Supporting Information

Rosewick et al. 10.1073/pnas.1213842110

SI Materials and Methods

Primary Ovine and Bovine Tissues. Primary leukemic B cells and lymphoid tumors examined in this study were collected from bovine leukemia virus (BLV)-infected sheep at the acute stage the disease. The latency period before tumor development ranged between 15 and 48 mo following BLV infection. All sheep were housed at the Centre de Recherches Vétérinaires et Agrochimiques (Brussels, Belgium). Experimental procedures were approved by the Comité d'Ethique Médicale de la Faculté de Médecine Université Libre de Bruxelles and were conducted in accordance with national and institutional guidelines for animal care and use. Sheep from which the tissues were collected had been infected either with a BLV infectious molecular clone (pBLV344), an infectious BLV variant (pBLVX3C), isogenic to the full-length wild-type 344 provirus (1, 2), or peripheral blood mononuclear cells (PBMCs) isolated from BLV-infected animals S17, S18, and S19 (M2531, M2532, M2241) as described earlier (3). Blood was collected in EDTA-containing tubes, and PBMCs were isolated using standard Ficoll-Hypaque separation. Lymphoid tumors were collected at necropsy and minced through a nylon mesh cell strainer (Becton Dickinson) to obtain single-cell suspensions. B-cell percentages were measured by fluorescence-activated cell sorting (ovine sIgM-specific mAb PIG45A; VMRD) and found to be ≥93% for all malignant samples used for further investigation. Purified B cells were generated from PBMCs isolated from BLV-infected sheep before tumor onset (preleukemic samples from animals M2531, M2532, M2241), and from three uninfected sheep (control libraries). The B-cell fraction was obtained using a magnetic bead-mediated negative selection procedure following PBMC depletion of CD14+, CD4+, CD8+, CD2+, and γ-δ TCR+ cells (mAbs 36F-18, 17D, BAT82A, 86D, MM61A from VMRD and MACS anti-mouse IgG1 and IgG2a microbeads from Miltenyi Biotec). Sample purity was verified by fluorescence-activated cell sorting (ovine sIgM-specific mAb PIG45A; VMRD). Bovine B-cell tumors were from our tumor collection stored at −80 °C and comprised samples from various geographical origins including Japan, France, United States, and Belgium. Tumor samples were collected from B-cell lymphoid masses developing following natural BLV infection and have been characterized previously (4). The latency period before tumor onset in these animals is not accurately documented.

HTLV-1 Adult T-Cell Leukemia Samples. PBMC samples were collected from three patients with an acute form of adult T-cell leukemia (ATL) (>80% circulating malignant cells) after informed consent was obtained in accordance with the Declaration of Helsinki and after institutional review board-approved protocol at the Centre Hospitalier de Fort de France (Guadeloupe, France). Total RNA was extracted using TRIzol reagent and precipitated with ethanol, with glycogene as carrier. All ATL samples were negative for Tax expression, as evidenced by quantitative RT-PCR (not shown).

Cell Lines and in Vitro Cell Culture Procedures. YR2 and L267 tumor B-cell lines were derived from ovine B-cell tumors M395 (B-cell leukemia) and T267 (B-cell lymphoma), respectively. Each pair (primary malignant B-cell clone/corresponding ex vivo-derived cell line) displays an identical BLV integration clonality profile and is characterized by the absence of viral mRNA and protein expression. $YR2_{LTaxSN}$ and $L267_{LTaxSN}$ result from the transduction of parental YR2 and L267 cells with the pLTaxSN retroviral vector and permanently express Tax and viral proteins (5). Lymphoid cell suspensions were cultured at a concentration of 10^6 cells/mL in OPTIMEM medium (Invitrogen) supplemented with 10% FCS, 1 mM sodium pyruvate, 2 mM glutamine, MEM nonessential amino acids solution 1X (Gibco), and 100 μg/mL kanamycin. The RNA Pol II inhibitor α-Amanitin (Sigma) was used at a final concentration of 1 μg/mL for 24 h. Poly-L-lysine hydrobromide (PLL) was purchased from Sigma-Aldrich (P6516). HeLaT were obtained from the American Type Culture Collection (ATCC). HCT-116 WT, HCT-116 $Dicer^{-/-}$, 293T WT, and 293T $TRBP^{-/-}$ were a gift from F. Kansanghi (George Mason University, Manassas, VA) and were generated as described previously (6). Adherent cells were cultured in DMEM supplemented with L-glutamine and penicillin/streptomycin with 10% FCS.

Deep Sequencing Analysis. The 36-bp sequencing raw reads were preprocessed to discard bad quality reads and trim the 3′ adapter (homemade Perl and R scripts). Ovine miRNA reads were annotated using the bovine miRNA database in miRBase 17 (7). Rfam 10 (8) was used for annotation of noncoding RNA families (snRNA, tRNA, rRNA). Mapping of the reads on the ovine (9) or the BLV proviral genome (10) was carried out using the short read aligner Bowtie 0.12.7 (11). For miRNA biogenesis product analysis, 76-bp sequencing data were preprocessed as above, mapped on the ovine or the BLV proviral genome, respectively, and analyzed using IntersectBed from the Bedtools suite (12) to extract the read numbers for each feature (loop, premiRNA, 5′ subproduct, 3′ subproduct, and moR). Small RNA libraries prepared from human HTLV-1–induced ATL samples were processed and sequenced following the same procedure. Human miRNA reads were annotated using the human miRNA database (miRBase 17). ATL-derived reads were mapped on the human hg19 (13) or the HTLV-1 proviral genome (NC_001436.1) using Bowtie.

Secondary Structure Prediction. Secondary structures of premiRNAs and miRNA clusters were predicted with RNAfold from the Vienna package (14)

Conservation Analysis. Analysis of the conservation of the miRNA sequence from different BLV isolates was conducted using (*i*) a set of seven isolates for which both miRNA and Tax sequences were available, obtained from National Center for Biotechnology Information (NCBI), and (*ii*) two nonpaired sets of 25 miRNA and 25 Tax sequences obtained from high-throughput sequencing (HTS), Sanger sequencing of ovine and bovine tumors, and from public database (NCBI). To perform the analysis, we first aligned the corresponding sequences using the multiple alignment tool clustalW (15). We then defined five sequence segments: (*i*) miR-mature, corresponding to the five mature BLV miRNAs; (*ii*) miR-star, corresponding to the five star BLV miRNAs; (*iii*) miR-rest, corresponding to 24 bp of 5′ flanking, 111 bp of 3′ flanking, 227 bp of intermiRNA, and 61 bp of loop sequences; (*iv*) Tax, corresponding to the 930-bp full-length Tax coding sequence; and (*v*) Tax-3rd, corresponding to 170 third codon position for the Tax region that does not overlap with Rex. For each nucleotide position within each segment, we computed the average nucleotide diversity, i.e., the proportion of sequence pairs for which the residue differs at that position, yielding a nucleotide diversity specific for position i in segment j, $\tilde{\pi}_{ij}$. For each segment j, we then computed a segment-specific nucleotide diversity $\tilde{\pi}_j$ as the average over all $\tilde{\pi}_{ij}$. This allowed us to compute differences in $\tilde{\pi}_{ij}$for all pairs of segments. The statis-

tical significance of the difference in segment-specific nucleotide diversities between two segments (say 1 and 2; Δ_{1-2}) was then calculated using a permutation test as follows. Assume that segment 1 has n positions, and segment 2, m positions. The corresponding $(n + m)$ $\tilde{\pi}_{ij}$ values would be randomly assigned to a first group with n values and a second with m values. We would then compute the average $\tilde{\pi}_j$ value for each of the two groups, compute the difference between these two values ($\Delta_{\psi 1-\psi 2}$), and store it. This operation was repeated 10,000 times, yielding 10,000 $\Delta_{\psi 1-\psi 2}$ values. The statistical significance of the difference obtained with the real data (Δ_{1-2}) was then estimated as the proportion of $\Delta_{\psi 1-\psi 2}$ values for which the absolute value would be equal or larger than Δ_{1-2} (i.e., a two-tailed test).

