020-8550, Japan

25

26 Running title: BLV clonality analysis by RAISING-CLOVA

27 *Correspondence: Satoru Konnai, konnai@vetmed.hokudai.ac.jp

28 #These authors contributed equally to this work.

29

30 **Abstract**

31 Bovine leukemia virus (BLV), a retrovirus, infects into B cells of ruminants and causes
32 aggressive leukemia or lymphoma in cattle, enzootic bovine leukosis (EBL). Clonal
33 expansion of BLV-infected cells is a promising marker for early detection and diagnosis
34 of EBL. Recently, we developed rapid amplification of the integration site without
35 interference by genomic DNA contamination (RAISING) and CLOVA, a software to
36 analyze clonality. RAISING-CLOVA could assess the risk of adult T-cell
37 leukemia/lymphoma development in human T-cell leukemia virus-I-infected individuals
38 through its clonality analysis. Thus, we herein examined the performance of
39 RAISING-CLOVA for the clonality analysis of BLV-infected cells and conducted a
40 comprehensive clonality analysis by RAISING-CLOVA in EBL and non-EBL cattle.
41 RAISING-CLOVA successfully distinguished EBL from non-EBL cattle with high
42 sensitivity and specificity. A longitudinal clonality analysis in BLV-infected sheep, an
43 EBL model, also confirmed the effectiveness of BLV clonality analysis with
44 RAISING-CLOVA for early detection of EBL development. Therefore, our study
45 emphasizes the usefulness of RAISING-CLOVA as a routine clinical test for monitoring
46 virus-related cancers.

47

48 **Introduction**

49 Enzootic bovine leukemia (EBL) is a B-cell lymphoma in cattle caused by
50 infection with bovine leukemia virus (BLV). BLV is a member of the genus
51 *Deltaretrovirus*, subfamily *Orthoretrovirinae*, family *Retroviridae*, and is genetically
52 related to human T cell leukemia virus type 1 (HTLV-1) (1). BLV infects the B cells of
53 ruminants, and its viral RNA is reverse-transcribed into double-stranded DNA, and then
54 integrated into the host genome as a provirus. Most infected cattle are asymptomatic
55 (aleukemic; AL) throughout their lifespan. About 30% of infected cattle develop
56 persistent lymphocytosis (PL), an abnormal proliferation of BLV-infected B cells. EBL
57 occurs in 1–5% of BLV-infected cattle, which present with B-cell malignant lymphoma
58 in lymph nodes and various other organs, leading to a poor prognosis and death (2).

59 Although BLV has been eradicated in some European countries through the
60 detection and culling of BLV-infected cattle, it is still endemic worldwide (3). The
61 incidence of BLV infection has been increasing during recent decades in Japan; a
62 nationwide survey of cattle conducted from 2009–2011 confirmed BLV infection in
63 40.9% of dairy cattle and 28.7% of beef cattle (4). The countermeasures against BLV are
64 urgently needed, but there is no commercially available vaccine or therapeutic drug
65 against BLV. Therefore, the herd management by detection, quarantine, and culling of
66 infected cattle is the most effective method of controlling BLV infection. However, this
67 approach is difficult to implement in endemic areas with large numbers of infected cattle.
68 To reduce the economic damage caused by EBL development in such endemic areas, it is
69 effective to detect and cull cattle at high risk of developing EBL in advance.

70 During the development of EBL, one or few clones of BLV-infected cells
71 undergo clonal expansion. These malignant cells possess identical integration sites of
72 BLV provirus. Therefore, the clonality of proviral integration sites in BLV-infected cells
73 has been considered as a useful marker to diagnose the onset of EBL (5–7). Several
74 methods have been recently developed to analyze transgene integration sites using
75 high-throughput sequencing (HTS), such as ligation-mediated PCR (5, 8), target capture
76 sequencing (6, 9, 10), inverse PCR (11), Linear amplification-mediated PCR
77 (LAM-PCR) (12), and non-restrictive LAM-PCR (13). However, it would be difficult to
78 analyze multiple specimens inexpensively, sensitively, and rapidly by these current
79 methods. Therefore, there is a need to develop a high-throughput method that overcomes
80 the problems of current methods and that can be applied in clinical testing for BLV
81 clonality analysis.

82 Recently, we developed Rapid Amplification of Integration Sites without
83 interference by genomic DNA contamination (RAISING) and a clonality analysis
84 software (CLOVA), a highly sensitive, rapid, inexpensive, and high-throughput method
85 to amplify and analyze random integration sites of transgenes in host genomes (14).
86 RAISING and CLOVA were originally developed for the risk assessment of adult T cell
87 leukemia/lymphoma (ATL) development in HTLV-1 carriers (14, 15). The clonality and
88 proviral integration site of HTLV-1-infected cells can be examined by analyzing the
89 sequence of the amplicon from RAISING using the CLOVA software. The clonality
90 value (Cv) of infected cells calculated by RAISING-CLOVA is an effective marker for
91 the prediction of the risk of ATL onset (14).

92 BLV is considered to utilize similar mechanisms for proviral integration and
93 tumorigenesis as HTLV-1 (9, 16). In our previous study, the proviral integration sites of
94 BLV-infected cells in AL and EBL cattle was successfully amplified using RAISING
95 (14). Sanger sequencing and HTS confirmed the clonal expansion of BLV-infected cells
96 only in the EBL specimen, even though the number of tested specimens was quite limited
97 ($n = 2$) (14). The clonality analysis of BLV-infected cells by RAISING-CLOVA could be
98 an effective method for the diagnosis and prediction of the EBL onset in cattle. Hence, in
99 this study, we performed the comprehensive clonality analysis of BLV-infected cells
100 using RAISING-CLOVA in EBL and non-EBL cattle, and examined its performance in
101 the diagnosis and prediction of the onset of lymphoma in cattle and sheep.

102

103 **Materials and methods**

104 **Blood and tumor samples**

105 Peripheral blood of BLV-infected cattle (Holsteins, Blacks, or crossbreds) was collected
106 from dairy and beef farms in Japan between 2017 and 2022. Peripheral blood, lymph
107 nodes, spleen, and various organs were also collected from cattle diagnosed with
108 lymphoma (Holsteins, Blacks, or crossbreds) at Livestock Hygiene Service Centers and
109 Meat Hygiene Inspection Centers in Japan between 2013 and 2022. The blood and tumor
110 samples were kept refrigerated until cell separation. The animal experiments were
111 approved by the Ethics Committee of the Faculty of Veterinary Medicine, Hokkaido
112 University (approval #17-0024). Verbal informed consent was obtained from the owners
113 for the participation of their animals in this study.

114 Experimental infection of BLV in sheep was conducted at the Research Farm in the Field
115 Science Center, Faculty of Agriculture, Iwate University. Sheep (Corriedale or Suffolk,
116 three months old) were intraperitoneally inoculated with 3.0×10^7 cells of BLV-infected
117 leukocytes isolated from BLV-infected cattle. After the viral challenge, peripheral blood
118 was collected from the BLV-challenged sheep. The procedures were approved by the
119 Iwate University Animal Care and Use Committee (approval no. A201703).

