The Nucleotide Sequence of Koala (Phascolarctos cinereus) Retrovirus: a Novel Type C Endogenous Virus Related to Gibbon Ape Leukemia Virus

JON J. HANGER,1 LINDELL D. BROMHAM,2 JEFF J. MCKEE,1,* TRACY M. O’BRIEN,1 AND WAYNE F. ROBINSON

Division of Veterinary Pathology and Anatomy1 and Department of Zoology,2 University of Queensland, 4072 Brisbane, Australia

Received 28 May 1999/Accepted 8 February 2000

A novel retrovirus, morphologically consistent with mammalian C-type retroviruses, was detected by electron microscopy in mitogen-stimulated peripheral blood mononuclear cell cultures from 163 koalas and in lymphoma tissue from 3 koalas. PCR amplified provirus from the blood and tissues of 17 wild and captive koalas, and reverse transcriptase-PCR demonstrated viral mRNA, viral genomic RNA, and reverse transcriptase activity in koala serum and cell culture supernatants. Comparison of viral sequences derived from genomic DNA and mRNA showed identity indicative of a single retroviral species—here designated koala retrovirus (KoRV). Southern blot analysis of koala tissue genomic DNA using labelled KoRV probes demonstrated banding consistent with an endogenous retrovirus. Complete and apparently truncated proviruses were detected in DNA of both clinically normal koalas and those with hematopoietic disease. KoRV-related viruses were not detected in other marsupials, and phylogenetic analysis showed that KoRV paradoxically clusters with gibbon ape leukemia virus (GALV). The strong similarity between GALV and KoRV suggests that these viruses are closely related and that recent cross-host transmission has occurred. The complete proviral DNA sequence of KoRV is reported.

Retroviruses comprise a large group of diploid RNA viruses that rely on reverse transcription as an essential part of their life cycle. Endogenous retroviruses (ERVs) and exogenous retroviruses have been demonstrated in a wide range of vertebrate species, and while some, particularly the exogenous viruses, may be pathogenic, many are not associated with disease (23, 37). ERVs, by definition, have become incorporated into the host genome by integration into germ line cells or early embryos and are transmitted as dominant Mendelian alleles. Exogenous (infectious) retroviruses are transmitted horizontally (3, 35).

The etiological role of retroviruses in a range of different diseases has been demonstrated in a number of vertebrate species (47). In some host species, interaction between ERVs and related exogenous retroviruses (helper viruses) is important in disease pathogenesis. For example, some feline leukemia virus (FeLV)-associated diseases of cats result from recombination between endogenous FeLV and exogenous FeLV strains (52). The presence of endogenous viruses may reduce the immune response to related pathogenic exogenous retroviruses, increasing the likelihood of disease, as demonstrated in chickens with avian leukosis or sarcoma virus infection (53). In contrast, expression of ERV genes may confer resistance to exogenous pathogenic viruses through receptor interference and other mechanisms (25, 45).

Lymphoma and leukemia have been long recognized as the most common form of neoplasia in both captive and free-living koalas. Mortality surveys of wild koalas indicate that these diseases account for around 3 to 5% of deaths in free-living koalas in the survey areas of New South Wales and southern Queensland (2, 7, 11, 24, 39, 55). However, anecdotal evidence from fauna parks in southeast Queensland suggests that up to 80% of mortalities in captive koalas may be attributable to lymphoma and a variety of leukemias (J. J. Hanger, unpublished data). The possible involvement of retroviruses in these diseases of koalas was suggested by a number of workers (7, 24). Type C retrovirus-like particles were seen by transmission electron microscopy (TEM) in tissues from a leukemic koala (8), in the blood of captive koalas abroad (61), and in mitogen-stimulated peripheral blood mononuclear cells (PBMCs) derived from 18 koalas of mixed clinical status (46). Retrovirus infection in koalas was confirmed with a report of the isolation and partial gene sequence of a retrovirus with homology to gibbon ape leukemia virus (GALV) and simian sarcoma virus (SSV) in both diseased and healthy koalas (41). Recently, sequence homologous to the pol region of murine leukemia virus (MLV)-related and GALV-related viruses was detected in the DNA of a wild-caught koala, but no clinical information was reported (37). We report the complete nucleotide sequence, phylogenetic analysis, and characterization of a novel type C ERV detected in koalas.