qRT-PCR Analysis of miRNA Expression. Total RNA was extracted using TRIzol Reagent (Ambion). The resultant RNA was treated with TURBO DNA-free (Ambion) to remove any contaminating DNA. miRNAs were quantified using either custom TaqMan stem–loop microRNA assays (ABI) following the manufacturer's recommendations or QuantiMir-based qRT-PCR assays (SBI; #RA420A-1). QuantiMir detection consisted of poly(A) tailing and incorporation of a 3′ tag sequence during reverse transcription, followed by PCR carried out using the 3′ universal primer and the exact sequence of each BLV miRNA as the 5′ primer for the quantification of each BLV miRNA (primers are listed in Table S8). Additional 5′ sequence-specific primers included hsa-miR-34a (Table S8) and the human snRNA U6 (SBI). TaqMan microRNA assays for hsa-miR 34a (assay ID 000426), hsa-miR 186 (assay ID 002285), RNA U6 (assay ID 001973), and 5S-rRNA (Hs03682751_gH) were purchased from ABI. Real-time qRT-PCR analysis was carried out using an ABI Prism HT7900 Sequence Detector System (Applied Biosystems), and amplified product levels were detected by real-time monitoring following the manufacturer's instructions. The level of expression of each miRNA is presented relative to the background amplification obtained with RNA isolated from either BLV-negative purified B cells or the relevant control cell population for each miRNA, and after normalization for 5S rRNA or snRNA U6. Data are presented as relative expression values calculated by the $\Delta\Delta C_T$ method. Reactions were carried out in triplicate. BLV miRNA-cluster expression was computed by calculating the average fold change of all BLV miRNAs. Copy numbers per cell were determined for each BLV miRNA by plotting average tumor-derived C_T values to standard curves generated by serial dilutions of the corresponding synthetic miRNA (Integrated DNA Technologies), assuming 30 pg of total RNA per tumor B cell. Error bars represent SD, and all of the figures were plotted using the statistical software R.

Statistical Analysis. Differential expression analysis of qRT-PCR data were performed using and independent t test. Mean fold change ± SD was used for the representation of qRT-PCR data. HTS data analysis was performed using the statistical software R and the bioconductor package DESEq (16). All data are presented as normalized read counts computed by DESeq. A value of $P \leq 0.05$ was considered statistically significant for differentially expressed miRNA analysis.

BLV miRNA Expression from Plasmid Lacking Pol II Promoter. Sets of primer pairs were designed to amplify each of the individual BLV miRNA hairpins and the entire region containing the miRNAs (listed in Table S7). PCR was carried out using Gotaq DNA Polymerase (Promega) with YR2 genomic DNA as the template. The resultant PCR products were cloned into a pCR2.1 plasmid (Invitrogen). Plasmid DNA was extracted using a Qiagen plasmid mini kit. The inserts in each plasmid were sequenced using the M13 forward and reverse primers with Big Dye, version 3.1 (Applied Biosystems). The purified reaction was analyzed on a ABI Prism 3730 DNA analyzer (Applied Biosystems), and resultant sequence compared with the BLV genome to ensure the correct region was incorporated in the plasmid.

Six-well plates of HeLaT and 293T cells were transfected with 1.6 μg of the appropriate plasmid using Lipofectamine 2000 reagent (Invitrogen). Twenty-four hours after transfection, RNA was extracted from the cells using TRIzol reagent (Ambion). The resultant RNA was treated with TURBO DNA-free (Ambion) to remove any contaminating DNA. qRT-PCR was carried out using custom TaqMan MicroRNA assays (ABI) for BLVmiR-B1-5p, BLVmiR-B2-5p, BLVmiR-B3-5p, BLVmiR-B4-5p, and BLVmiR-B5-5p using a 9700HT (ABI) instrument.

BLV miRNA Expression from Lentiviral Vector. The entire BLV miRNA region was amplified as described in the pCR plasmid cloning method and the PCR product inserted in the HIV-based lentiviral vector pEZX-MR03 (Genecopeia) downstream of EGFP. Control pEZX-MR03 and pEZX-BLVmiR lentiviral expression plasmids were cotransfected into 293T cells with the Lenti-Pac HIV Packaging mix (Genecopeia) and lentivirus-containing supernatants were harvested 48 h after transfection. Viral particles were stored at −80 °C. HCT-116 WT, HCT-116 Dicer$^{-/-}$, 293T WT, and 293T TRBP$^{-/-}$ were transduced in 24-well plates using pEZX-MR03 and pEZX-BLVmiR lentiviral particles with 8 μg/mL polybrene (Sigma) and cultured in the presence of 500 ng/mL puromycin (Sigma). EGFP expression of puromycin-selected transduced cells was assessed by flow cytometry and showed ≥95% purity. Total RNA was extracted from EGFP-positive cells using TRIzol reagent (Ambion) and further processed as described for transfected cells.

Southern Blot and BLV Proviral Load Analysis. Genomic DNA was extracted using Genomic Tip 100 (Qiagen) and analyzed by Southern blot analysis as previously described (5). The nylon-bound EcoRI-digested genomic DNAs were hybridized with a ^{32}P-labeled BLV full-length proviral DNA probe. Total BLV proviral load was measured in DNA extracted from PBMCs by real-time quantitative PCR using *tax*-specific and β-actin primers and probes as described earlier (2). Integration site-specific quantitative PCR was carried out with DNA extracted from PBMCs of M2531 using primer pairs designed for amplifying the tumor-specific integration site of provirus M2531. Each forward primer matches a 5′ genomic proviral flanking sequence located within intron 3–4 of Myc binding protein 2 (MYCBP2), and the reverse primer is within the U3 region of the BLV LTR. Forward primers were as follows: 5′-CAGTGGGGAGTACTTGGGAA-3′ and 5′-AATCCAGTTGGTGTCATAGCA-3′ [Genome Bos Taurus: UCSC UMD 3.1/bosTau6 (17), coordinates 52,722,890–52,722,829 and 52,722,483–52,722,447, respectively]. Reverse primer was as follows: 5′-TCTGGCAGCTGACGTCTCTGT-3′, coordinates 51–71 according to Sagata (10). Amplification and data acquisition for all real-time PCR and qRT-PCR experiments were carried out using an ABI Prism HT7900 Sequence Detector System (Applied Biosystems).

Northern Blot Analysis. Northern blot analysis was carried out using 10 μg of total RNA and hybridization with a BLV Tax probe as previously described (18).

