Fate of Premalignant Clones during the Asymptomatic Phase Preceding Lymphoid Malignancy

Vincent Moulés,[1] Carole Pomier,[1] David Sibon,[1,2] Anne-Sophie Gabet,[1] Michal Reichert,[3] Pierre Kerkhofs,[4] Luc Willems,[5] Franck Mortreux,[1] and Eric Wattel[1,2]

[1]Oncovirologie et Biothérapies, UMR5537 CNRS-Université Claude Bernard, Centre Léon Bérard; [2]Service d'Hématologie, Pavillon E, Hôpital Edouard Herriot, Place d'Arsonval, Lyon, France; [3]Department of Pathology, National Veterinary Research Institute, Pulawy, Poland; [4]Department of Virology, Veterinary, and Agrochemical Research Centre, 1180 Uccle, Belgium; and [5]Faculté Universitaire des Sciences Agronomiques, Gembloux, Belgium

Abstract

Almost all cancers are preceded by a prolonged period of clinical latency during which a combination of cellular events helps move carcinogen-exposed cells towards a malignant phenotype. Hitherto, investigating the fate of premalignant cells *in vivo* remained strongly hampered by the fact that these cells are usually indistinguishable from their normal counterparts. Here, for the first time, we have designed a strategy able to reconstitute the replicative history of the bona fide premalignant clone in an animal model, the sheep experimentally infected with the lymphotropic bovine leukemia virus. We have shown that premalignant clones are early and clearly distinguished from other virus-exposed cells on the basis of their degree of clonal expansion and genetic instability. Detectable as early as 0.5 month after the beginning of virus exposure, premalignant cells displayed a two-step pattern of extensive clonal expansion together with a mutation load ~6 times higher than that of other virus-exposed cells that remained untransformed during the life span of investigated animals. There was no fixation of somatic mutations over time, suggesting that they regularly lead to cellular death, partly contributing to maintain a normal lymphocyte count during the prolonged premalignant stage. This equilibrium was finally broken after a period of 18.5 to 60 months of clinical latency, when a dramatic decrease in the genetic instability of premalignant cells coincided with a rapid increase in lymphocyte count and lymphoma onset. (Cancer Res 2005; 65(4): 1234-43)

Introduction

Almost all cancers are preceded by a prolonged period of clinical latency during which it is assumed that cellular, chromosomal, genetic, epigenetic, and molecular aberrations help move carcinogen-exposed cells towards a malignant phenotype (1). This premalignant phase could include clinical signs such as colonic polyps for some colonic adenocarcinomas, Fanconi anemia or myelodysplastic syndromes for some acute myelogenous leukemias, or actinic dermatosis for some nonmelanoma skin cancers. However, in many other cases, the prolonged period that precedes tumor onset is clinically latent. The presence of clone-specific markers can be detected at birth in patients who develop acute lymphoblastic or myelogenous leukemia after several years (2). This has been evidenced for t(8 ;21) (3) and t(12 ;21) (4) translocations as well as *IGH* and *TCR* gene rearrangements (5). Whether these abnormalities actually correspond to the premalignant clone or rather represent the signature of several candidate clones of which only one will become malignant remains unknown. In the case of tumors with identified carcinogen, a brief or prolonged carcinogen exposure triggers the growth of a large number of cells. However, among those, a unique clone will ultimately undergo malignant transformation. Getting an early access to the clone from which the tumor will develop (i.e., from the bona fide premalignant clone) would strongly help to understand malignant transformation. However, premalignant cells are usually indistinguishable from their normal counterparts, thereby ruling out the possibility to tackle the events governing early oncogenesis or leukemogenesis *in vivo*. By using a bovine leukemia virus (BLV) infectious molecular clone, it is possible to infect sheep that is not a natural host for this virus (6). Experimentally infected animals develop BLV-associated tumors after a 1- to 4-year period of latency. In the present report, we first show that the preleukemic phase of BLV infection includes the persistent clonal expansion of infected cells whereas malignant cells are characterized by their specific BLV flanking sequences. Using sensitive quantitative PCR-based methods that specifically amplify tumor and nontumor BLV integration sites, the clone from which the tumor had developed could be retrospectively monitored over time during the premalignant phase that preceded lymphoma onset. The fate of the premalignant clone (i.e., its duration and degree of clonal expansion together with its level of genetic instability) was compared over time with that of other BLV-positive clones, permitting for the first time to reconstitute the natural history of malignant lymphoma in an animal model. Together with the duration and the degree of clonal expansion, the level of genetic instability clearly discriminated premalignant cells from other virus-exposed cells that remained untransformed during the whole life span of investigated animals (i.e., that did not generate malignancy during the course of experimental infection). From these differences it was possible to depict the clonal history of premalignant cells during the latency period preceding tumor onset.

Materials and Methods

Animals and Experimental Infection. All sheep used in the study were maintained under controlled conditions at the Veterinary and Agrochemical Research Centre (Uccle, Belgium). Sheep were infected by direct inoculation of a cloned BLV provirus, as previously described (6). At regular time intervals, blood samples were collected on different occasions by jugular venipuncture. Total leukocyte counts were determined by using a Coulter

Note: V. Moulés and C. Pomier contributed equally to this work.

A-S. Gabet is currently in the Unit of Infection and Cancer, IARC, WHO, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France.

Requests for reprints: Eric Watel, Oncovirologie et Biothérapies, UMR5537 CNRS-Université Claude Bernard, Centre Léon Bérard, 28, rue Laënnec, 69373 Lyon cedex 08, France. Phone: 33-4-78-78-26-69; Fax: 33-4-78-78-27-17; E-mail: wattel@lyon.fnclcc.fr.

counter ZN, and the corresponding lymphocyte numbers were calculated from the complete blood cell count after examination under the microscope. In parallel, the corresponding sera were analyzed for BLV seropositivity using immunodiffusion and ELISA techniques. Furthermore, blood samples were mixed with EDTA, then stored at −80°C for further molecular experiments. The onset of overt tumor was defined as the time at which the animal developed clinical symptoms of lymphoma, mainly adenopathy.

Quantification of Bovine Leukemia Virus Circulating Proviral Load. To estimate the entire population of cells undergoing carcinogenesis resulting from BLV infection, we measured the circulating amount of BLV proviral sequences by quantitative PCR. Circulating and tumor BLV proviral loads were assessed by an accurate and reproducible quantitative LightCycler real time PCR using 0.3 μg of DNA from blood lymphocytes or tumor (lymphoma) diluted to a final volume of 20 μL. DNA was extracted with phenol/chloroform (1:1), then precipitated with 100% ethanol. The reaction mixture included polymerase (LightCycler Kit Fast Start DNA Master Hybridization Probes; Roche Diagnostics, Mannheim, Germany), 2 mmol/L $MgCl_2$, 500 nmol/L primer BLVQF (5′-TTGACGCTATGCCCAG-3′) positioned at 7,819 to 7,834, 500 nmol/L primer BLVQR (5′-AGTCTTAGTTTCCCGCT-3′) positioned at 8,114 to 8,098, 100 nmol/L donor probe 3′-end labeled with fluorescein (5′-GTGGTCCAGTCCTCAGGC-FL) positioned at 8,028 to 8,045, and 200 nmol/L acceptor probe 5′-end labeled with LC Red640 (5′-LC Red640-TACAACGCTTCCTCCATGACC-p) positioned at 8,048 to 8,069 [nucleotides are numbered according to the BLV GenBank reference: K02120 (7)]. Amplification conditions were 95°C for 8 minutes to activate the polymerase, followed by 55 cycles at 95°C for 10 seconds, 60°C for 10 seconds, and 72°C for 10 seconds. In order to standardize the amount of DNA subjected to quantification, we used the sheep β-globin gene as an internal standard, as described by Tajima et al. (8). The standard curve for β-globin was generated using DNA extracted from BLV-negative sheep blood cells. Circulating BLV proviral loads were plotted as a function of time. To estimate the fluctuation of BLV proviral loads during the asymptomatic phase of the infection, the proviral load corresponding to the last harvested sample having a normal lymphocyte count was divided by that of the first sample with detectable BLV sequences.