MATERIALS AND METHODS

PCR. Genomic DNA was extracted from the blood and tissues of four wild and four captive koalas by methods previously described (50). Total RNA or polyadenylated RNA was extracted from tissues and blood by using either Trizol reagent (Life Technologies) or the Quick-Prep Micro mRNA Purification kit (Pharmacia Biotech), respectively. Koalas from which tissues were taken had either died or had been euthanized because of serious disease or injuries. Four koalas (one wild and three captive) from which tissues or blood were taken had leukemia or lymphoma. DNA for PCR and Southern hybridization controls was extracted from tissue samples from a mediastinal lymphoma in a dog, bone marrow from an eastern grey kangaroo (Macropus giganteus), and blood from a common wombat (Vombatus ursinus) and a southern hairy-nosed wombat (Lasiorhinus latifrons).

Initial PCR amplification of a 3-kb koala retrovirus (KoRV) sequence from genomic DNA was accomplished with an upstream primer, tRNA5′ (CAT TTG GGG CTC GTC CGG GAT), which annealed to the tRNA5′ primer binding

* Corresponding author. Mailing address: Division of Veterinary Pathology and Anatomy, University of Queensland, 4072 Brisbane, Australia. Phone: 61 7 33652118. Fax: 61 7 33651355. E-mail: j.mckee@mailbox.uq.edu.au.
FIG. 1. Graphic representation of KoRV provirus showing positions of primer pairs used to amplify proviral DNA. The primer pair USF-USR flanks the site, and a downstream degenerate pro-pol primer, JPU2r [GC(ACC) GCC A(GA)(CTG) A(GA)(GT) A(GA)(GT) A(GA)G TC(AG) TC] (bases in parentheses represent degenerate bases, described previously (10)). PCR was performed in 50-μl reaction mixtures with the Expand Long Template PCR system (Boehringer, Mannheim, Germany) in accordance with the manufacturer’s recommendations. Cycling conditions were as follows: initial denaturation at 94°C for 1 min, 1 cycle; denaturation at 94°C for 30 s, annealing at 60°C for 1 min, and extension at 68°C for 2 min, 10 cycles; followed by cycles of denaturation at 94°C for 30 s, annealing at 60°C for 1 min, and extension at 68°C for 45 s plus 20 s for each consecutive cycle up to 2 min. A further 10 cycles with the same temperatures and 2-min extension times were performed followed by a final extension at 68°C for 7 min.

Touchdown PCR (46) was performed in 50-μl reaction mixtures using the Expand Long Template PCR system (Boehringer, Mannheim) in accordance with the manufacturer’s recommendations. Two pairs of primers were used to amplify the putative koala provirus. One pair was designed to amplify the entire provirus, including both 5′ and 3′ long terminal repeats (LTRs): the forward primer USF (AA(T)GAA GGC AGA AAT CATG AGC) was designed to anneal to the first 26 bases of the U3 region of the LTR, and the reverse primer USR (GTT AGT CCT CTG ACC TGG AGA) annealed to the last 21 bases of the U5 region (Fig. 1). The second pair of primers was designed to exclude repeated regions from the amplicon: the forward primer KrV1 (GC(G)CAG GCC CTG CTC CTG GAA GAC(T)G) annealed to the first 24 bases of the 5′-repeated region (H) and the reverse primer KrVr (GCC CCTG TAT ATT CTG TCT ACT TT) annealed to the last 21 bases of the 3′ U3 region (Fig. 1). Cycling conditions for both primer pairs were as follows: denaturation at 94°C for 2 min (1 cycle); denaturation at 94°C for 30 s, annealing at 68°C for 1 min, and extension at 68°C for 4 min (2 cycles); followed by cycles of similar cycles, each with annealing at 66, 64, 62, and 60°C; and then 10 cycles with annealing at 58°C. This was followed by five similar cycles with annealing at 60°C with 3-min extension times for each of the cycles, and then four cycles at the same temperatures with a 4-min extension time. The final extension at 68°C was for 7 min.

Cloning and sequencing. PCR bands were resolved on a 0.9% agarose gel and then excised and extracted with the QiAquick Gel Extraction kit (Qiagen). The fragments were cloned with the pGEM-T Easy Vector system (Promega), and automated sequencing was performed with M13 forward and reverse primers (Promega) or virus sequence-specific primers by using the BigDye Terminator method described previously (60). PBMCs were cultured in RPMI (Trace Sciences International), supplemented with 10% fetal bovine serum (FBS), 10% phytohemagglutinin (PHA), and 1,000 IU of human recombinant interleukin-2 per ml. Supernatants were harvested at between days 8 and 12 of culture.