Argonaute Immunoprecipitation. Immunoprecipitation using anti-hAGO2 antibody was performed based on procedures in (19–21). Briefly 12×10^6 ovine YR2 tumor B cells and 8×10^6 HeLa cells (negative control) were resuspended in 850 μL of ice-cold lysis buffer [2.5 mM $MgCl_2$, 200 mM NaCl, 20 mM Tris·HCl, 0.25% Igepal CA-630, 5 μL/mL RNasin (Promega), 5 μL/mL SUPERaseIn (Ambion)]. Cells were then lysed on ice with a glass-Teflon homogenizer in the presence of 1× EDTA-free Complete protease inhibitors (Roche). The solution was then centrifuged for

10 min at 3,000 × *g* to pellet cell debris. Precleaning of the supernatant was carried out with Dynabead Protein A magnetic beads (Invitrogen) coated with rabbit anti-mouse IgG (Jackson ImmunoResearch) and nonimmune mouse IgG (Invitrogen). The mixture was then centrifuged for 10 min at 10,000 × *g*. A portion of the supernatant was retained and stored at +4 °C. The lysate was then immunoprecipitated with Dyna Protein A beads that had been coated with anti-mouse IgG and anti-hAGO2 antibody (Diagenode; 2A8) or nonimmune mouse IgG (Invitrogen) to act as a negative control. Both were incubated overnight at 4 °C with constant rotation. Beads were then separated with a magnet, and the retained cell lysate was kept as a supernatant sample after immunoprecipitation. Beads were washed with lysis buffer containing 0.5% Igepal followed by centrifugation at 500 × *g* for 2 min. RNA was then extracted from the beads with RNeasy FFPE kit (Qiagen) following the manufacturer's instructions. The resultant RNA was treated with Turbo DNA free kit (Ambion) and used for BLV miRNA or miR-34a qRT-PCR analysis. Total RNA isolated from YR2 and BLV-negative PBMCs were used as positive and negative RT-PCR controls, respectively. 5S rRNA was assayed for normalization. The $\Delta\Delta C_T$ method subtracts the ΔC_T for IgG background from the Ago2 ΔC_T, and relative expression refers to the relative enrichment of the target miRNA by Argonaute (Ago) immunoprecipitation (IP) over background IgG IP.

Microarray Analysis of BLV-Induced B-cell Tumors. Gene expression profiles were obtained from cross-species microarray analysis using a previously described procedure (18). Ovine probes were synthesized from total RNA extracted from malignant tissues (B-cell leukemia or lymphoma) or the pooled B-cell fractions purified from PBMCs collected from three healthy sheep as a reference B-cell population for assessing differential expression. Microarray analysis was carried out in triplicate with an additional dye swap for each of the 12 ovine tumor tissues analyzed. HMG-box transcription factor 1 (HBP1) and peroxydasin homolog (PXDN) differential expression was calculated as the mean expression from 12 tumor samples relative to a pool of three control B-cell samples for spotted cDNAs corresponding to human HBP1 or PXDN sequences (NCBI AK025284 and AF200348), respectively. Differential gene expression analysis was performed using an independent *t* test.

1. Willems L, et al. (1992) In vivo transfection of bovine leukemia provirus into sheep. *Virology* 189(2):775–777.
2. Van den Broeke A, et al. (2010) Cytotoxic responses to BLV tax oncoprotein do not prevent leukemogenesis in sheep. *Leuk Res* 34(12):1663–1669.
3. Van Den Broeke A, et al. (1999) In vivo rescue of a silent tax-deficient bovine leukemia virus from a tumor-derived ovine B-cell line by recombination with a retrovirally transduced wild-type tax gene. *J Virol* 73(2):1054–1065.
4. Kettmann R, et al. (1982) Leukemogenesis by bovine leukemia virus: Proviral DNA integration and lack of RNA expression of viral long terminal repeat and 3′ proximate cellular sequences. *Proc Natl Acad Sci USA* 79(8):2465–2469.
5. Merimi M, 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(11):5929–5939.
6. Cummins JM, et al. (2006) The colorectal microRNAome. *Proc Natl Acad Sci USA* 103(10):3687–3692.
7. Griffiths-Jones S (2004) The microRNA Registry. *Nucleic Acids Res* 32(Database issue): D109–D111.
8. Gardner PP, et al. (2011) Rfam: Wikipedia, clans and the "decimal" release. *Nucleic Acids Res* 39(Database issue):D141–D145.
9. Archibald AL, et al.; International Sheep Genomics Consortium (2010) The sheep genome reference sequence: A work in progress. *Anim Genet* 41(5):449–453.
10. Sagata N, et al. (1985) Complete nucleotide sequence of the genome of bovine leukemia virus: Its evolutionary relationship to other retroviruses. *Proc Natl Acad Sci USA* 82(3):677–681.
11. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25.
12. Quinlan AR, Hall IM (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.
13. Lander ES, et al.; International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921, and correction (2001) 411(6838):720..
14. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The Vienna RNA websuite. *Nucleic Acids Res* 36(Web Server issue):W70–W74.
15. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21): 2947–2948.
16. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
17. Zimin AV, et al. (2009) A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol* 10(4):R42.
18. Klener P, et al. (2006) Insights into gene expression changes impacting B-cell transformation: Cross-species microarray analysis of bovine leukemia virus tax-responsive genes in ovine B cells. *J Virol* 80(4):1922–1938.
19. Nelson PT, et al. (2007) A novel monoclonal antibody against human Argonaute proteins reveals unexpected characteristics of miRNAs in human blood cells. *RNA* 13(10):1787–1792.
20. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460(7254):479–486.
21. Takeda H, Charlier C, Farnir F, Georges M (2010) Demonstrating polymorphic miRNA-mediated gene regulation in vivo: Application to the g+6223G->A mutation of Texel sheep. *RNA* 16(9):1854–1863.