Semiquantitative Inverse PCR Amplification of Bovine Leukemia Virus Extremities and Their 3′ Flanking Sequences. BLV integration was analyzed by inverse PCR, as shown in Fig. 1. This method consists in amplifying the 3′ extremity of the proviruses together with their flanking sequences. Two micrograms of DNA were digested by 20 units of *Nla*III and 20 units of *Mfe*I in 1 × *Nla*III-*Mfe*I buffer for 3 hours at 37°C. *Mfe*I digestion was done to avoid the amplification of a 536 bp segment of the 5′ long terminal repeat complementary to the set of 3′ inverse PCR primers. The completion of the digestion was controlled by 1% agarose gel electrophoresis. DNA was extracted with phenol/chloroform (1:1), then precipitated with 100% ethanol. One microgram of digested DNA was circularized for 16 hours at 16°C with 20 units of T4 DNA ligase in 600 L of 1× T4 DNA ligase buffer and 1 mmol/L ATP, final concentration. As there is a stochastic component to the detection of retrovirus integration sites using inverse PCR (9), samples were analyzed in quadruplicate, as previously described for HTLV-1 (9): 4 × 500 ng of circularized DNA were amplified for 39 cycles using 200 mol/L of the primer pair BLV3′S 5′-GGCTAGAATCCCCG-TACCTC-3′ at position 8,278 to 8,297 and BLV3′AS 5′-GACGTCT-CTGTCTGGTTTACGG-3′ at position 8,245 to 8,224. Amplifications were done using 3.5 units of the Pfu DNA polymerase. Thermal cycling parameters were 95°C 10 minutes, 35 × (95°C 1 minute, 60°C 1 minute, 72°C 3 minutes), followed by a final elongation step of 10 minutes at 72°C.

Runoff Analysis of Amplified Products. The length polymorphism of 3′ or 5′ BLV flanking sequences was assessed by runoff analysis. This method consists in the linear PCR amplification of the provirus 3′ extremities together with their flanking sequences. Two microliters of amplified inverse PCR products were submitted to 10 cycles of linear PCR with 2 μmol/L of 5′-^{32}P-radiolabeled primer BLV3′RO (5′-CGCGCTTGTTTCCTGTCTTA-3′; position 8,626-8,645), 2 units of Stoffel fragment of the Taq DNA polymerase, and 200 μmol/L of each deoxynucleotide triphosphate in a final volume of 20 μL. Thermal cycling parameters were 95°C 10 minutes, 10 × (95°C 1 minute, 59°C 1 minute, 72°C 3 minutes), followed by a final elongation step of 10 minutes at 72°C. After boiling, 2 L of runoff products were analyzed on 6% sequencing gel.

Cloning and Sequencing. Purified inverse PCR products were phosphorylated by the T4 polynucleotide kinase, then ligated with *Sma*I-digested and dephosphorylated M13mp18 replicative form DNA, as previously described (10). After transformation of *Escherichia coli* XL1 by electroporation, recombinant M13 plaques were screened by hybridization with the BLV3′RO or the BLV5′RO long terminal repeat-specific ^{32}P-labeled oligonucleotide. Single-stranded templates were sequenced using

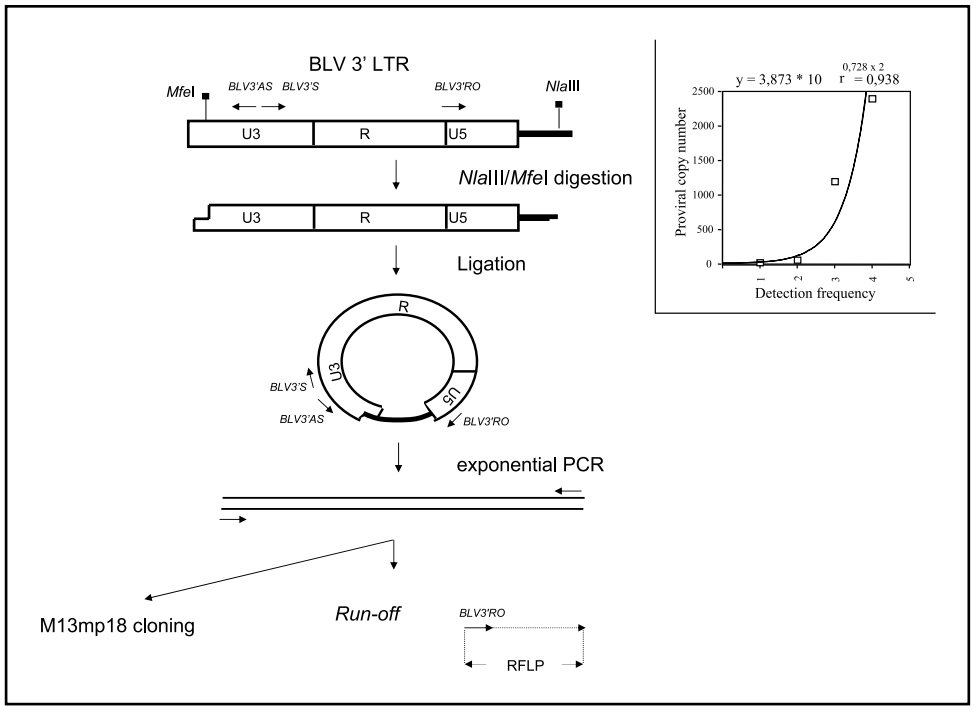

Figure 1. Inverse PCR protocols used for amplifying 3′ BLV integration sites. *Inset,* stochastic nature of BLV 3′ integration site detection by quadruplicate inverse PCR. BLV 3′ U3RU5 region integrated within the cellular DNA. As detailed in Materials and Methods, DNA is first digested by *Nla*III and *Mfe*I, then circularized. Exponential PCR and runoff analyses are done with BLV-specific oligonucleotide permitting to detect BLV integration site polymorphism. Alternatively, amplified products can be cloned.