RT-PCR. Synthesis of cDNA from mRNA extracted from the blood or tissues was performed with the Universal Riboclone cDNA Synthesis kit (Promega). PCR of cDNA was performed with the primer pairs RNAi (described above)—gag, pol, and env—PCR (Fig. 2), and the cycling conditions were those described above for initial amplification of 3-kb koala provirus products. Viral RNA was extracted from cell-free culture supernatant and koala serum by using the Viral RNA Extraction kit (Qiagen) in accordance with the manufacturer’s recommendations. PBMCs were seeded at 1,000 cells per ml of culture medium containing 100 IU of human recombinant interleukin-2 per ml. PBMCs were cultured in 24-well plates in 1 ml of culture medium containing 100 IU of human recombinant interleukin-2 per ml for 20 min. Hundred fifty microliters of virus lysis buffer (0.5% Triton-X 100, 20% glycerol, 0.5 mM phenylmethylsulfonyl fluoride, 800 mM NaCl, 50 mM Tris [pH 7.5]) was added to each precipitate, which was then incubated at 4°C for 30 min. Samples were vortexed briefly, and 40 μl was added to 90 μl of RT reaction buffer [final concentration: 0.1 M Tris (pH 8.3), 0.02 M dithiothreitol, 1 mM MgCl₂, 0.1% Nonidet P-40, 0.1 M NaCl, 5 mM dNTP, 1 U of pol(α)– oligo(dT)₁₄–₁₈ primer (Promega). Samples were incubated in quadruplicate in 96-well plates for 4 h at 37°C. Trichloroacetic acid (12.5 μl) was added, and samples were stored at 4°C for 30 min to precipitate the H⁻TTP. Samples were then harvested onto glass fiber discs, placed into 2 ml of scintillation fluid, and counted.

Conclusively, a more sensitive RT-PCR-based RT assay (38) was used by substituting the pol(α)– oligo (dT) with 40 ng of kanamycin resistance gene mRNA (Promega), added as a template for reverse transcription, and 30 ng of oligo(dT)₁₂–₁₈ primer (Promega). Serum samples (25 μl) from 31 animals were tested in this incubation for 4 h. Ten Aliquots were used in a PCR with kanamycin template-specific primers and reagents in the Access RT-PCR kit (Promega) in accordance with the manufacturer's instructions.
FIG. 2. Complete nucleotide sequence of a KoRV provirus. The boundaries of the major sections of the genome are shown by arrowheads: 5' LTR (U3-R-U5, boundaries based on GALV [GenBank accession no. M26927]), gag, pro-pol, env, and 3' LTR. Sequences important in transcription and reverse transcription are shown in boxes: CAAT box, TATA box, Cys-His box, polyadenylation signal (PAS), polypurine tract (PPT) and tRNA pro primer binding site (PBS). Primer sequences described in the text are underlined. Note that nucleotides 1 to 8411 were sequenced from one clone. Nucleotides 8412 to 8431 were inferred from the 5' LTR.
turer’s instructions. PCR bands were resolved on a 2% agarose gel, purified with the Qiagquick Gel Extraction kit (Qiagen), and sequenced as described above.

Electron microscopy. Tissues for TEM were fixed in 3% glutaraldehyde in cacodylate buffer, postfixed in 1% osmium tetroxide, prepared by standard methods, and examined electron microscopically. PBMCs from the cell cultures were fixed in 3% glutaraldehyde in cacodylate buffer, embedded in 3% agarose, and then prepared as described above.

Nucleotide sequence accession number. The complete provirus sequence reported here is contained in the GenBank nucleotide sequence database under accession no. AF151794.

RESULTS

Molecular characterization. Initial PCR amplification of a 3-kb KoRV sequence from genomic DNA was accomplished with an upstream primer (tRNAf) which annealed to the tRNApro primer binding site and a downstream degenerate universal pro-pol primer (JPU2r) described previously by Colbatzky and Jacobs (10) (data not shown). Sequence analysis of this amplimer demonstrated close similarity to GALV sequences (13), which enabled the design of primers flanking other regions of the putative koala provirus. Definitive KoRV primers, which flank the entire provirus, were designed based upon sequences detected by these methods. All 17 koalas amplified KoRV sequence from blood or tissue with a variety of KoRV-specific primers. PCR amplification of sequences with primers U3F and U5R resulted in expected bands of approximately 8.4 kb, representing putative full-length proviral amplimers, and 0.5 kb, representing amplification of the LTRs only. Primers KRVf and KRVr amplified products of approximately 8 kb, representing full-length provirus amplimers without repeated regions of the LTRs. However, a number of bands of intermediate size (4 to 7 kb) were consistently amplified with both primer pairs in PCRs. PCR analysis of multiple tissues from one koala (without hematopoietic disease) showed identical banding patterns between tissues. However, among koalas, all showed 8-kb (primers KRVf and KRVr) or 8.4-kb (primers U3F and U5R) bands, but the sizes and number of smaller bands varied. Southern hybridization of PCR blots with a KoRV pro-pol probe detected the 8.4- and 8-kb bands as well as some of the smaller bands, but not the 0.5-kb bands representing LTR-only amplimers (Fig. 3B). No PCR bands were obtained from tissues taken from a dog or a kangaroo (Fig. 3A). With primers tRNAf and JPU2r, a weak 3-kb band was amplified from DNA derived from the blood of a common wombat and a southern hairy-nosed wombat. Subsequent sequencing showed that these amplimers were not derived from retrovirus.