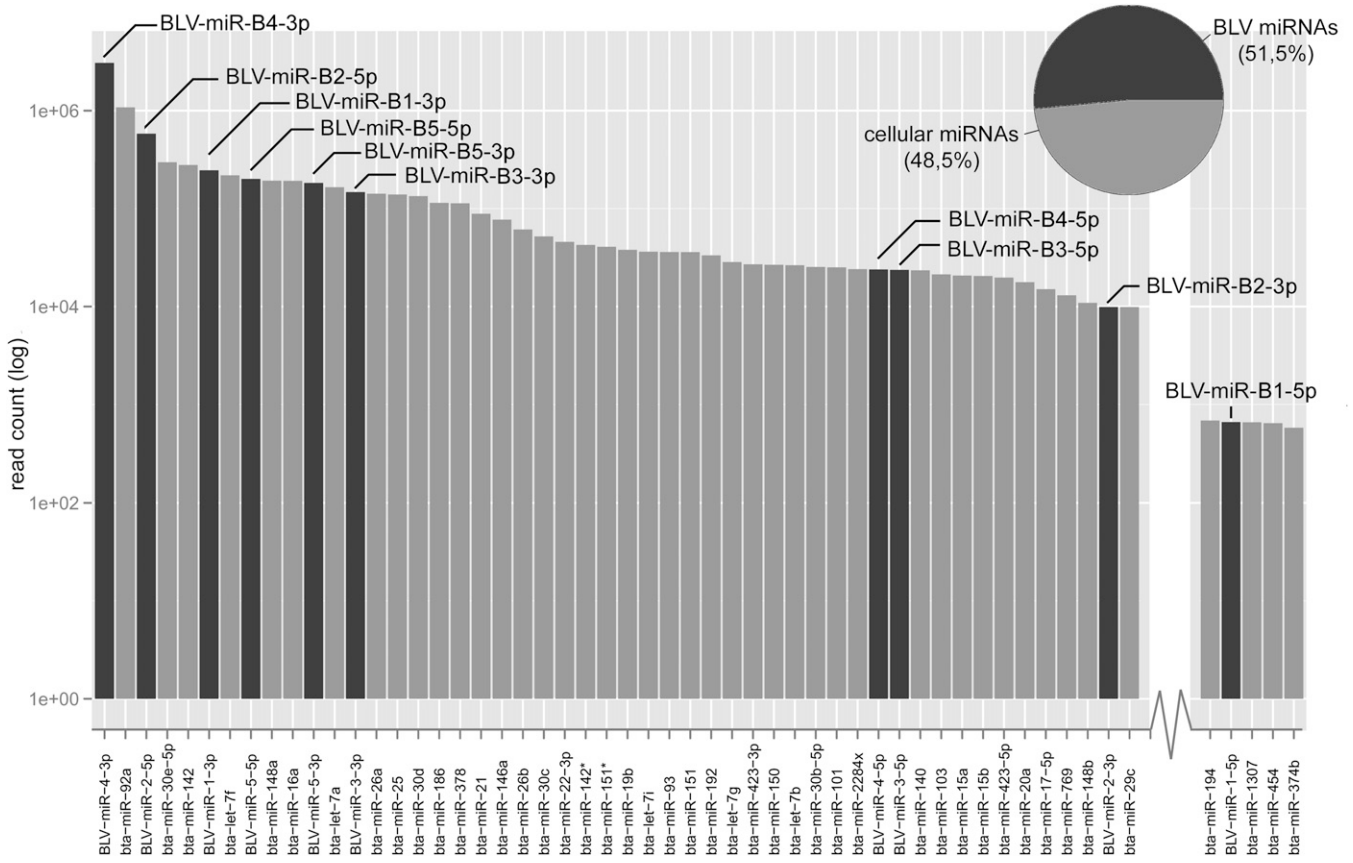

Fig. S1. Relative expression of viral and cellular miRNAs in ovine malignant B-cell lines. miRNA expression is represented by the mean HTS read counts for each miRNA in BLV-carrying tumor B-cell lines YR2 and L267. The 120 most abundant miRNAs are shown. Black bars: BLV miRNAs; gray bars: ovine cellular miRNAs.

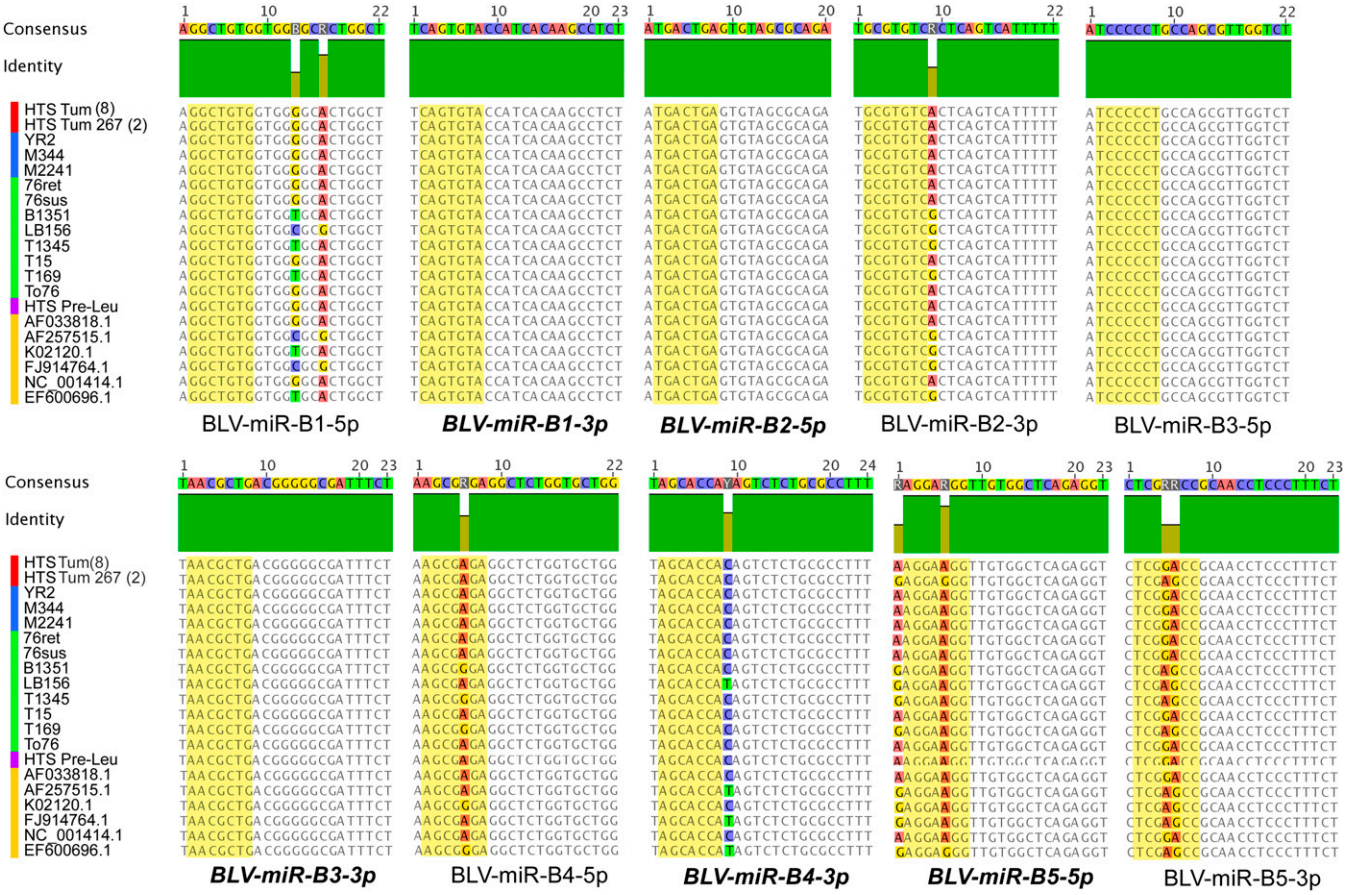

Fig. S2. BLV miRNAs are conserved across ovine and bovine tumor isolates. Alignments of the most commonly sequenced miRNAs in deep sequenced primary tumor samples (HTS Tum), in Sanger sequenced ovine (blue) and bovine (green) primary tumors, in deep sequenced preleukemic B cells (HTS Preleukemia), and in publicly available sequences (yellow). For HTS, sample numbers are shown in brackets. Presumed mature strands are represented in bold. The highlighted regions indicate theoretical miRNA seed regions (positions 2–7). Of note, sequence conservation of the mature BLV miRNAs extends over their entire length, implying that target recognition is not seed-determined only.