120

121 **Cell isolation**

122 Peripheral blood mononuclear cells (PBMCs) were separated from blood samples by
123 density gradient centrifugation using Percoll (GE Healthcare, Chicago, IL, USA). Whole
124 blood was lysed with ACK lysing buffer to separate white blood cells. Separated cells
125 were then washed twice with phosphate buffered saline (PBS, pH 7.4) and filtered

126 through a 40 μ m cell strainer (BD Biosciences, San Jose, CA, USA). Tissue specimens
127 were shredded with scissors and filtered using a 40 μ m or 100 μ m cell strainer (BD
128 Biosciences) to obtain cell suspensions, and washed twice with PBS. Cells were stained
129 with Trypan Blue Stain (Thermo Fisher Scientific, Waltham, MA, USA) and the number
130 of viable cells was measured using a Countess II FL Automated Cell Counter (Thermo
131 Fisher Scientific). Cells were either used immediately or frozen at -80°C until used in
132 experiments.

133

134 **Cell lines**

135 Three bovine leukemic cell lines BLV-infected and uninfected cell lines were used in this
136 study: a BLV-infected B-cell line BL3.1 (17) and a BLV-uninfected T cell line BTL26
137 (18). All cell lines were cultured in RPMI 1640 medium (Sigma-Aldrich, St. Louis, MO,
138 USA) supplemented with 10% heat-inactivated fetal bovine serum (Thermo Fisher
139 Scientific), 100 IU/mL penicillin, 100 μ g/mL streptomycin, and 2 mM L-glutamine
140 (Thermo Fisher Scientific) at 37°C and 8% CO₂.

141

142 **Preparation of genomic DNA**

143 Genomic DNA from whole blood samples, PBMCs, and tissue specimens of cattle was
144 extracted using Wizard Genomic DNA Purification Kits (Promega, Madison, WI, USA)
145 or Quick-DNA Miniprep Kits (Zymo Research, Irvine, CA, USA). Genomic DNA from
146 whole blood samples of sheep was extracted using MagDEA Dx SV (Precision System
147 Science, Matsudo, Japan) with a magLEAD 12gC instrument (Precision System Science).
148 Genomic DNA was extracted from mixtures of KU-1 or BL3.1 and BTL26 for a total of 1

149 $\times 10^6$ cells using Quick-DNA Miniprep Kits (Zymo Research). The DNA concentrations
150 of the samples were measured by UV absorbance at 260 nm using a NanoDrop 8000
151 Spectrophotometer (Thermo Fisher Scientific). Agarose gel electrophoresis was
152 performed to check for the degradation of each DNA sample.

153

154 **Quantification of BLV proviral load (PVL)**

155 The BLV *pol* gene was measured in the genomic DNA samples of blood and tissue
156 samples of cattle using real-time PCR with a BLV Detection Kit (Takara Bio, Otsu,
157 Japan) with a LightCycler 480 System II (Roche Diagnostics, Mannheim, Germany). A
158 serial dilution series of the positive control DNA for each kit were used to generate
159 calibration curves to determine the copy number of the BLV provirus. Each DNA sample
160 was tested in duplicate. For the blood and tumor samples of sheep, the BLV *tax* gene was
161 measured in the DNA samples using real-time PCR, as described previously (19).

162

163 **Diagnosis of BLV infection and EBL**

164 BLV infection in cattle was diagnosed by confirmed by the detection of anti-BLV
165 antibodies using a commercial ELISA kit (JNC, Tokyo, Japan), and by the detection of
166 BLV provirus using real-time PCR. Plasma samples were screened for anti-BLV antibody.
167 Seropositive samples were further tested for the presence of BLV provirus to confirm
168 BLV infection, as described above. Samples that tested positive for the provirus were
169 diagnosed as “BLV-infected.” The number of lymphocytes in blood samples was counted
170 using an automated hematology analyzer (Celltac α ; Nihon Kohden, Tokyo, Japan).

171 BLV-infected cattle were classified as aleukemic (AL) or PL based on the lymphocyte
172 counts as follows: AL < 8,000 cells/ μ L; PL > 8,000 cells/ μ L.

173 Blood and tissue samples from cattle with lymphoma were diagnosed as B-cell
174 lymphoma based on immunophenotyping analysis by flow cytometry and/or B-cell
175 clonality analysis by PCR targeting bovine immunoglobulin heavy chain (20) combined
176 with the quantification of BLV provirus by real-time PCR, as described above.

177

178 **Amplification of integration sites of BLV provirus by RAISING**

179 The integration sites of BLV provirus were amplified by RAISING as previously
180 described (14), with some modifications. The primers and reagents used in each step are
181 shown in Supplemental Tables 1 and 2. The reaction conditions of each step are shown in
182 Supplemental Table 3. Single-stranded DNA (ssDNA) of the 3' LTR region of the BLV
183 provirus and the downstream region of the host genome was synthesized from the
184 extracted genomic DNA using the primer BLV-F1 and KOD-Plus-Neo DNA Polymerase
185 (Toyobo, Osaka, Japan). The synthesized ssDNA was purified using a Monarch PCR &
186 DNA Cleanup Kit (New England Biolabs, Ipswich, MA, USA) and was eluted in
187 ultrapure water. Then, poly(A) and poly(G) tails were added at the 3' end of the purified
188 ssDNA by terminal transferase (New England Biolabs). The double-stranded DNA was
189 then synthesized and amplified by PCR from the poly(AG)-tailed ssDNA using the
190 primers BLV-F2 and NV-oligo-dT-ADP1 and Q5 Hot Start High-Fidelity DNA
191 Polymerase (New England Biolabs). The second PCR was performed using diluted PCR
192 products, the primers BLV-F3 and ADP1-HTS-R1, and KOD-Plus-Neo DNA Polymerase
193 (Toyobo).

194

195 **Sequencing and clonality analysis using CLOVA**

196 The products of the second PCR were purified using AMPure XP (Beckman Coulter,
197 Fullerton, CA, USA) and analyzed using Sanger sequencing with a BigDye Terminator
198 v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) on a 3130Xl or 3730Xl DNA
199 Analyzer (Thermo Fisher Scientific). Clonality analysis of BLV-infected cells was
200 performed using the sequencing signal data of each sample by a CLOVA software (14),
201 an R program that automatically analyzes the clonality value (Cv) of transgene-integrated
202 cells by dividing the average of signal peak area values of 20 nucleotides at 5' terminal of
203 host genome sequence of the dominant clone by that at 3' terminal of BLV proviral
204 sequence.

205

206 **Statistical analysis**

207 Significant differences among multiple groups were identified using Kruskal-Wallis
208 one-way analysis of variance followed by Dunn's test. Association of two values were
209 tested using Receiver operating characteristics (ROC) curve analyses were performed to
210 determine the optimal cutoff values, where sensitivity approximates specificity for each
211 risk factor. All statistical tests were performed using GraphPad Prism 6 (GraphPad
212 Software, San Diego, CA, USA). Differences were considered statistically significant at
213 $P < 0.05$.

214 **Results**

215 **Performance of clonality analysis by RAISING-CLOVA targeting BLV**

216 To examine the detection limit of RAISING targeting BLV, RAISING was performed
217 using a series of DNA samples from mixtures of different percentages of a BLV-infected
218 B-cell line (BL3.1) and a BLV-uninfected strain (BTL26) and the amplified products
219 were analyzed by Sanger sequencing. BL3.1 harbors multiple copies of BLV provirus in
220 the genome (7) and PVL of BL3.1 was 240 copy/100 cells. Thus, DNA samples extracted
221 from specimens including 0.0001% to 100% BL3.1 contain different PVL ranging from
222 0.0024 to 240 copy/100 cells. The amplified fragments were observed in the
223 electrophoresis and the combined sequences of BLV provirus and host genome were
224 detected in the DNA samples containing 0.012–240 copy/100 cells of provirus (Fig. 1A
225 and Supplemental Fig. 1). The detection limit for RAISING targeting BLV was 0.012
226 copy/100 cells in PVL, which is comparable with that for RAISING targeting HTLV-1
227 (0.032% in HTLV-1 PVL) (14). These results suggest that RAISING targeting BLV is
228 sensitive enough to detect proviral insertion sites, even in specimens with low PVL.

229 Our previous study has shown that RAISING-CLOVA targeting HTLV-1 fails to
230 accurately measure Cv in samples with PVL less than 0.5% (14). Here, we examined the
231 sensitivity of clonality analysis by RAISING-CLOVA targeting BLV. The Cv of BL3.1
232 was 0.56 when measured using the specimen with PVL 240 copy/100 cells (Fig. 1B and
233 Supplemental Fig. 1). The values were comparable (Cv: 0.58) when measured using the
234 specimens with PVL 1.2 copy/100 cells, but it differed by more than 0.1 for specimens
235 with PVL lower than 0.24 copy/100 cells (Fig. 1B and Supplemental Fig. 1). This result
236 indicates that RAISING-CLOVA targeting BLV can accurately measure Cv when the

237 PVL is at least 1.2 copy/100 cells and is feasible enough for the analysis of most
238 BLV-infected specimens.

239 To further examine the performance and accuracy of the clonality analysis by
240 RAISING-CLOVA, we tested DNA samples obtained from BLV-infected cattle with or
241 without EBL. Representative results of RAISING-CLOVA in blood specimens of
242 BLV-infected cattle with different values of Cv are shown in Fig. 1C. The Cv reflected
243 the intensity of the sequence signal peaks of integration sites at the host genome (Fig. 1C).
244 These results are consistent with our previous results of RAISING-CLOVA using
245 specimens of HTLV-1-infected patients as well as a preliminary analysis using specimens
246 of BLV-infected cattle (14). Furthermore, the clonality analyses of the identical
247 BLV-infected specimens ($n = 32$) in the two different laboratories showed a high
248 interrater agreement in Cv (Fig. 1D). These results indicate that RAISING-CLOVA is
249 very accurate and reproducible method for measuring the clonality of BLV-infected cells.
250

251 **Comprehensive clonality analysis of BLV-infected cells in EBL and non-EBL cattle**

252 The clonality analysis by RAISING-CLOVA was performed on peripheral blood
253 samples from AL ($n = 107$), PL ($n = 79$), and EBL cattle ($n = 101$), and tissue samples
254 including tumors from EBL cattle ($n = 175$) collected from farms throughout Japan. In
255 addition, BLV PVL was also measured in these samples, because previous studies
256 proposed that PVL is a candidate marker for EBL diagnosis (21, 22). Representative
257 results of RAISING-CLOVA in EBL and non-EBL blood specimens are shown in Fig.
258 2A. In an EBL cattle (EBL015), Sanger sequencing analysis detected monoclonal
259 patterns of proviral integration in peripheral blood and tumor specimens and their values

260 of Cv were calculated as 1.00 (Fig. 2A, indicating clonal expansion of BLV-infected cells).
261 In contrast, in non-EBL cattle (AL002 and PL002), the sequencing analysis detected
262 polyclonal patterns of proviral integrations with Cv 0.09 and 0.06, respectively (Fig. 2A).
263 Among the blood samples, EBL showed significantly higher Cv than non-EBL specimens,
264 such as AL and PL (median: 0.63, 0.09, and 0.08, respectively) (Fig. 2B). In contrast,
265 there was no significant difference in the PVL of blood between PL and EBL (median:
266 35.10 and 25.78 copy/100 cells, respectively), although tumors of EBL showed higher
267 PVL (median: 59.06 copy/100 cells) than blood samples at all disease stages (Fig. 2C). In
268 addition, the Cv of tumors in EBL was higher (median: 0.81) than that of their blood
269 (median: 0.63) (Fig. 2B). Tumors of EBL were observed in a variety of tissues in cattle,
270 including multiple lymph nodes and non-lymphoid tissues such as heart, kidney, uterus,
271 digestive, and respiratory organs. There were no significant differences in the Cv and
272 PVL of tumors from lymphoid and non-lymphoid tissues in EBL (Fig. 2D and E).
273 Furthermore, identical integration sites of BLV provirus were detected by Sangar
274 sequencing among blood and tumor samples from same EBL cattle, even in the blood
275 specimens with low Cv (EBL049 and EBL174) (Supplemental Fig. 2). These results
276 indicate that the Cv of blood samples from RAISING-CLOVA is an effective marker for
277 distinguishing EBL from non-EBL cattle.

278

279 **Clinical utility of clonality analysis by RAISING-CLOVA targeting BLV**

280 To determine whether the Cv or PVL of BLV-infected cells was valuable for the
281 diagnosis of EBL, we performed ROC analysis and examined the specificity and
282 sensitivity of EBL diagnosis by Cv and PVL in the blood samples. The area under the

283 ROC curve (AUC) for Cv (0.9368) was higher than that for PVL (0.5759) for the blood
284 samples of EBL and non-EBL cattle (Fig. 4A and B). The cutoff value of Cv in blood for
285 EBL diagnosis was 0.17, which could distinguish EBL cases with 87.1% sensitivity and
286 93.0% specificity (Fig. 4C and D). Out of 101 blood specimens of EBL, the Cv less than
287 0.17 was detected in 13 specimens (12.9%) (Fig. 4D). These animals showed clonal
288 expansion of malignant cells in tumors, but not in blood. In addition, 13 cases (7.0%) of
289 AL and PL cattle were identified with the Cv > 0.17 in blood (Fig. 4D). Blood samples
290 may also be suitable for follow-up and prognostic studies of non-EBL cattle with
291 intermediate Cv. On the other hand, the cutoff value of PVL in blood was 32.10 copy/100
292 cells, which was not a suitable marker, because more than half of the samples had lower
293 values than the cutoff (Fig. 4D). Thus, these experiments indicate that Cv is a better
294 diagnostic marker for EBL compared to PVL.

295

296 **Longitudinal clonality analysis of experimental lymphoma model of sheep**

297 Sheep infected with BLV develop lymphoma at higher frequencies after shorter latency
298 periods than cattle (2). To further investigate the usefulness of Cv as a predictive marker
299 for the development of EBL, longitudinal clonality analysis was conducted in an
300 experimental infection model of sheep with BLV. In Cv remained low immediately after
301 infection, but increased before or at the onset of lymphoma (Fig. 5A). In three of the four
302 tested animals (#24, #93, and #114), Cv peaked earlier than PVL in blood (Fig. 5A).
303 Sequence analysis of the integration sites in blood and tumor tissues also revealed that the
304 identical integration sites were detected before and at the onset of lymphoma in all tested

305 sheep (Fig. 5B). These results indicate that clonality analysis by RAISING-CLOVA is a
306 promising method for early prediction of lymphoma onset in BLV infection.

307 **Discussion**

308 When a retrovirus infects a host, it may persist in the host throughout its lifespan, causing
309 severe disease in some carriers. In the cattle industry, BLV infection is widespread in all
310 parts of the world except Western Europe (3), and causes major economic losses in the
311 production of milk and beef (23, 24). However, there is no effective treatment or
312 vaccination to control BLV infection. In BLV-endemic countries such as Japan and the
313 United States, the control measure of eradicating BLV-infected cattle is completely
314 impractical. Therefore, the development of a novel method for early prediction of tumor
315 development in the carrier stages would contribute to the reduction of economic losses in
316 livestock production. In this study, we applied a novel molecular method called
317 RAISING with the clonality analysis software CLOVA (14) to amplify BLV proviral
318 integration site in host genome and analyze the clonality of BLV-infected cells. We
319 further examined the usefulness of the clonality analysis by RAISING-CLOVA as a
320 method for EBL diagnosis and the early prediction of lymphoma onset.

321 For BLV clonality analysis, several molecular methods have been developed to
322 analyze proviral integration sites of BLV-infected cells, including ligation-mediated PCR
323 (5), target capture sequencing (6, 9), and inverse PCR (7, 11, 25). Most of these methods
324 requires HTS analysis for the detection of integration sites, but the high cost of the
325 analysis makes it unsuitable for clinical diagnosis with large numbers of samples.
326 Additionally, these current methods raise concerns about the sensitivity and bias of
327 detection, which use restriction enzymes or ultrasound sonication for DNA
328 fragmentation.

329 RAISING is a highly sensitive, highly accurate, rapid, inexpensive, and
330 high-throughput method to overcome the problems of conventional methods (14). In this
331 study, we analyzed Cv of BLV-infected cells in EBL and non-EBL cattle using
332 RAISING-CLOVA, and found that Cv discriminated between non-malignant and
333 malignant samples successfully but BLV PVL did not. In previous studies, PL and
334 EBL-infected cows are known to have higher provirus levels than AL cows (21, 26).
335 Therefore, proviral levels are considered to be an important marker in EBL diagnosis
336 (Ohno et al., 2015; Kobayashi et al., 2019). However, PVL was not a suitable marker for
337 EBL diagnosis in this study. This inconsistency is presumably because a large number of
338 samples was examined in this study, including PL cows with high PVL. The clonality
339 analysis by RAISING-CLOVA should be conducted with another cohort of clinical
340 samples to confirm the reproducibility of the present analysis.

341 Among EBL cases analyzed in this study, 13 specimens (12.9%) did not show
342 tumorigenesis in peripheral blood, a finding which is consistent with our previous report
343 (Nishimori et al., 2017). In EBL cattle, lymphocytosis and/or the presence of atypical
344 lymphocytes is reportedly observed in the peripheral blood, but this does not appear to be
345 the case in all cases. However, more detailed EBL typing may be possible by examining
346 the clonality of tumor cells in the blood and tissues of a larger number of EBL cattle. In
347 addition, recent studies suggest that tumor cells of EBL sometimes harbor defective
348 proviruses (6, 11). EBL diagnosis using qPCR targeting provirus would miss those
349 defective cases. In this study, several tumor cells with proviruses deficient in the *pol* gene,
350 which was the target of qPCR in this study, were identified in EBL tumors (data not
351 shown). In BLV-specific forward primers targeting the 3'LTR and its upstream regions,

352 these regions seemed to be less likely to be deleted in the provirus of EBL tumors (6, 11)
353 and could be an optimal target of primers.

354 In BLV infection, antisense transcripts, called *AS*, were found to be constantly
355 expressed in EBL tumors through promoter activation of the 3' LTR region (27). A recent
356 study reported that activation of antisense transcription originating from the 3' LTR forms
357 a chimeric transcript of the *AS* gene and the host driver gene upstream of the provirus,
358 resulting in enhanced transcription of driver genes during BLV infection and following
359 tumorigenesis (9). Therefore, it is important to analyze the integration site of BLV
360 provirus in the bovine genome in order to investigate its contribution to tumorigenesis of
361 EBL. Further studies are warranted to address this issue by HTS analysis of the amplicon
362 of RAISING in non-EBL and EBL cattle.

363 To further validate the usefulness of clonality analysis by RAISING-CLOVA as a
364 diagnostic method, a comprehensive follow-up study of non-EBL cattle is required in a
365 clinical setting, focusing on carrier animals identified to have higher Cv in the screening
366 test. Previous studies have demonstrated that clonality analysis of HTLV-1-infected
367 patients showed that the Cv of infected cells increased earlier than the increase in
368 peripheral blood provirus levels before the onset of ATL, suggesting the applicability of
369 this method to the risk assessment of ATL (14, 15). In this study, similar results were
370 confirmed in a sheep lymphoma model. Sheep can be a good model for validating the
371 usefulness of this method, because they develops lymphoma at high frequencies and with
372 shorter period compared to cattle (2). We will also test additional sheep samples for
373 further validation. Taken together, further studies of BLV clonality analysis by
374 RAISING-CLOVA will contribute to the establishment of a new control measure to

375 predict disease prognosis of non-EBL cattle and prevent EBL onset by exposing high-risk
376 animals.

377

378 **Acknowledgments**

379 We are grateful to all farmers and veterinarians who provided clinical samples used in
380 this study. We thank Enago (<http://www.enago.jp>) for the English language review.

381

382 **Author contributions**

383 TO, SK, MS, NM, SM, and KO designed the work. TO, HS, MS, TM, NN, and SY
384 performed the experiments. MS, TM, NN, SY, and KM provided intellectual input, field
385 samples, laboratory materials, reagents, and/or analytic tools. TO, HS, SK, MS, TM, NN,
386 SY, and KM acquired, analyzed, and interpreted the data. TO and HS wrote the
387 manuscript. SK, MS, SY, KM, NM, SM, and KO revised the manuscript. All authors read
388 and approved the final manuscript.

389

390 **Data availability statement**

391 The datasets used and analyzed in this study are available from the corresponding author
392 on reasonable request.

393

394 **Additional Information**

395 MS, TM, and NN have a patent pending for materials and techniques described in this
396 paper (application number PCT/JP2020/30907). The other authors declare no competing
397 interests. This work was supported by grants from Ito Memorial Foundation (to SK), the
398 Science and Technology Research Promotion Program for Agriculture, Forestry,
399 Fisheries, and Food Industry, Japan (number 26058B; to SK), the NARO, Bio-oriented
400 Technology Research Advancement Institution (the special scheme project on regional

401 developing strategy; grant 16817557 to SK), and grants-in-aid for Scientific Research
402 (project numbers 19KK0172 and 22H02503 to SK, 19K15993 and 22K15005 to TO,
403 17H03594 to MS). The funders had no role in study design, data collection and
404 interpretation, or the decision to submit the work for publication.

405

406 **References**

407 1. Sagata, N. *et al.* Complete nucleotide sequence of the genome of bovine leukemia
408 virus: its evolutionary relationship to other retroviruses. *Proc. Natl. Acad. Sci. U.*
409 *S. A.* **82**, 677–681 (1985).

410 2. Schwartz, I. & Levy, D. Pathobiology of bovine leukemia virus. *Vet. Res.* **25**,
411 521–536 (1994).

412 3. Polat, M., Takeshima, S. & Aida, Y. Epidemiology and genetic diversity of
413 bovine leukemia virus. *Virol. J.* **14**, 209 (2017).

414 4. Murakami, K., Kobayashi, S., Konishi, M., Kameyama, K. & Tsutsui, T.
415 Nationwide survey of bovine leukemia virus infection among dairy and beef
416 breeding cattle in Japan from 2009–2011. *J. Vet. Med. Sci.* **75**, 1123–1126 (2013).

417 5. Gillet, N. A. *et al.* Massive depletion of bovine leukemia virus proviral clones
418 located in genomic transcriptionally active sites during primary infection. *PLoS*
419 *Pathog.* **9**, e1003687 (2013).

420 6. Ohnuki, N. *et al.* A target enrichment high throughput sequencing system for
421 characterization of BLV whole genome sequence, integration sites, clonality and
422 host SNP. *Sci. Rep.* **11**, 4521 (2021).

423 7. Nishimori, A., Andoh, K., Matsuura, Y., Kumagai, A. & Hatama, S.
424 Establishment of a simplified inverse polymerase chain reaction method for
425 diagnosis of enzootic bovine leukosis. *Arch. Virol.* **166**, 841–851 (2021).

426 8. Firouzi, S. *et al.* Development and validation of a new high-throughput method to
427 investigate the clonality of HTLV-1-infected cells based on provirus integration
428 sites. *Genome Med.* **6**, 29–32 (2014).

429 9. Rosewick, N. *et al.* Cis-perturbation of cancer drivers by the HTLV-1/BLV
430 proviruses is an early determinant of leukemogenesis. *Nat. Commun.* **8**, (2017).

431 10. Katsuya, H. *et al.* The Nature of the HTLV-1 Proivirus in Naturally Infected
432 Individuals Analyzed by the Viral DNA-Capture-Seq Approach. *Cell Rep.* **29**,
433 724–735 (2019).

434 11. Artesi, M. *et al.* PCIP-seq: simultaneous sequencing of integrated viral genomes
435 and their insertion sites with long reads. *Genome Biol.* **22**, 97 (2021).

436 12. Schmidt, M. *et al.* High-resolution insertion-site analysis by linear
437 amplification-mediated PCR (LAM-PCR). *Nat. Methods.* **4**, 1051–1057 (2007).

438 13. Gabriel, R., *et al.* Comprehensive genomic access to vector integration in clinical
439 gene therapy. *Nat. Med.* **15**, 1431–1436 (2009).

440 14. Wada, Y., *et al.* RAISING is a high-performance method for identifying random
441 transgene integration sites. *Commun. Biol.* **5**, 535 (2022).

442 15. Saito, M. *et al.* A high-throughput detection method for the clonality of Human
443 T-cell leukemia virus type-1-infected cells in vivo. *Int. J. Hematol.* **112**, 300–306
444 (2020).

445 16. Lairmore, M. D. Animal Models of Bovine Leukemia Virus and Human
446 T-Lymphotropic Virus Type-1: Insights in Transmission and Pathogenesis. *Annu.*
447 *Rev. Anim. Biosci.* **2**, 189–208 (2014).

448 17. Harms, J. S. & Splitter, G. A. Loss of MHC I transcription trans-activator in the
449 bovine B-LCL, BL3.1. *Vet. Immunol. Immunopathol.* **51**, 39–54 (1996).

450 18. Komori, H., Ishiguro, N., Horiuchi, M., Shinagawa, M. & Aida, Y. Predominant
451 p53 mutations in enzootic bovine leukemic cell lines. *Vet. Immunol.
452 Immunopathol.* **52**, 53–63 (1996).

453 19. Ochiai, C. *et al.* Characterization of microRNA expression in B cells derived from
454 Japanese black cattle naturally infected with bovine leukemia virus by deep
455 sequencing. *PLoS One* **16**, e0256588 (2021).

456 20. Nishimori, A. *et al.* Identification of an atypical enzootic bovine leukosis in Japan
457 by using a novel classification of bovine leukemia based on immunophenotypic
458 analysis. *Clin. Vaccine Immunol.* **24**, e00067-17 (2017).

459 21. Ohno, A., Takeshima, S., Matsumoto, Y. & Aida, Y. Risk factors associated with
460 increased bovine leukemia virus proviral load in infected cattle in Japan from
461 2012 to 2014. *Virus Res.* **210**, 283–290 (2015).

462 22. Kobayashi, T. *et al.* Increasing Bovine leukemia virus (BLV) proviral load is a
463 risk factor for progression of Enzootic bovine leucosis: A prospective study in
464 Japan. *Prev. Vet. Med.* **178**, 104680 (2020).

465 23. Ott, S. L., Johnson, R. & Wells, S. J. Association between bovine-leukosis virus
466 seroprevalence and herd-level productivity on US dairy farms. *Prev. Vet. Med.* **61**,
467 249–262 (2003).

468 24. Gillet, N. *et al.* Mechanisms of leukemogenesis induced by bovine leukemia
469 virus: prospects for novel anti-retroviral therapies in human. *Retrovirology* **4**, 18
470 (2007).

471 25. Murakami, H. *et al.* Bovine leukemia virus integration site selection in cattle that
472 develop leukemia. *Virus Res.* **156**, 107–112 (2011).

473 26. Somura, Y., Sugiyama, E., Fujikawa, H. & Murakami, K. Comparison of the copy
474 numbers of bovine leukemia virus in the lymph nodes of cattle with enzootic
475 bovine leukosis and cattle with latent infection. *Arch. Virol.* **159**, 2693–2697
476 (2014).

477 27. Durkin, K. *et al.* Characterization of novel Bovine Leukemia Virus (BLV)
478 antisense transcripts by deep sequencing reveals constitutive expression in tumors
479 and transcriptional interaction with viral microRNAs. *Retrovirology* **13**, 33
480 (2016).

481

482

483 **Figure legends**

484 **Fig. 1. Performance of BLV clonality analysis by RAISING-CLOVA.**

485 (A, B) Detection limit of RAISING targeting BLV and sensitivity of its clonality analysis
486 by RAISING-CLOVA. RAISING-CLOVA was performed using genomic DNA from the
487 mixture of the BLV-infected cell line BL3.1 with the BLV-uninfected cell line BTL26.
488 Samples were prepared by mixing BL3.1 with BTL26 in proportions ranging from 100%,
489 50%, 10%, 5%, 1%, 0.5%, 0.1%, 0.05%, 0.01%, 0.005%, and 0.001% of BL3.1. (A)
490 Products of RAISING using these mixed samples were visualized by electrophoresis on
491 2% agarose gel. BLV PVL contained in each mixed sample was shown above the photo.
492 (B) The minimum BLV PVL to measure a reliable Cv with RAISING-CLOVA was
493 examined. Sangar sequence spectrum of amplicons of RAISING using a dilution series of
494 BL3.1 were shown in Supplemental Figure 1. (C) Comparison of sequence spectra
495 patterns and signal plots of CLOVA analysis in blood samples of BLV-infected cattle
496 with different clonality values. RAISING-CLOVA was performed using genomic DNA
497 of peripheral blood from EBL and non-EBL cattle. In the plots of CLOVA analysis,
498 signals of dominant clones and the others were shown in purple and orange lines,
499 respectively. The Cv was calculated by dividing the average of signal peak area values of
500 host genome sequence of the dominant clone (red lines) by that of proviral sequence
501 (blue lines). (D) Comparison of Cv analyzed in two different laboratories. RAISING was
502 independently performed at Hokkaido University (HU) and the National Institute of
503 Infectious Diseases (NIID) using the specimens of BLV-infected cattle ($n = 32$) and Cv
504 was determined using CLOVA. Spearman's rank correlation coefficient was used for
505 statistical analysis.

506

507 **Fig. 2. Comprehensive clonality analysis by RAISING-CLOVA targeting BLV using**
508 **specimens of EBL and non-EBL cattle.**

509 (A–E) RAISING-CLOVA was performed using genomic DNA of blood samples of AL
510 ($n = 107$), PL ($n = 79$), and EBL cattle ($n = 101$) and tumor samples of EBL cattle ($n =$
511 175). (A) Representative Sanger sequence spectrum of blood and tumor samples of
512 BLV-infected cattle (blood samples of AL002, PL002, and EBL015 and a tumor sample
513 from a lymph node of EBL015). (B, C) Cv (B) and BLV PVL (C) was measured in blood
514 samples of AL, PL, and EBL cattle and tumor samples of EBL cattle. (D, E) Cv (D) and
515 BLV PVL (E) among blood and tissue categories in EBL cattle. (B–E) Median values for
516 each group are indicated by black bars. Dunn's tests were used for statistical analysis.
517 *** $p < 0.001$, **** $p < 0.0001$, n.s., not significant.

518

519 **Fig. 3. Clinical utility of BLV clonality analysis by RAISING-CLOVA using blood**
520 **and tumor samples.**

521 (A, B) ROC analysis using Cv (A) and PVL (B) of blood samples to distinguish EBL in
522 the blood samples. AUC, area under the curve; CI, confidence interval. (C) Frequencies
523 of blood specimens from EBL ($n = 101$, purple), PL ($n = 79$, orange), and AL cattle ($n =$
524 107, green) per Cv at intervals of 0.05. A dotted line indicates the proposed cut-off value
525 to classify BLV-infected cattle into EBL. (D) Scatter plots of Cv and BLV PVL of blood
526 samples from EBL ($n = 101$, purple diamond), AL ($n = 107$, green circle), and PL cattle
527 ($n = 79$, orange triangle). Dotted lines indicate proposed cut-off values to classify
528 BLV-infected cattle into EBL.

529

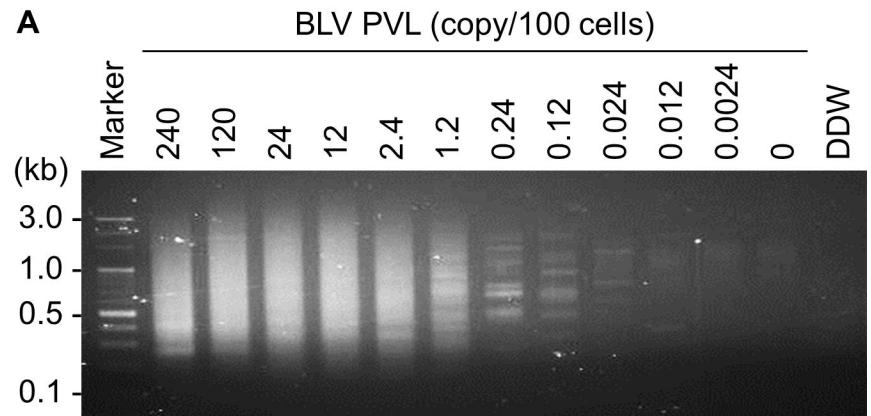
530 **Fig. 4. Longitudinal clonality analysis of BLV-infected sheep with lymphoma.**

531 (A, B) RAISING-CLOVA was performed using genomic DNA of blood and tumor
532 samples of BLV-challenged sheep ($n = 4$). (A) Kinetics of BLV PVL (blue circle) and Cv
533 (orange diamond) in blood samples of BLV-infected sheep from viral challenge to
534 lymphoma onset. (B) Sequence analysis of the proviral integration sites in blood and/or
535 tumor samples of the challenged sheep.

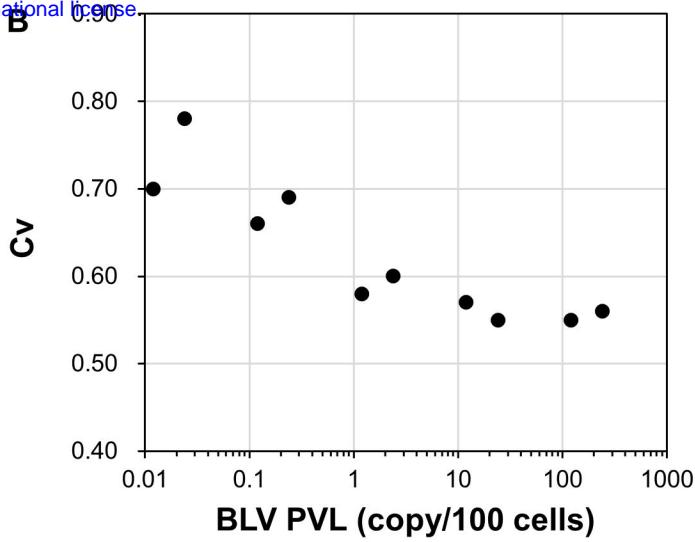
536

Figure 1

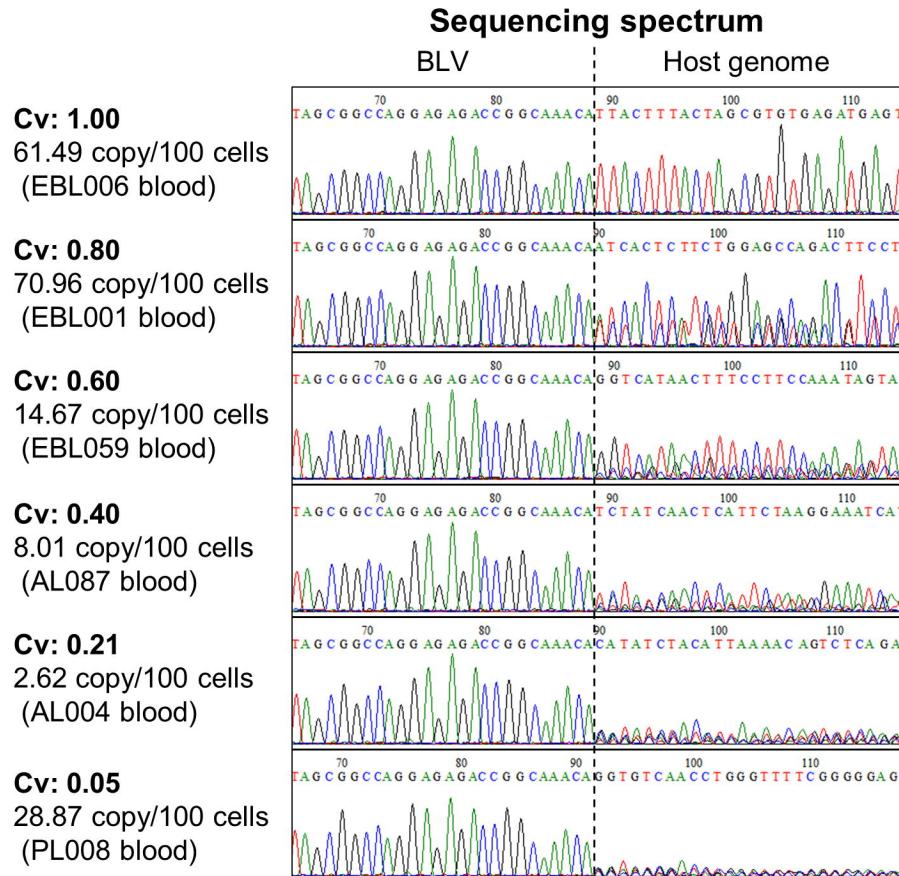
A



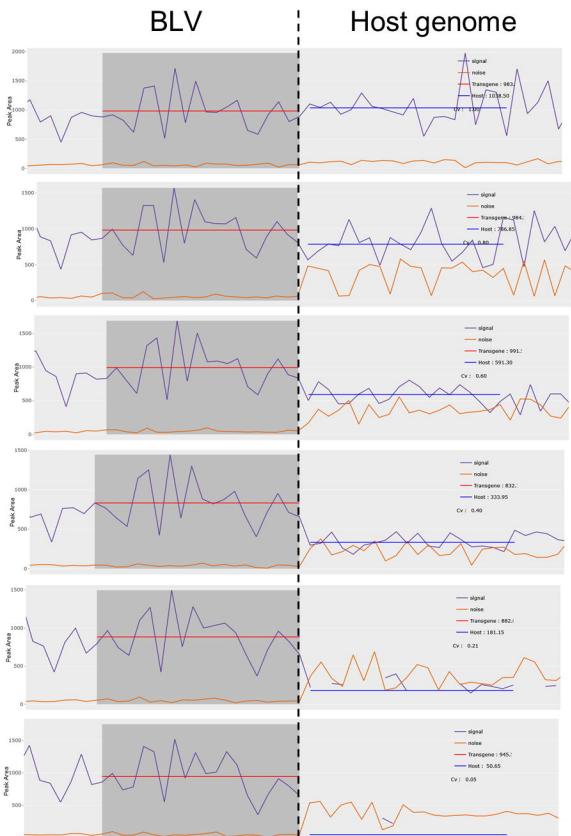
B



C



CLOVA images



D

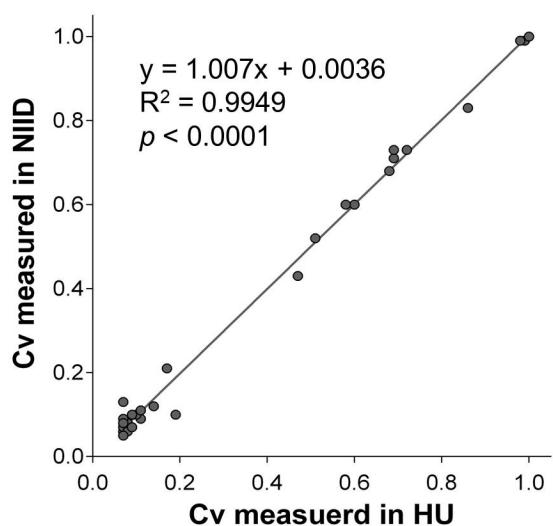
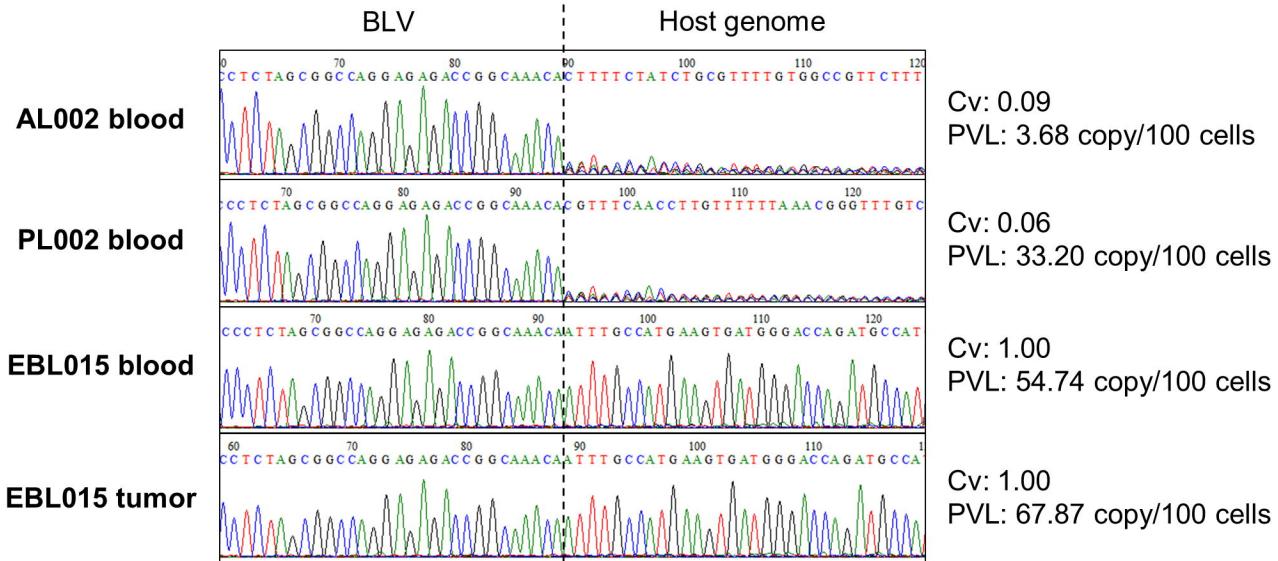
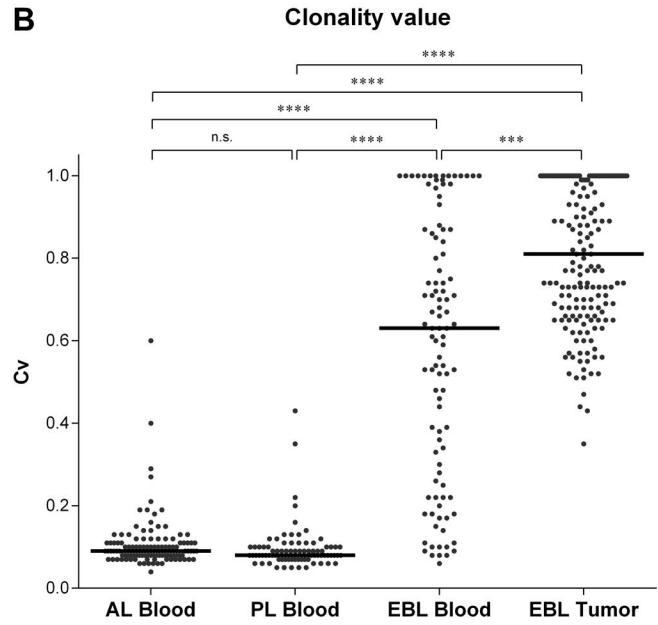


Figure 2

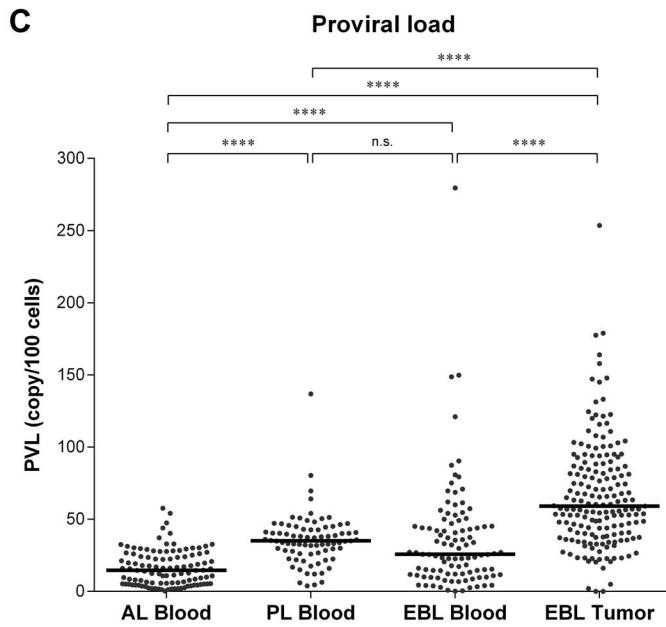
A



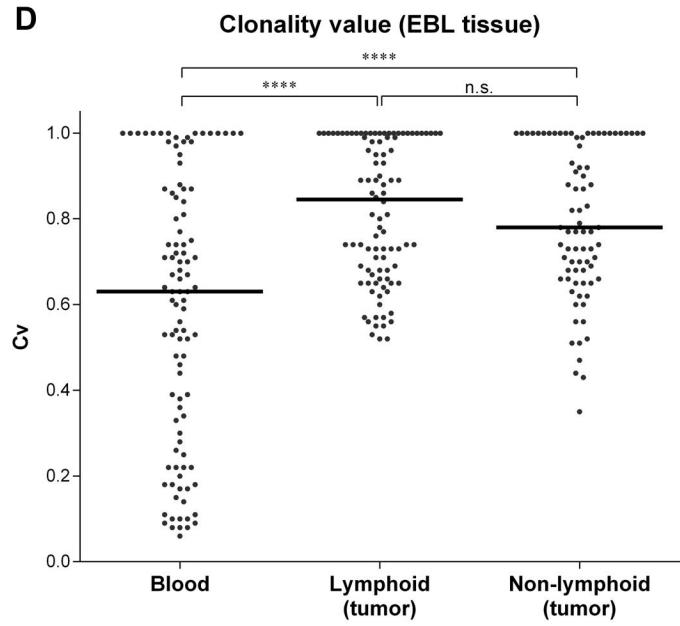
B



C



D



E

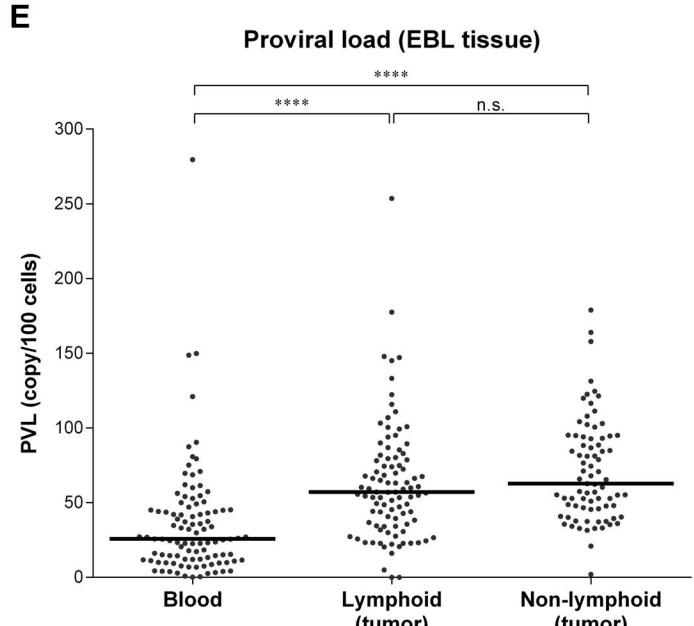


Figure 3

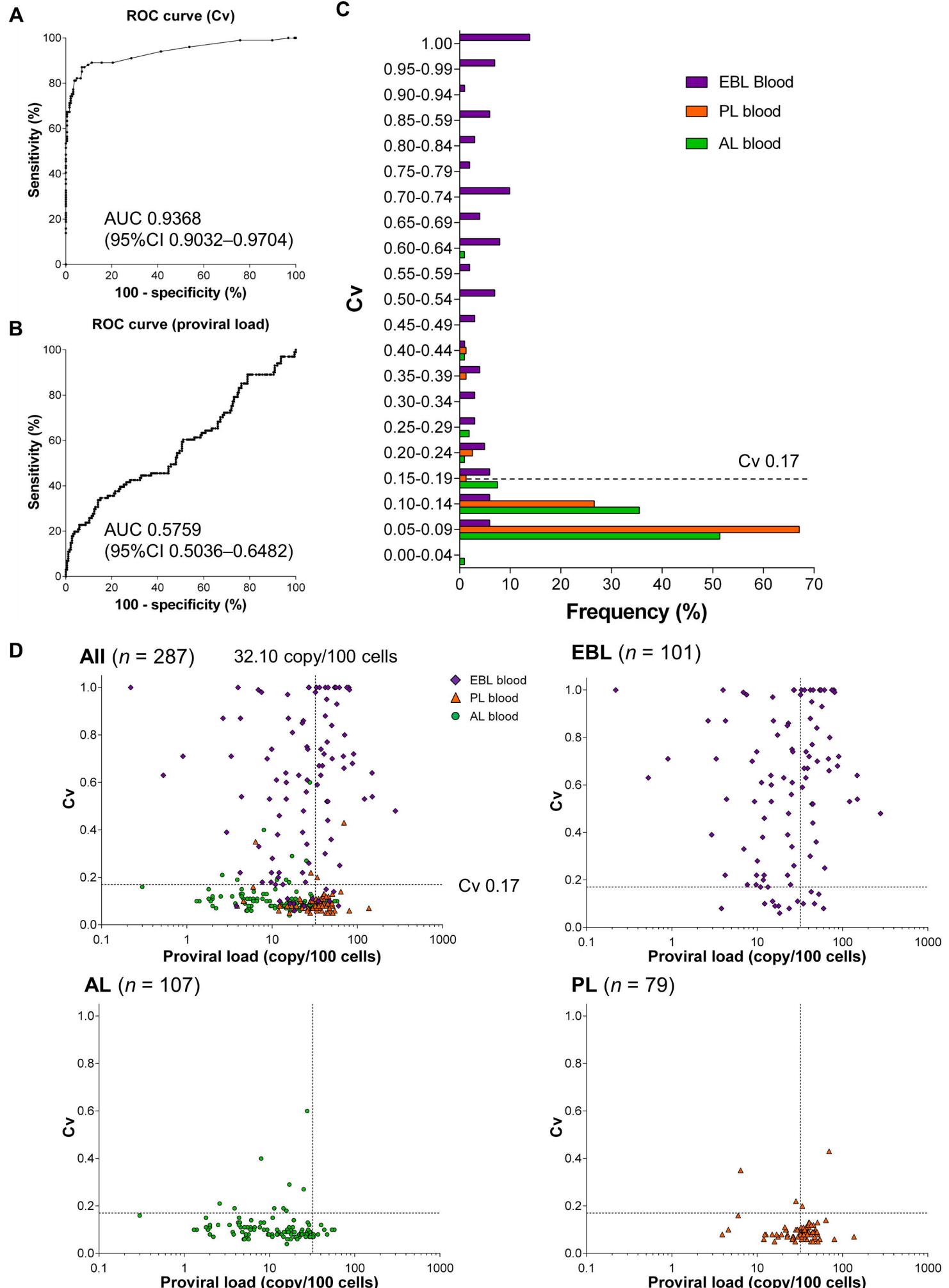


Figure 4