Sequence analysis of clones of the 8.4- and 8-kb fragments revealed a complete provirus of 8,431 bp with the structure of a simple type C mammalian retrovirus. Common features of

![FIG. 3. (A) PCRs using primer pairs KRVf-KRVr and U3F-USR resolved on a 1% agarose gel stained with ethidium bromide. Lanes: 1 and 9, 1-kb-ladder molecular weight markers; 2, no-DNA PCR control; 3, kangaroo DNA PCR control; 4, lymphoma DNA amplified with primer pair KRVf-KVRr; 5, blood DNA from the same koala amplified with the same primers; 7 and 8, the same samples amplified with primer pair U3F-USR, showing amplimers representing full-length and truncated proviruses, as well as LTRs only (0.5-kb bands); 10 and 11, KRVf-KRVr amplimers from blood DNA from two different koalas. (B) Southern blot of a similar PCR gel probed with a KoRV pol probe. Lanes: 1 digoxigenin-labelled molecular weight marker; 2, kangaroo DNA PCR control; 3, dog lymphoma DNA PCR control; 4, koala DNA U3F-USR PCR products; 5, koala DNA KRVf-KRVr PCR products.](http://jvi.asm.org/)

<table>
<thead>
<tr>
<th>Sequence compared</th>
<th>Length (nucleotides)</th>
<th>Nucleotide % identity</th>
<th>Amino acid residue no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provirus KoRV</td>
<td>8,431</td>
<td>78</td>
<td>NA*</td>
</tr>
<tr>
<td></td>
<td>GALV</td>
<td>8,513</td>
<td></td>
</tr>
<tr>
<td>LTR KoRV</td>
<td>505</td>
<td>69</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>GALV</td>
<td>494</td>
<td></td>
</tr>
<tr>
<td>gag KoRV</td>
<td>1,566</td>
<td>82</td>
<td>522</td>
</tr>
<tr>
<td></td>
<td>GALV</td>
<td>1,560</td>
<td>520</td>
</tr>
<tr>
<td>pol KoRV</td>
<td>3,384</td>
<td>82</td>
<td>1128</td>
</tr>
<tr>
<td></td>
<td>GALV</td>
<td>3,384</td>
<td>1128</td>
</tr>
<tr>
<td>env KoRV</td>
<td>1,980</td>
<td>74</td>
<td>660</td>
</tr>
<tr>
<td></td>
<td>GALV</td>
<td>1,995</td>
<td>665</td>
</tr>
</tbody>
</table>

* NA, not applicable.
retroviral genomes were identified, including 5' and 3' LTRs, a tRNA\textsuperscript{PRO} primer binding site; a CAAT box; a TATA box; gag, pro-pol, and env coding sequences; a Cys-His box (protein translation, CAYCKERGHWAREC), a polypurine tract, and a polyadenylation signal (Fig. 2). There are three ORFs on two forward-reading strands corresponding to gag, pol, and env ORFs with the typical structure of gag and pro-pol in the same frame, while the 5' end of env overlaps the 3' end of pol in an adjacent frame. A splice donor site was located at bases 564 to 570, and a presumptive splice acceptor was located at bases 5701 to 5709 (based upon sequencing of a spliced env transcript). There are several small ORFs, with predicted protein lengths greater than 30 amino acids, whose significance, if any, has yet to be determined. Partial sequence analysis of some of these clones was not performed, and the deleted regions were not defined.