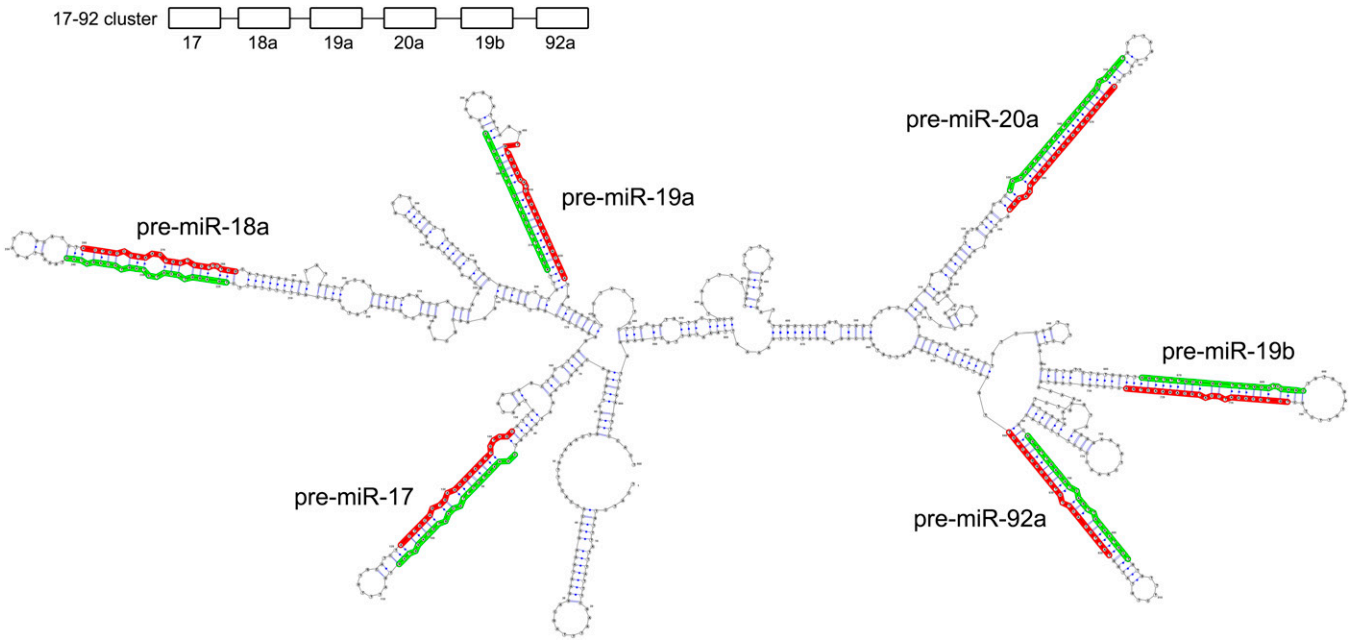

Fig. S3. Mapping of the canonically processed miR 17–92 cluster reads detected in ovine B-cell tumors on the predicted precursor hairpins. RNA secondary structure prediction of the cellular miR cluster 17–92 region and alignments of the ovine B-cell tumor-derived HTS reads (green: 5p strands; red: 3p strands) on the precursor hairpins. Prediction was generated using RNAfold (1).

1. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The Vienna RNA websuite. *Nucleic Acids Res* 36(Web Server issue):W70–W74.