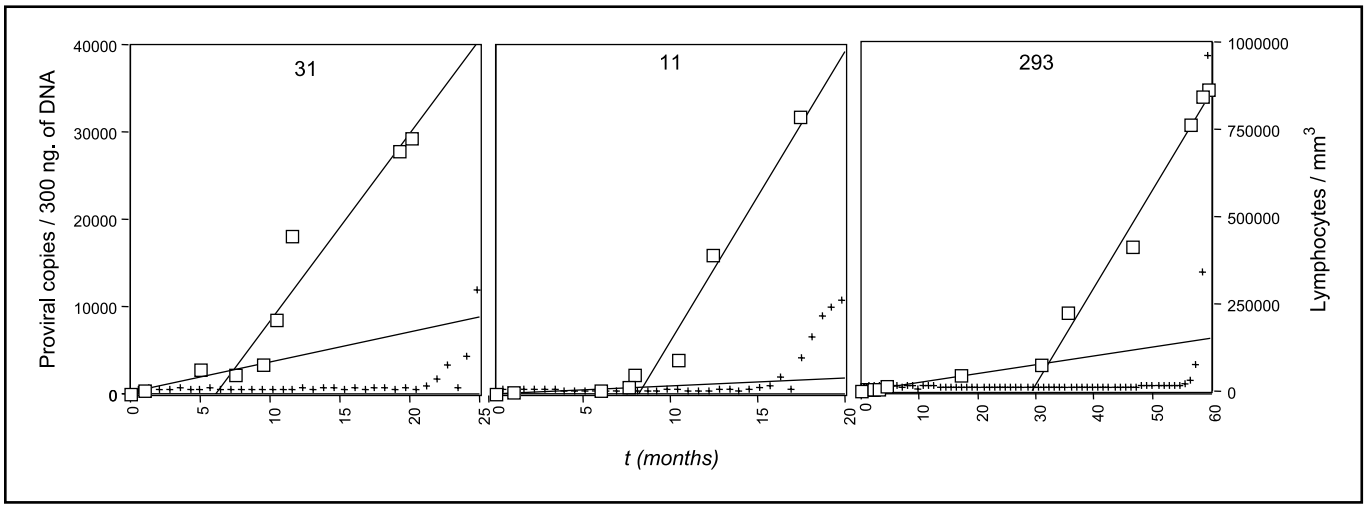

Figure 2. Fluctuation in BLV proviral loads over time in experimentally infected sheep. Twenty-eight blood samples were harvested over time from the three experimentally infected animals and analyzed for BLV proviral load by real-time quantitative PCR, as detailed in Materials and Methods. (□) BLV circulating proviral load at each time. (+++) fluctuation of circulating lymphocyte counts. For each animal, the two curves represent the two phases of circulating proviral load fluctuation over time (see Results for details).

fluorescent dideoxynucleotides. The products were resolved on an Applied Biosystems 377A DNA sequencer with 377A software. Sequence alignments were done with Sequence Navigator Software.

Semiquantitative Clonotypic Runoff Analysis of Malignant and Nonmalignant Integration Sites over Time. For the three sheep followed over time, oligonucleotides specific to nine 3′ BLV flanking sequences derived from samples obtained on different occasions were synthesized. These primers were subsequently radiolabeled and used for the clonotypic runoff analysis of inverse PCR products from all harvested samples derived from the corresponding animals. Primer extension was carried out as for conventional runoff experiments described above.

Statistical Analysis. SPSS statistical software version 11 was used for analysis. Descriptive statistics were used to describe proviral load, clonality, and somatic mutation frequency. We used the two-sided χ^2 test to compare two categorical variables, and the Fisher's exact test for 2×2 tables. The correlation of data was assessed by Spearman's ρ nonparametric method. $P < 0.05$ was considered significant in all analyses.

Results

Two-Step Nature of Bovine Leukemia Virus–Positive Cell Expansion *In vivo* during the Preleukemic Phase of Experimental Bovine Leukemia Virus Infection in Sheep. Three animals were first infected by direct inoculation of a cloned wild-type BLV provirus, then followed over time. For animals 31, 11, and 293, anti-BLV antibodies were detected 1.8, 0.5, and 2.1 months after experimental infection, respectively. Peripheral blood lymphocyte counts remained stable over time until 2.5 to 8 months before tumor onset, when animals developed rapid lymphocytosis. Circulating BLV-infected cells could be detected by quantitative PCR from the very first harvested samples. Hence, animals 31, 11, and 293 displayed positive BLV PCR results 1.6, 0.5, and 1 month after experimental infection. It is evident from Fig. 2 that, for each of the three sheep, the circulating BLV proviral loads increased as a function of time. For each animal, temporal fluctuations of the BLV proviral loads seemed to follow two distinct linear phases, including a weak slope followed by an accelerated phase ($P < 0.02$ and Spearman′s $\rho > 0.94$ for each phase in each animal). Both curves were found to be roughly parallel in sheep 11, 31, and 293 (Fig. 2), and intersection occurred at 8.4, 7.3, and 32.3 months from experimental infection, respectively. BLV proviral loads also measured in the tumor DNA (Figure 2) showed no significant differences between the three animals. Fig. 2 shows that there was a strong discordance between the temporal fluctuation of proviral loads and that of lymphocytosis. Indeed, from the first analyzed sample, circulating proviral loads of animals 31, 11, and 293 could be increased by factors of 59, 177, and 507, respectively, in the absence of hyperlymphocytosis.

Monoclonal Bovine Leukemia Virus Integration within the Tumor DNA of Experimentally Infected Animals. There is a stochastic component to the detection of retrovirus integration sites using inverse PCR. When samples are analyzed in quadruplicate, the most frequently detected signals correspond to the most abundant ones, as previously described (9). In the present work, the stochastic nature of BLV inverse PCR was found to appear at BLV integration site frequencies ranging between 25 and 2400 copies of the BLV provirus per microgram of blood DNA (Fig. 1). A quadruplicate inverse PCR analysis of *Mfe*I-*Nla*III–digested plasmids bearing an integrated BLV long terminal repeat showed that, above 2400 copies (diluted in 1 g), detection was 4/4. At copy numbers ranging from 1200 to 2400, 62.5 to 1200, and 25 to 62.5, detection was 3/4, 2/4, and 1/4, respectively. Accordingly, DNA samples from BLV-infected animals were analyzed in quadruplicate (4×0.5 g). Inverse PCR was done using the tumor DNA as shown in Fig. 1. A unique dominant signal was obtained with the tumor DNA from animals 11 and 293 (Fig. 3), meaning that these two sheep had monoclonal lymphoma with respect to BLV provirus integration. In contrast, three bands were present after runoff analysis of inverse PCR products derived from the tumor DNA of animal 31. Dilutions of DNA from this sample were subsequently analyzed by inverse PCR; the signal at ~269 bp disappeared at 50 ng, whereas those at 276 and 131 bp were repeatedly detectable (four times) at low concentration (10 ng, not shown). Cloning and sequencing allowed to show that the signal at 131 bp corresponded to a large internal deletion of the former provirus (not shown). Therefore, as for the other two sheep, a unique integration site represented the molecular signature of the malignant clone derived from animal 31.