KoRV sequence comparison (GCG) demonstrated a 78% nucleotide similarity across the complete genome between KoRV (8,431 bp) and GALV (8,513 bp; M26927) and strong similarity across the three major coding regions with GALV and KoRV (8,431 bp) and GALV (8,513 bp; M26927) and strong similarity across the three major coding regions with GALV (8,513 bp; M26927) and strong similarity across the complete genome between GALV (8,513 bp; M26927) and strong similarity across the complete genome between GALV (8,513 bp; M26927) and strong similarity across the complete genome between KoRV (8,431 bp) and GALV (8,513 bp; M26927) and strong similarity across the three major coding regions with GALV

All phylogenetic analyses grouped the KoRV sequence with GALV sequences, to the exclusion of PERV and the MLV group. Figure 4 shows the maximum-likelihood tree for the pol gene. The maximum likelihood trees for the env and gag sequences were the same as that shown in Fig. 4. Separate analysis of an alignment of only the 1st and 2nd codon positions of pol also gave the tree in Fig. 4, as did an alignment of only 3rd codon positions. Spectral analysis indicated strong support for a KoRV-GALV clade, with little conflicting signal (Fig. 5).

Multiple sequence alignment of two full and two partial (spanning 3 kb of gag-pro-pol) provirus sequences from four different koalas showed greater than 98.5% identity between all sequences (data not shown). Three of the koalas were captive koalas from southeast Queensland, but were unrelated, and one wild koala was a resident in a geographically isolated region of New South Wales. RT-PCR of mRNA extracted from the blood of a clinically normal captive koala, with three primer pairs spanning the retroviral genome from the tRNA primer binding site to the 3' LTR, resulted in bands of the expected length with 98.5% sequence identity to proviral sequences (data not shown). RT-PCR of the same sample with primer pair KRVf-KRVr (Fig. 1) amplified a band of approximately 2.6 kb, which sequencing showed to represent spliced env transcripts.

Southern hybridization of genomic DNA digested with PstI from a range of tissues taken from two koalas and hybridized with a gag-pro-pol-specific probe demonstrated a banding pattern, as well as smearing in the 3- to 12-kb range (Fig. 6). One koala was affected by lymphoid neoplasia, which involved multiple tissues; the other was not affected by neoplastic or hematopoietic disease. Probes did not bind to DNA from lymphoma tissue from a dog or bone marrow from an Eastern Grey kangaroo.

![FIG. 4. Maximum likelihood phylogeny for the pol gene of KoRV, GALV or GALVX, MLV-type viruses (MMLV, FMLV, and FeLV), and PERV. The transition/transversion ratio (3.1) and gamma shape parameter (0.44) were estimated from the data [D. L. Swofford, PAUP*: Phylogenetic Analysis Using Parsimony (and other methods), version 4. Sinauer Assoc., Sunderland, Mass., 1999]. The results are the same for the gag and env genes and for all alternative alignments of the pol gene (see Materials and Methods).](http://jvi.asm.org/)

![FIG. 5. Spectral plot showing the relative support for branches in the pol phylogeny (Fig. 4). The frequency of patterns in the sequence data that support a partition (clade) is given above the x axis, and the frequencies of patterns that support a conflicting bipartition are given below. The KoRV-GALV clade has strong support and little conflicting signal. Spectral analysis was performed with the Spectrum package (9).](http://jvi.asm.org/)

![FIG. 6. Southern blot of koala genomic DNA digests. Lanes: 1, digoxigenin-labelled molecular weight marker; 2, dog genomic DNA digest (control); 3, kangaroo genomic DNA digest (control); 4 to 7, DNA digests from a koala with lymphoma: 4, 10 μg of bone marrow DNA digested with PstI; 5, 10 μg of bone marrow DNA digested with HindIII; 6, 10 μg of lymphoma DNA digested with HindIII; 7, 10 μg of lung DNA digested with HindIII; 8, 2 μg of bone marrow DNA digested with HindIII from a koala without lymphoma.](http://jvi.asm.org/)
**Viral RNA.** Viral RNA extracted from either serum or cell-free PBMC culture supernatant was reverse transcribed and analyzed by PCR with the primer pairs gagI-gagII, polI-polII, and envI-envII (Fig. 2). Of the 10 koalas examined, all three primer pairs amplified bands of the expected lengths in 2 koalas (in both serum and PBMC culture supernatant), but in 3 koalas, only amplifiers of gag and env or only env was detected. Viral RNA was not detected by RT-PCR in serum or supernatant in one koala; however, in another koala, pro-pol transcripts were detected in bone marrow, but not in blood or serum.