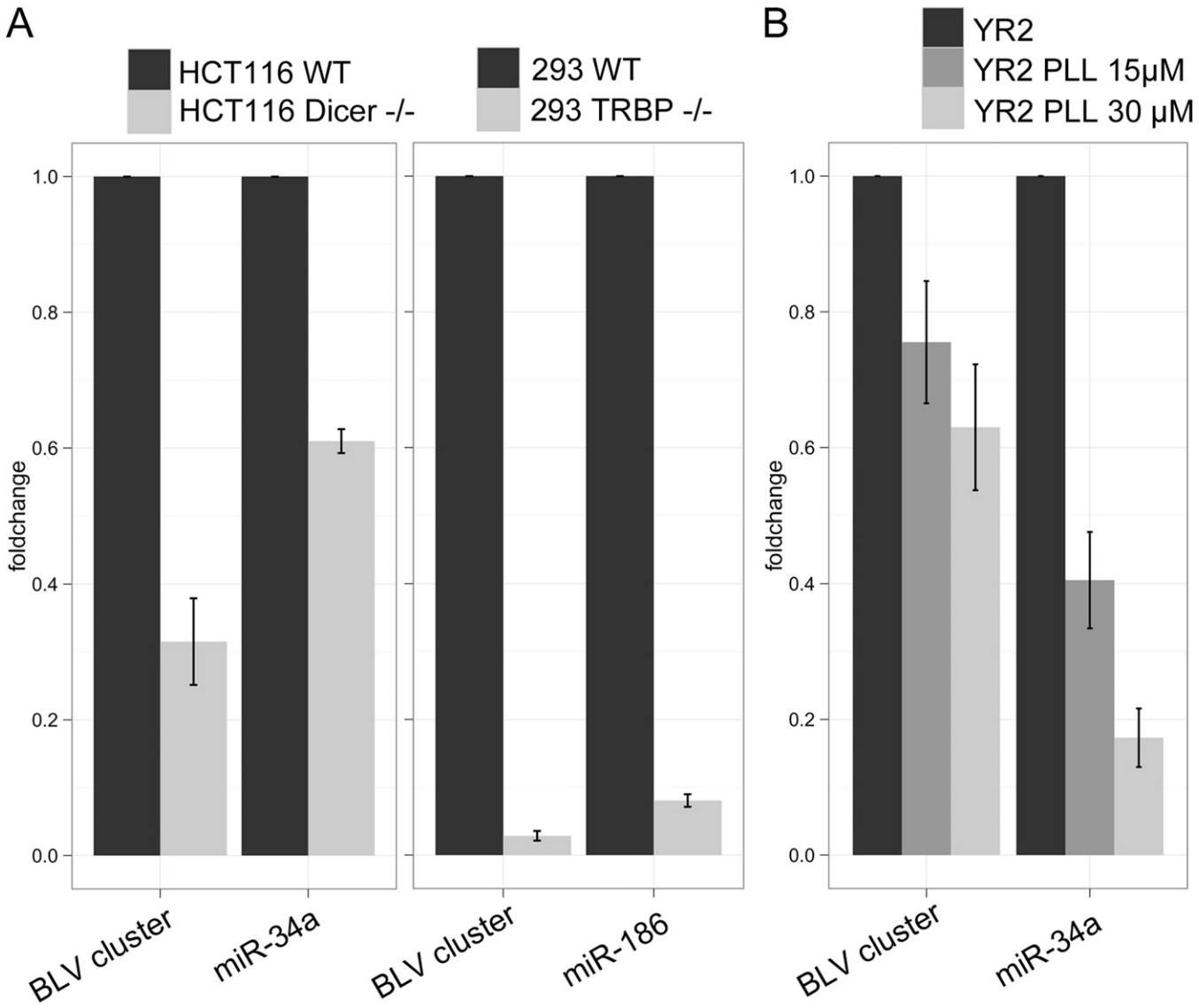

Fig. S4. Tumor-derived BLV miRNAs depend on Dicer cleavage. (*A*) BLV miRNA processing is Dicer-dependent: qRT-PCR analysis of the BLV miRNA cluster delivered into Dicer$^{-/-}$ and TRBP$^{-/-}$ cells (1). The tumor-derived BLV miRNA cluster was cloned within the EGFP 3′-UTR of lentiviral vector pEZX-MR03 (Genecopeia) (*SI Materials and Methods*). Delivery into wild-type HCT116 and 293T cells resulted in significant BLV miRNA expression, creating a context analogous to BLV-infected B cells in vivo. Stable integration of the cluster in either Dicer- or TRBP-deficient cells was associated with a decrease of BLV miRNA levels, confirming a role for Dicer in their processing. Expression in EGFP-positive lentiviral vector-transduced HCT116 Dicer$^{-/-}$ and 293T TRBP$^{-/-}$ cells is shown relative to BLV miRNA-transduced wild-type HTC116 and 293T cells, respectively. Cellular miRNA 186 and 34a were examined as control miRNAs for Dicer dependence in HCT116 and 293T cells, respectively. BLV miRNA-cluster expression was computed by calculating the average fold-change of all examined BLV miRNAs; (*B*) Poly-L-lysine (PLL) treatment confirms Dicer dependence of BLV miRNA processing in tumor B cells: qRT-PCR quantification of BLV miRNAs or cellular miR 34a in YR2 tumor B cells treated with increasing doses of PLL (0, 15, and 30 μM) for 48 h. BLV miRNA cluster expression is shown relative to untreated YR2 cells. BLV miRNA abundance is decreased in a dose-dependent manner. Shown is mean fold change ± SD. See also Table S8 for primer description.

1. Chendrimada TP, et al. (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436(7051):740–744.

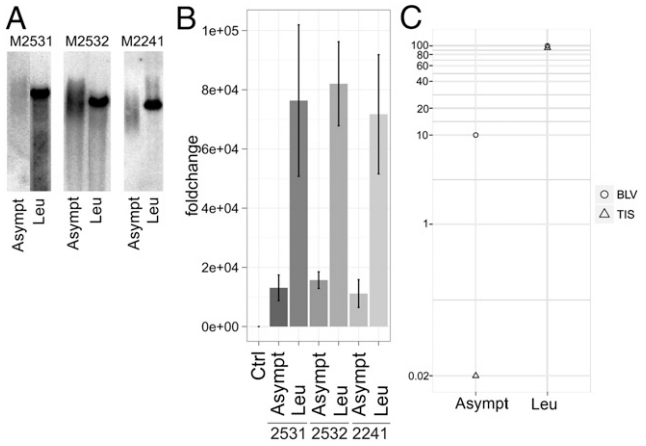

Fig. S5. BLV miRNA expression in asymptomatic animals. (*A*) Southern blot analysis confirms BLV polyclonal integration before tumor development. EcoRI-digested DNA from leukemic and asymptomatic PBMCs (animals M2531, M2532, M2241) probed with a BLV SacI fragment. (*B*) qRT-PCR reveals BLV miRNA expression before tumor development: miRNA levels in B cells from asymptomatic stages relative to matched leukemic B-cell sample (mean fold change ± SD). (*C*) Quantification of the pretumoral clone in asymptomatic M2531 B cells: qPCR of M2531 asymptomatic PBMCs using BLV-specific primers for total viral load or tumor integration site (TIS)-specific primers for pretumoral clone quantification. In leukemic cells, the tumor clone accounts for 95% of total proviral load. Asymptomatic viral load is shown relative to leukemic (set to 100%).

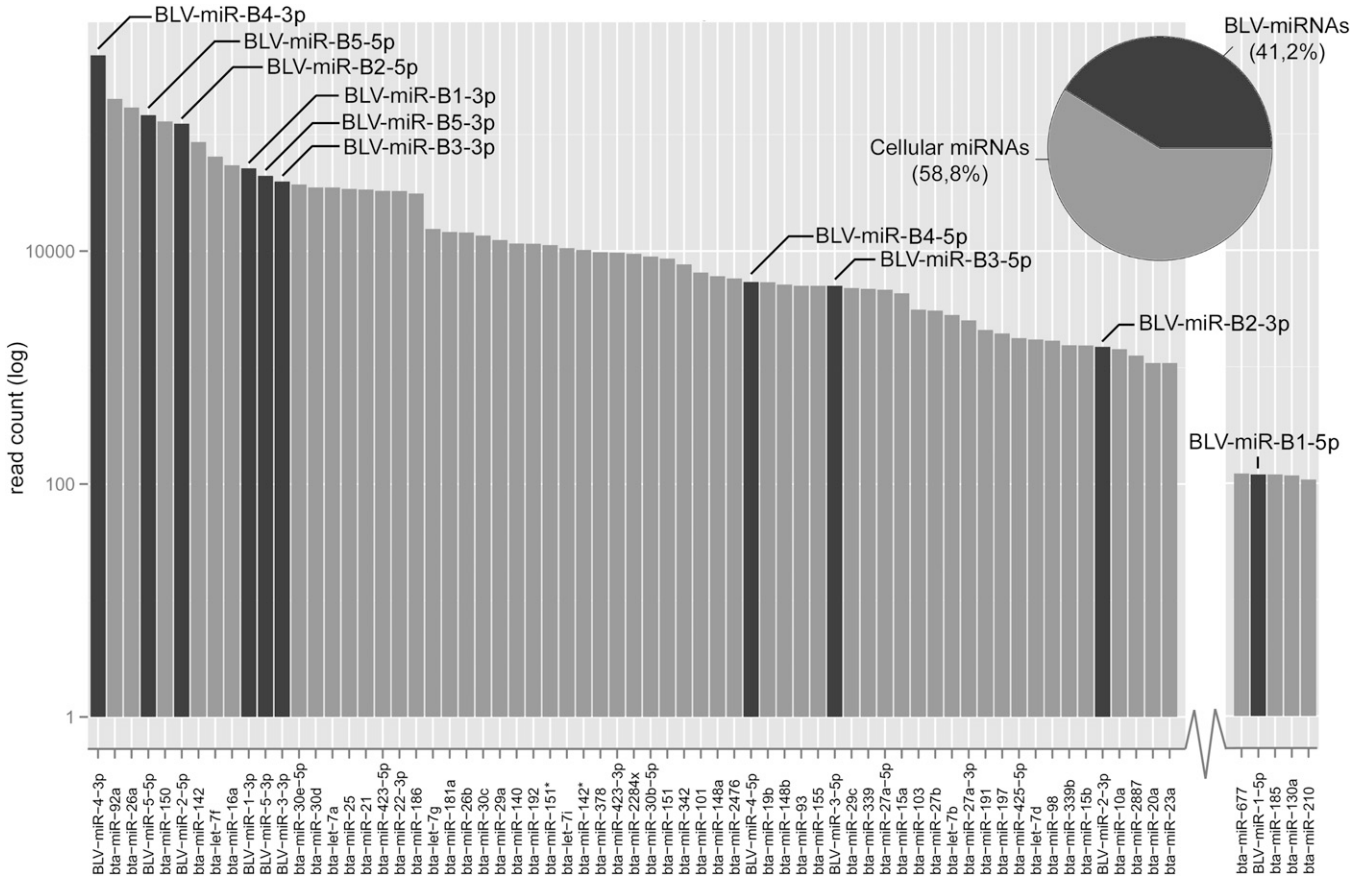

Fig. S6. Relative expression of viral and cellular miRNAs in B cells isolated from asymptomatic animals. Expression is represented by the mean HTS read counts for each miRNA in the purified B-cell fraction of PBMCs collected from three asymptomatic BLV-infected sheep (Table S5). The 120 most abundant miRNAs are shown. Black bars: BLV miRNAs; gray bars: ovine cellular miRNAs.