Clonal Distribution of Bovine Leukemia Virus–Infected Cells over Time during the Course of Experimental Infection. We next investigated the clonal distribution of BLV-positive cells over

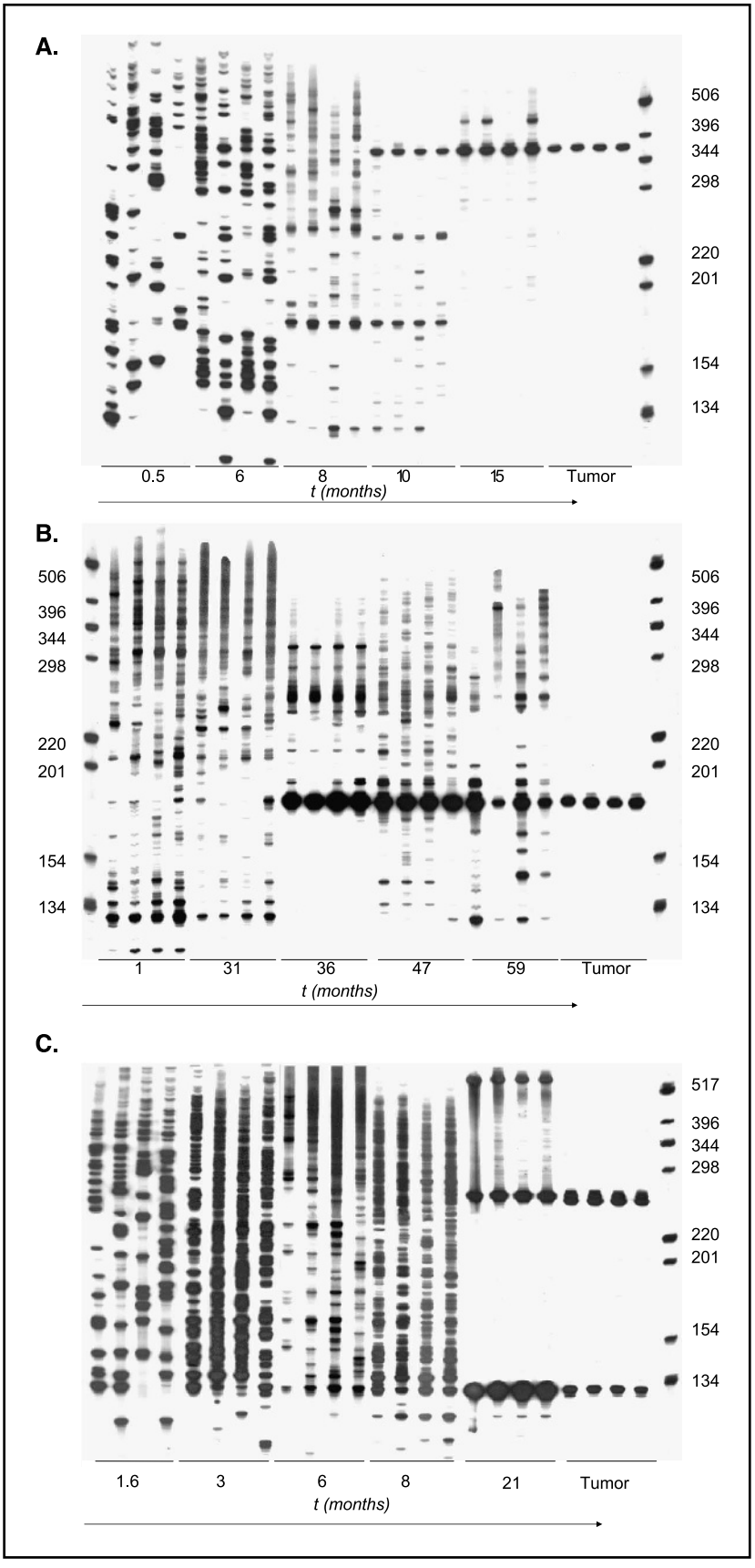

Figure 3. Clonality of BLV-infected cells over time in animals 11 (*A*), 293 (*B*), and 31 (*C*). Inverse PCR was done with samples harvested over time as shown in Fig. 1 and as detailed in Materials and Methods. For each animal, the BLV proviral integration pattern within the tumor DNA is also shown on the right. *Horizontal arrows,* premalignant and malignant clones.

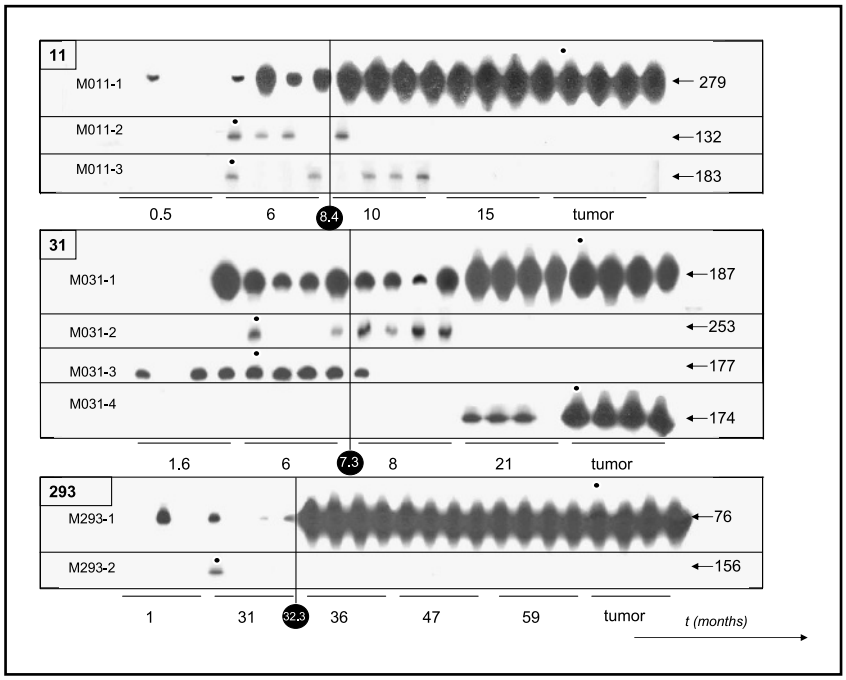

Figure 4. Clonotypic malignant and nonmalignant integration site-specific primer extension over time. After cloning and sequencing of inverse PCR products derived from samples harvested over time, nine integration site-specific primers were used for runoff analysis of quadruplicate inverse PCR products derived from the three sheep harvested on different occasions. As malignant clones derived from tumor DNA, premalignant clones were detected and semi-quantified with malignant clone integration site-specific primers, whereas non-premalignant clones were detected and semi-quantified with primers complementary to other BLV integration sites derived from samples harvested during the asymptomatic phase of the infection. Primer extension was done as detailed in Materials and Methods. For each animal, the first line represents the temporal fluctuation of the premalignant integration site, with the last analyzed sample corresponding to the tumor DNA. *Black plain circle,* time at which the inverse PCR products were cloned for integration site sequencing and specific primer synthesis, for each of the nine clones analyzed over time. Each sample was analyzed in quadruplicate. *Vertical line,* intersection between the two lines [representing the progressive then accelerated phases of proviral load (Fig. 2) and clonal expansion] for each animal; time of intersection is indicated (*bottom*). Sequences of integration site-specific primers for clones M011-1, M011-2, M011-3, M031-1, M031-2, M031-3, M031-4, M293-1, and M293-2 were 5′-TGCAATTATTGTAATTGATGAC3′, 5′-TCACAGACTTGGAGAA-CAAACTG-3′, 5′-TCATGGAATCACAAAGAGTCAGA-3′, 5′-TTAAGCCATCCAGTGGTGTG-3′, 5′-TACCATGTATGGATGTGAGAGTTG-3′, 5′-TTGCACACTT-TAAAAGTCGGAG-3′, 5′-TCTGGGCCCAGCTGTACG-3′, 5′-TCTGGGCCCAGCTGTACG-3′, 5′-TGCTCTATATAAACATCTTAGTCTTTCC-3′, and 5′-TCTGGACCCTGCAGCTCTT-3′, respectively.