**RT assay.** RT assays with supernatants from PBMCs cultured in vitro from nine koalas, with and without lymphoid neoplasia, showed optimal RT activity in the presence of manganese. When the conventional 3H-TTP incorporation RT assay was used, weak activity levels were demonstrated; culture supernatants from PBMCs cultured in vitro from nine koalas, showed a mean activity of 5,663 dpm (standard deviation [SD], 2,749; range, 1,965 to 10,432) compared to negative controls (924 dpm [n = 4; SD, 539]), positive control FFV supernatant (42,198 dpm [n = 4; SD, 11,821]) and positive control raw MMLV RT 24,630 dpm (n = 4; SD, 5,674). When the more sensitive qualitative RT-PCR-based RT assay was applied to koala serum, 9 of 31 (29%) reactions amplified the expected size fragment of kama-mycin cDNA, indicating in vivo RT activity in the serum of these animals.

**Electron microscopy.** TEM demonstrated characteristic type C retrovirus-like particles associated with lymphoma in three koalas or with mitogen-stimulated cultured PBMCs derived from 163 of 166 koalas. Particles were 80 to 110 nm in diameter, enveloped, and contained an electron-dense spherical core. Budding forms were observed, with characteristic C-shaped capsid assembly at the cytoplasmic membrane. For three koalas, PBMCs grew poorly in cell culture and virus particles were unable to be demonstrated by TEM.

**DISCUSSION**

From the data presented, it is clear that a provirus (KoRV) related to the simian type C retroviruses is present in the DNA of koalas. The observation that retrovirus-like particles were produced by both tumor cells in vivo and cultured PBMCs in vitro suggests that proviral transcription and virus assembly occur. This is supported by the detection of retroviral transcripts homologous to proviral sequences in the blood of koalas and the detection of viral RNA and RT activity in serum and in PBMC culture supernatants.

Proivirus was detected in all koalas and in all tissues by PCR, and virus-like particles were demonstrated by TEM in 98% of mitogen-stimulated PBMC cultures. Southern hybridization of genomic DNA demonstrated similar intensity and banding patterns across a range of tissues in koalas with and without lymphoid neoplasia. These data suggest that KoRV is an ERV, although exogenous forms may also exist. Full-length provirus sequences were detected and contained all of the basic genetic elements common to type C retroviruses (44, 58), namely, flanking LTRs; gag, pro-pol, and env genes; a tRNA primer binding site; a polypurine tract; a CAAT box; a TATA box; a Cys-His box; and a polyadenylation signal. The sequences showed the most similarity to the GALV-related simian retroviruses and also showed ORFs and spliced env transcripts consistent with the type C mammalian group of retroviruses.

While full-length provirus is present in koalas, there is also strong evidence for the presence of truncated proviruses. Undersized PCR products derived from primers that span the complete proviral genome were detected in all animals, although the molecular size ranges of these products differed between animals. These amplifiers hybridized to labelled KoRV pro-pol, and partial sequencing confirmed their KoRV derivation, suggesting that they represent truncated proviruses. Tissue-, animal-, or population-specific truncated KoRV variants may have arisen in the koala genome either via recombination excision events or novel superinfection or reinsertion events. Considering the marked variation between koalas in the lengths of apparently truncated proviruses detected by PCR, it appears that either KoRV is not stable, in contrast to most other endogenous viruses (3), or PCR detection is confounded by the presence of related exogenous virus.

Detection of some viral transcripts containing gag and/or env sequences and failure to detect viral transcripts containing pro-pol sequences by RT-PCR of serum and cell-free culture supernatants from some koalas may be due to transcription and expression of truncated proviruses and assembly of virions containing defective genomic RNA. In humans, there are numerous examples of expression of defective or truncated endogenous proviruses in both normal and transformed tissues (1, 18, 34, 35, 57, 59). Furthermore, the expression of the retroviral Gag protein alone may be sufficient to cause assembly and budding of virus-like particles from the cell (58). Clearly, in the koala genome, there are multiple copies of proviruses, including both full-length and truncated proviruses, at least some of which are expressed both in vivo and in vitro.

The inability to detect strong RT activity in the serum or PBMC culture supernatants of some koalas may be explained by a number of possibilities. First, inhibition of RT activity by tissue-derived specific inhibitors has been demonstrated with some human ERVs (59). A second possibility is that virions may not contain detectable RT because of a defective or absent pro-pol gene. This is consistent with the observed expression of defective or truncated proviruses and has been previously demonstrated for HERV-K, which encodes a functional enzyme with weak activity (57). A third possibility is that assay conditions may not have been optimal for the detection of KoRV RT activity.

