JOURNAL OF VIROLOGY, June 1989, p. 2629–2637
0022-538X/89/062629-09$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 63, No. 6

The Palindromic LTR-LTR Junction of Moloney Murine Leukemia Virus Is Not an Efficient Substrate for Proviral Integration

LESLIE I. LOBEL, JOHN E. MURPHY, AND STEPHEN P. GOFF*

Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, 630 West 168th Street, New York, New York 10032

Received 9 December 1988/Accepted 14 February 1989

We generated viral constructs to test the hypothesis that the major substrate on retroviral DNA that is utilized for proviral DNA integration is the palindromic sequence, termed the LTR-LTR junction, normally present in circular molecules formed by joining the two termini of linear proviral DNA. Recombinant viral genomes were built which carried a selectable marker and an extra copy of the LTR-LTR junction from a cloned circular provirus. The junction sequence in each case was positioned such that its use during integration would lead to an easily detected, aberrantly integrated proviral DNA. Analysis of DNA from cells infected with the virus constructs showed that the introduced junction sequence is used at least 1,000-fold less efficiently than the natural sequences at the ends of the genome. This suggests that a linear or more exotic DNA intermediate is most likely the true precursor for the integration reaction.

The life cycle of the replication-competent retroviruses begins with the synthesis of a linear double-stranded proviral DNA from the incoming RNA genome. The sequences at the ends of the linear DNA are distinctive, consisting of long terminal repeats (LTRs) in direct orientation (9, 29) and containing at the very termini short inverted repeats (6, 31). Soon thereafter, two other duplex DNA forms appear: a circle containing two tandem LTRs, of the structure expected from the blunt-end ligation of the termini of the linear DNA; and a circle containing one LTR, of the structure expected from homologous recombination between the two LTRs of the linear DNA (30, 41). During the next few hours, one or more of these forms is inserted into the cellular DNA to generate the stably integrated provirus, always arranged such that sequences at the edges of the LTRs are joined to host sequences (6, 17, 18, 31, 37). Since all three of the unintegrated DNA forms coexist in recently infected cells, it has remained uncertain which of these proviral DNA forms might serve as the precursor for the formation of the integrated provirus (38).

Experiments in one retroviral system, that of the cytopathic spleen necrosis virus, have suggested that the LTR-LTR junction contained on the two-LTR circle is at least a major precursor structure for integrative recombination (26). A viral genome carrying an extra copy of the LTR-LTR junction was shown to replicate for many rounds of infection with wild-type helper virus and to give rise to structures indicative of utilization of the ectopic junction for recombination. In this system, the issue of whether other structures could also serve as precursors was uncertain; and the generality of the results to the many other diverse retrovirus families also remained unclear. To determine whether a very different virus, the Moloney murine leukemia virus (Mo-MuLV), could also utilize the LTR-LTR junction sequence efficiently, we prepared similar retroviral genomes carrying a selectable marker and an extra, internal copy of the LTR-LTR junction. Transfer of these genomes into cells through a single round of infection by using helper virus particles did not result in efficient use of the internal junction. Inclusion of a crippling mutation in the natural LTR tip of some of these constructs reduced the frequency of normal integration profoundly; even in these circumstances the intact internal LTR-LTR junction was not well used. These results suggest that the palindromic LTR-LTR sequence is not, in itself, an efficient substrate for integrative recombination in the murine retroviruses.

MATERIALS AND METHODS

Cells and viruses. NIH 3T3 cells, ψ2 cells (22), and CV-1 cos cells (10) were grown in Dulbecco modified Eagle medium (DMEM) containing 10% calf serum. Wild-type Mo-MuLV and mutant *dl*587 were as previously described (5). Virus infections were performed in the presence of 8 μg of Polybrene per ml. Transient expression of DNAs was by the DEAE-dextran method (24). Stable transformation of mammalian cells with DNA was by the calcium phosphate procedure (40), and transformants were isolated by selection with G418 (400 μg/ml; GIBCO Laboratories, Grand Island, N.Y.).

DNA analysis. Genomic DNA from infected cell clones or pools was prepared as described previously (11), digested by various enzymes according to the manufacturer (New England BioLabs, Inc., Beverly, Mass.), and analyzed by blot (35) hybridization (39) with a *neo*-specific probe (pSV2neo [36]) labeled by nick translation (28). Sequence determination was by the chemical degradation method (23).

Bacteriophage libraries. Clones of integrated proviruses of the pIS5 and pVNBJ2 genomes were made by inserting *Bam*HI fragments of genomic DNA into EMBL3, packaging in phage coats by using Gigapack extracts (Stratagene), and screening (2) by hybridization with a *neo*-specific probe.

Plasmid rescue of proviral clones. The CV-1 cos cell fusion rescue method (4) was used with minor modifications. Approximately 10^6 cells carrying the SV2neo-marked provirus were plated with an equal number of CV-1 cos cells (10) and allowed to grow for 24 h in DMEM containing 10% calf serum and 10% fetal calf serum. The cells were washed two times with DMEM, treated with 50% polyethylene glycol 1500 in DMEM for 60 s, and then washed extensively with DMEM. The cultures were fed daily for 3 days, and the low-molecular-weight DNA was isolated (16). The DNA was

* Corresponding author.