time. Sheep 11 developed leukemia/lymphoma 18.5 months after experimental infection. For this animal (Fig. 3*A*), the first blood sample analyzed for BLV integration was harvested 0.5 month after experimental infection. This date corresponded to the first detection of circulating BLV antibodies. At that time, the proviral load was 88 copies/300 ng of DNA. Figure 3*A* shows that these sequences were distributed among 45 clones of variable abundance, as evidenced by runoff analysis of quadruplicate inverse PCR products. Hence, 26, 11, and 8 clones were detected one, two, and three times after quadruplicate inverse PCR, respectively, meaning that they had a clonal frequency of 25 to <62.5, 62.5 to <1200, and 1200 to <2400 copies in 1 g of circulating DNA. None was detected four times. Similarly, clones of BLV-bearing cells could be detected in subsequently harvested samples, some of which being persistently detected over time (Fig. 3*A*). However, the figure shows that the overall number of detected clones tended to decrease as a function of time whereas the proportion of abundant clones [i.e., those detected four times after quadruplicate experiment (≥2400 copies in 1 g of circulating DNA)] was found to increase as a function of time. A signal at 369 bp, corresponding to the size of the tumor integration site, could be detected in the circulating DNA during the asymptomatic phase of the infection (Fig. 3*A*). This signal was clearly distinguishable in samples harvested 0.5, 6, 8, 10, and 15 months after experimental infection (Fig. 3*A*), meaning that the premalignant clone could be detected early. In earlier samples, the background of polyclonally distributedBLV flanking sequences hampered its detection. Animal 293 was characterized by a prolonged period of latency owing to its development of leukemia/lymphoma 60 months after experimental infection only. As for sheep 11, the degree of persistent clonal expansion of BLV-positive clones seemed to increase as a function of time. Figure 3*B* shows that the overall number of detected clones tended to decrease as a function of time whereas the proportion of those detected four times after quadruplicate experiment (≥2400 copies in 1 g of circulating DNA) increased. Figure 3*B* shows that the signal at 180 bp corresponding to the malignant clone could be detected in early harvested samples. Animal 31 developed leukemia/lymphoma 23.5 months after experimental infection. A pattern of persistent clonal expansion also characterized BLV replication over time in this sheep (Fig. 3*C*). Furthermore, the two signals corresponding to the molecular signature of the malignant clone could be detected in the DNA derived from earlier samples (Fig. 3*C*). Overall, for each animal, a pattern reminiscent of that observed in blood samples from HTLV-1 infected individuals with adult T-cell leukemia/lymphoma (ATLL) or pre-ATLL clearly appeared during the late stages of experimental infection, indicative of the extensive expansion of a restricted number of clones on a background of oligo/polyclonal expansion of BLV-bearing cells (10–14). There was a significant correlation between the duration of the infection and the number of abundant clones [i.e., those detected four times ($P < 0.023$ and Spearman's $\rho = 0.504$)], but not that of less expanded clones. Similarly, the proviral load significantly correlated with the number