The DNA sequence data from gag, pol, and env unambiguously support a grouping of KoRV and GALV, to the exclusion of PERV and the MLV group. A similar grouping was also obtained from an independent phylogenetic study that compared pol fragments derived from a wide range of host taxa (37). This is an intriguing grouping, because gibbons (Hylobates spp.) and koalas are taxonomically distant, making virus acquisition from a common ancestor unlikely and because gibbons and koalas are from remote biogeographic regions, making direct natural transmission improbable. A preliminary investigation of small numbers of animals failed to detect KoRV-like retroviruses in other marsupials, including an eastern grey kangaroo (Macropus giganteus) and common wombats (Vombatus ursinus) and southern hairy-nosed wombats (Lasiorhinus latifrons) (data not shown). Wombats are the closest living relatives of koalas, although still relatively distantly related. Other studies have also failed to demonstrate KoRV-like viruses in marsupials other than koalas (23, 37, 40). This suggests that endogenization of KoRV in koalas postdates the divergence of koalas from other marsupials.

There are two possible explanations for the close relationship between KoRV and GALV. Either KoRV and GALV share a more recent common ancestor than all other ERVs sequenced to date, and therefore the grouping reflects the true evolutionary history of these sequences, or the grouping is an artifact of sequence analysis. The potential causes of spurious phylogenetic signal include recombination between strains, base composition bias, sequence saturation, and molecular convergence. Each of these artifacts can be discounted on the
basis of detailed DNA sequence analysis, suggesting that the KoRV-GALV grouping represents a true biological relationship.

Recombination between virus lineages can cause unusual phylogenetic groupings, because different regions of the genome have different phylogenetic histories. However, there is no evidence that recombination is causing the KoRV-GALV grouping, because all viral genes gave the same phylogenetic signal. There is little evidence of saturation (true signal obscured by multiple hits) in the DNA sequences (Fig. 5). Nor is there any evidence for base composition bias pulling KoRV and GALV together, because the base composition of these sequences does not differ noticeably from the base compositions of the Friend MLV, MMLV, and PERV sequences. The KoRV-GALV grouping is unlikely to be caused by molecular convergence, which would require independent acquisition of many identical substitutions in both KoRV and GALV lineages and would expect to result in a strong conflicting signal in the data set. However, spectral analysis shows little conflicting signal (Fig. 5). Strong support for the KoRV-GALV grouping from both the 1st and 2nd codon positions, which predominately cause nonsynonymous changes, and 3rd codon positions, which are mostly “silent,” provides further evidence against molecular convergence, since a convergent signal would be expected only in the nonsynonymous changes.

Some ERV phylogenies closely match the phylogeny of their hosts, indicating that the viruses are evolving as an endogenous element of the host genome, rarely infecting other hosts. Examples include the MLV-related retroviruses in reptiles and amphibians (23, 37) and some ERVs of great apes (27). Because the phylogeny does not match the order of branching of the mammalian tree (37) (Fig. 4), KoRV and GALV could only be considered to be strictly host tracking if the KoRV-GALV ancestor represents a novel insertion into the ancestral mammalian genome independently of the other mammalian ERVs sequenced to date. The similarity of KoRV and GALV to FeLV, the MLVs, and PERV would then be due to the similarity of the free-living viruses that inserted into the mammalian genome multiple times at least 130 million years ago, at the time of metatherian and eutherian mammal divergence (26). This could be confirmed by identifying the host genome insertion points for these viruses; if insertion is random, then independent acquisitions of ERVs will occur at different loci (27). However, the depth of molecular divergence in the KoRV-GALV clade is less than half that observed in the MLV clade, so the host-tracking hypothesis would require that the rate of molecular evolution in the KoRV-GALV lineage was dramatically slower than that of the MLV genomes. Given that GALV is exogenous, it seems highly unlikely that its rate of molecular evolution would have decreased to such an extent. Indeed, it would be expected that the rate of molecular evolution in an actively infectious virus would be greater than that of an endogenous virus (49). A host-tracking explanation for the KoRV-GALV grouping is therefore inconsistent with the sequence data.

There seems to be no reason to cast doubt on the phylogenetic relationship between KoRV and GALV to the exclusion of other retroviruses sequenced to date. There are two possible explanations for the relationship: inheritance from the genome of the last common ancestor of gibbons and koalas or cross-species transmission. The former explanation is unlikely, given the DNA sequence data and the preliminary assessment that KoRV-like viruses are absent in other marsupials. It seems most likely that the KoRV-GALV grouping is the result of a relatively recent cross-species transmission of a KoRV-like virus into gibbon and/or koala populations. Because gibbons and koalas are from distinct biogeographical regions with very different faunas, direct cross-infection in the wild is improbable. However, natural transfer via an intermediate host, for example, a mobile species such as rodents, bats, or birds or an arthropod vector, cannot currently be ruled out. A scenario suggested by Martin and coworkers (37) is that transmission was dependent upon an intermediate host, possibly a species of Asian rodent in which ERVs related to the primate type C viruses have been detected (6, 30).