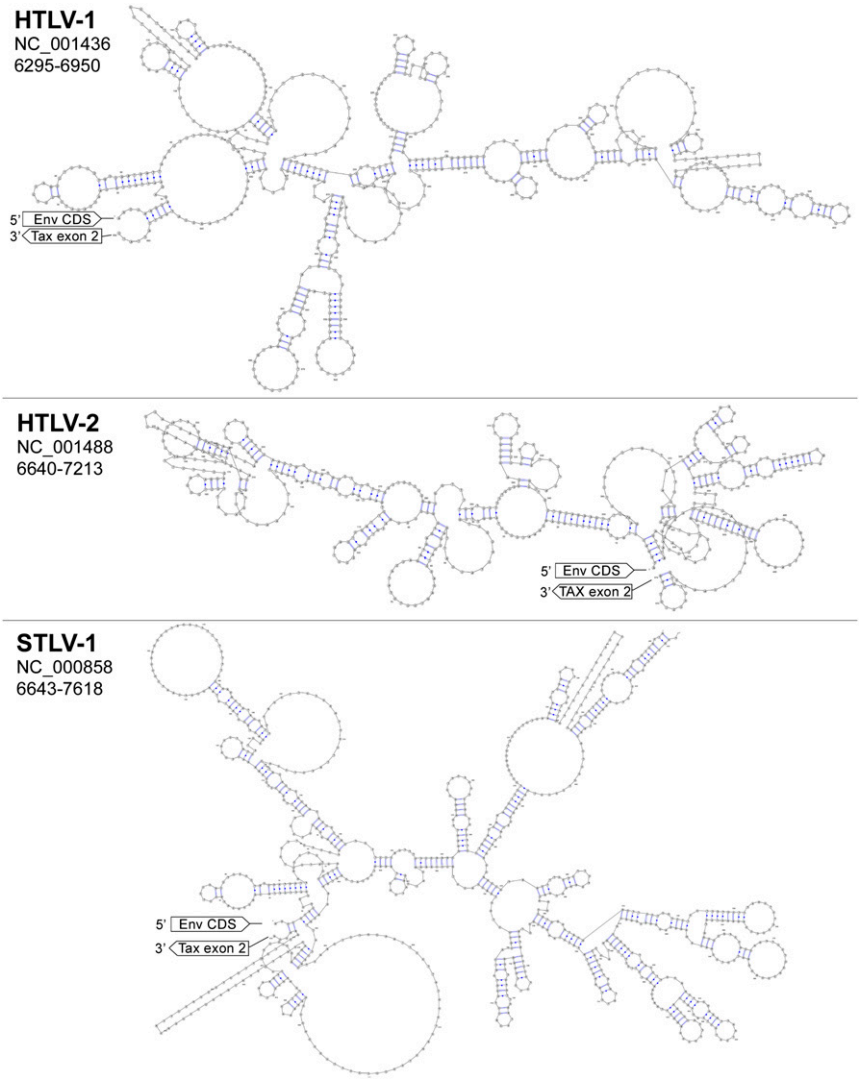

Fig. S7. Predicted RNA secondary structure of BLV miRNA-corresponding regions in human and simian deltaretroviruses. Predictions in HTLV-1, HTLV-2, and STLV, respectively, were generated using RNAfold (1). The genome coordinates of the sequences used for secondary structure prediction are indicated. The prediction does not reveal typical miRNA precursor-like hairpin structures.

1. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The Vienna RNA websuite. *Nucleic Acids Res* 36(Web Server issue):W70–W74.

Table S1. Small RNA sequence analysis of B-cell tumors: High-throughput sequencing (HTS) summary

HTS summary	Read count	%
Total reads	31,982,397	100
Aligned on ovine genome	23,093,532	72.2
Noncoding RNAs	7,737,222	24.19
Cellular miRNAs	3,131,930	9.79
BLV miRNAs	3,113,071	9.73

Summary of HTS data of small RNA libraries obtained from eight primary B-cell leukemia/lymphoma collected from BLV-infected sheep and tumor B-cell lines derived from two of these primary tumors. Read counts correspond to the mean read numbers for the 10 samples. Alignment on the ovine genome was performed using Bowtie. Noncoding RNA sequences and cellular miRNAs were derived from RFAM 10 and miRBase 17, respectively.

Table S2. Small RNA sequence analysis of B-cell tumors: BLV miRNAs

	Primary B-cell tumors (n = 8)				Tumor B-cell lines (n = 2)			
BLV miRNAs	Min	Max	Mean	%	Min	Max	Mean	%
BLV-miR-B1-5p	17	573	220	0.01	299	1,046	659	0.01
BLV-miR-B1-3p	47,177	204,681	97,840	5.63	189,274	331,668	245,100	5.46
BLV-miR-B2-5p	132,807	560,313	266,095	15.3	415,337	825,213	579,736	12.92
BLV-miR-B2-3p	2,225	30,400	11,748	0.68	6,974	12,498	9,871	0.22
BLV-miR-B3-5p	3,628	17,442	7,462	0.43	10,834	47,347	23,679	0.53
BLV-miR-B3-3p	32,197	134,743	74,677	4.29	72,865	245,134	146,827	3.27
BLV-miR-B4-5p	2,952	7,232	4,782	0.28	18,852	32,631	23,938	0.53
BLV-miR-B4-3p	409,398	2,451,420	1,086,331	62.47	2,477,670	4,007,100	3,074,105	68.52
BLV-miR-B5-5p	79,129	196,535	134,381	7.72	192,384	209,342	200,863	4.48
BLV-miR-B5-3p	22,598	104,719	55,399	3.19	178,605	186,053	182,329	4.06
Total			1,738,935	100			4,487,107	100

Shown is a summary of HTS data of BLV-mapping reads. Alignment on the BLV genome was performed using Bowtie and the viral sequence NCBI K02120. Mean read numbers for each subgroup and the minimum/maximum read numbers observed among all samples are listed. For each precursor miRNA, the most abundant processed derivative (mature miRNA) is shown in bold.