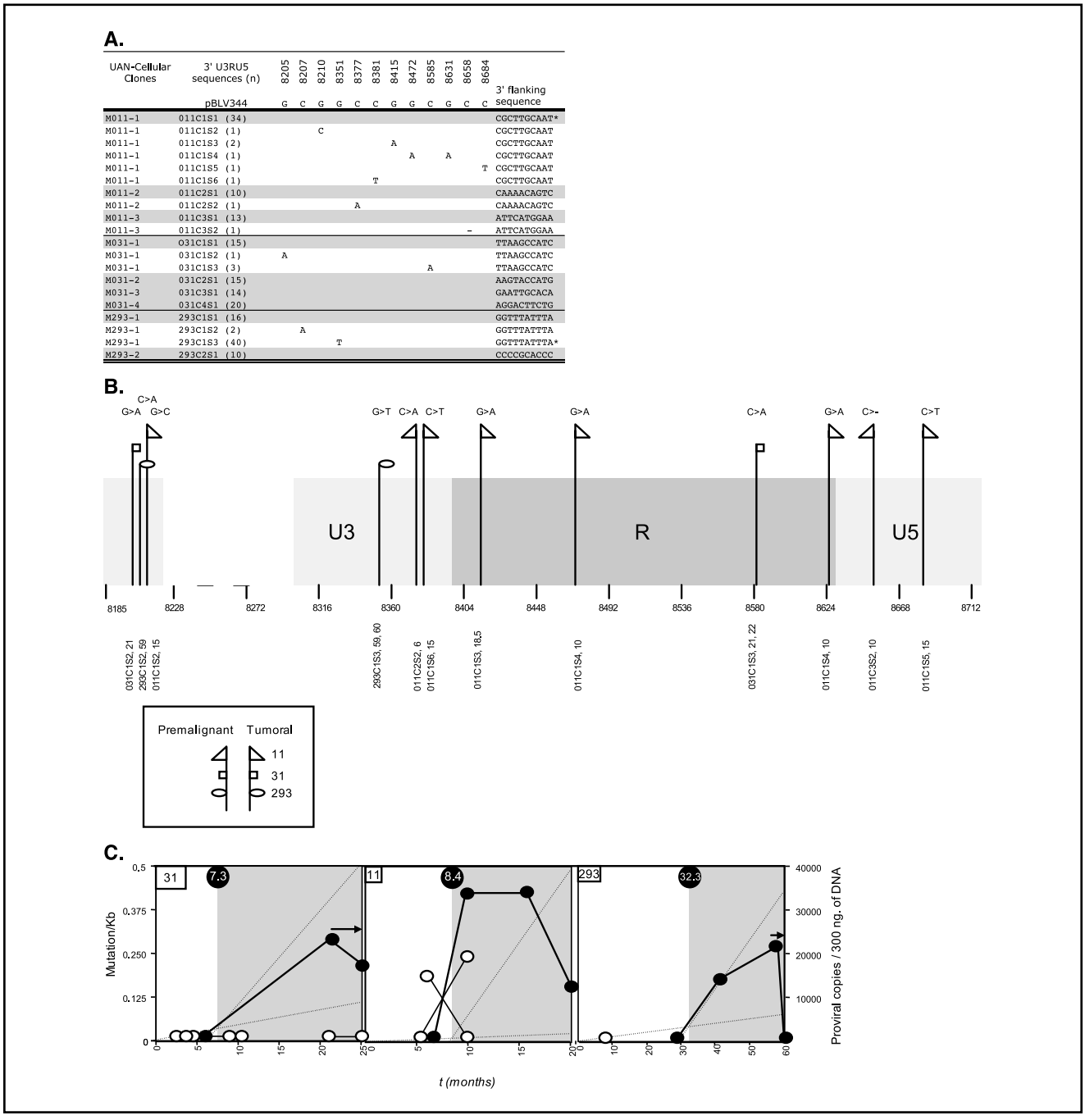

Figure 5. Somatic mutations of the BLV provirus *in vivo*. *A,* overall, nine distinct BLV 3′ integration sites were isolated over time; *right,* the 10 bp of the corresponding flanking cellular sequence. U3RU5 sequences were aligned according to pBLV344 plasmid sequence. (–) deletions. Coordinates of the sequence are those of the K02120 sequence (7). *Asterisk,* sequences harboring mutations in their flanking sequences. Sheep are identified by their UAN. Each cluster of U3RU5 sequences sharing a common integration site, and therefore belonging to a unique clone of expanded BLV-positive cells, is identified by its cellular clone number. *Bold horizontal bar,* separation of the clusters of sequences derived from each of the three sheep DNA. *Gray lanes,* consensus U3RU5 sequences identical for the three animals. For animals 11, 31, and 293, clones 11-1, 31-1, and 293-1 corresponded to premalignant or malignant clones. For each cellular clone, the numbers of consensus and, if present, mutated U3RU5 sequences are indicated between brackets. *B,* random distribution of somatic mutations along the BLV 3′ long terminal repeat U3RU5 sequence. *Hatched portion,* the region in the U3 sequence that was not encompassed by inverse PCR specific primers. *Bottom,* the corresponding clone number and time point (in days post-infection) for each substitution. *C,* temporal fluctuation of the somatic mutation loads of URU5 sequences and integration sites over time in the three experimentally infected sheep. For each point, the number of mutations was divided by the length of the analyzed sequence (U3RU5 + flanking sequence). *Gray area,* second, accelerated phase of increased proviral load and clonal expansion (this figure and Figs. 3 and 6). ○, mutation loads of the BLV-positive clones that remained untransformed during the life span of the three sheep; ●, those of premalignant clones. For the three animals, the last sample corresponds to the tumor. The two horizontal arrows in animals 31 and 293 represent two persistently detected substitutions. They include the C→A mutation at position 8,585 of the premalignant clone derived from sheep 31 (sequence 031C1S3, *A*) and the G→T mutation at position 8,351 of the premalignant clone derived from sheep 293 (sequence 293C1S2, *A*).

of clones detected four times ($P < 0.002$ and Spearman′s $\rho = 0.643$) but not with the frequency of polyclonal forms. Overall, although the duration of the latency period preceding tumor onset was significantly different between animals, these three experimentally infected sheep displayed a roughly identical route of BLV replication. It included a pattern of persistent clonal expansion that increased as a function of time, with the suggestion of an early detection of premalignant clones.

Clonal History of Premalignant versus Non-Premalignant Bovine Leukemia Virus–Bearing Cells over Time. In addition to the three tumor samples, inverse PCR products derived from 23 additional samples harvested at various time points were cloned and sequenced. This permitted to sequence six additional, distinct BLV integration sites. Nine integration site-specific oligonucleotides, including those complementary to the tumor-specific flanking sequences, were synthesized and used for clonotypic runoff analysis of quadruplicate inverse PCR products obtained over time in the three sheep (Fig. 4). For each animal, the premalignant clone was defined as that detected during the asymptomatic phase of the infection and having the same integration site as the corresponding tumor clone, whereas other BLV-positive clones corresponded to persistently BLV-positive expanded cells that remained untransformed during the course of the infection (i.e., that did not produce tumors). As BLV integration is at random, with no preference for any specific region of the genome, tumor or malignant BLV-positive cells were assumed to derive from their premalignant counterparts. By clonotypic PCR, it was possible to compare, using the same method, the temporal fluctuation in the degree of clonal expansion of BLV-positive cells over time between the three premalignant and six non-premalignant clones. For the three animals, the premalignant clone could be specifically detected as early as in the first harvested samples (i.e., at 0.5 to 1.6 months from experimental infection; Fig. 4). By contrast, only one of the six non-premalignant clones could be detected early ($P = 0.025$). For animal 11, the first detection of the premalignant clone coincided with the time of seroconversion, whereas it preceded it for animals 31 and 293. The second major difference between the two categories of clones was their temporal fluctuation. Premalignant clones were regularly detected in all analyzed samples and their degree of clonal expansion increased as a function of time (Fig. 4). In contrast, although some non-premalignant clones were persistently detected over periods of 5 (Fig. 4, clone M011-3) to 6 (Fig. 4, clone M031-2/3) months, all these forms were transiently detected without time-dependent increase in their degree of clonal expansion. It is evident from Figs. 2 and 4 that the degree of premalignant clonal expansion significantly correlated with both circulating proviral

Table 1. Analyses of somatic mutations in U3RU5 region of BLV provirus and flanking cellular sequences

	U3RU5	Flanking sequences
Substitutions	11	2
Transitions	6	0
Purine:purine	4	0
Pyrimidine:pyrimidine	2	0
Transversions	5	2
T > N	0	2
C > N	3	0
A > N	0	0
G > N	2	0
Single-base deletions	1	0
ΔT	0	0
ΔC	1	0
ΔG	0	0
ΔA	0	0

NOTE: >, sense of substitution; N, any nucleotide; Δ, deletion.

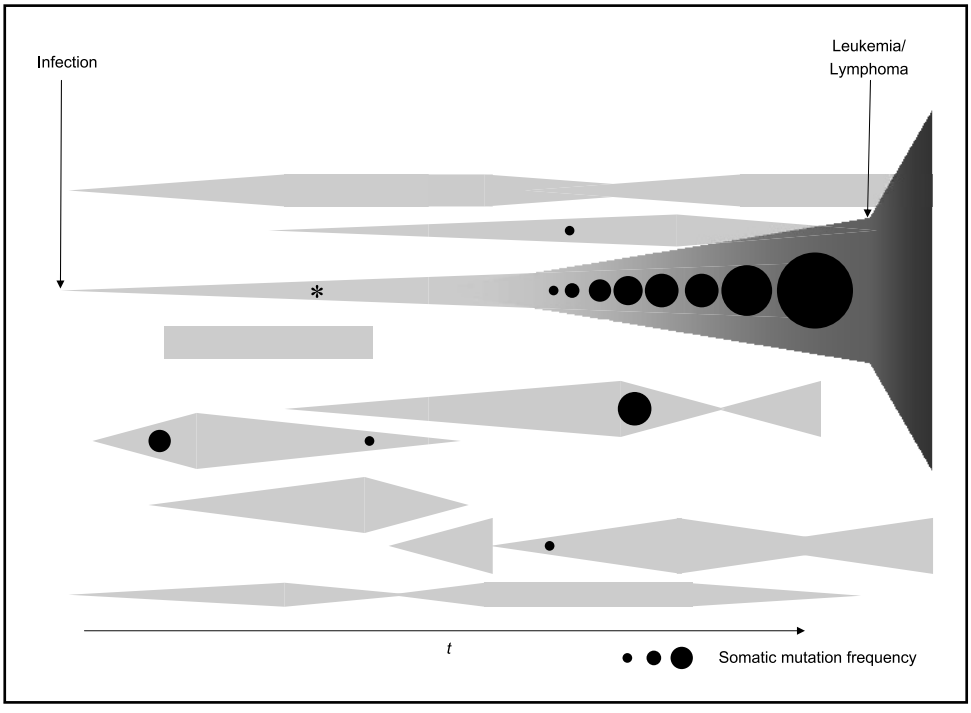

Figure 6. Clonal history of BLV-associated leukemia/lymphoma in experimentally infected sheep as a model of early oncogenesis. The figure represents, in one animal, the temporal fluctuations of BLV-infected cells together with their mutation frequency during the premalignant phase of experimental infection (i.e., from the time of infection to tumor stage). *Vertical arrows,* time of experimental infection and the onset of overt lymphoma. *Horizontal lines,* clones; line thickness correlates with the degree of clonal expansion. *Black circles,* somatic mutation frequency; circle size correlates with mutation frequency. *Asterisk,* premalignant clone.

loads ($P < 10^{-4}$ and Spearman's $\rho = 0.871$) and the duration of experimental infection ($P < 10^{-4}$ and Spearman's $\rho = 0.806$). No such correlation was observed for non-premalignant clones. Therefore, by contrast to the background of BLV-positive cells that remained untransformed during the entire life span of the three animals, premalignant cells are characterized by early and repeated detection with a pattern of extensive clonal expansion.