KoRV has been isolated from both captive and wild koalas, but to our knowledge, GALV has been detected only in captive gibbons. This observation suggests that iatrogenic infection of captive gibbons with a KoRV-like virus derived either from live captive koalas, from tissue or fomites, is also a possible mechanism of recent cross-species transfer. If this observation holds true, it would imply that GALV infection in gibbons is a recent artifact of captivity. Alternatively species mixing and opportunities for either direct or indirect cross-infection between koalas and gibbons may have been promoted by indigenous trade between southeast Asia and mainland Australia in precolonial times. This mechanism has been proposed as a transmission method of an Australian marsupial arthropod parasite, Heterodoxus spiniger, to Asian dogs (16).

Whether iatrogenic or natural transmission has occurred, pinpointing the date of such a host jump through “molecular clock” dating may help to resolve the path of cross-species infection. However, dating virus divergence times from molecular data is made difficult by lack of a suitable calibration rate. There are no currently known divergence dates that can be used to calibrate the rate of change in KoRV or GALV. Given that substitution rates can vary enormously for different virus lineages (49) and between genes within a lineage (32), it is prudent to be wary of assuming a molecular clock to date virus divergences. However, the relatively low degree of divergence between the env genes of KoRV and the various GALV strains supports a recent divergence. In view of the apparently widespread endogenous infection of koalas and the suggestion that the exogenous GALV viruses are relatively recently acquired by simians (33), the possibility remains that the GALV group originated in koalas and spread to gibbons by a yet to be discovered mechanism. The parallel finding that macropodid herpesviruses paradoxically cluster with the simian herpesviruses (36) also supports the possibility of an occult epidemiological link between Australian marsupials and old-world primates.

Lymphoid malignancy is a relatively common disease of both free-living and captive koalas (11, 55), and its clinical similarity to retrovirus-associated disease in other mammals has been noted (7, 24). The ability of GALV and SSV to cause neoplasia in their host species has been demonstrated in both natural (28, 54, 56) and experimental infections (12, 29). In view of the relatively high prevalence of hematopoietic neoplasia and related diseases in koalas (4, 11, 61) and anecdotal evidence for lymphoid neoplasia epizootics in captive koala colonies, it is reasonable to suspect a pathogenic role for KoRV. However, KoRV was detected in all diseased and healthy koalas tested and appears to be endogenous. Although this lack of clinical correlation does not rule out a pathogenic role for the virus, it implies that other factors are necessary for disease to occur. Potential cofactors include superinfection with KoRV exogenous pathogenic variants, recombination with other retroviruses, interactions with host haplotypes, interactions with somatic mutations, or coinfection with other viruses, such as herpesviruses.

Serious consideration should be given to the zoonotic potential of the GALV-KoRV group. Early experimental work
on GALV and more recent investigation of GALV as a vector for gene therapy have demonstrated multiple host and tissue tropisms (17). The broad in vitro host range of GALV, its apparently well-conserved receptor (42), and the evidence presented for taxonomic cross-transmission in vivo suggest that the members of the GALV-KoRV group have the potential for future host jumping. Furthermore, simian type C-related virus sequences and cross-reactive proteins have been detected in humans (14, 17, 20, 31), and a novel strain of GALV (GALVX) has been isolated from a human cell line (5, 43). Comparisons are aimed at refining the epidemiology of the KoRV-GALV group and at clarifying the pathogenic and zoonotic potential of this virus.

ACKNOWLEDGMENTS

We thank Steven Barker for his intellectual input, Kevin Bradley and Anne Williams (Dreamworld, Gold Coast, Australia), Ray and Jennifer Chafer (Wild World, Cairns, Australia), Vere Nicholson (Currumbin Sanctuary, Gold Coast, Australia), Paul O’Callagan and Galit Tzipori (Lone Pine Sanctuary, Brisbane, Australia), Anne Williams and Rosie (Currumbin Sanctuary, Gold Coast, Australia), Paul O’Callagan and Anne Williams (Dreamworld, Gold Coast, Australia), Ray and

REFERENCES


