Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 2699–2703, May 1983
Genetics

Endogenous murine leukemia proviral long terminal repeats contain a unique 190-base-pair insert

(mink cell focus-forming virus/xenotropic provirus/recombination/transposon-like feature)

ARIFA S. KHAN AND MALCOLM A. MARTIN

Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205

Communicated by Wallace P. Rowe, February 2, 1983

ABSTRACT We have determined the nucleotide sequence in the U3–R regions of the long terminal repeat (LTR) associated with NFS-Th-1 xenotropic murine leukemia virus (MuLV) DNA and the LTR components of five endogenous proviruses cloned from BALB/c mouse chromosomal DNA. The five endogenous MuLV LTRs contained the regulatory signals thought to be important in viral transcription, such as "TATAA" and CCAAT-like boxes. A unique feature in four of the endogenous LTRs was the presence of a highly conserved 190-base-pair (bp) insert bounded by 6-bp direct repeats located 48 bp upstream from the C-C-A-A-T sequence. This segment was absent from LTRs associated with ecotropic, xenotropic, or mink cell focus-forming (MCF) MuLV proviruses. All five endogenous LTR segments also contained a 14-bp duplication of a sequence located near the 5′ end of the first component of the long (>72-bp) direct repeat of ecotropic and MCF MuLV LTRs. An evolutionary scheme relating LTRs associated with endogenous MuLV proviral DNAs to those found in ecotropic or xenotropic proviruses is presented. Nucleotide sequence analysis also suggested that the U3 region of the MCF247 MuLV LTR is derived from an NFS xenotropic related MuLV.

The RNA genome of type C retroviruses contains terminally redundant sequences (R) that are located adjacent to segments unique to the 5′ (U5) and 3′ (U3) termini (1). After infection, the viral RNA is reverse transcribed into double-stranded DNA, which subsequently becomes integrated into cellular chromosomal DNA (2, 3). Both the unintegrated and integrated copies of proviral DNA contain terminally duplicated sequences derived from each end of the viral genome forming a structure (U3–R–U5) referred to as the long terminal repeat (LTR). LTRs contain important regulatory signals for the promotion, initiation, and processing of viral mRNAs and may play a role in the integration of proviral DNA into the host chromosome (3). Furthermore, nucleotide sequence analyses show that LTRs possess structural features characteristic of transposable elements found in bacteria, yeast, and *Drosophila* (4–11).

In several vertebrate species, DNA copies of type C retroviruses have been identified as genetically stable integral components of chromosomal DNA (12). Such endogenous proviruses have been shown to be vertically transmitted, can be mapped to specific chromosomal loci, and, in some cases, may be expressed as infectious retroviruses (12). We have recently described the molecular cloning and characterization of endogenous murine leukemia virus (MuLV) proviral DNAs isolated from AKR/J and BALB/c mouse DNAs (13). In that study, we focused on cloned retroviral segments containing regions that hybridized to the LTR of Harvey sarcoma virus DNA. Our analyses indicated that the U3 segment of 11 endogenous LTRs was larger [525–570 base pairs (bp)] than the corresponding region (390–490 bp) in ecotropic, xenotropic, or mink cell focus-forming (MCF) MuLV proviruses (13). Furthermore, restriction enzyme mapping of the *gag*, *pol*, and *env* regions of these cloned endogenous proviral DNAs revealed the existence of highly conserved and unique restriction sites not present in analogous segments of infectious MuLV proviruses (13). The failure to detect the expression of most of these endogenous proviral DNAs as infectious MuLVs could be due to a number of factors including point mutations, deletions, or insertions in the *gag*, *pol*, or *env* genes. Structural alterations of the endogenous LTRs might also affect the expression of these proviruses as a result of their role in modulating the synthesis and processing of viral mRNAs.

Because we had previously shown that the LTRs associated with endogenous MuLV proviruses could be distinguished from the LTR segments associated with known infectious viruses (13) and since the numerous copies of these elements present in mammalian chromosomal DNA could potentially regulate the expression of adjacent cellular genes, we determined the nucleotide sequences in the U3 R regions of five endogenous LTRs. The results of these sequence analyses and comparisons with LTRs associated with infectious xenotropic, ecotropic, and MCF proviruses are presented.