Table S3. Sequence conservation of BLV miRNA and Tax sequences

	miR-mat	miR-star	miR-rest	Tax	Tax-3rd
$n = 7$					
miR-mat	0.0068 (113)	0.0721	0.0061	0.0562	0.0007
miR-star		0.0281 (112)	0.2986	0.8356	0.1089
miR-rest			0.0420 (428)	0.0154	0.2665
Tax				0.0258 (930)	0.0028
Tax-3rd					0.0557 (170)
$n = 25$					
miR-mat	0.0028 (113)	0.0214	—	0.0028	<0.0001
miR-star		0.0241 (112)	—	0.8452	0.0121
miR-rest			—	—	—
Tax				0.0256 (930)	<0.0001
Tax-3rd					0.0573 (170)

Conservation analysis using sequence data from seven BLV isolates with both miRNA and Tax sequences ($n = 7$; *Upper*) or 2×25 BLV isolates with either miRNA or Tax sequences ($n = 25$; *Lower*). miR-mat, sequence segments corresponding to the mature BLV miRNAs; miR-rest, sequence segments corresponding to 5′ and 3′ flanking, intermiRNA, and loop sequences in the BLV miRNA region (not available in the second series); miR-star, sequence segments corresponding to the star BLV miRNAs; Tax, full-length Tax coding sequence; Tax-3rd, third codon positions of the segment of Tax that does not overlap with Rex. Diagonal: nucleotide diversity (= average difference per nucleotide site across all pairs of isolates) for the corresponding sequence segment, and—between brackets—number of corresponding nucleotide positions. Off-diagonal: *P* value of the differences between corresponding nucleotide diversities.

Table S4. BLV miRNA and cellular cluster miR 17–92 biogenesis products in B-cell tumors

	5′ mature			3′ mature			5′ subproduct			Loop			3′ subproduct			premiRNA			moR		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
BLV-premiR-B1	17	1,046	440	47,177	331,668	171,470	0	20	4	19	73	47	1	7	3	0	5	2	0	0	0
BLV-premiR-B2	132,807	825,213	422,916	2,225	30,400	10,810	11	256	72	151	1,074	467	0	19	7	0	6	4	0	0	0
BLV-premiR-B3	3,628	47,347	15,570	32,197	245,134	110,752	12	313	78	65	215	135	3	20	12	0	1	1	0	0	0
BLV-premiR-B4	2,952	32,631	14,360	409,398	4,007,100	2,080,218	2	35	13	126	16,776	4,263	1	22	9	0	2	1	0	0	0
BLV-premiR-B5	79,129	209,342	167,622	22,598	186,053	118,864	69	1,350	763	0	6	2	14	73	42	0	2	1	0	0	0
premiR-17	3,844	24,754	13,194	303	1,883	953	0	1	1	5	35	18	0	1	1	0	0	0	51	983	288
premiR-18a	319	6,933	2,771	35	380	177	0	1	1	0	2	1	0	4	1	0	12	4	0	16	5
premiR-19a	16	114	52	2,396	12,759	7,400	0	0	0	0	3	1	0	0	0	0	0	0	39	257	127
premiR-20a	4,434	24,395	14,433	33	308	144	0	1	1	4	24	9	0	1	1	0	0	0	323	853	533
premiR-19b	21	138	72	11,298	52,351	31,198	0	1	1	11	128	47	0	3	1	0	0	0	27	130	76
premiR-92a	161	2,254	878	301,970	1,343,465	750,204	0	3	2	4	43	14	0	6	3	0	0	0	6	42	18

BLV miRNA and cellular cluster miR 17–92 processing products were detected by deep sequencing of small RNA libraries prepared from ovine B-cell tumors using a modified method (*SI Materials and Methods*). Sequencing runs were extended to 76 cycles and broader windows of small RNA sizes were captured. Numbers represent minimum, maximum, and mean read counts from 10 B-cell tumor samples.

Table S5. Characterization of B-cell samples collected from asymptomatic BLV-infected sheep

Sheep no.	Tumor onset	Time of PBMC collection		Clonality	B cells in PBMC, %	B cells in purified fraction, %	Viral load in PBMC	Viral load in purified B-cell fraction, %
M2531	20 m	11 m pi	9 m pt	Poly	55	89	10	16.1
M2532	29 m	20 m pi	9 m pt	Poly	62	93	15	22.5
M2241	48 m	20 m pi	28 m pt	Poly	66	95	13	18.7

Time of tumor onset is shown in months (m) postinoculation (pi). PBMC collection in asymptomatic animals (WBC counts in normal range) is indicated in months pi and prior tumor onset (pt). % B cells corresponds to the fraction of sIgM-positive cells as determined by fluorescence-activated cell sorting. B cells were purified using a magnetic bead-mediated negative selection procedure (*SI Materials and Methods*). Viral load was estimated by Southern blot and q-PCR as previously described (1). Small RNA libraries were prepared from RNA isolated from the purified B-cell fraction of PBMCs from asymptomatic BLV-infected sheep.

1. Van den Broeke A, et al. (2010) Cytotoxic responses to BLV tax oncoprotein do not prevent leukemogenesis in sheep. *Leuk Res* 34(12):1663–1669.

Table S6. B-cell samples collected from asymptomatic BLV-infected sheep: Deep sequencing read counts

Deep sequencing	Read count
BLV-miR-B1-5p	167
BLV-miR-B1-3p	76,655
BLV-miR-B2-5p	184,512
BLV-miR-B2-3p	2,188
BLV-miR-B3-5p	7,423
BLV-miR-B3-3p	57,904
BLV-miR-B4-5p	8,024
BLV-miR-B4-3p	703,298
BLV-miR-B5-5p	215,877
BLV-miR-B5-3p	64,354

Deep sequencing read counts represent the average read numbers for the three nonmalignant samples. Mature miRNAs are represented in bold.

Table S7. BLV miRNA cloning primers

Primer name	Forward	Reverse	BLV genome
BLV-miRs	CCCCTAAACCCGATTCTGAT	GGGCTTGTTACATGGGAAGA	6,053–6,845
BLV-miR-1	CCCGATTACATCAACCTTCG	CATGCACTCAGAGAGGCTCA	6,130–6,289
BLV-miR-2	GAACACAGCCCTACCCTGAG	GGGATGGTGAAGCAGAAGTG	6,265–6,430
BLV-miR-3	TTGCCTGTGACACGGTTAAG	TCTTATCCCCAGCACCAGAG	6,383–6,524
BLV-miR-4	GCGATTTCTTGCAGCTGTG	GTGCAACAGGGCGTAAAAAG	6,474–6,607
BLV-miR-5	ATGCAGCTTCCGCTTTTTAC	AACCTCCGTTAGGGGGTCTA	6,576–6,760

Table S8. BLV miRNA QuantiMir qRT-PCR primers (DNA)

Primer name	Sequence
BLV-miR-B1-5p	AGGCTGTGGTGGGGCACTGGCT
BLV-miR-B1-3p	TCAGTGTACCATCACAAGCCTCT
BLV-miR-B2-5p	ATGACTGAGTGTAGCGCAGA
BLV-miR-B2-3p	TGCGTGTCACTCAGTCATTTT
BLV-miR-B3-5p	ATCCCCCTGCCAGCGTTGGTCT
BLV-miR-B3-3p	TAACGCTGACGGGGGCGATTTCT
BLV-miR-B4-5p	AAGCGAGAGGCTCTGGTGCTGG
BLV-miR-B4-3p	TAGCACCACAGTCTCTGCGCCTTT
BLV-miR-B5-5p	AAGGAAGGTTGTGGCTCAGAGGT
BLV-miR-B5-3p	CTCGGACCGCAACCTCCCTTTCT
hsa-miR-34a-5p	TGGCAGTGTCTTAGCTGGTTGT