Somatic Mutation in the Bovine Leukemia Virus Provirus and Flanking Cellular Sequences during Clonal Expansion *In vivo*. From inverse PCR products obtained over time (Fig. 4), a total of 202 molecular clones were sequenced, encompassing the 3′ U3RU5 sequences of BLV (441 bp) and their flanking cellular sequences (range 46 to 235 bp). The sequences could be arranged into nine cellular clones, including the three malignant forms, depending on the base composition of cellular integration sites. Five of the nine clones (~56%) were not homogeneous (Fig. 5*A*) and a total of 11 single-base substitutions and one single-base deletion were scored with no evidence of reverse transcription-associated substitution (Table 1). The average 3′ U3RU5 mutation frequency was $\sim 3.3 \times 10^{-3}$ substitution per base (12 distinct substitutions in 9 RU5 sequences of 441 bp) with no evidence of hotspots (Fig. 5*B*). Two additional somatic mutations were also found in the flanking sequences of a couple of clones derived from animals 11 and 293 (Fig. 5*A*). Overall, these results evidence that a high level of genetic instability resulting from somatic mutations accompanies BLV replication via clonal expansion.

At the Pre-Tumor Stage, Genetic Instability Distinguishes the Premalignant Clone from Other Infected Cells, Correlates with the Degree of Clonal Expansion, and Increases as a Function of Time. Figure 5*C* represents the temporal fluctuation of the mutation loads in the three sheep. In animal 11, genetic instability was not restricted to premalignant and malignant clones. The two clones that remained untransformed during the course of the infection were also found to harbor somatic mutations; for both clones, mutations were observed on a single occasion only. Their mutation loads (0.17 substitution/kb for clone M011-2 at 6 months and 0.25 for clone M011-3 at 10 months) were not significantly different from those observed with the corresponding premalignant clone, and even exceeded that of the corresponding tumor (0.16 mutation/kb). As animal 11, sheep 31 displayed somatic mutations of both premalignant and tumor clones. However, none of the other three untransformed clones analyzed on five occasions were found to be mutated over time. Similar results were observed in animal 293, with a temporal increase in the mutation load of the premalignant clone contrasting with the high level of genetic stability of other clonally expanded cells. During the asymptomatic phase of the infection, the mutation frequencies per time point of premalignant and non-premalignant clones were 5/8 and 2/12, respectively ($P = 0.04$). Premalignant clones displayed a significantly higher frequency of somatic mutations than other BLV-positive clones (0.2 versus 0.035 substitution/kb, $P = 0.015$; Fig. 5*C*). Mutated clones were characterized by a high level of clonal expansion; the mutation load correlated with the frequency of integration site detection after quadruplicate clonotypic runoff analysis ($P < 0.024$ and Spearman's $\rho = 0.48$). As the degree of clonal expansion increased as a function of time during the course of infection, it was not surprising that the somatic mutation load of premalignant clones was a time-dependent variable ($P < 0.015$ and Spearman's $\rho = 0.51$). Furthermore, the two-step increase over time of circulating proviral loads and clonal expansion strongly correlated with a two-step increase in the genetic instability of BLV premalignant clones (Figs. 3, 4, and 5*C*). None of the 14 detected substitutions was found to persist over time during the asymptomatic phase that precedes tumor onset. However, two substitutions in two premalignant clones (animals 31 and 293) detected at last blood sample harvest were also present in the corresponding tumor that developed a few weeks later. As shown in Fig. 5*A* and *C*, they included the CA mutation at position 8,585 of the premalignant clone derived from sheep 31 (sequence 031C1S3, Fig. 5*A*) and the GT mutation at position 8,351 of the premalignant clone derived from sheep 293 (sequence 293C1S2, Fig. 5*A*). Overall, premalignant clones could be clearly distinguished from other infected cells on the basis of their mutation load. As for proviral loads and clonal expansion, two clearly distinguishable phases characterize the dynamics of somatic mutations. However, the time-associated increase of mutation loads was found to be restricted to circulating premalignant lymphocytes, as their tumor counterparts displayed significantly lower mutation frequencies despite having extremely high proviral loads (Fig. 5*C*).

Discussion

Using the sheep model of BLV infection, our study has reconstituted the clonal history of premalignant clones during the asymptomatic phase preceding BLV-associated malignancies. Whereas other BLV-positive cells remain untransformed during the course of experimental infection, premalignant cells are characterized by their early detection, persistent clonal expansion, and high level of genetic instability with respect to somatic mutation loads. Furthermore, both the degree of clonal expansion and the genetic instability of premalignant cells increased as function of time through a two-step process, with the suggestion of a decrease in the level of genetic instability at tumor stage.

In the present model, the first hallmark of premalignant cells was their early detection. In each of the three sheep, the premalignant clones could be detected in the first harvested samples (i.e., as early as 0.5 to 1.6 months after experimental infection). For two animals, these dates preceded the detection of anti-BLV specific antibodies whereas both events coincided in the third sheep. As BLV spread is thought to combine reverse transcription-associated replication with clonal expansion, this first result implies that the generation of a premalignant clone includes an early infection of its progenitor cell, the infected progeny of which will constitute the premalignant clone. Therefore, a prolonged period of virus exposure, corresponding here to a prolonged clonal expansion of BLV-positive cells, seems to be a prerequisite for subsequent malignant transformation. The second hallmark of premalignant clones was their pattern of clonal expansion. By contrast to other BLV-positive clones, the degree of clonal expansion of premalignant cells was found to increase as a function of time (Figs. 3 and 4). The temporal increase in premalignant clonal expansion contributed to the overall increase in circulating BLV proviral loads (Fig. 2). It is of note that during a prolonged period of time, the dramatic increase in premalignant cells did not influence the overall lymphocyte count, meaning that premalignant cells replaced normal circulating lymphoid cells. It is clear from Figs. 3 and 4 that after a progressive phase of clonal expansion came a dramatic increase in premalignant cell clonal expansion. The duration of the first phase corresponded to about one third (animal 31) to one half (animals 11 and 293) of the asymptomatic premalignant period. Such temporal distribution suggests that, after a necessary early

infection, the premalignant clones underwent a second event that specifically altered their growth kinetics, as compared with all other BLV-positive clones investigated.

The third characteristic of premalignant cells was their degree of genetic instability, as assessed by monitoring their somatic mutation loads. Mutations were not restricted to the premalignant cells due to some non-premalignant BLV-infected cells also found to harbor mutated 3′ U3RU5 sequences. However, the mean mutation frequency of premalignant clones was significantly higher than that of other cells. More importantly, when considering the temporal fluctuation of the mutation loads for the three animals, this frequency is found to strongly increase as a function of time and to correlate with the degree of clonal expansion in premalignant cells. Such correlation between mutation loads and cellular proliferation has previously been observed in human HTLV-1 infection (10). However, the present model made it possible to investigate the two parameters over time, clearly demonstrating that, as for overall proviral loads and clonal expansion, the mutation process of premalignant cells only occurred during the accelerated phase of premalignant cell dissemination (Fig. 5*C*). This strongly suggests that the same factor triggers both the clonal expansion and the genetic instability of premalignant cells. It is of note that none of the detected mutations persisted during the premalignant phase of the infection. It is highly improbable that all observed substitutions could have been repaired, suggesting that, in our model, the genetic instability of premalignant cells is frequently incompatible with cell survival. Such correlation is not surprising if we assume that the entire genome of an infected cell undergoes the same mutation process as that of currently analyzed sequences (U3RU5 sequences + integration sites), providing the opportunity for frequent alterations of key genes involved in host cell homeostasis. Alternatively, somatic mutations might increase the formation of viral or cellular neoantigens, allowing control by host cell-mediated immunity.

Two major genetic changes seem to characterize the shift from premalignant stage to overt tumor. The first one includes the fixation of some somatic mutations, as observed for C8585A and G8551T substitutions in animals 31 and 293, respectively. The second change corresponds to a significant decrease in the mutation frequency (i.e., in the occurrence of newly detected substitutions within the tumor DNA as compared with premalignant cells; Fig. 5*C*). These strong changes in the dynamics of genetic variability coincided with the development of clinical tumors and a dramatic increase in circulating lymphocyte counts. Hence, for animals 11, 31, and 293, the shift from last premalignant stage to tumor stage is accompanied by a decrease in somatic mutation frequencies by 64%, 17%, and 100%, respectively (Fig. 5*C*), whereas the corresponding circulating lymphocyte counts increase by factors of 71, 87, and 40, respectively (Fig. 2). Thus, at least in animals 11 and 293, the last phase of tumor development includes a dramatic decrease in cellular genetic instability, thereby reducing mutation-associated cellular death and allowing for a massive increase of tumor burden. This means that, for these animals, a unique mutated clone acquired a sufficient selective advantage that precluded any need for an ongoing mutator phenotype. Such stabilization of the cellular genome during tumor onset is reminiscent of the shift in chromosomal stability that accompanies the biphasic implication of telomerase activity in cancer initiation and development. In this example, there are experimental evidences of suboptimal premalignant telomerase activity that, in combination with p53 inactivation, leads to telomere dysfunction-associated chromosomal instability (15). Strong selection for cellular survival at this stage leads to telomerase reactivation, endowing cells with immortal growth potential and stabilizing their genome (16). Present results suggest that the genetic instability associated with the somatic mutation process, as that resulting from large chromosomal abnormalities, needs to be stabilized to permit the final development of overt malignancy. Acquired mutations of genes involved in the control of cellular genetic stability might have contributed to this observation. Alternatively, as recently observed with HTLV-1 (17), somatic mutations might generate viral and/or cellular mutants that can escape CTL recognition thereby promoting leukemogenesis.

Focusing our investigations on two of the major hallmarks of cancer cells (i.e., altered growth kinetics and genetic instability), we could show that in the present model of virus-induced hematologic malignancy, premalignant clones can be early and clearly distinguished from other virus-exposed cells on the basis of their degree of clonal expansion and genetic instability. A unique event seems to trigger both clonal expansion and genetic instability that, at the premalignant stage, evolve through a two-step process. Somatic mutations seem to regularly lead to cellular death, contributing in part to maintaining a normal lymphocyte count during the premalignant stage. This equilibrium is finally broken after a 18.5- to 60-month period of clinical latency, when a dramatic decrease in the genetic instability of premalignant cells occurs. For the three animals, this coincides with lymphoma onset and results, in 1 to 2 weeks, in the development of extremely high tumor burdens (Fig. 6). In addition to these results, the present strategy opens the possibility to further investigate additional molecular and cellular characteristics of premalignant cells and to specifically evaluate therapeutic interventions during early carcinogenesis.

Acknowledgments

Received 5/25/2004; revised 11/30/2004; accepted 12/7/2004.

Grant support: Ligue Nationale Contre le Cancer (équipe labellisée, 2003); bursaries from Comité Départemental de la Loire de la Ligue Nationale Contre le Cancer (V. Moulés), Comité Départemental de l'Ain de la Ligue Nationale Contre le Cancer (C. Pomier), Association Française pour la Recherche sur le Cancer (D. Sibon), and Centre Léon Bérard (A-S. Gabet).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Denis Gerlier for fruitful discussions.

References

1. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57–70.
2. Hjalgrim LL, Madsen HO, Melbye M, et al. Presence of clone-specific markers at birth in children with acute lymphoblastic leukaemia. Br J Cancer 2002;87:994–9.
3. Wiemels JL, Xiao Z, Buffler PA, et al. *In utero* origin of t(8;21) AML1-ETO translocations in childhood acute myeloid leukemia. Blood 2002;99:3801–5.
4. Spathas DH, Stewart J, Singer IO, et al. Detection of t(12;21) in childhood acute lymphoblastic leukemia by fluorescence *in situ* hybridization. Cancer Genet Cytogenet 1999;110:7–13.
5. Yagi T, Hibi S, Tabata Y, et al. Detection of clonotypic IGH and TCR rearrangements in the neonatal blood spots of infants and children with B-cell precursor acute lymphoblastic leukemia. Blood 2000;96:264–8.
6. Willems L, Kettmann R, Dequiedt F, et al. *In vivo* infection of sheep by bovine leukemia virus mutants. J Virol 1993;67:4078–85.

7. Sagata N, Yasunaga T, Tsuzuku-Kawamura J, et al. Complete nucleotide sequence of the genome of bovine leukemia virus: its evolutionary relationship to other retroviruses. Proc Natl Acad Sci U S A 1985; 82:677–81.
8. Tajima S, Takahashi M, Takeshima SN, et al. A mutant form of the tax protein of bovine leukemia virus (BLV), with enhanced transactivation activity, increases expression and propagation of BLV *in vitro* but not *in vivo*. J Virol 2003;77:1894–903.
9. Cavrois M, Wain-Hobson S, Wattel E. Stochastic events in the amplification of HTLV-I integration sites by linker-mediated PCR. Res Virol 1995;146:179–84.
10. Mortreux F, Leclercq I, Gabet A, et al. Somatic Mutation in Human T-Cell Leukemia Virus Type 1 Provirus and Flanking Cellular Sequences During Clonal Expansion *In vivo*. J Natl Cancer Inst 2001;93: 367–77.
11. Cavrois M, Wain-Hobson S, Gessain A, Plumelle Y, Wattel E. Adult T-cell leukemia/lymphoma on a background of clonally expanding human T-cell leukemia virus type-1-positive cells. Blood 1996;88: 4646–50.
12. Leclercq I, Cavrois M, Mortreux F, et al. Oligoclonal proliferation of human T-cell leukaemia virus type 1 bearing T cells in adult T-cell leukaemia/lymphoma without deletion of the 3′ provirus integration sites. Br J Haematol 1998;101:500–6.
13. Gabet AS, Mortreux F, Talarmin A, et al. High circulating proviral load with oligoclonal expansion of HTLV-1 bearing T cells in HTLV-1 carriers with strongyloidiasis. Oncogene 2000;19:4954–60.
14. Mortreux F, Gabet AS, Wattel E. Molecular and cellular aspects of HTLV-1 associated leukemogenesis *in vivo*. Leukemia 2003;17:26–38.
15. Artandi SE, Chang S, Lee SL, et al. Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. Nature 2000;406:641–5.
16. Artandi SE, DePinho RA. Mice without telomerase: what can they teach us about human cancer? Nat Med 2000;6:852–5.
17. Furukawa Y, Kubota R, Tara M, Izumo S, Osame M. Existence of escape mutant in HTLV-I tax during the development of adult T-cell leukemia. Blood 2001; 97:987–93.