MATERIALS AND METHODS

Recombinant DNA Clones. The *Pst* I/*Sma* I segment derived from the U3–R region of the LTRs associated with the endogenous BALB/c MuLV proviral DNA clones B-56, B-73, B-14 (5′ LTRs), and B-34 (3′ LTR) (13) and with NFS-Th-1 xenotropic DNA cloned from MuLV-infected cells (14) was ligated into M13 mp8 and M13 mp9 replicative form vector DNAs (15) previously cleaved with *Pst* I/*Sma* I. Because of the absence of a *Sma* I site in the 3′ LTR associated with endogenous clone B-77 (13), a segment [2.8 kilobases (kb)] extending from the *Pst* I site in the U3 region to an *Eco*RI site in the flanking cellular sequences was cloned into M13 mp8 replicative form DNA previously cleaved with *Pst* I/*Eco*RI.

DNA Preparation and Sequence Analysis. *Escherichia coli* JM103 bacterial cells were transformed with recombinant replicative form DNA, and phage containing a DNA insert (white plaques) were isolated and propagated in 2-ml cultures for 8 hr. Twenty-microliter aliquots from each phage preparation were lysed and subjected to 0.6% agarose gel electrophoresis at 40 V for 16 hr. The single-stranded DNA obtained was transferred to nitrocellulose filters and hybridized, as described (16), to a ^{32}P-labeled LTR segment isolated from Harvey sarcoma virus DNA (17). Single-stranded recombinant M13 phage DNA was prepared from LTR-reactive clones and its sequence was determined as described (18).

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

Abbreviations: bp, base pair(s); LTR, long terminal repeat; MuLV, murine leukemia virus; MCF MuLV, mink cell focus-forming MuLV; Mo-MuLV, Moloney murine leukemia virus; kb, kilobase(s).

RESULTS

Nucleotide Sequence of the LTR Associated with a Xenotropic MuLV Provirus. As a part of the characterization of 12 endogenous MuLV proviruses cloned from AKR/J and BALB/c mouse DNAs, the associated LTR regions were compared with analogous segments of ecotropic, xenotropic, and MCF MuLV proviral DNAs (13). In 11 of 12 DNA clones, the U3 region of the endogenous LTR (525–570 bp) was detectably larger than the corresponding (390- to 490-bp) segment associated with known infectious MuLVs. Because further analysis of the endogenous LTRs would require the determination of their nucleotide sequences and comparison with analogous regions of known infectious proviruses and the nucleotide sequences of the LTRs associated with MCF and ecotropic MuLV proviruses were already known (19–21), we initially determined the sequence of the LTR present in NFS-Th-1 xenotropic proviral DNA. Furthermore, the sequence determination of the LTR associated with xenotropic MuLV proviral DNA might shed light on the unique biological properties of this class of murine retrovirus since alterations in the LTR regions could affect the expression of viral gene products in different cell types and thus specify the host range of the virus (22).

The strategy used to determine the sequences of the exogenous NFS-Th-1 xenotropic MuLV LTR (LTR-xeno) and the endogenous MuLV LTRs present in mouse chromosomal DNA was to analyze a segment defined by a *Pst* I site located within 40 nucleotides of the 5′ end of the U3 region and a *Sma* I site mapped in the R region of most MuLV LTRs and then determine the sequences of M13 recombinants containing such *Pst* I/*Sma* I DNA inserts. The nucleotide sequence of the 385-bp *Pst* I/*Sma* I fragment of the LTR-xeno and a comparison of it with the LTRs of AKR MCF247, AKR ecotropic, and Moloney (Mo) ecotropic MuLVs is presented in Fig. 1. The U3–R portion of the LTR-xeno contains many of the structural features of the other three MuLV LTRs, such as the "TATAA" box at position 683–690, the C-C-A-A-T sequence at position 624–628, and the nucleotide guanosine, thought to be important in capping, at position 714. The LTR-xeno lacks the 74- to 106-bp direct repeat present in the U3 region of the ecotropic and MCF MuLV LTRs shown in Fig. 1. Nonetheless, the LTR-xeno does contain a single copy of a 59-bp "core" sequence that is an integral element of the 74- to 106-bp direct repeat and is located between positions 133 and 195 and 253 and 315 in Fig. 1. Despite the absence of a large direct repeat in LTR-xeno, its similarity to the LTR associated with the MCF MuLV is quite striking (92% homology). The analysis presented in Fig. 1 indicates greater divergence between the LTR-xeno and the LTRs of Mo- and AKR ecotropic MuLVs (76% and 83% polynucleotide sequence homology, respectively) with a majority of the changes located adjacent to both termini of the core sequence between positions 50–108 and 316–379. In contrast, these portions of the MCF MuLV LTR and the LTR-xeno are highly conserved.