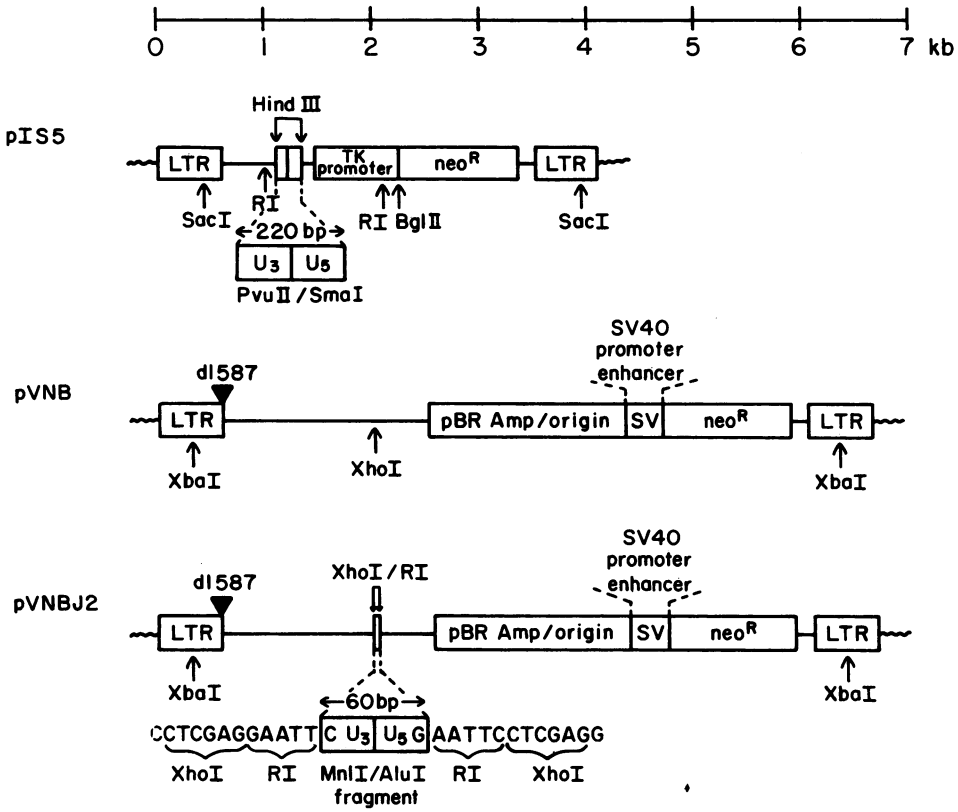

FIG. 1. Maps of the retroviral constructs used in this study. Regions containing the LTRs, promoters, and genes are indicated by boxes; other retroviral sequences are indicated by straight lines. Flanking sequences are indicated by wavy lines. The position of the *dl*587 deletion mutation is indicated by the black triangle. The details of the LTR-LTR junction regions are shown in the expanded sections. Selected restriction sites are indicated. TK, Thymidine kinase; RI, *Eco*RI; SV40, simian virus 40.

then used to transform *Escherichia coli* HB101 to kanamycin resistance.

Plasmid constructions. (i) Formation of pIS5. Plasmid pIS is a fresh isolate of plasmid pMV5-tk*neo* described previously (19). The LTR-LTR junction region of Mo-MuLV was obtained from plasmid pMo600, generously provided by S. W. Mitra (Merck Sharp & Dohme). This plasmid consisted of the 600-base-pair (bp) fragment released by *Xba*I cleavage of p8.8 (33) inserted into the *Pst*I site of pBR322 by GC tailing. A 220-bp fragment containing the LTR-LTR junction region was isolated after cleavage of pMo600 with *Sma*I plus *Pvu*II; *Hin*dIII linkers were added to the termini and activated by cleavage with *Hin*dIII, and the fragment was inserted into the unique *Hin*dIII site of pIS, yielding pIS5.

(ii) Formation of pVNB. A full-length proviral clone containing an 8-bp deletion in the U5 region of the 5′ LTR, termed pNCNB2, was obtained from J. Colicelli. This plasmid was made by transfer of the *Cla*I-*Xho*I fragment of a permuted clone of mutant *dl*587 (5) into the 5′ LTR region of the wild-type proviral clone pNCA (5). Plasmid pV4neo (21) was cleaved with *Eco*RI plus *Xho*I to remove the 5′ LTR region, and this portion of the provirus was replaced by the corresponding fragment of pNCNB2, yielding the vector construct pVNB.

(iii) Formation of pVNBJ2. A 61-bp fragment containing the LTR-LTR junction was isolated after cleavage of pMo600 with *Mnl*I plus *Alu*I; *Eco*RI linkers were added to the termini and activated by cleavage with *Eco*RI, and the fragment was cloned into pBR322. The fragment was excised with *Eco*RI, the termini were filled in with Klenow fragment, and *Xho*I linkers were ligated to recreate the *Eco*RI sites at the termini. The linkers were activated by cleavage with *Xho*I, and the fragment was inserted into the unique *Xho*I site of pVNB, yielding pVNBJ2.

RESULTS

Construction of retrovirus genomes carrying the LTR-LTR junction. Three retroviral genomes designed to test the ability of an LTR-LTR junction to serve as an integration substrate were generated by standard recombinant techniques. The simplest construct, pIS5, contained a selectable cassette, the bacterial neomycin resistance gene (Neo^r), driven by the promoter of the herpesvirus thymidine kinase gene, in place of the coding regions of the Mo-MuLV provirus (Fig. 1, top). The Neo^r cassette was inserted in the same transcriptional orientation as the viral transcription. Expression of the Neo^r gene was mediated by mRNAs initiated at the thymidine kinase promoter and polyadenylated at sequences in the 3′ LTR of the provirus. In this way, sequences could be inserted into the vector between the 5′ LTR and the thymidine kinase promoter without affecting expression of the marker. As a potential integration site, a 220-bp fragment containing half of the R region and all of U5 from one LTR and abutting part of U3 from the other LTR was excised from a cloned copy of the two-LTR circular viral DNA (33). This fragment was inserted in inverted orientation relative to the sequences in the normal viral LTRs into the vector pIS, forming the construct pIS5 (Fig. 1, top).

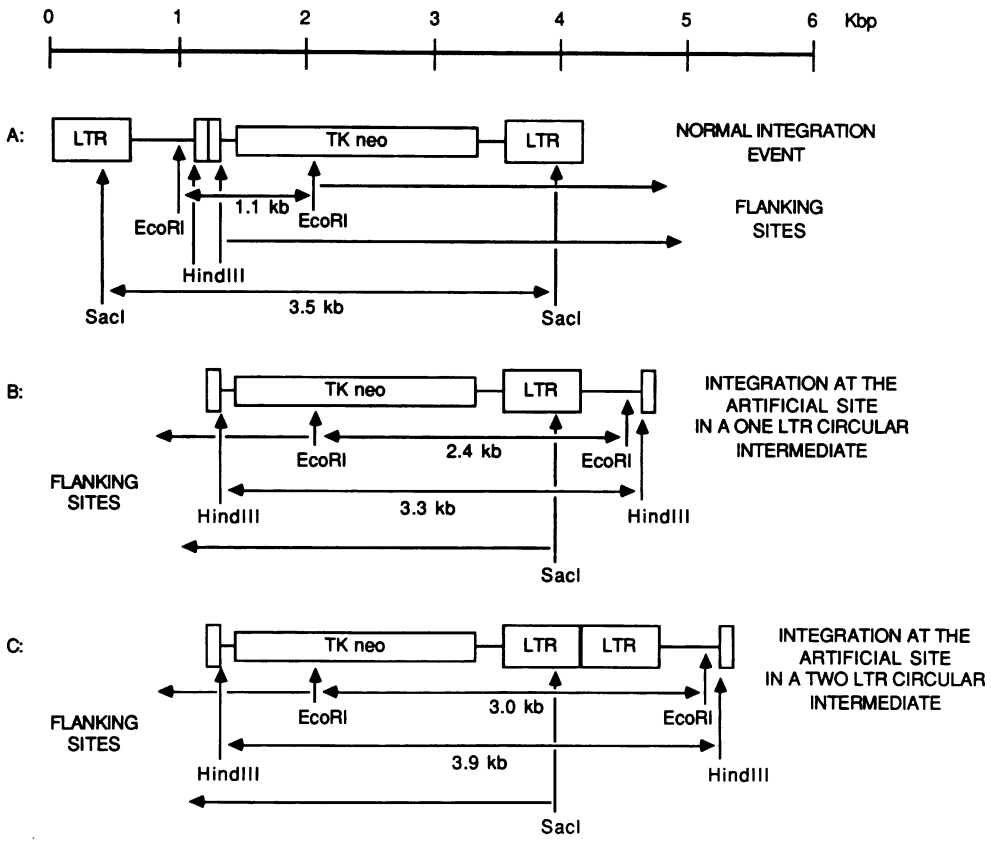

FIG. 2. Proviral structures expected from the integration of pIS5 and predicted fragment sizes after cleavage with selected restriction enzymes. (A) Product of integration at normal LTR edges. (B) Product of integration of a single-LTR circle at the ectopic LTR-LTR junction. (C) Product of integration of a double-LTR circle at the ectopic LTR-LTR junction. Kbp, Kilobase pairs.

The other constructs contained the same selectable gene, but under the control of the simian virus 40 early region promoter. They contained in addition the simian virus 40 origin of replication and a functional origin of replication from pBR322 (Fig. 1, middle and bottom). The added sequences permitted the viral DNA to act as a shuttle vector, replicating in *E. coli*, and facilitated the recovery of cloned proviruses in bacteria. As before, candidate integration sequences could be inserted between the 5′ LTR and the selectable cassette. In an attempt to reduce utilization of the natural integration sequences in the complete LTRs, we replaced the 5′ LTR with a mutant LTR bearing an 8-bp deletion at the 3′ edge of the U5 region. The presence of this mutation, *dl*587, in an otherwise wild-type virus has been shown to reduce the integration of proviral DNA profoundly, with minimal effect on synthesis of the viral DNA (5). We reasoned that the inclusion of this mutation would reduce the background of normal integrants and permit detection of even rare use of the internal, ectopic integration site.

As a potential substrate for integration, a fragment containing 30 bp of U5 abutting 31 bp of U3 was isolated from a cloned viral genome. This sequence was bounded by *Eco*RI linkers and *Xho*I linkers and was then inserted in inverted orientation relative to the normal LTRs into the vector pVNB to form the virus pVNBJ2 (Fig. 1, bottom). Integration at the natural LTRs would yield a normal provirus expressing the Neo^r marker internally, and integration at the LTR-LTR junction would give a permuted provirus that should also express this autonomous marker. To confirm that such a permuted provirus could indeed express the marker and be recovered, we transformed cells with a test DNA (clone pB2J2 digested with *Xho*I; see below) that mimicked this structure. This DNA yielded Neo^r colonies at levels comparable to those of pVNB and pVNBJ2, indicating that integrants resulting from use of the LTR-LTR junction could indeed be recovered.

Generation and analysis of cell lines carrying integrated pIS5 proviruses. To produce virus particles carrying the pIS5 viral RNA, the cloned DNA was introduced into the ψ2 producer cell line (22) and G418-resistant cells were pooled and grown to mass cultures. These cultures constitutively produce particles and selectively package only the vector RNA, without packaging the helper virus RNA. Virion particles were harvested from the culture medium, filtered, and used to infect fresh NIH 3T3 cells. Recipient cells were again selected with G418, and individual clones were picked and grown into large cultures. Genomic DNA was isolated from 21 such cloned lines and used in Southern blot analyses to determine the structure of the integrated proviruses.

Integration of the pIS5 genome at the normal LTR edges would produce proviruses of a defined structure (Fig. 2A). Cleavage of cellular DNA containing such a provirus with *Sac*I would release a full-length proviral fragment containing the Neo^r marker irrespective of flanking host sequences; cleavage with *Eco*RI or *Hin*dIII would release fragments of various sizes, depending in each case on the flanking sequences. Integration of circular DNAs at the internal LTR-LTR junction would give one of two permuted viral structures, containing either one or two of the natural LTRs (Fig.

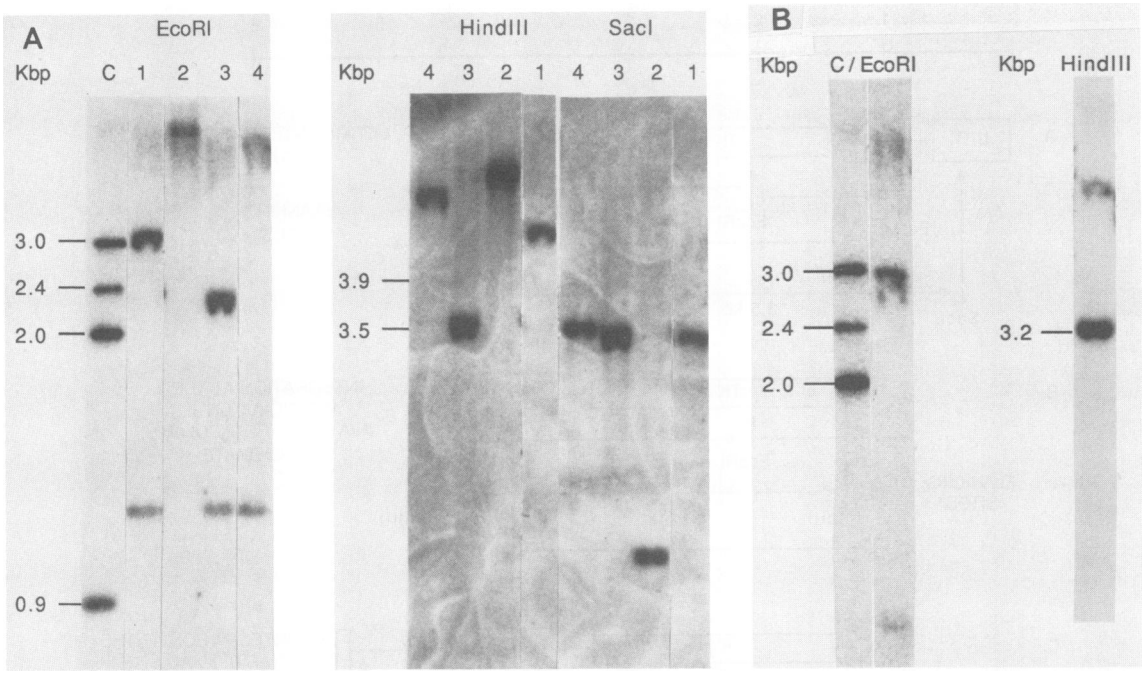

FIG. 3. Southern blots of genomic DNA isolated from cloned cell lines carrying pIS5 proviruses hybridized with a *neo*-specific DNA probe. The sizes of marker DNA fragments are indicated in kilobases (lanes C). (A) DNA samples of four cloned lines (lanes 1 to 4) digested with the indicated restriction enzymes. (B) DNA of one line, clone 111, digested with *Eco*RI and *Hin*dIII as indicated. This DNA was used to obtain molecular clone IS5λA1 (see Fig. 4A). Kbp, Kilobase pairs.

2B and C). Cleavage of cellular DNAs containing these proviruses with *Eco*RI or *Hin*dIII would release fragments of known sizes, independent of flanking host DNA; cleavage with *Sac*I would release fragments of variable size. Thus, analysis of Southern blots probed for *neo* gene homology, barring coincidental placement of flanking sites, could allow us to determine which viral sequences were used for integration.

The results of Southern blot analyses showed that the majority of the proviruses (14 of 21 examined) were intact and had unambiguously been inserted at the natural LTRs (examples shown in Fig. 3A, lanes 1 and 4). Analysis of four proviruses yielded maps inconsistent with the presence of intact genomes; in some cases, gross deletions were apparent (Fig. 3A, lanes 2). In the three remaining cases, the maps were ambiguous owing to the positions of flanking sites potentially consistent with integration at either site (Fig. 3A, lanes 3; Fig. 3B). To resolve the ambiguity for one of these cases, we isolated a proviral clone from a lambda phage genomic DNA library made from such a cloned line. Analysis of this phage, IS5λA1, revealed the presence of a full-length provirus integrated at the natural LTRs with a coincidentally placed flanking *Hin*dIII site (Fig. 4A, top). Thus, none of the 21 integration events examined had clearly utilized the internal LTR-LTR junction; only two proviruses could possibly have employed the junction, and the analysis was consistent with even these two utilizing the natural LTRs. We conclude that, at least in the context of the pIS5 genome, the natural LTRs were strongly preferred (at least 10-fold) over the junction as an integration substrate.

Generation and analysis of cell lines carrying pVNBJ2 proviruses. Experiments similar to those described above were performed with a retroviral vector, pVNB, deficient in proviral integration at its natural LTR edges by virtue of a deletion mutation in U5 (see above). A small fragment containing the LTR-LTR junction was inserted into this vector to form the construct pVNBJ2 (Fig. 1). To ensure that the helper virus in these experiments could not recombine with the construct and restore the missing sequences at the natural LTR edge, a helper genome (*dl*587 [5]) was used that also contained the same deletion in U5. A mixture of the helper genome and the selectable vector, either with or without the LTR-LTR junction, was used to transfect NIH 3T3 cells to permit transient expression, and the resulting virus mixture was allowed to spread to neighboring cells in the culture. Cloned cells stably retaining the Neo^r marker on the provirus were selected in medium with G418. As expected, the pVNB vector parent generated very few proviruses, yielding only about 100 colonies per μg of mixed DNAs applied; this low level (about 1,000-fold less than constructs carrying wild-type LTRs) reflects the limited ability of viral genomes carrying the *dl*587 deletion to form integrated proviruses (5). The construct carrying the LTR-LTR junction, pVNBJ2, yielded approximately the same low number of colonies, about 100/μg of applied DNA. Thus, the presence of the LTR-LTR junction did not rescue the integration defect of the parent vector, suggesting that if it was used at all, it was only used at a very low efficiency.

Twenty clones of cells carrying the pVNB or the pVNBJ2 proviruses were pooled, grown into mass cultures, and used to prepare genomic DNA. The DNA was cleaved, and the structure of the proviruses was analyzed by Southern blots hybridized with Neo^r sequences. A normal provirus would exhibit a well-defined structure (Fig. 5A); cleavage with *Xba*I or *Sac*I would release a constant 6.1-kilobase (kb) fragment irrespective of flanking DNA, and cleavage with *Xho*I or *Eco*RI would yield fragments of variable size. Proviruses integrated at the LTR-LTR junction would give different structures (Fig. 5B and C); here cleavage with *Xho*I

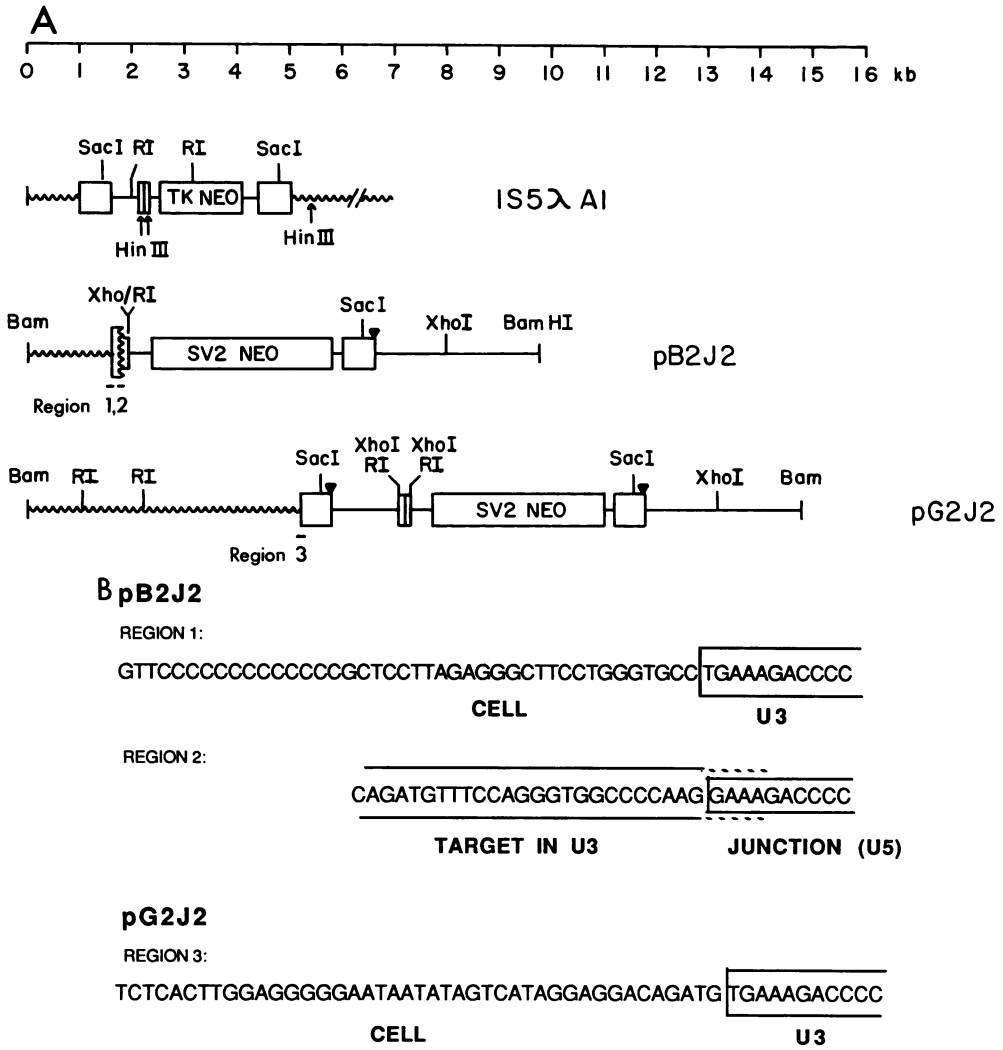

FIG. 4. Structure of cloned proviruses from cell lines infected with pIS5 and pVNBJ2 viral genomes. (A) Maps of proviral DNAs. Wavy lines indicate flanking host sequences; boxes and straight lines indicate proviral sequences. Regions 1, 2, and 3 indicate regions where nucleotide sequences were determined. RI, *Eco*RI. (B) Nucleotide sequences of selected regions of proviral clones pB2J2 and pG2J2. Large boxes indicate LTRs, and a smaller box indicates the ectopic LTR-LTR junction. Region 1, The host cell-U3 junction of the 5′ LTR of pB2J2; the U3 edge was joined to host DNA at the normal sequence. Region 2, The junction sequence formed by the autointegrative deletion between a point in the middle of U3 of the 5′ LTR and the U5 edge of the LTR-LTR junction of pB2J2; the U5 edge was joined 1 bp from the normal site. Region 3, The host cell-U3 junction of the 5′ LTR of pG2J2; joining was at the normal site.

or *Eco*RI would release fragments of specific sizes, while cleavage with *Xba*I and *Sac*I would give variable fragments.

The analysis of pooled DNAs carrying the control pVNB proviruses showed that most were intact and integrated normally, i.e., *Xba*I and *Sac*I digestion yielded a single 6.1-kb band, while *Eco*RI and *Xho*I digestion yielded many variable bands (Fig. 6). There were some prominent bands, especially after digestion with *Eco*RI; these may represent clones carrying a particular provirus that came to predominate in the pooled cell population. The parallel analysis of pooled DNAs carrying pVNBJ2 gave a different result. Cleavage with *Xba*I and *Sac*I did yield the 6.1-kb band indicative of normal proviruses but also produced other unexpected bands. We estimate that about one-half of the proviruses yielded the band indicative of normal integration. Cleavage with *Eco*RI and *Xho*I yielded some variable bands indicative of normal proviruses, but unlike the pVNB control, often gave prominent bands near 6 kb, especially with *Xho*I. These fragments could represent proviruses integrated at the LTR-LTR junction, and the relative intensity suggests that this class consists of one-half the proviruses at most. Thus, by this analysis, the pVNBJ2 proviruses might use the LTR-LTR junction at a frequency comparable to that of the *dl*587 mutant LTRs, or about 1,000-fold less well than the wild-type LTR edges. Further analysis (below) suggests that even this estimate is too high.

Cloning and characterization of two aberrant pVNBJ2 proviruses. To determine the structure of the aberrant proviruses formed by pVNBJ2, we prepared genomic DNA libraries in the phage vector λEMBL3 after cleavage with *Bam*HI, and phage containing Neo^r sequences were identified by in situ hybridization to plaque lifts. Two such phage were identified of 10^6 screened. The inserts in the phage were excised with *Bam*HI, cyclized, and used to transform bacteria to kanamycin resistance; the presence of the pBR322 sequences and the *neo* gene in the provirus permit-

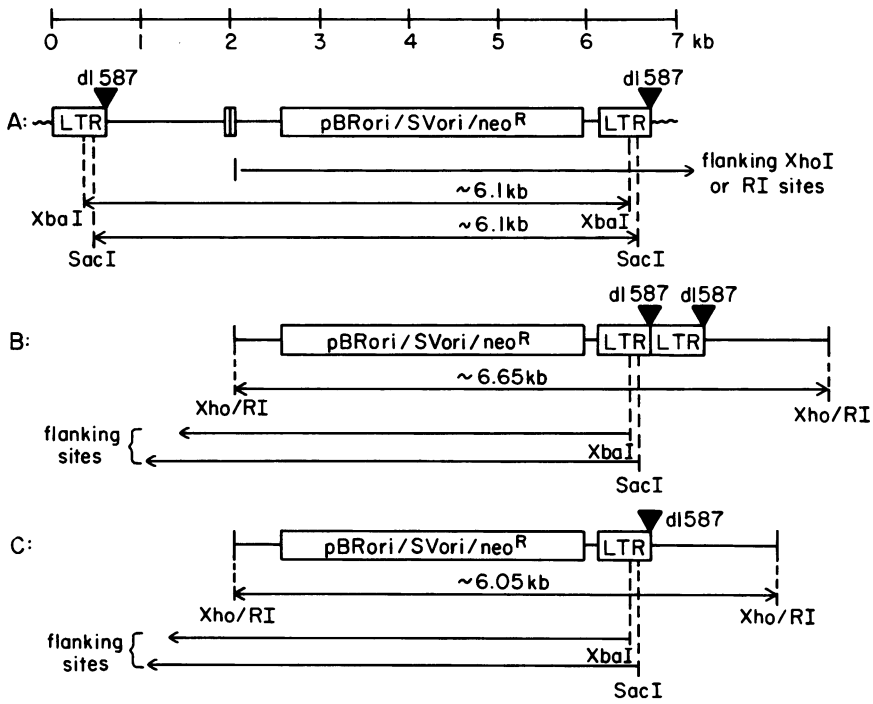

FIG. 5. Proviral structures expected from the integration of pVNBJ2 provirus and predicted fragment sizes after digestion with selected restriction enzymes. (A) Product of integration at normal LTRs. (B) Product of integration of a circular DNA containing two LTRs at the ectopic LTR-LTR junction sequence. (C) Product of integration of a circular DNA containing one LTR at the ectopic LTR-LTR junction sequence. SV, Simian virus 40; RI, *Eco*RI.

ted the circularized inserts to replicate autonomously in bacteria and confer drug resistance. The resulting plasmids, designated pB2J2 and pG2J2, were subjected to mapping, hybridization tests, and selective nucleotide sequence analysis.

One plasmid, pB2J2, contained a 6.1-kb *Xho*I fragment indicative of integration at the LTR-LTR junction but did not contain the requisite *Eco*RI fragment. Further mapping revealed a complex structure (Fig. 4A, middle). From the *Bam*HI site at one end of the fragment, DNA apparently of cellular origin extended for 2 kb to abut the 5′ half of a wild-type LTR, forming a normal host-U3 boundary at the left edge of the provirus (Fig. 4B). This LTR, however, was broken near its middle (32) and joined to the U5 edge at the LTR-LTR junction present in the pVNBJ2 provirus; sequence analysis revealed that this joining occurred by a nearly normal integration reaction that had utilized the junction (Fig. 4B). Apparently, the proviral sequences from the middle of the 5′ LTR to the LTR-LTR junction had been deleted by an integrative recombination reaction. The rest of the provirus was intact through the 3′ LTR, which retained its original *dl*587 mutation. But beyond the 3′ LTR no host sequences were found; rather, a tandem copy of the helper provirus was present, extending to the first *Bam*HI site present in the *pol* gene of Mo-MuLV.

The deletion present in pB2J2 is intriguing and represents the only use of the LTR-LTR junction for recombination that we detected. The sequence analysis showed that the reaction product is not perfect: a total of three bases (AAT) rather than the usual two were removed from the LTR edge. The other, reciprocal half of the product was not present in the clone, precluding a determination of the number of target bases duplicated, if any, in the reaction. The site of the breakpoint suggests that the LTR-LTR junction can perhaps

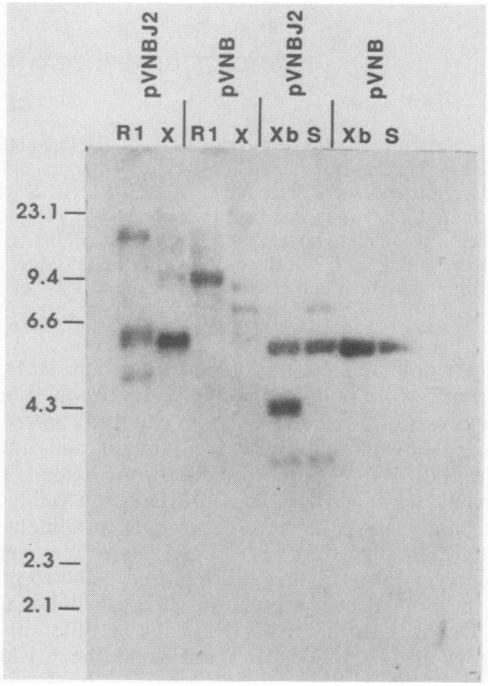

FIG. 6. Southern blot of genomic DNAs of pooled populations of cell lines derived after infection with the pVNB and pVNBJ2 proviruses, hybridized with a *neo*-specific probe. DNAs were digested with the following restriction endonucleases: R1, *Eco*RI; X, *Xho*I; Xb, *Xba*I; and S, *Sac*I. The positions of migration of marker DNAs of the indicated sizes (in kilobases) are shown at the left.

be very rarely recognized in integrative reactions. Even here, of course, we cannot exclude the possibility that the breakpoint position is coincidentally close to the junction and that this was not a true integrative reaction.

The second plasmid, designated pG2J2, had a similar but somewhat simpler structure (Fig. 4A, bottom). The DNA contained host sequences at one end, joined correctly to a 5′ LTR with the *dl*587 mutation; in this case, the provirus was complete and retained the intact LTR-LTR junction (Fig. 4B). As before, a tandem copy of the helper virus lay beyond the 3′ LTR, which again retained the *dl*587 mutation. The structure of these proviruses accounted for the presence of the 6.1-kb *Xho*I fragment detected by blot analysis (Fig. 6); i.e., the tandem vector and helper genomes contained a fragment of the exact size expected for utilization of the LTR-LTR junction during integration. But the LTR-LTR junctions in the two cloned proviruses were in fact not joined to host DNA. Similar events could account for the 6.1-kb *Eco*RI fragment seen at lower levels (Fig. 6); two tandem copies of the pVNBJ2 provirus itself would yield such a fragment. The formation of tandem copies of proviral DNA has been documented previously in a related context, the aberrant integration of proviruses lacking the retroviral IN region encoding the *trans*-acting function required for integration (15). We surmise that the *dl*587 mutation, affecting the *cis*-acting sequences required for integration, similarly resulted in the recovery of tandemly integrated proviral DNAs. In both the previous work and in this case, it remains unexplained why the elimination of natural integration mechanisms yields tandem integrants.

Generation and analysis of pVNBJ2 proviruses without helper genomes. To preclude the recombination between vector and helper genomes that formed tandem proviruses, we generated new infected cell lines using a ψ^- (packaging defective) helper virus (22). The pVNB control plasmid and the pVNBJ2 plasmid were introduced into the ψ2 cell line by calcium phosphate-mediated transformation, and stable producer lines were isolated by selection with G418. Virions were harvested and used to infect fresh NIH 3T3 cells at low multiplicity, and recipient clones were isolated after a second round of selection. Three pools of 30 to 40 colonies were made for each virus construct, grown into mass cultures, and used to prepare DNA. The structure of the proviruses was determined by Southern blots after cleavage with various restriction enzymes as before (examples shown in Fig. 7). The results show that cleavage with *Sac*I and *Xba*I released a common 6-kb fragment in perhaps one-fifth of the cases, indicative of normal proviruses; most of the fragments were of various sizes, indicative of aberrant proviruses. The patterns were similar for the vector with and without the added LTR-LTR junction. Cleavage with *Xho*I and *Eco*RI did not release the common 6-kb fragment indicative of use of the LTR-LTR junction; rather, all the fragments were of various sizes. There was thus no apparent utilization of the LTR-LTR junction, nor formation of tandem proviruses. There was not a single cell population containing a potential provirus formed by use of the LTR-LTR junction. We conclude that the LTR-LTR junction is an extremely poor substrate for integrative recombination, at least in the context of the pVNBJ2 construct, and that its efficiency must be well below that of the LTRs bearing the *dl*587 mutation.

Retroviral replication is notoriously error prone (1, 12), and the inserted LTR-LTR junctions in these constructs might be subject to pressure to undergo inactivating mutations. To determine whether the LTR-LTR junction sequence was correctly transmitted during reverse transcription and faithfully retained, we rescued one provirus from each of two of the populations of pooled colonies resulting from infection with pVNBJ2. Each cell population was fused to CV-1 cos cells to induce excision and amplification of the proviruses, and the low-molecular-weight DNA was harvested and used to transform *E. coli* to kanamycin resistance. Plasmids carrying complete proviruses were identified, and the sequence of the junction in each provirus was determined. Both junctions were intact, containing the perfect palindromic inverted repeats of the original U3-U5 joint.

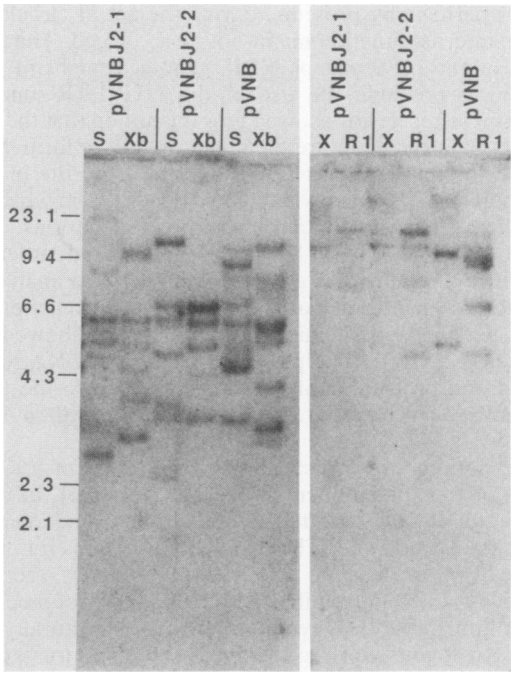

FIG. 7. Examples of Southern blots of genomic DNAs of pooled populations of cell lines derived as described in the legend to Fig. 6. DNAs were derived from one pool of colonies arising after infection with pVNB virus (lanes pVNB) and from two pools of colonies arising after infection with pVNBJ2 (lanes marked pVNBJ2-1 and pVNBJ2-2). Each sample was derived from a pool of 30 to 40 cloned cell lines. DNAs were digested with the following restriction endonucleases: S, *Sac*I; Xb, *Xba*I; X, *Xho*I; R1, *Eco*RI. The positions of migration of marker DNAs of the indicated sizes (in kilobases) are shown at the left.

DISCUSSION

Our analysis of proviruses formed by MuLV retroviral constructs carrying ectopically placed LTR-LTR junctions showed that this palindromic sequence by itself is not sufficient to act as the viral integration site for insertion into host DNA. The results suggest but do not prove that the LTR-LTR junction normally formed by circularization of the linear proviral DNA is also not the normal integrating structure and that the circular intermediate is not the natural precursor to the integrated provirus. There are, however, many alternative explanations. First, the two LTR-LTR junction fragments we used might not be large enough, or might be too large, for proper recognition. We believe that this possibility is remote, based on mutational studies showing that only short sequences at each natural LTR edge are essential (25; J. E. Murphy and S. P. