

Massive Depletion of Bovine Leukemia Virus Proviral Clones Located in Genomic Transcriptionally Active Sites during Primary Infection

Nicolas A. Gillet^{1,2*}, Gerónimo Gutiérrez^{3,9}, Sabrina M. Rodriguez^{1,2,9}, Alix de Brogniez^{1,2,9}, Nathalie Renotte^{1,2}, Irene Alvarez³, Karina Trono³, Luc Willems^{1,2*}

1 Molecular and Cellular Epigenetics, Interdisciplinary Cluster for Applied Genoproteomics (GIGA) of University of Liège (ULg), Sart-Tilman, Liège, Belgium, **2** Molecular and Cellular Biology, Gembloux Agro-Bio Tech, University of Liège (ULg), Gembloux, Belgium, **3** Instituto de Virología, Centro de Investigaciones en Ciencias Veterinarias y Agronómicas, INTA, Castelar, Argentina

Abstract

Deltaretroviruses such as human T-lymphotropic virus type 1 (HTLV-1) and bovine leukemia virus (BLV) induce a persistent infection that remains generally asymptomatic but can also lead to leukemia or lymphoma. These viruses replicate by infecting new lymphocytes (i.e. the infectious cycle) or via clonal expansion of the infected cells (mitotic cycle). The relative importance of these two cycles in viral replication varies during infection. The majority of infected clones are created early before the onset of an efficient immune response. Later on, the main replication route is mitotic expansion of pre-existing infected clones. Due to the paucity of available samples and for ethical reasons, only scarce data is available on early infection by HTLV-1. Therefore, we addressed this question in a comparative BLV model. We used high-throughput sequencing to map and quantify the insertion sites of the provirus in order to monitor the clonality of the BLV-infected cells population (i.e. the number of distinct clones and abundance of each clone). We found that BLV propagation shifts from cell neo-infection to clonal proliferation in about 2 months from inoculation. Initially, BLV proviral integration significantly favors transcribed regions of the genome. Negative selection then eliminates 97% of the clones detected at seroconversion and disfavors BLV-infected cells carrying a provirus located close to a promoter or a gene. Nevertheless, among the surviving proviruses, clone abundance positively correlates with proximity of the provirus to a transcribed region. Two opposite forces thus operate during primary infection and dictate the fate of long term clonal composition: (1) initial integration inside genes or promoters and (2) host negative selection disfavoring proviruses located next to transcribed regions. The result of this initial response will contribute to the proviral load set point value as clonal abundance will benefit from carrying a provirus in transcribed regions.

Citation: Gillet NA, Gutiérrez G, Rodriguez SM, de Brogniez A, Renotte N, et al. (2013) Massive Depletion of Bovine Leukemia Virus Proviral Clones Located in Genomic Transcriptionally Active Sites during Primary Infection. *PLoS Pathog* 9(10): e1003687. doi:10.1371/journal.ppat.1003687

Editor: Susan R. Ross, University of Pennsylvania School of Medicine, United States of America

Received: June 6, 2013; **Accepted:** August 22, 2013; **Published:** October 3, 2013

Copyright: © 2013 Gillet et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was supported by the “Fonds National de la Recherche Scientifique” (FNRS), the Télévie, the Interuniversity Attraction Poles (IAP) Program BELVIR initiated by the Belgian Science Policy Office, the “Action de Recherche Concertée Glyvir” of the “Communauté française de Belgique”, the Belgian Foundation Against Cancer, the Sixth Research Framework Programme of the European Union (project INCA LSHC-CT-2005-018704), the “Neoangio” excellence program and the “Partenariat Public Privé” PPP INCA of the “Direction générale des Technologies, de la Recherche et de l’Energie/DG06” of the Walloon government, the “Centre anticancéreux près ULg” (CAC), the “Synbiofor” and “Agricultureislife” projects of GxABT, the “ULg Fonds Spéciaux pour la Recherche”, the “Plan Cancer” of the “Service Public Fédéral” and INTA projects AESA 203992/PNLEC 1601. The costs of high throughput sequencing were supported by the “Roche and Illumina award” with the help of the GIGA technology platforms and Diagenode bioruptor. NAG (postdoctoral researcher) and LW (Research Director) are members of the FNRS while IA and GG are researchers from CONICET. The research funders and sponsors had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: n.gillet@ulg.ac.be (NAG); luc.willems@ulg.ac.be (LW)

9 These authors contributed equally to this work.

Introduction

The deltaretrovirus genus includes human T-lymphotropic viruses (HTLVs), simian T-lymphotropic viruses (STLVs) and the bovine leukemia virus (BLV). These viruses induce a life-long persistent infection that remains generally asymptomatic (reviewed by [1–3]). Nevertheless, HTLV-1 and BLV cause leukemia or lymphoma in a minority of infected hosts after a long period of latency [3,4]. Viral spread within the host uses two distinct processes. First, the infectious cycle results from virion attachment to target lymphocytes, entry of viral single-stranded RNA, reverse-transcription and integration as provirus into the host genome

(also known as the infectious cycle) [5–7]. The second strategy of replication relies on driving cell proliferation using viral regulatory proteins such as Tax (i.e. the mitotic cycle) [8,9]. These two viral replication routes thus generate a series of infected cell populations that are composed of numerous distinct clones (i.e. a population of cells carrying the provirus at a given site of the host genome). Animal models using experimental inoculation of squirrel monkey with HTLV-1 or sheep with BLV demonstrated that the infectious cycle dominates early infection and finishes 1 to 8 months later [10,11]. Thereafter, the proviral load (PVL) is mainly maintained by mitotic replication of infected cells [12–15]. In HTLV-1 infected individuals, the majority of the infected clones are indeed

Author Summary

Human T-lymphotropic Virus 1 (HTLV-1) induces a persistent infection that remains generally asymptomatic. Nevertheless, in a small proportion of individuals and after a long latency, HTLV-1 infection leads to leukemia or lymphoma. Onset of clinical manifestations correlates with a persistently elevated number of infected cells. Because the vast majority of cells are infected at early stages, primary infection is a crucial period for HTLV-1 persistence and pathogenesis. Since HTLV-1 is transmitted through breast feeding and because systematic population screenings are rare, there is a lack of available samples at early infection. Therefore, we addressed this question in a closely related animal model by inoculating cows with Bovine Leukemia Virus (BLV). We show that the vast majority of cells becoming infected during the first weeks of infection and do not survive later on. We also demonstrate that the initial host selection occurring during primary infection will specifically target cells that carry a provirus inserted in genomic transcribed regions. This conclusion thus highlights a key role exerted by the host immune system during primary infection and indicates that antiviral treatments would be optimal when introduced straight after infection.

relatively stable during many years [16]. Remarkably, using BLV-sheep experimental infection, it has been shown that the leukemic clone can be detected as early as one month after inoculation [17]. Thus, efficient virus replication via production of virions and infection of new target cells occurs mostly during a very short period following viral inoculation (so-called primary infection). As a consequence, the vast majority of the infected clones were created during this crucial period of primary infection.

Unfortunately, these important early times of HTLV-1 infection cannot be studied due to the paucity of available samples. Therefore, very little is known about the modes of viral replication during this primary phase of infection. For instance, what proportion of clones generated during primary infection will establish in the long term? Is there a negative selection against particular clones? And if so, is there a role of genomic integration sites in clonal selection? This is particularly important because clones are not equal regarding their proliferative potential. In HTLV-1 infected patients, the abundance of a given clone is enhanced by the integration of its provirus in an actively transcribed area of the genome [16,18]. Furthermore, it is well described that increased proviral load (PVL) and clonal expansion of infected cells are key events during the process of leukemogenesis in HTLV-1 carriers. Indeed, HTLV-1 PVL correlates with the risk of developing adult T-cell leukemia or lymphoma (ATLL) [19] and HTLV-1 clonality of ATLL patients is characterized by the presence of massively expanded clones [12,13,16].

In this context, the aim of this study was to evaluate clonal evolution of the infected cells with longitudinal samples during primary infection with respect to clonal diversity and host-dependent negative selection. Due to the difficulty at accessing samples from newly HTLV-1 infected patients, we addressed this question in a closely related animal model by inoculating cows with BLV. We selected animals whose BLV PVL set point values encompassed the whole range of natural variability. We then used a newly developed method to map and quantify the insertion sites of the provirus in order to monitor the clonality of the BLV-infected cells population (i.e. the number of distinct clones and

abundance of each clone) during primary infection. We demonstrated that (1) BLV initially targets transcribed regions of the genome for integration; (2) later, a massive clone selection occurs during primary infection disfavoring proviruses located nearby genes; (3) nevertheless, the abundance of the long term maintained clones benefits from the transcriptional activity of the genomic region surrounding the provirus.

Results

Antibody response against BLV and BLV proviral load evolution during primary infection

Five cows were inoculated with a cloned BLV provirus (strain 344) as described in the Materials and Methods section. Figure 1A shows that all animals developed a humoral response directed against BLV SU, the viral surface envelope protein. Anti-SU antibodies were detected at 16 and 30 days post-inoculation of animals #21 and #23/31/492/535, respectively. Since then, all animals remained sero-positive for SU antibodies.

Figure 1B shows the PVL (number of proviral copies per 100 peripheral blood mononuclear cells, PBMCs) following BLV inoculation of cows at day 0. The PVLs sharply increased with a maximal value reached between 30 to 68 days post-inoculation (dpi) (30 days for #21, 43 days for #23, #31 and #535 and 68 days for #492, arrows on figure 1B). PVL peak values differed widely among animals ranging from 1.6 proviral copies per 100 PBMCs for animal #21 up to 236 proviral copies per 100 PBMCs for animal #535. Then, the PVLs subsequently decreased and reached a level that differed widely between animals. Cows #21, #23 and #31 (green lines) presented a low PVL set point (respectively 0.2, 0.5 and 0.1 copies per 100 PBMCs at day 252) as observed in the majority of infected animals kept in herds [20]. Animals #492 and #535 (red lines) developed a very high BLV burden (respectively 43.3 and 117.6 copies per 100 PBMCs at day 252) as measured in a small proportion of infected animals ([21] and INTA experimental facilities, data not shown). None of the five animals progressed to disease within two years post inoculation.

BLV inoculation into 5 cows thus resulted in the onset of an antibody response at days 16–30 which slightly preceded the maximal proviral burden (at days 30–68). This experimental setting is thus representative of BLV infection in herd conditions.

Polyclonal propagation during primary infection

To trace BLV clonality during primary infection, proviral insertion sites were selectively amplified, identified by high-throughput sequencing and quantified as schematized in Supporting Figure S1 in Text S1. Briefly, genomic DNA was extracted from PBMCs and sonicated. The end of the BLV 3'LTR and a fragment of bovine genomic DNA were amplified by linker-mediated PCR and the products subjected to high-throughput sequencing. Proviral insertion sites were determined by alignments of sequences downstream of the 3' LTR (read 1; Figure S1 in Text S1). A clone was defined as a population of cells carrying the BLV provirus at a given insertion site in the bovine genome. The abundance of a clone was quantified by counting the number of different shear sites for that particular clone using read 2 alignments. A complete description of the procedure is given in the Material and Methods section.

During primary infection, a broad range of clones were generated within each animal: from 264 in #21 to 12906 in #535 (Table S1 in Text S1). The distributions of clone abundance for each animal and time point are depicted in Figure 2A by pie charts where each slice represents the relative abundance of a

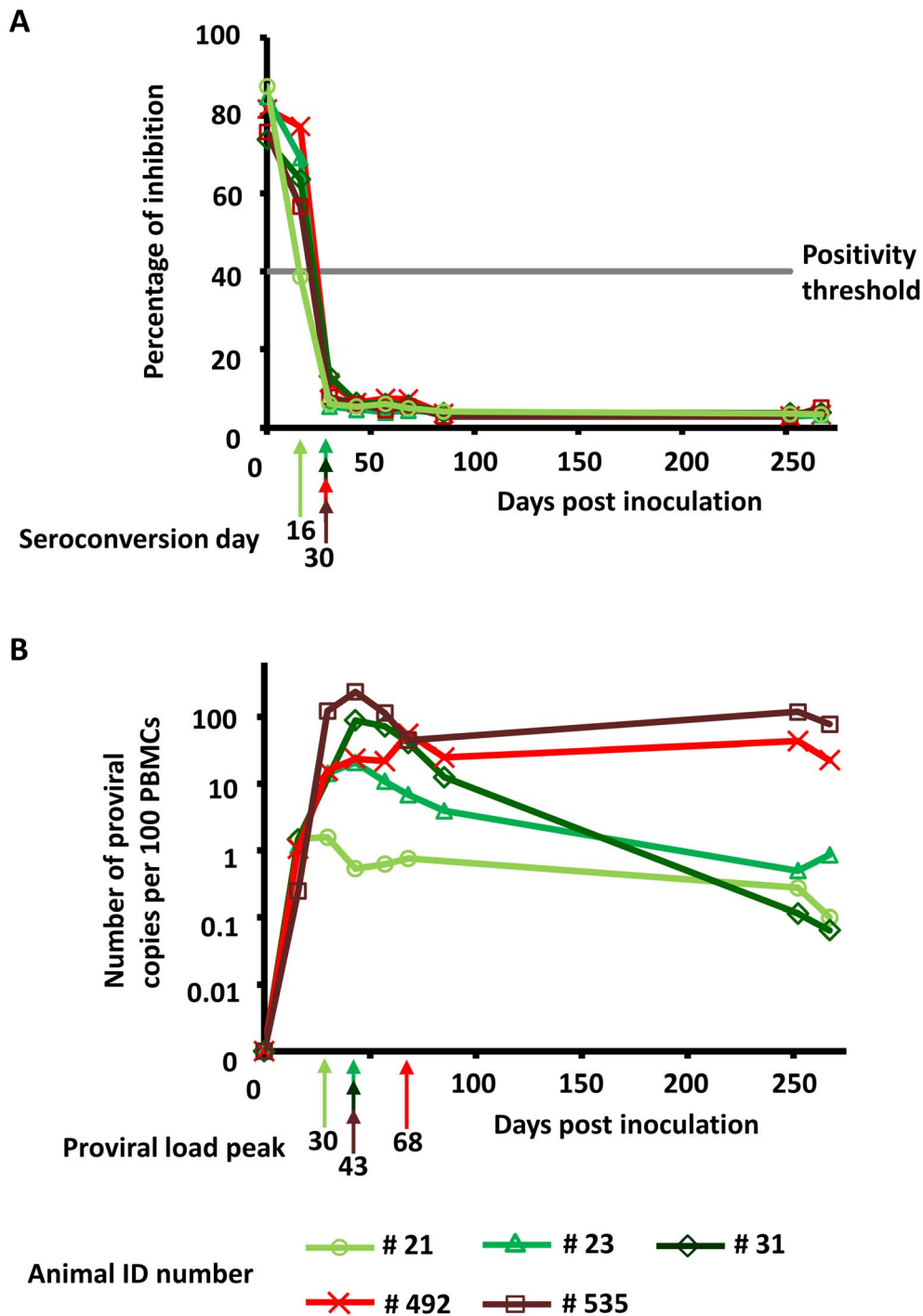


Figure 1. Antibody responses against BLV and BLV proviral loads. **A.** BLV was inoculated at day zero. Antibodies against BLV SU were detected with a competitive ELISA. Sero-positivity (arrows) occurred when percentage of inhibition dropped below a threshold of 40% (defined as the seroconversion day). **B.** Proviral loads (PVLs) were determined by quantitative real time PCR. Arrows indicate the date when PVL reached maximum value within each animal.
doi:10.1371/journal.ppat.1003687.g001

given clone. Based on this data, we calculated an oligoclonality index which is a measure of the non-uniformity of the clone abundance distribution (i.e. an index close to 1 corresponds to an unevenly-distributed population composed almost of a single clone, see Supplemental Materials and Methods in Text S1). The oligoclonality index was extremely low at 16 dpi, showing that all

infected clones had approximately the same abundance (Figure 2B). The oligoclonality index then rose and reached a peak value after 43 days for animals #21, #23, #31 and #535 and after 57 days for animal #492. These data thus demonstrate that early infection is characterized by a very polyclonal propagation.

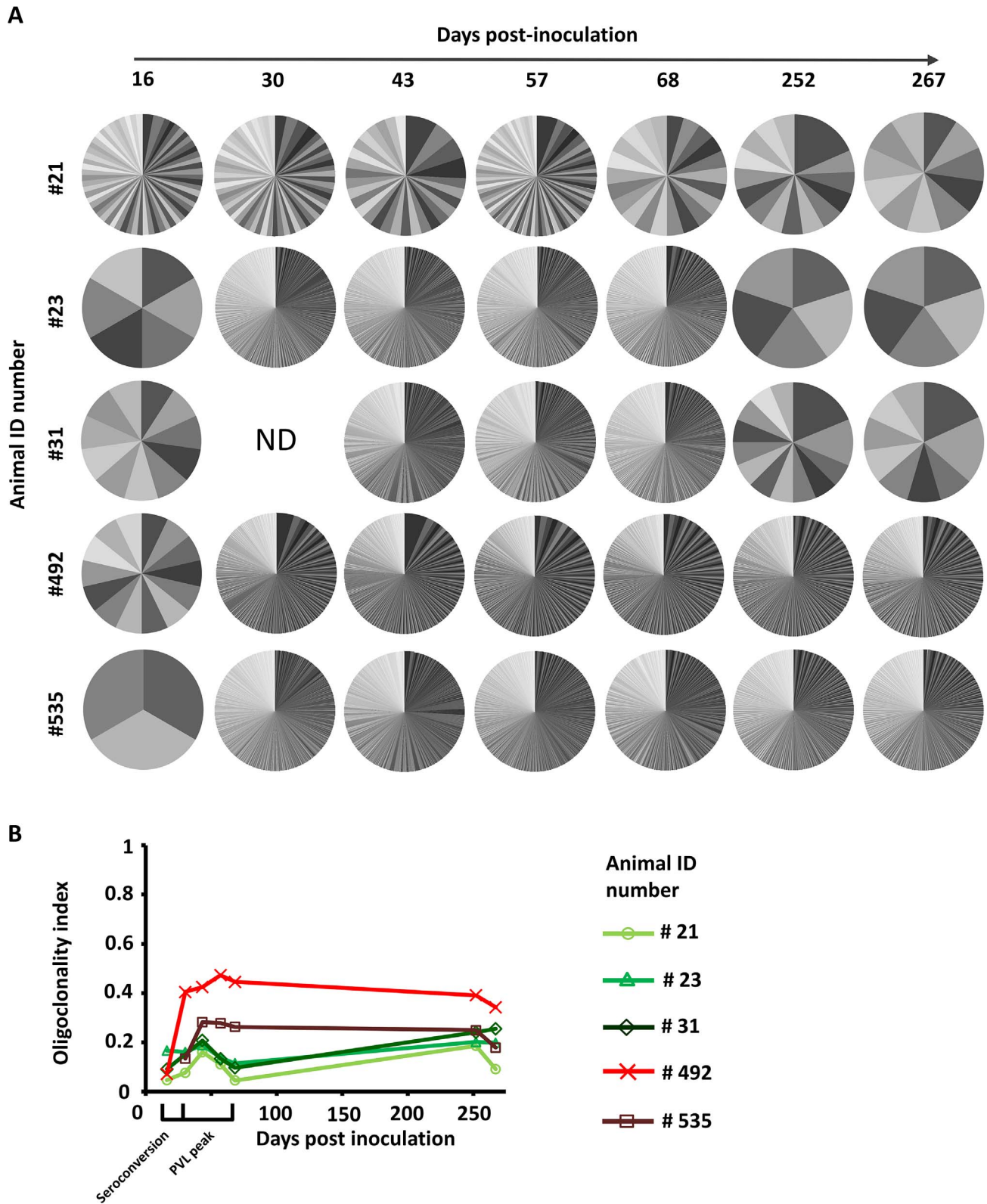


Figure 2. Distribution of clone abundance determined by high throughput sequencing. A. The pie-charts depict the distribution of clone abundance for each animal and time point. Each slice represents a particular BLV-infected clone and the size of the slice represents the relative abundance of that clone. ND: not done due to sample unavailability. Note that although the gray tones are identical at different dates, the clones are different. **B.** Oligoclonality index measures the non-uniformity of the clone abundance distribution and ranges from 0 (perfect polyclonality) to 1 (perfect monoclonality).

doi:10.1371/journal.ppat.1003687.g002

Massive depletion of BLV-infected clones after primary infection

At later times post-infection, a massive depletion of BLV-infected clones occurred (Figure 3A). The left-hand dot plot shows that, on average, 97% of the clones identified at seroconversion were no longer detected at 267 dpi. Similarly, the other dot plot shows that, on average, 92% of the proviral load at seroconversion resulted from clones that disappeared at 267 dpi. The pie chart on the right of Figure 3A illustrates this massive clonal depletion, the red slices representing the clones of #492 at dpi = 30 that were not detected in the longer term. We conclude that primary infection is thus characterized by a massive depletion of proviral clones.

BLV replication shifts from neo-infection to clonal proliferation

To quantify the clonal overlap between two successive time points, a similarity index was calculated (see Supplemental Materials and Methods in Text S1). Figure 3B shows the evolution of the similarity index calculated between two successive populations separated by a period of two weeks. Soon after seroconversion, the similarity indexes were close to zero showing that, between 2 successive time points, the infected cell populations shared few clones. Rapidly, the similarity indexes shown on Figure 3B rose to reach values close to 1 in all animals except in #21. Infected cell populations thus stabilized in terms of clonal composition and abundance. In the long term (at dpi = 252), the similarity indexes remained close to 1 in all animals except #21. Figure 3C illustrates the kinetics of appearance of clones that were detected in the long term. It appeared that a small percentage of the long term proviral load originated from clones detected at the seroconversion period. Nevertheless, the cumulative curves rose sharply from day 30 and then reached a plateau at day 68 in all animals except #21 (Figure 3C). This shows that most clones detected in the long term originated from the PVL peak period. The pie chart on the right illustrates that a large proportion of clones established in the long term originated from the beginning of the infection. The red slices symbolize the clones present in the long term and already detected at day 68 or before. In contrast, all clones detected in the long term in animal #21 were novel, resulting in a similarity index of 0 at day 252 (Figure 3B) and a flat cumulative curve (no clone at day 267 was detected at earlier time points, Figure 3C). So, long term proviral load in animal #21 might have been maintained either by newly infected cells or by pre-existing clones not detected beforehand. It should be mentioned that the numbers of clones detected in this particular animal were low at every time point compared to the 4 other cows (see Table S1 in Text S1) due to a constantly low proviral load.

The kinetics of the appearance of long term clones and of the similarity indexes thus demonstrate that BLV propagation shifts from cell neo-infection to proliferation of pre-existing clones (i.e. clonal expansion) generally soon after seroconversion.

BLV proviral integration initially favors pol II and pol III transcribed genomic regions

The upper panel of Figure 4A shows the favored genomic consensus sequence surrounding the BLV provirus insertion site. As comparison, Figure 4B depicts the preferred insertion site of HTLV-1 in the human genome [22]. It thus appears that BLV and HTLV-1 favor highly similar genomic sequences for insertion.

BLV provirus integration did not occur randomly as illustrated on Figure 4C. Indeed, sequence analyses demonstrate that BLV proviruses initially inserted more frequently than expected by chance inside Refseq genes (Figure 4C, second row and first

column: 0 bp). Insertion was also favored nearby CpG islands, Refseq genes, tRNA genes and tRNA pseudogenes (Figure 4C, first 4 rows of the heat map). These different DNA sequences are associated with transcribed regions of the genome. Indeed, CpG islands are genomic regions with relatively high CG content generally associated with promoters [23]. Genes, whose promoters are rich in CpG sequences, tend to be expressed in most tissues (housekeeping genes). On the other hand, tRNA (transfer RNA) genes are constitutively transcribed by RNA polymerase III (pol III) unlike mRNA coding genes (which are like Refseq genes) transcribed by RNA polymerase II (pol II). Finally, tRNA pseudogenes are tRNA-derived repeats that no longer produce functional tRNA but that are still associated with polIII activity [24].

We conclude that BLV initial integration favors pol II and pol III transcribed regions of the genome.

In contrast, BLV disfavors interspersed repeats such as LINE, SINE and ERVs (Figure 4C). These elements are the largest class of sequences in mammalian genomes, making for about 50% of their total length. Most common repeats are retrotransposons that can replicate and reinsert in another site of the genome. Nevertheless, transposition does not occur in somatic cells where retrotransposons are silenced, thereby preventing insertional mutagenesis. Retrotransposons can be divided in several classes with notably SINE (Short Interspersed Nuclear Element), LINE (Long Interspersed Nuclear Element) and LTR-ERVs (Endogenous RetroVirus that are also retrotransposons with Long Terminal Repeats, LTR). Figure 4C demonstrates that insertion occurred less frequently than expected nearby retrotransposons LINE BovB, SINE BOV-A2, SINE ART2A and LTR ERVs.

We conclude that BLV initial integration disfavors transcriptionally silenced regions of the genome.

Negative selection against proviruses located nearby transcribed regions during primary infection

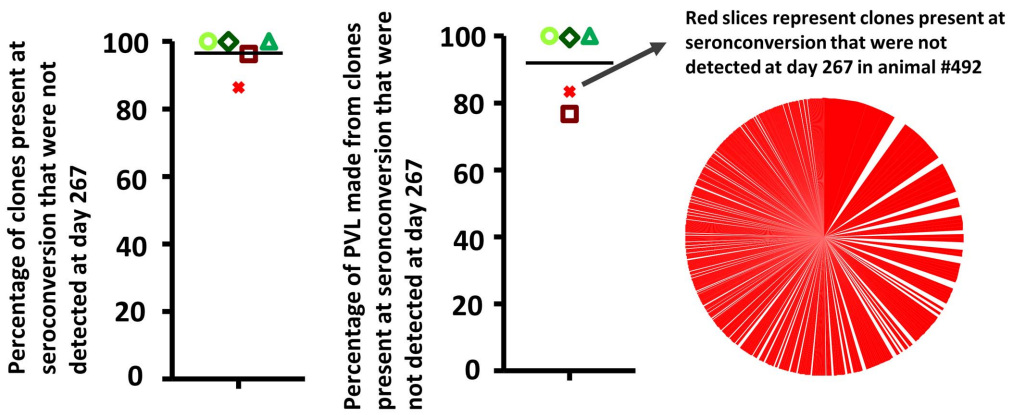
Figure 5 illustrates the characteristics of the genomic integration site during clonal evolution from seroconversion up to day 68. The proviral environment at seroconversion serves as reference point (Figure 5, first column of the heat map). The proportion of clones carrying a provirus located next to CpG islands, Refseq genes, tRNA genes and pseudogenes significantly decreased from the seroconversion date up to 68 days after inoculation (Figure 5, first 4 lanes). In contrast, there was a trend to increase the proportion of clones containing a provirus next to genomic repeats (Figure 5, six last lanes: LINE, SINE and ERVs). The proviral environment of clones detected at days 252 or 267 was not statistically different from that observed at day 68 (data not shown).

We conclude that massive depletion of clones during primary infection is characterized by a preferential selection against proviruses inserted in transcribed regions.

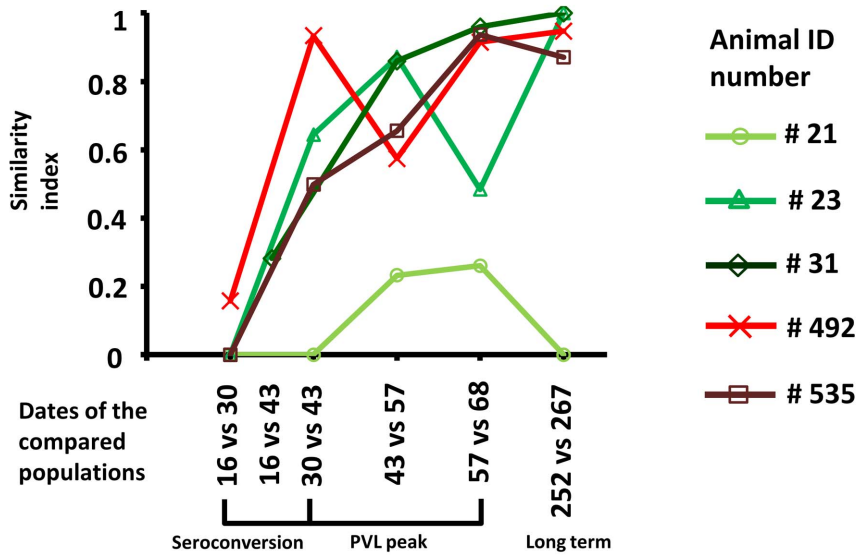
The abundance of established clones benefits from a provirus inserted in a transcribed genomic environment

Finally, we characterized the genomic environment of the proviruses that thrived in the long term. Therefore, we analyzed clonality in a series of cows infected for more than 2 years and harboring a wide range of PVL (Figure 6A). These animals presented a polyclonal distribution of BLV-infected cells (Figure 6B). The numbers of detected clones per animal are shown in Table S2 in Text S1. The distributions of clone abundance for each animal are depicted in Figure 6C. Figure 6D compares the genomic environment of proviruses from clones of increasing abundance. The first column of the heat map, which

A



B



C

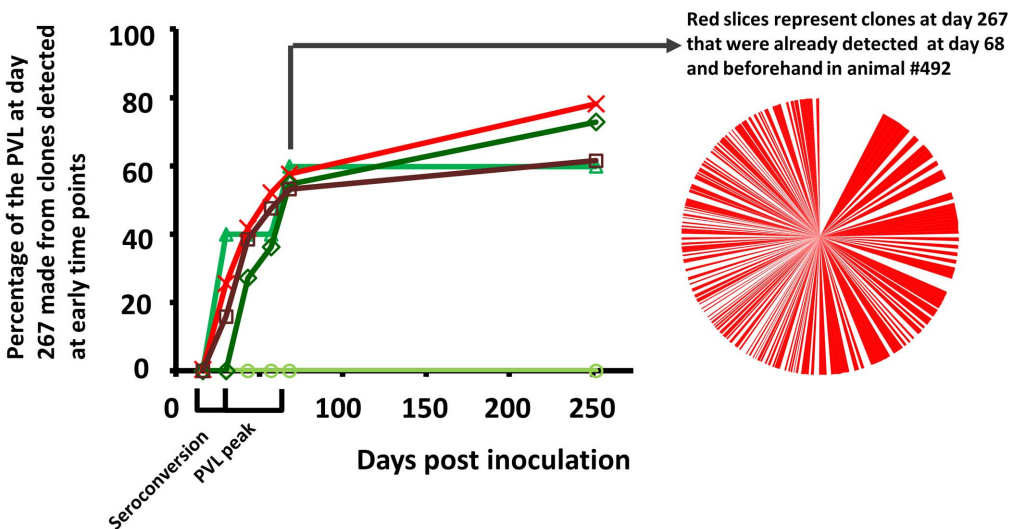


Figure 3. Evolution of the clonal composition. A. Dot-plots show the number of clones present at seroconversion that will no longer be detected at day 267. The means were indicated by the horizontal black lines. The pie-chart illustrates the massive clonal depletion observed in animal #492. Each slice represents a clone detected at seroconversion and the red ones correspond to clones that will not be detected in the long term. **B.** Similarity indexes quantify the clonal overlap between two successive populations and ranges from 0 (no clones are shared between the two populations) to 1 (the 2 populations are totally similar). The compared populations were separated by a period of about 2 weeks time. **C.** Cumulative curves showing the proportion of the PVL at day 267 that is composed from clones detected at earlier time points. The pie chart on the right illustrates the amount of clones at day 267 that were already detected at day 68 and beforehand. Each slice represents a clone detected at day 267 and the red ones correspond to clones that were detected at day 68 or before.
doi:10.1371/journal.ppat.1003687.g003

shows the less represented clones (below 1 cell per 10^3 PBMCs) and was used as reference. The proportion of clones carrying a provirus located next to CpG islands or Refseq genes significantly increased with clone abundance (Chi-squared test for trend, respectively $p = 0.0008$ and $p = 0.001$). Although not significant, there was a similar trend for proximity with tRNA genes and tRNA pseudogenes. We conclude that the abundance of a long term established BLV-infected clone is enhanced by the integration of its provirus nearby a transcribed unit.

Discussion

HTLV-1 induces a persistent infection that is fortunately generally asymptomatic. Nevertheless, in a small proportion of individual and after a long latency, the infection leads to leukemia or lymphoma. These clinical manifestations are correlated with a persistently elevated proviral load and an oligo- or mono-clonal distribution of the infected cells. The same mechanism applies to BLV that induces leukemia or lymphoma in a minor fraction of infected animals. BLV also induces a persistent lymphocytosis in about 40% of the infected cows whereas this is far less common in HTLV-1 infected individuals. In fact, BLV and HTLV-1 are related deltaretroviruses sharing a similar genomic organization and infecting cells of the hematopoietic system (respectively mainly B cells and T cells). In this respect, BLV infection might be useful to address important questions unanswered in the HTLV-1 system. In particular, early infection appears to be critical at determining the different populations of cells (i.e. clones) that will subsequently thrive and expand during pathogenesis [10,11,17]. The mechanisms undergoing during this initial period of HTLV-1 infection cannot be addressed because of lack of samples (e.g. when transmission occurs via breast feeding) or due to the absence of systematic screening of populations at risk. In this report, we aimed at identifying the clones created during primary infection (i.e. position of the proviral insertion site in the host genome) and quantifying their abundance (i.e. number of cells per clone). With this objective, we performed high throughput sequencing of proviral insertion sites in BLV-inoculated cows during the early days of infection.

We first show that thousands of clones are generated during primary infection indicating that initial replication occurs through the infectious cycle. None of the newly generated clones massively expanded during this early period revealing a polyclonal pattern of viral expansion and confirming observations made in the BLV-infected sheep model [11] and in squirrel monkey experimentally inoculated with HTLV-1 [10]. A second contribution of this report concerns the genomic sequence of the insertion sites that are strikingly similar in the BLV and HTLV-1 systems. Thirdly, BLV provirus insertion occurs nearby transcribed genomic regions, as reported for HTLV-1 [16,18,25]. In fact, most (if not all) retroviruses target a characteristic weak palindromic consensus nucleotide sequence at the site of integration [26,27] and favor insertion within a particular genomic environment [28–30]. These preferences originate from the structure of the pre-integration complex given by the viral integrase [31,32], its possible cellular

partners [33] and from the relative accessibility of the genome [30]. Taken together, our observations thus reinforce the similarities between BLV and HTLV-1.

Therefore, the BLV model can be informative to understand mechanisms of early infection by deltaretroviruses, a period that cannot be easily addressed in the HTLV-1 system as discussed previously due to the logistic and ethical problems of collecting sequential neonatal samples. In BLV-inoculated sheep, the time period delimiting the infectious cycle was estimated to last the first 8 months [11]. Sheep are not natural hosts for BLV and develop leukemia/lymphoma at higher frequencies after shorter latency periods compared to cattle. In our report, we now show that this period generally lasts about 2 months in the bovine species. It is possible that a shorter infectious cycle period also limits the probability of leukemia/lymphoma in the natural host. Indeed, oncogenesis, which is a rare event that occurs in a single or in a limited number of cells, correlates with the number of infected cells. Even though we showed that BLV propagation rapidly shifts to clonal expansion to maintain the bulk of the infected cells, we cannot exclude that infection of new cells by virions has totally been stopped. Because analysis is based on the detection of proviral insertion sites in PBMCs, intermittent burst of virus replication with clearance of the newly infected cells and/or replication in tissues other than the peripheral blood may occur. Alternatively, it is also possible that the infectious cycle is ongoing permanently but that the immune response is sufficiently efficient to destroy all newly infected cells. Also, our method is limited by the number of infected cells present in the collected blood sample and so, the clonality analysis is constrained by the representative sample, making difficult the interpretation of clonal succession data from low proviral load animals such as #21.

Currently, there are no available data pertaining to the early viral replication in HTLV-1 infected patients. Based on this study and previous publications using animal models [10,11], it is likely that the HTLV-1 infectious cycle is also mainly restricted to the first months post-infection. This prediction has important clinical applications in terms of viral transmission. Since the infectious cycle requires production of virions and reverse transcription, anti-retroviral treatments could be instrumental to avoid mother-to-child transmission during delivery or breast feeding. In several endemic regions, breastfeeding cannot be prohibited because of societal and sanitary reasons. In particular, formula milk requires access to clean water and does not provide passive immunity to local pathogens. Based on this report, we propose that an early and short anti-retroviral treatment soon after birth might reduce HTLV-1 transmission and even possibly limit the long term HTLV-1 load. This option has recently been tested in another retrovirus, HIV that propagates preferentially via neo-infection of lymphocytes. Indeed, HIV-infected patients that initiated cART (combination antiretroviral therapy) during primary infection remained post-treatment controllers with very low viremia for many years [34].

A major contribution of our report is to demonstrate the massive depletion of the early generated clones. In Figure 3, we indeed show that the primary infection period is characterized by

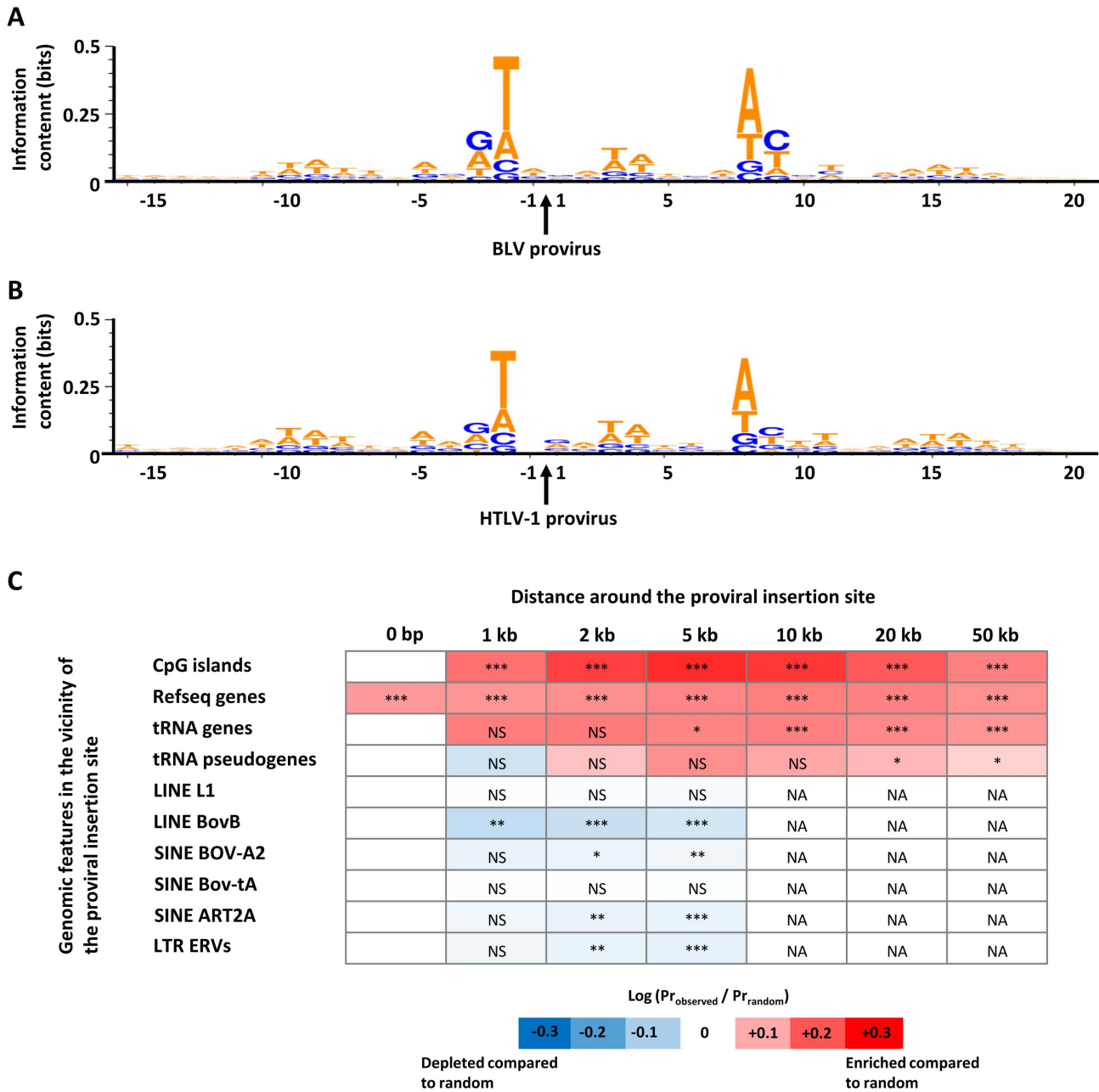


Figure 4. Genomic site preferences for insertion at seroconversion. **A.** Favored bovine genomic DNA sequence for BLV integration analyzed using WebLogo 3. The site of BLV integration in the target DNA sequence occurs between positions -1 and 1 with the provirus in forward orientation. Y-axis represents the information content at each target base position (perfect conservation would have a score of 2 bits). The height of the letter reflects the degree of base conservation. Data are representative for the all 5 animals. **B.** For comparison purposes, the favored human genomic DNA sequence for HTLV-1 integration is shown below. Human genomic sequences surrounding the HTLV-1 insertion sites are obtained from Gillet NA, Cook L, Laydon DJ, Hlela C, Verdonck K, et al. (2013) Strongyloidiasis and Infective Dermatitis Alter Human T Lymphotropic Virus-1 Clonality in vivo. PLoS Pathog 9(4): e1003263. doi:10.1371/journal.ppat.1003263. **C.** Genomic heat map of integration preferences at the time of seroconversion compared to the random distribution generated *in silico*. Genomic features (promoter marks, genes and interspersed repeats) analyzed are shown at the left of the heat map. For each genomic feature, the integration frequency was examined over several genomic length intervals around the proviral insertion site as shown at the top of the heat map. A colored scale is shown along the bottom of the panel with increasing shades of blue indicating depletion compared to random and increasing shades of red indicating enrichment. P-values calculated by Chi-squared test show significance of the departure from the random distribution (NS for non significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; NA for non applicable). Because of the high number of interspersed repeats in the genome, all proviruses lie nearby 10 kb of such repeats and enrichment/depletion is no longer applicable (NA in the heat map). Data are representative for the all 5 animals. doi:10.1371/journal.ppat.1003687.g004

the disappearance of the vast majority of clones. Considering the current knowledge in the BLV/HTLV-1 systems [35,36], it is likely that this clonal selection is due to counter selection by the

host immune response. Importantly, the clones carrying a provirus in a transcribed genomic environment were even more susceptible to be destroyed. We speculate that the stronger selection against

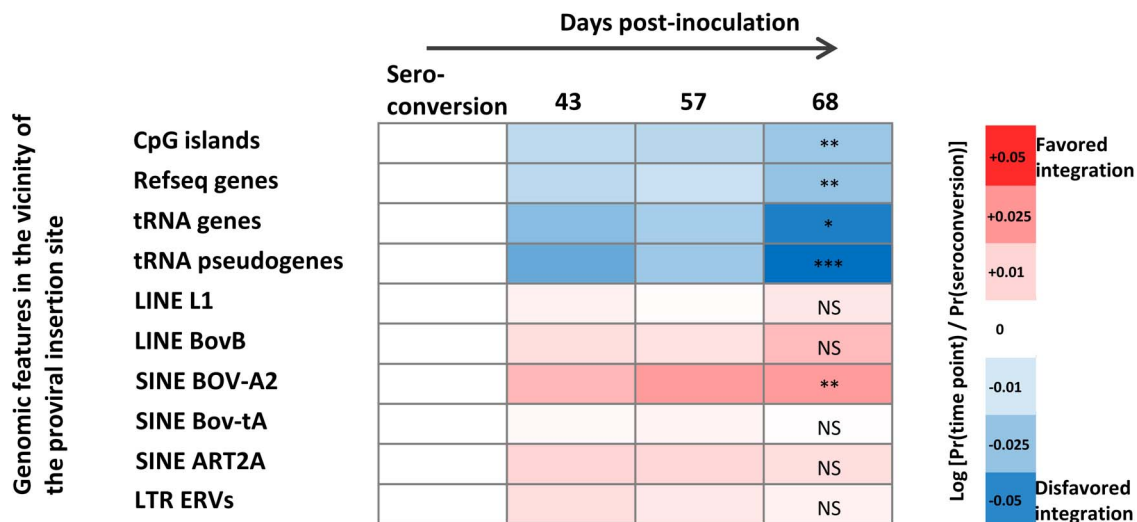


Figure 5. Evolution of provirus environment during primary infection. Genomic heat map of integration preferences compared to the distribution observed at the time of seroconversion. Genomic features analyzed are shown to the left of the heat map. Gene and CpG proximity was calculated with a genomic interval of 50 kb. An interval of 5 kb was used for interspersed repeats. For each genomic feature, the integration frequency was examined at different time points after the day of inoculation as shown at the top of the heat map. A colored scale is shown along the right hand side of the panel with increasing shades of blue indicating depletion compared to the distribution at the seroconversion and increasing shades of red indicating enrichment. In the column at day 68, p-values are calculated by Chi-squared test for trend and show significance of the increasing or decreasing trend from the seroconversion date to day 68 post-inoculation. Data are representative for the all 5 animals. doi:10.1371/journal.ppat.1003687.g005

proviruses located nearby genes might be the result of an increased viral expression and a higher exposure to the host immune system [37]. In this context, HDAC inhibitors have been shown to induce BLV [38,39] and HTLV-1 expression [40–42] and consequently increase the exposure of the infected cells to the host immune system. Remarkably, treatment of STLV-1 infected monkeys with a combination of the HDAC inhibitor valproate and the reverse transcriptase inhibitor AZT induced a persistent decrease of the PVL [43]. Thus, one might consider testing such combination of an antiretroviral drug and an HDAC inhibitor in the early days of the infection.

It has been previously shown that HTLV-1 infected clones are not equal regarding their proliferative potential. Indeed, the abundance of a given clone is enhanced by the integration of its provirus in an actively transcribed area of the genome [16,18]. We demonstrated here that this mechanism is also present in BLV-infected cows where clones that thrive and proliferate in the long term carry a provirus in a transcribed environment. We should recall that transcriptional activity of the genomic environment of the provirus is inferred from genomic features like CpG islands and genes and so, it is not a direct experimental measurement of the transcription. It thus appears that two opposite forces will act during primary infection and dictate the fate of the long term clonal composition: (1) BLV initially favors integration into genes or promoters and (2) host negative selection disfavors proviruses located next to transcribed regions. The outcome of these two forces will determine the PVL set point value as clonal abundance will benefit from carrying a provirus in transcribed regions but will concomitantly also be reduced by the immune response. Infected hosts able to more efficiently eliminate clones carrying a provirus integrated nearby transcribed units will thus have lower PVLs. Differences in immune response efficiencies have been identified in BLV-infected animals as well as in HTLV-1 infected individuals [44–47]. These differences translate into disease susceptibilities and were related at least partially to the genotype of the host [48–53]. Our preliminary data show that negative selection against

clones carrying a provirus nearby transcribed regions is more efficient in animals that will subsequently present low proviral load set point (data not shown).

At a first glance, the observation that clone abundance positively correlates with provirus proximity with transcribed units appears to conflict with the extremely low levels of viral expression measured in vivo [54,55]. Indeed, structural viral transcripts from peripheral blood lymphocytes or tumors can only be amplified by the means of very sensitive techniques such as in situ hybridization or RT-PCR. With progression from persistent lymphocytosis to leukemia, BLV expression levels even tend to decrease [56]. In tumors, BLV proviruses can be completely silent regarding viral messenger RNAs albeit carrying an intact genomic sequence [57]. Similarly, 5'-LTR directed transcription of HTLV-1 can be completely abrogated in leukemic clones by methylation and by deletion or mutation of viral genes [58–60]. Nevertheless, proviral silencing of BLV and HTLV-1 appeared to be selective, leaving some specific transcript untouched. This is the case for HBZ (HTLV-1 b-ZIP factor) that remains strongly expressed in ATLL cells [61]. Recently, BLV has been shown to produce microRNAs that are also actively transcribed in BLV malignant cells [62]. Mechanistically, the silencing of BLV structural genes can be achieved by repressive histone marks deposited onto the promoter [63] without interfering with pol III dependent transcription of the BLV microRNAs [62,64]. Consistently, we detected constitutive expression of BLV-miR-B4-3p despite very low levels of structural viral transcripts (data not shown). In this context, it is important to note that BLV favors pol III transcribed regions (tRNA genes and pseudogenes) upon insertion (Figure 4C). Furthermore, abundance of these established clones benefits from a provirus inserted in pol III transcribed regions (Figure 6). We might speculate that this genomic environment could in turn favor BLV microRNA expression. We propose that the abundance of successful clones benefit from the transcriptional activity of the genomic region surrounding the provirus to maintain robust expression of BLV microRNAs. Alternatively, BLV insertion nearby pol III regulated

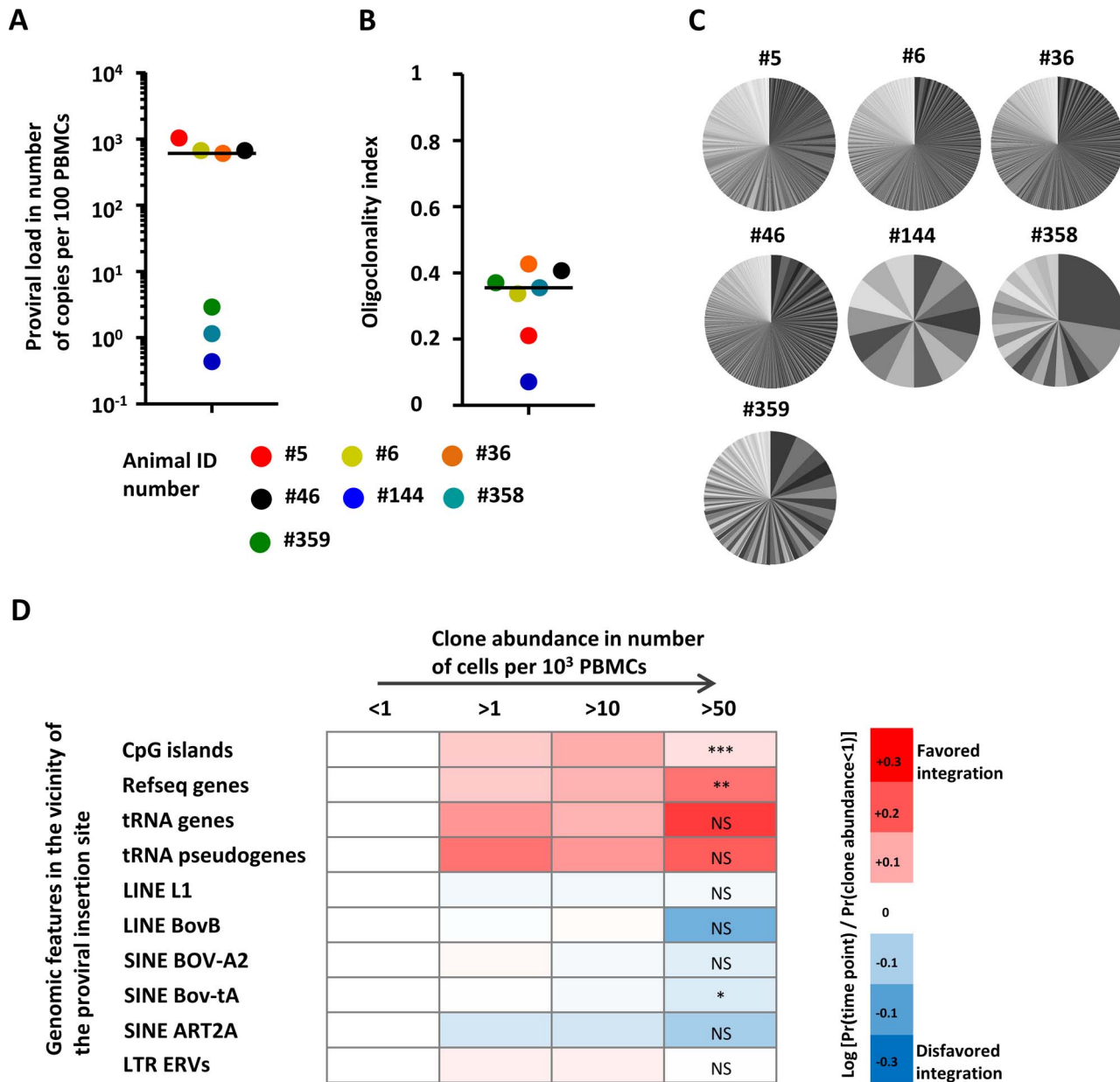


Figure 6. Evolution of provirus environment of clones with increasing abundance in long term infected animals. **A.** Proviral load in long-term infected animals (i.e. infected for more than 2 years). Median is indicated by the horizontal black line. **B.** Oligoclonality index in the long-term infected animals. Median is indicated by the horizontal black line. **C.** The pie-charts depict the distribution of clone abundance for each animal. The size of each slice in the pie-charts represents the relative abundance of each BLV-infected clone. Note that although the gray tones are identical between different animals, the clones are different. **D.** Genomic heat map of integration preferences compared to the distribution observed in the lowest abundance clones (clone abundance less than 1 cell per 1000 PBMCs). Genomic features analyzed are shown to the left of the heat map. Gene and CpG proximity was calculated with a genomic interval of 50 kb. An interval of 5 kb was used for interspersed repeats. For each genomic feature, the integration frequency was examined for clones having a given abundance as shown at the top of the heat map. A colored scale is shown along the right hand side of the panel with increasing shades of blue indicating depletion compared to the distribution of the lowest abundance clones and increasing shades of red indicating enrichment. In the column at the highest clone abundance (clone abundance more than 50 cells per 1000 PBMCs), p-values are calculated by Chi-squared test for trend and show significance of the increasing or decreasing trend from the smallest clones. Data are representative for all 7 animals. doi:10.1371/journal.ppat.1003687.g006

genes such as tRNAs, 5S rRNA and other small RNAs may affect their activity and modify the cell metabolism. In particular, pol III directed transcription of housekeeping genes is a key parameter of cell growth and replication.

In summary, we have characterized BLV clonality during primary infection using high throughput sequencing of the

proviral insertion sites in 7 sequential samples of 5 BLV-inoculated cows and as well as in long term infected animals. We demonstrate that BLV proviruses initially integrate into transcribed regions of the genome but are massively depleted later on. This mechanism may have important outcomes for HTLV-1 prevention and treatment.

Materials and Methods

Experimental setting of BLV infection

Cows were experimentally infected with the wild-type BLV strain 344. Two 15 cm-diameter dishes containing subconfluent HeLa cells were transfected with 80 µg of plasmid pBLV344 containing the BLV provirus, recovered in 5 ml PBS at day 3 and injected subcutaneously. The presence of anti-BLV antibodies was determined using a competitive ELISA test (IDEXX Leukosis Blocking Ab Test). Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll Hypaque density gradient centrifugation (Sigma-Aldrich), washed and cryopreserved in fetal calf serum with 10% DMSO (Sigma-Aldrich).

Ethics statement

Animal experimentations were conducted in accordance with national and international guidelines for animal care and use described in the Manual for use and care of experimental animals emitted by INTA. Handling of cows and experimental procedures were reviewed and approved by INTA's Institutional Committee for Care and Use of Experimental Animals (CICUAE-INTA) under protocol number 35/2010.

BLV proviral load measurement

DNA was extracted from PBMCs using DNeasy Blood and Tissue kit (Qiagen). BLV DNA was PCR amplified using pol gene sequence-specific primers 5'-GAAACTCCAGAGCAATGGCA-TAA-3' and 5'-GGTTCCGGCCATCGAGACA-3'. As reference for genomic DNA, β-actin was amplified with oligonucleotides 5'-TCCCTGGAGAAGAGCTACGA-3' and 5'-GGCAGACT-TAGCCTCCAGTG-3'. Three dilutions of DNA (100 ng, 33 ng and 11 ng) were amplified by real-time quantitative PCR in a Roche light cycler using MESA green master mix (Eurogentec). The thermal protocol used started with a 95°C 5 min denaturation step; then 45 cycles as follows (95°C 15 sec, 60°C 20 sec, 72°C 40 sec) and terminated with a melting curve. PCR efficacies were calculated for each sample using the three dilutions. Standard curves were generated using PCR4topo vectors (Life Technologies) containing the corresponding pol or actin amplicon. Proviral load was calculated, as an average of the three dilutions, from the number of proviral copies divided by half of the number of actin copies and expressed as number of proviral copies per 100 of PBMCs.

Selective amplification and quantification of BLV proviral insertion sites

Ten micrograms of genomic DNA extracted from PBMCs were sheared by sonication with the *Diagenode* Bioruptor instrument using the following protocol (15 sec ON, 90 sec OFF, 4 cycles in a 4°C water bath). DNA ends were then end-repaired using 15 units of T4 DNA polymerase (New England Biolabs), 5 units of DNA polymerase I Klenow fragment (NEB), 50 units of T4 polynucleotide kinase (NEB) and 0.8 mM of dNTP (Sigma) in T4 DNA ligase buffer (NEB) at 20°C during 30 min. DNA was then cleaned using a Qiaquick PCR purification kit (Qiagen) and eluted in 64 µl of EB buffer. Addition of an adenosine at the 3' ends of the DNA was performed by adding 0.2 mM of dATP (Sigma) and 15 units of Klenow Fragment 3' to 5' exo- (NEB) in NEB2 buffer (NEB) at 37°C for 30 min. DNA was then cleaned using a Qiaquick PCR purification kit and eluted in 40 µl of EB. One hundred pmol of a partially double stranded DNA linker was ligated to the DNA ends using a Quick ligation kit (NEB). Twenty four different linkers were designed, each one with a specific 8 bp tag (see Primer list in Text S1) to allow multiplexing of DNA samples during the

sequencing. DNA was cleaned using a Qiaquick PCR purification kit and eluted in 60 µl of EB. The 60 µl of ligated product was then split into 3 aliquots of 20 µl and each aliquot was used in a separate PCR1 reaction (see Figure S1). For each PCR reaction, 20 µl of ligated product was mixed with 0.2 mM of dNTP (Sigma), 50 pmol of BLV_LMP1 primer (binds BLV LTR), 10 pmol of VU primer (which anneals to the strand of the Vectorette Unit linker generated by the amplification from BLV_LMP1), 1 unit of Phusion DNA polymerase in High Fidelity buffer (Finnzyme, NEB). The following thermal protocol was used: denaturation for 30 sec at 98°C; then 30 cycles (5 sec at 98°C, 10 sec at 62°C, 30 sec at 72°C); followed by 10 min at 72°C; and finally cooled at 4°C. The 3 PCR1 products, derived from the same sample were then pooled. The DNA was cleaned using a Qiaquick PCR purification kit and eluted in 150 µl of EB. To perform PCR2, 1 µl of the cleaned PCR1 product was mixed with 0.2 mM of dNTP (Sigma), 25 pmol of P5_BLV_LMP2 primer binding the BLV LTR, 25 pmol of P7 primer binding the linker, 1 unit of Phusion DNA polymerase in High Fidelity buffer (NEB). The following thermal protocol was used: 98°C for 30 sec; 30 cycles (5 sec at 98°C, 10 sec at 62°C, 30 sec at 72°C); 10 min at 72°C; 4°C until user stops. DNA was then cleaned using a Qiaquick PCR purification kit and eluted in 50 µl of EB. A library was constructed by pooling the different PCR2 products (each one possessing a specific tag). Quantification of the libraries was made by qPCR using primers P5 and P7 (see Primer list in Text S1) and a MESA green master mix (Eurogentec) in a Roche light cycler 480 instrument. Three dilutions of the library (200 pg, 66 pg and 22 pg) were amplified. Standard curves were generated using a library quantified on a titration flow cell previously run on a Genome Analyzer II (Illumina). Stock libraries were diluted down to 8 pM and clustered on the flow cell. Paired-end reads (read1 and read2 each 50 bp) plus a 8 bp tag read (read 3) were acquired on a GA II or a Myseq Illumina instrument. Read 1 and read 2 were mapped against the bovine genome (build *Bos_taurus_UMD_3.1/bosTau6*) and the proviral insertion site and the shear site were deduced.

“Sister cells” are cells where the BLV provirus is inserted at the same site in the cellular genome and a “clone” is a population of sister cells. For each unique insertion site, the number of amplicons of different length (i.e. different shear sites) give an estimate of the number of sister cells of that infected clone [65]. The absolute abundance of a given clone *i* (number of cells per 100 PBMCs) was calculated from the number of sister cells and the measurement of the proviral load as follows:

$$A_{abs}(i) = \frac{X_i}{\sum_{i=1}^D X_i} * PVL$$

where X_i is the number of sister cells of the *i*th clone, D the number of observed clones and PVL the proviral load.

The relative abundance of a given clone *i* (in percent of the proviral load) was expressed as follows:

$$A_{rel}(i) = \frac{X_i}{\sum_{i=1}^D X_i} * 100$$

Oligoclonality index

To measure the clonality of the infected cell population, i.e. the non-uniformity of the clone abundance distribution, we used the oligoclonality index based on the Gini coefficient [66]. Oligoclon-

ality index ranges from 0 (all the BLV-infected clones having the same abundance, i.e. perfect polyclonality) to 1 (only one BLV-infected clone constitutes the total proviral load, i.e. perfect monoclonality). Details of the calculations are given in supplemental file Text S1 (Supplemental Materials and Methods).

Similarity index

To measure the likeness or overlap between two populations of BLV-infected cells, we calculated a similarity index. This number ranges from 0 to 1, with 0 indicating that no clones are shared between the two populations and 1 corresponding to a complete identity (all the clones present in population 1 were also present in population 2 and *vice versa*). Because this index takes clone abundance into account, populations that contain the same clones but have different clone abundance will have an index value of less than 1. Similarity indexes were calculated between 2 successive time points separated by two weeks. Details of the calculations are given in supplemental file Text S1 (Supplemental Materials and Methods).

Genetic environment around the proviral insertion site and generation of random sites

Favored genomic DNA sequences for BLV or HTLV-1 integration were represented using WebLogo 3 [67]. Genomic annotations flanking the proviral insertion sites were retrieved using Galaxy [68] which is a web-based genome analysis tool. DNA sequences (read 1 and read 2 like sequences) from 100,000 random sites in the bovine genome were generated using Galaxy

and back-aligned to the bovine genome using the same pipeline to eliminate any potential bias due to alignment limitations.

Statistics

Statistical tests were performed using GraphPad Prism, Microsoft Excel and R softwares. The symbol *** was used when $p < 0.001$, ** when $p < 0.01$, * when $p < 0.05$, NS (Non Significant) when $p > 0.05$ and NA for Non Applicable.

Supporting Information

Text S1 The supporting file Text S1 contains supporting tables S1 and S2, supporting material and methods and supporting figure S1.
(PDF)

Acknowledgments

We thank the GIGA geno-transcriptomic platform and more particularly Benoit Hennuy and Wouter Coppieters for their help and skills for the high-throughput sequencing and mapping. We thank Dr Lucy Cook (Imperial College London) for her careful editing of the manuscript.

Author Contributions

Conceived and designed the experiments: NAG KT LW. Performed the experiments: NAG GG SMR AdB NR. Analyzed the data: NAG. Contributed reagents/materials/analysis tools: GG IA KT. Wrote the paper: NAG LW. Critically revised the manuscript and approved the final submitted version: NAG GG SMR AdB NR IA KT LW.

References

- Peeters M, Delaporte E (2012) Simian retroviruses in African apes. *Clin Microbiol Infect* 18: 514–520.
- Cook LB, Elemans M, Rowan AG, Asquith B (2013) HTLV-1: persistence and pathogenesis. *Virology* 435: 131–140.
- Gillet N, Florins A, Boxus M, Burteau C, Nigro A, et al. (2007) Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in human. *Retrovirology* 4: 18.
- Matsuoka M, Jeang KT (2011) Human T-cell leukemia virus type 1 (HTLV-1) and leukemic transformation: viral infectivity, Tax, HBZ and therapy. *Oncogene* 30: 1379–1389.
- Pais-Correia AM, Sachse M, Guadagnini S, Robbiati V, Lasserre R, et al. (2010) Biofilm-like extracellular viral assemblies mediate HTLV-1 cell-to-cell transmission at virological synapses. *Nat Med* 16: 83–89.
- Nejmehdine M, Bangham CR (2010) The HTLV-1 Virological Synapse. *Viruses* 2: 1427–1447.
- Jones KS, Lambert S, Bouttier M, Benit L, Ruscetti FW, et al. (2011) Molecular aspects of HTLV-1 entry: functional domains of the HTLV-1 surface subunit (SU) and their relationships to the entry receptors. *Viruses* 3: 794–810.
- Lairmore MD, Anupam R, Bowden N, Haines R, Haynes RA, 2nd, et al. (2011) Molecular determinants of human T-lymphotropic virus type 1 transmission and spread. *Viruses* 3: 1131–1165.
- Edwards D, Fenizia C, Gold H, de Castro-Amarante MF, Buchmann C, et al. (2011) Orf-I and orf-II-encoded proteins in HTLV-1 infection and persistence. *Viruses* 3: 861–885.
- Mortreux F, Kazanji M, Gabet AS, de Thoisy B, Wattel E (2001) Two-step nature of human T-cell leukemia virus type 1 replication in experimentally infected squirrel monkeys (*Saimiri sciureus*). *J Virol* 75: 1083–1089.
- Pomier C, Alcaraz MT, Debacq C, Lancon A, Kerkhofs P, et al. (2008) Early and transient reverse transcription during primary deltaretroviral infection of sheep. *Retrovirology* 5: 16.
- Cavrois M, Wain-Hobson S, Gessain A, Plumelle Y, Wattel E (1996) Adult T-cell leukemia/lymphoma on a background of clonally expanding human T-cell leukemia virus type-1-positive cells. *Blood* 88: 4646–4650.
- Etoh K, Tamiya S, Yamaguchi K, Okayama A, Tsubouchi H, et al. (1997) Persistent clonal proliferation of human T-lymphotropic virus type I-infected cells in vivo. *Cancer Res* 57: 4862–4867.
- Cavrois M, Leclercq I, Gout O, Gessain A, Wain-Hobson S, et al. (1998) Persistent oligoclonal expansion of human T-cell leukemia virus type 1-infected circulating cells in patients with Tropical spastic paraparesis/HTLV-1 associated myelopathy. *Oncogene* 17: 77–82.
- Gabet AS, Gessain A, Wattel E (2003) High simian T-cell leukemia virus type 1 proviral loads combined with genetic stability as a result of cell-associated provirus replication in naturally infected, asymptomatic monkeys. *Int J Cancer* 107: 74–83.
- Gillet NA, Malani N, Melamed A, Gormley N, Carter R, et al. (2011) The host genomic environment of the provirus determines the abundance of HTLV-1-infected T-cell clones. *Blood* 117: 3113–3122.
- Moules V, Pomier C, Sibon D, Gabet AS, Reichert M, et al. (2005) Fate of premalignant clones during the asymptomatic phase preceding lymphoid malignancy. *Cancer Res* 65: 1234–1243.
- Melamed A, Laydon D, Gillet N, Tanaka Y, Taylor G, et al. (2013) Genome-wide determinants of proviral targeting, clonal abundance and expression in natural HTLV-1 Infection. *PLoS Pathogens* 9: e1003271. doi:10.1371/journal.ppat.1003271.
- Iwanaga M, Watanabe T, Utsunomiya A, Okayama A, Uchimaru K, et al. (2010) Human T-cell leukemia virus type I (HTLV-1) proviral load and disease progression in asymptomatic HTLV-1 carriers: a nationwide prospective study in Japan. *Blood* 116: 1211–1219.
- Gutierrez G, Carignano H, Alvarez I, Martinez C, Porta N, et al. (2012) Bovine leukemia virus p24 antibodies reflect blood proviral load. *BMC Vet Res* 8: 187.
- Gutierrez G, Alvarez I, Politzki R, Lomonaco M, Dus Santos MJ, et al. (2011) Natural progression of Bovine Leukemia Virus infection in Argentinean dairy cattle. *Vet Microbiol* 151: 255–263.
- Gillet N, Cook L, Laydon D, Hlela C, Verdonck K, et al. (2013) Strongyloidiasis and Infective Dermatitis Alter Human T Lymphotropic Virus-1 Clonality in vivo. *PLoS Pathogens* 9: e1003263. doi:10.1371/journal.ppat.1003263.
- Larsen F, Gundersen G, Lopez R, Prydz H (1992) CpG islands as gene markers in the human genome. *Genomics* 13: 1095–1107.
- Moqtaderi Z, Wang J, Raha D, White RJ, Snyder M, et al. (2010) Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat Struct Mol Biol* 17: 635–640.
- Meekings KN, Leipzig J, Bushman FD, Taylor GP, Bangham CR (2008) HTLV-1 integration into transcriptionally active genomic regions is associated with proviral expression and with HAM/TSP. *PLoS Pathog* 4: e1000027.
- Wu X, Li Y, Crise B, Burgess SM, Munroe DJ (2005) Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *J Virol* 79: 5211–5214.
- Holman AG, Coffin JM (2005) Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *Proc Natl Acad Sci U S A* 102: 6103–6107.
- Wu X, Li Y, Crise B, Burgess SM (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300: 1749–1751.

29. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, et al. (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2: E234.
30. Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD (2007) HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res* 17: 1186–1194.
31. Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, et al. (2006) Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog* 2: e60.
32. Derse D, Crise B, Li Y, Prinler G, Lum N, et al. (2007) Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *J Virol* 81: 6731–6741.
33. Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, et al. (2005) A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med* 11: 1287–1289.
34. Saez-Cirion A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, et al. (2013) Post-Treatment HIV-1 Controllers with a Long-Term Virological Remission after the Interruption of Early Initiated Antiretroviral Therapy ANRS VISCONTI Study. *PLoS Pathog* 9: e1003211.
35. Bangham CR (2009) CTL quality and the control of human retroviral infections. *Eur J Immunol* 39: 1700–1712.
36. Florins A, Gillet N, Asquith B, Boxus M, Burteau C, et al. (2007) Cell dynamics and immune response to BLV infection: a unifying model. *Front Biosci* 12: 1520–1531.
37. Florins A, de Brogniez A, Elemans M, Bouzar AB, Francois C, et al. (2012) Viral expression directs the fate of B cells in bovine leukemia virus-infected sheep. *J Virol* 86: 621–624.
38. Merezak C, Reichert M, Van Lint C, Kerkhofs P, Portetelle D, et al. (2002) Inhibition of histone deacetylases induces bovine leukemia virus expression in vitro and in vivo. *J Virol* 76: 5034–5042.
39. Achachi A, Florins A, Gillet N, Debacq C, Urbain P, et al. (2005) Valproate activates bovine leukemia virus gene expression, triggers apoptosis, and induces leukemia/lymphoma regression in vivo. *Proc Natl Acad Sci U S A* 102: 10309–10314.
40. Mosley AJ, Meekings KN, McCarthy C, Shepherd D, Cerundolo V, et al. (2006) Histone deacetylase inhibitors increase virus gene expression but decrease CD8+ cell antiviral function in HTLV-1 infection. *Blood* 108: 3801–3807.
41. Lezin A, Gillet N, Olindo S, Signate A, Grandvaux N, et al. (2007) Histone deacetylase mediated transcriptional activation reduces proviral loads in HTLV-1 associated myelopathy/tropical spastic paraparesis patients. *Blood* 110: 3722–3728.
42. Olindo S, Belrose G, Gillet N, Rodriguez S, Boxus M, et al. (2011) Safety of long-term treatment of HAM/TSP patients with valproic acid. *Blood* 118: 6306–6309.
43. Afonso PV, Mekaouche M, Mortreux F, Toulza F, Moriceau A, et al. (2010) Highly active antiretroviral treatment against STLV-1 infection combining reverse transcriptase and HDAC inhibitors. *Blood* 116: 3802–3808.
44. Meiom R, Moss S, Brenner J, Heller D, Trainin Z (1997) Levels and role of cytokines in bovine leukemia virus (BLV) infection. *Leukemia* 11 Suppl 3: 219–220.
45. Asquith B, Mosley AJ, Barfield A, Marshall SE, Heaps A, et al. (2005) A functional CD8+ cell assay reveals individual variation in CD8+ cell antiviral efficacy and explains differences in human T-lymphotropic virus type 1 proviral load. *J Gen Virol* 86: 1515–1523.
46. Kattan T, MacNamara A, Rowan AG, Nose H, Mosley AJ, et al. (2009) The avidity and lytic efficiency of the CTL response to HTLV-1. *J Immunol* 182: 5723–5729.
47. Hilburn S, Rowan A, Demontis MA, MacNamara A, Asquith B, et al. (2011) In vivo expression of human T-lymphotropic virus type 1 basic leucine-zipper protein generates specific CD8+ and CD4+ T-lymphocyte responses that correlate with clinical outcome. *J Infect Dis* 203: 529–536.
48. Jeffery KJ, Usuku K, Hall SE, Matsumoto W, Taylor GP, et al. (1999) HLA alleles determine human T-lymphotropic virus-I (HTLV-I) proviral load and the risk of HTLV-I-associated myelopathy. *Proc Natl Acad Sci U S A* 96: 3848–3853.
49. Jeffery KJ, Siddiqui AA, Bunce M, Lloyd AL, Vine AM, et al. (2000) The influence of HLA class I alleles and heterozygosity on the outcome of human T cell lymphotropic virus type I infection. *J Immunol* 165: 7278–7284.
50. Macnamara A, Rowan A, Hilburn S, Kadolsky U, Fujiwara H, et al. (2010) HLA class I binding of HBZ determines outcome in HTLV-1 infection. *PLoS Pathog* 6: e1001117.
51. Seich Al Basatena NK, Macnamara A, Vine AM, Thio CL, Astemborski J, et al. (2011) KIR2DL2 enhances protective and detrimental HLA class I-mediated immunity in chronic viral infection. *PLoS Pathog* 7: e1002270.
52. Nagaoka Y, Kabeya H, Onuma M, Kasai N, Okada K, et al. (1999) Ovine MHC class II DRB1 alleles associated with resistance or susceptibility to development of bovine leukemia virus-induced ovine lymphoma. *Cancer Res* 59: 975–981.
53. Juliarena MA, Poli M, Sala L, Ceriani C, Gutierrez S, et al. (2008) Association of BLV infection profiles with alleles of the BoLA-DRB3.2 gene. *Anim Genet* 39: 432–438.
54. Jensen WA, Rovnak J, Cockerell GL (1991) In vivo transcription of the bovine leukemia virus tax/rex region in normal and neoplastic lymphocytes of cattle and sheep. *J Virol* 65: 2484–2490.
55. Kinoshita T, Shimoyama M, Tobinai K, Ito M, Ito S, et al. (1989) Detection of mRNA for the tax1/rex1 gene of human T-cell leukemia virus type I in fresh peripheral blood mononuclear cells of adult T-cell leukemia patients and viral carriers by using the polymerase chain reaction. *Proc Natl Acad Sci U S A* 86: 5620–5624.
56. Merimi M, Klener P, Szyal M, Cleuter Y, Bagnis C, et al. (2007) Complete suppression of viral gene expression is associated with the onset and progression of lymphoid malignancy: observations in Bovine Leukemia Virus-infected sheep. *Retrovirology* 4: 51.
57. Van den Broeke A, Cleuter Y, Chen G, Portetelle D, Mammerickx M, et al. (1988) Even transcriptionally competent proviruses are silent in bovine leukemia virus-induced sheep tumor cells. *Proc Natl Acad Sci U S A* 85: 9263–9267.
58. Tamiya S, Matsuoka M, Etoh K, Watanabe T, Kamihira S, et al. (1996) Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia. *Blood* 88: 3065–3073.
59. Takeda S, Maeda M, Morikawa S, Taniguchi Y, Yasunaga J, et al. (2004) Genetic and epigenetic inactivation of tax gene in adult T-cell leukemia cells. *Int J Cancer* 109: 559–567.
60. Miyazaki M, Yasunaga J, Taniguchi Y, Tamiya S, Nakahata T, et al. (2007) Preferential selection of human T-cell leukemia virus type I provirus lacking the 5' long terminal repeat during oncogenesis. *J Virol* 81: 5714–5723.
61. Satou Y, Yasunaga J, Yoshida M, Matsuoka M (2006) HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc Natl Acad Sci U S A* 103: 720–725.
62. Rosewick N, Momont M, Durkin K, Takeda H, Caiment F, et al. (2013) Deep sequencing reveals abundant noncanonical retroviral microRNAs in B-cell leukemia/lymphoma. *Proc Natl Acad Sci U S A* 110: 2306–2311.
63. Merimi M, Klener P, Szyal M, Cleuter Y, Kerkhofs P, et al. (2007) Suppression of viral gene expression in bovine leukemia virus-associated B-cell malignancy: interplay of epigenetic modifications leading to chromatin with a repressive histone code. *J Virol* 81: 5929–5939.
64. Kincaid RP, Burke JM, Sullivan CS (2012) RNA virus microRNA that mimics a B-cell oncomiR. *Proc Natl Acad Sci U S A* 109: 3077–3082.
65. Berry CC, Gillet NA, Melamed A, Gormley N, Bangham CR, et al. (2012) Estimating abundances of retroviral insertion sites from DNA fragment length data. *Bioinformatics* 28: 755–762.
66. Gini (1914) Sulla misura della concentrazione e della variabilita dei caratteri: Transactions of the Real Istituto Veneto di Scienze.
67. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188–1190.
68. Goecks J, Nekrutenko A, Taylor J (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11: R86.