Nucleotide Sequence Determination of LTRs Associated With Endogenous MuLV Proviral DNAs. Restriction mapping of the LTR components of endogenous MuLV proviruses (LTR-endog) with *Pst* I and *Kpn* I indicated the existence of at least two classes that contained larger (525- and 570-bp) U3 regions than ecotropic, xenotropic, or MCF MuLV proviruses (13). One of the clones (B-77) contained a U3 region that was similar in size (400 bp) to its analogue in infectious MuLVs. With the same strategy used to determine the sequences of the LTR-xeno, the sequences of the *Pst* I/*Sma* I fragments from endogenous MuLV clones containing 586 bp (clones B-14 and B-34), 543 and 545 bp (clones B-73 and B-56, respectively), and 419 bp (clone B-77) U3 regions were determined and compared with the LTRs associated with xenotropic, ecotropic, and MCF MuLVs.

As shown in Fig. 1, the five LTR-endog segments examined retained many of the characteristic landmarks present in retrovirus LTRs. All had an intact T-A-T-A-A-A-A-A sequence (position 683–690) located 21 to 22 nucleotides upstream from the R region as well as the nucleotide guanosine at position 714. All of the LTR-endog segments harbored a C-C-A-A-T sequence but, in contrast to the MuLV LTR standards, this signal was altered to C-C-A-A- in four of the LTR-endog fragments and in the fifth (B-77) was present as C-C-A-C-T. Although none of the endogenous LTRs contained a 74–106 bp direct repeat, all of them had an identifiable core sequence (in Fig. 1, extending from position 133 to 195 or from 253 to 315). One clone (B-77) had a partial duplication of this core sequence. The LTR-endog segments were more than 90% homologous to the xenotropic and MCF MuLV LTRs and shared about 80% of their nucleotide sequence with LTRs associated with ecotropic MuLVs. Furthermore, the five endogenous LTRs and the xenotropic MuLV LTR contained a 6-bp direct repeat (4 and 4*) located adjacent to the 5′ terminus of the TATAA box. Segment 4 was also present in MCF and ecotropic LTRs as a single copy. The functional significance of this duplication adjacent to the TATAA box is currently unclear.

The most striking feature of four of the five endogenous LTRs was the presence of a 190-bp insert (designated A in Fig. 1) located between the core sequence and the CCAAT box (between positions 379 and 570). This DNA segment was missing from the U3 region of endogenous clone B-77. As shown in Fig. 1, the nucleotide sequence of insert A was highly conserved among the four LTR-endog segments analyzed. Insert A contained an adenosine-rich 10-bp direct repeat (designated 3* and 3) and was bordered by an imperfect 6-bp direct repeat (designated 2* and 2) at both termini. Although segment A was totally absent from the MuLV LTR standards, the xenotropic and MCF MuLV LTRs contained recognizable copies of sequences 2* and 2 while the ecotropic MuLV LTRs harbored only sequence 2 (Fig. 1). Comparison of the nucleotide sequence of insertion A with those present in the nucleic acid query program available from the Georgetown University Computer Center (as of Sept. 1982) revealed no homology after allowing a 10-bp mismatch.

All five endogenous LTRs contained a 14-bp duplication of sequence 1 (designated sequence 1*) located between positions 79 and 94. Sequence 1 was readily identifiable in the LTRs associated with both MCF and xenotropic MuLVs (positions 94–108) and was not reiterated. Sequence 1 was not highly conserved in the LTRs of the two ecotropic MuLVs examined but instead formed the 5′-terminal components of the long (>72-bp) direct repeat.

The analysis presented in Fig. 1 indicates that the nucleotide sequences of the different LTR-endog segments are highly conserved. The sequences of the LTRs associated with clones B-14 and B-34 are identical. The LTRs of clones B-56 and B-73 belong to a class containing an internal 525-bp *Pst* I/*Kpn* I segment (13). The U3 regions in the LTRs of these two clones contain a 37-bp deletion extending from position 321 to position 365, accounting for their smaller size relative to the LTR-endog segments of clones B-14 and B-34.

DISCUSSION

We have shown that endogenous MuLV LTRs contain several of the structural elements present in the LTRs of known infectious ecotropic, xenotropic, and MCF MuLVs such as the TATAA and CCAAT boxes and the core component of the 74- to 106-bp direct repeat. The LTR-endog segments studied, however, could be distinguished from the MuLV LTR standards on the basis of at least two characteristic DNA insertions.

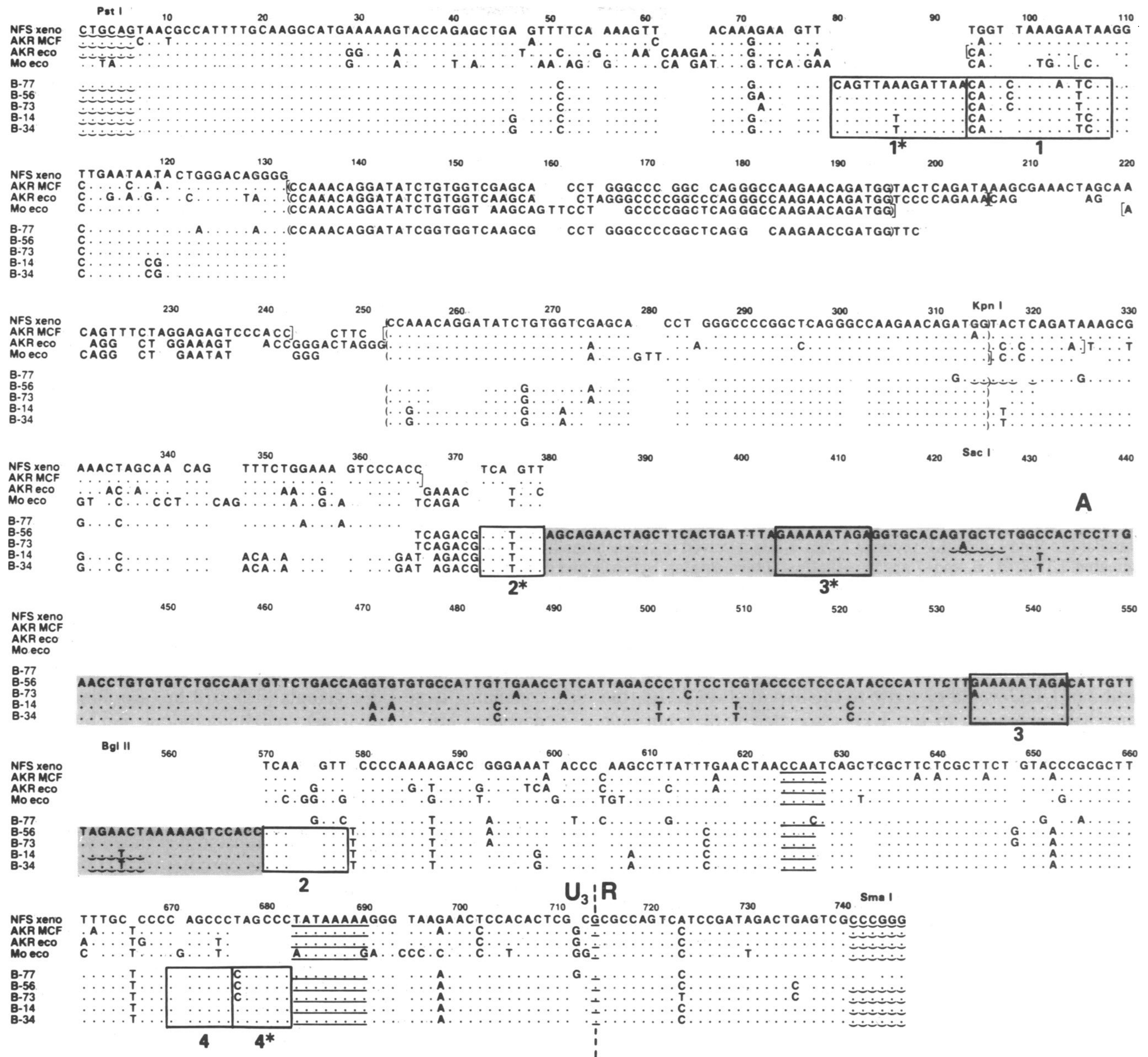

FIG. 1. Comparison of MuLV LTR nucleotide sequences. The nucleotide sequences of five endogenous MuLV LTR segments are compared with analogous regions of xenotropic, MCF, and ecotropic MuLV LTRs. The sequence of the *Pst* I/*Sma* I segment of the NFS-Th-1 xenotropic MuLV LTR was used as the "standard". Dots indicate nucleotide identity; blank space, absence of a nucleotide; ⌣, restriction endonuclease cleavage site (indicated above the nucleotide position); ----, boundary between the U3 and R regions; underlined nucleotide, a regulatory signal. Direct repeats associated with NFS xenotropic (NFS xeno), AKR MCF247-derived MCF2 (20) (AKR MCF), AKR 614 ecotropic (21) (AKR eco), and Mo-MuLV ecotropic MLV 1A (21) (Mo eco), are enclosed by brackets and the associated internal core sequences are within parentheses. Direct repeats present in the endogenous LTRs are boxed; these elements are designated 1 and 1*, 2 and 2*, 3 and 3*, and 4 and 4*. The shaded area indicates the 190-bp insert A that is unique to endogenous LTRs.

Four of the five cloned endogenous LTRs contained a highly conserved 190-bp DNA segment (A) inserted between the core component in the long (>72-bp) direct repeat and the CCAAT box. This large insertion was flanked by a 6-bp direct repeat, a structural characteristic of transposable elements. Furthermore, the deletion in clones B-56 and B-73 of sequences adjacent to the 190-bp insert is also compatible with the transposon-like features associated with segment A. In addition, Southern blot analysis showed that sequences in insert A are highly reiterated in the BALB/c mouse chromosomal DNA and can be distinguished from LTR-reactive segments (unpublished data). All five of the endogenous LTRs contained a unique 14-bp duplicated sequence (1* in Fig. 1), a single copy of which was found in the LTRs associated with infectious MuLVs. Furthermore, all of the endogenous MuLV LTRs also contained a 6-bp duplication (4* in Fig. 1) immediately upstream from the TATAA box that was also present in xenotropic but absent in MCF and ecotropic MuLV LTRs.

The unique structural features of the five LTR-endog segments raise several interesting possibilities regarding their evolutionary relationship to the LTRs associated with known infectious MuLVs. In the scheme proposed in Fig. 2, LTR*1* or -*3* could represent the progenitor of the LTR regions found in xenotropic or ecotropic MuLV proviruses. The hypothetical LTR3 could give rise to LTR*1* (present in endogenous MuLV clones B-14 and B-34) following the insertion of the transposon-like element A and resulting in the concomitant duplication of segment 2. Conversely, if LTR*1* is the precursor to all MuLV LTRs, LTR3 could be generated by the deletion of the 190-bp segment A and adjoining region 2*. LTR2, found in the endoge-

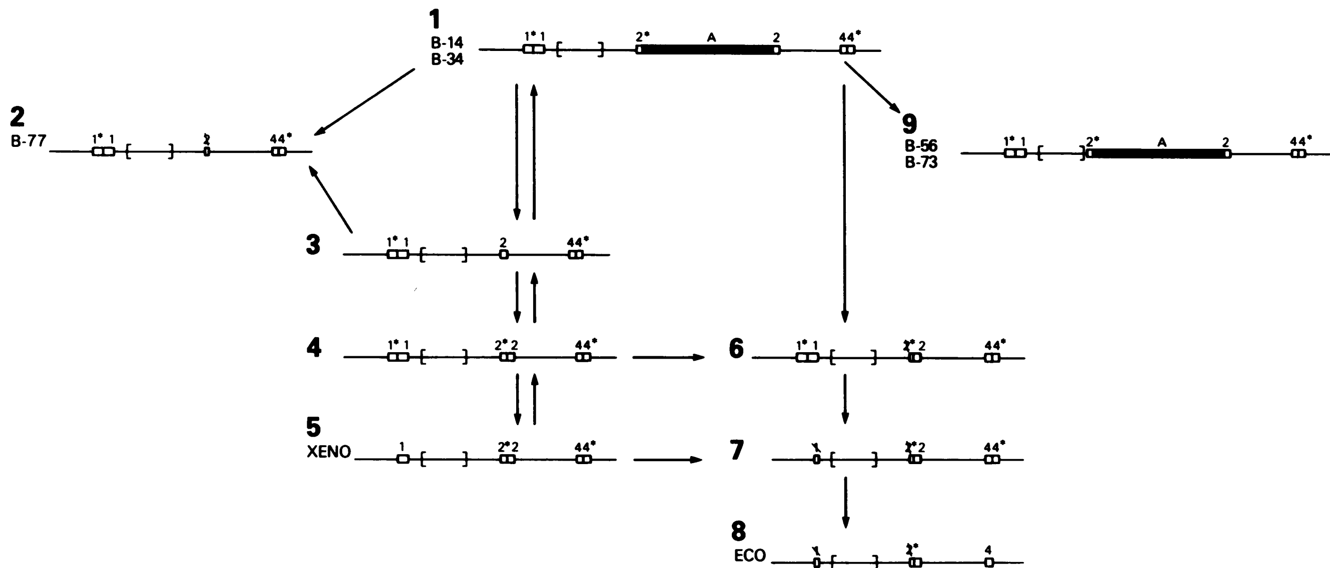

FIG. 2. A possible evolutionary scheme relating endogenous and infectious MuLV LTRs. LTR*1*–LTR*9* represent the region in U3–R extending from *Pst* I to *Sma* I of MuLV proviral LTRs. LTRs involved in this scheme that have been identified as components of specific proviral DNAs are indicated. LTR*3*, -*4*, -*6*, and -*7* are hypothetical intermediates. Symbols and designations are as in Fig. 1. The solid bar represents segment A; open boxes indicate direct repeats present in the LTRs and are numbered as in Fig. 1. 1̸, 2̸, and 2̸*, deletions in sequences 1, 2, and 2*, respectively. [ ], Core sequence.

nous MuLV clone B-77, could be formed in a similar fashion with the deletion extending to region 2. A duplication of segment 2 in LTR*3* would give rise to the hypothetical LTR*4*, the immediate precursor to the LTR associated with xenotropic MuLVs (LTR5). Although these analyses of LTR nucleotide sequences provide little definitive information about the relationship of the endogenous LTRs to known infectious MuLV LTRs, the greater structural (and sequence) similarity of the LTR-endog segments to LTR-xeno regions than to ecotropic MuLV LTRs suggests a more direct evolutionary lineage. Accordingly, the LTR-xeno (LTR5) has been positioned closer to the putative MuLV LTR progenitor (LTR*1* or LTR3) in the pathway presented in Fig. 2 than to the ecotropic MuLV LTR. The latter (LTR8) may have evolved from LTR-endog structure *1* via LTR-xeno or its immediate precursor LTR*4* or, alternatively, the ecotropic MuLV LTR could have originated from a separate lineage.

Of the 12 cloned endogenous LTRs examined, 8 had a structure similar to the LTR segments of clones B-56 and B-73 (13), which are characterized by the presence of the 190-bp insertion and an adjoining 37-nucleotide deletion. This LTR-endog (LTR9, Fig. 2) is obviously related to LTR*1* but is not in the lineage giving rise to xenotropic or ecotropic MuLV LTRs, both of which have preserved sequences located between the core and region 2*.

The one endogenous LTR (B-77) of the 12 examined that had a U3 region similar in size (400 bp) to LTRs associated with infectious MuLVs possessed several structural features of the putative LTR progenitor, LTR3 (Fig. 2). Although the LTR of endogenous clone B-77 (LTR2) was the only LTR-endog segment analyzed that did not contain insertion A, it still contained the 14-bp 1* region that is a hallmark of endogenous LTRs. Of the five endogenous MuLV clones containing an *env* gene, four, including clone B-34, had restriction sites and hybridization properties typical of MCF MuLV proviral DNAs (13). In contrast, the nucleotide sequence of the clone B-77 *env* region was typical of a xenotropic provirus (23). Clone B-77 thus represents an endogenous MuLV provirus containing a "mature" LTR (missing insertion A) associated with a xenotropic (rather than the more abundant MCF) MuLV *env* gene.

One of the unanswered questions arising from this study is whether the endogenous LTRs are functional. We have previously reported that virtually all of the LTR-endog segments containing large U3 segments (525–570 bp) are joined to endogenous MuLV proviruses containing unique and highly conserved restriction sites such as *Sac* I and *Xba* I in the *gag* region and *Xba* I in the *pol* gene (13). To date, no infectious MuLV proviral DNAs with these characteristic cleavage sites have been isolated. Since our previously reported examination of cloned endogenous MuLV proviruses indicated that they had been inserted at multiple loci into mouse chromosomal DNA (on the basis of different flanking cellular DNA) (13), we concluded that an ancestral MuLV-like retrovirus with the typical LTR-endog structure containing a large U3 region and unique *gag* and *pol* genes was at one time infectious and capable of replicating in mouse cells. The failure to detect the expression of these endogenous MuLV proviral DNAs could simply reflect the presence of point mutations, deletions, or hypermethylation of viral gene sequences with little if any alteration to the LTR-endog segments. The results of preliminary experiments evaluating the functional activity of endogenous LTRs directly in a simian virus 40 vector lacking the 72-bp direct repeat but containing the chloramphenicol acetyltransferase gene (22) indicate that the LTR-endog segment in clone B-56 is weakly active (unpublished data).

Unexpectedly, the results obtained from the nucleotide sequence determinations of LTR-endog segments and their comparison to LTRs associated with other known infectious MuLVs have implications on the recombinational events associated with the generation of MCF MuLVs. We previously reported (13) that an endogenous provirus similar to one of the cloned MuLV segments (A-12) participates in the recombinational events associated with the generation of AKR MCF247 because both of them contain indistinguishable and characteristic *env* gene segments. Clone A-12 cannot be the direct precursor of infectious MCF viruses because it has unique restriction sites in the *gag* and *pol* regions that are present in other endogenous MuLV proviruses and absent in MCF proviral DNAs. Because the *gag* and *pol* regions in the majority of MCF and ecotropic MuLV proviral DNAs are virtually identical, we (13) and others (24) have proposed that MCF viruses could be generated after the recombination of an ecotropic MuLV with endogenous *env* seg-

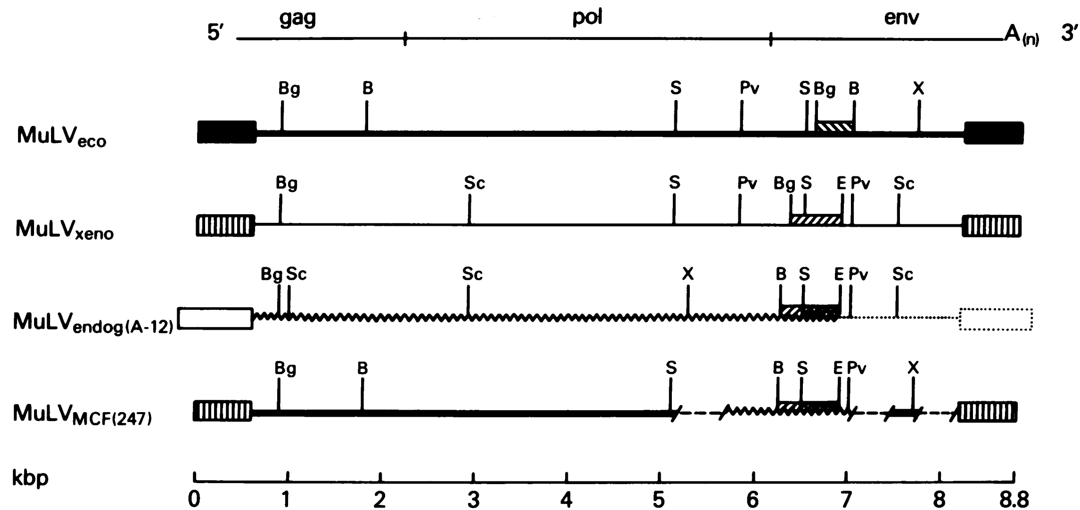

FIG. 3. Sources of nucleotide sequences present in AKR MCF247 proviral DNA. Structures of three MuLV proviruses [ecotropic, $MuLV_{eco}$; xenotropic, $MuLV_{xeno}$; endogenous AKR DNA clone A-12, $MuLV_{endog(A-12)}$] thought to be involved in the generation of AKR MCF247 [$MuLV_{MCF(247)}$] are indicated diagrammatically. Restriction sites relevant to defining the recombinant regions were Bg, *Bgl* II; B, *Bam*HI; S, *Sma* I; Pv, *Pvu* II; X, *Xba* I; Sc, *Sac* I; E, *Eco*RI. ▧ and ▨, reactivity in the *env* region to ecotropic- and xenotropic-specific probes as determined previously; ▬, hybridization to both *env* specific probes; ----, possible regions of recombination during the generation of AKR MCF247; ······, uncloned segment in the endogenous A-12 MuLV proviral DNA. Restriction sites in the uncloned region of A-12 DNA have been located based on their conservation in analogous segments in other endogenous MuLV proviral DNAs (13). kbp, Kilobase pair(s).

ments that preexisted in mouse chromosomal DNA. However, as shown in Fig. 1, since the LTR associated with AKR MCF247 MuLV has a nucleotide sequence more closely related to LTR-xeno than to the LTRs of ecotropic MuLVs or of the LTR-endog associated with MCF-type *env* segments, the recombinational events giving rise to MCF viruses must also involve xenotropic MuLV elements. This is reflected in the structure of MCF247 proviral DNA presented in Fig. 3, which diagrammatically shows the possible sources of different MCF segments. Thomas and Coffin (25) have recently proposed that the generation of polytropic (MCF) MuLVs involves similar recombinational partners. Recombination of an ecotropic MuLV with the endogenous A-12 MuLV provirus between the *Sma* I site at 5.1 kb in the *pol* region (ecotropic MuLV) and the *Bam*HI site at 6.25 kb (endogenous MuLV) would link a MCF-specific *env* segment to ecotropic *gag* and *pol* sequences. A second recombinational event between the 7.0-kb *Pvu* II site in the endogenous MuLV *env* region and the 7.8-kb *Xba* I site of ecotropic MuLVs would add ecotropic *env* sequences to the 3′ side of the MCF-specific *env* segment. Finally, recombination with a xenotropic MuLV segment would introduce LTR-xeno-specific sequences to the 3′ end of the MCF MuLV provirus. Assuming a high conservation of U5 sequences between xenotropic and ecotropic LTRs, reverse transcription would generate a recombinant MCF provirus possessing LTR-xeno segments at both termini. These recombinational events would result in the generation of class 1 MCF MuLVs, which are characterized by a unique LTR structure (26). The class 2 MCFs, which appear to retain the ecotropic LTR (26), would presumably not have engaged in the recombination with xenotropic MuLV sequences. This is a particularly important hypothesis, because the class 1 MCF viruses are highly lymphomagenic while the class 2 MCFs are virtually nonpathogenic.

We are grateful to Dr. Jerry Weir for introducing us to the M13 sequence analysis procedure. We thank Dr. Wallace P. Rowe for helpful discussion and critical review of the manuscript.

1. Coffin, J. M., Hageman, T. C., Maxam, A. M. & Haseltine, W. A. (1978) *Cell* **13**, 761–773.
2. Temin, H. M. (1981) *Cell* **27**, 1–3.
3. Varmus, H. E. (1982) *Science* **216**, 812–820.
4. Calos, M. & Miller, J. (1980) *Cell* **20**, 579–595.
5. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3937–3941.
6. Ju, G. & Skalka, A. M. (1980) *Cell* **22**, 379–386.
7. Hishinuma, F., DeBona, P. J., Astrin, S. & Skalka, A. M. (1981) *Cell* **23**, 155–164.
8. Majors, J. E. & Varmus, H. E. (1981) *Nature (London)* **289**, 253–258.
9. Reddy, E. P., Smith, M. J., Canaani, E., Robbins, K. C., Tronick, S. R., Zain, S. & Aaronson, S. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5234–5238.
10. Shimotohno, K., Mizutani, S. & Temin, H. M. (1980) *Nature (London)* **285**, 550–554.
11. van Beveren, C., Goddard, J. C., Berns, A. & Verma, I. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3307–3311.
12. Coffin, J. (1982) in *Molecular Biology of Tumor Viruses: RNA Tumor Viruses*, eds. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed., pp. 1109–1203.
13. Khan, A. S., Rowe, W. P. & Martin, M. A. (1982) *J. Virol.* **44**, 625–636.
14. Buckler, C. E., Hoggan, M. D., Chan, H. W., Sears, J. F., Khan, A. S., Moore, J. L., Hartley, J. W., Rowe, W. P. & Martin, M. A. (1982) *J. Virol.* **41**, 228–236.
15. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Rol, B. A. (1980) *J. Mol. Biol.* **143**, 161–178.
16. Israel, M. A., Vanderryn, D. F., Meltzer, M. L. & Martin, M. A. (1980) *J. Biol. Chem.* **255**, 3798–3805.
17. Martin, M. A., Bryan, T., McCutchan, T. F. & Chan, H. W. (1981) *J. Virol.* **39**, 835–844.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
19. Kelly, M., Holland, C. A., Lung, M. L., Chattopadhyay, S. K., Lowy, D. R. & Hopkins, N. H. (1983) *J. Virol.* **45**, 291–298.
20. Khan, A. S., Repaske, R., Garon, C. F., Chan, H. W., Rowe, W. P. & Martin, M. A. (1982) *J. Virol.* **41**, 435–448.
21. van Beveren, C., Rands, E., Chattopadhyay, S. K., Lowy, D. R. & Verma, I. M. (1982) *J. Virol.* **41**, 542–556.
22. Laimins, L. A., Khoury, G., Gorman, C., Howard, B. & Gruss, P. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6453–6457.
23. Repaske, R., O'Neill, R. R., Khan, A. S. & Martin, M. A. (1983) *J. Virol.* **46**, 204–211.
24. Chattopadhyay, S. K., Cloyd, M. W., Linemeyer, D. L., Lander, M. R., Rands, E. & Lowy, D. R. (1982) *Nature (London)* **295**, 25–31.
25. Thomas, C. Y. & Coffin, J. M. (1982) *J. Virol.* **43**, 416–426.
26. Lung, M. L., Hering, C., Hartley, J. W., Rowe, W. P. & Hopkins, N. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **44**, 1269–1274.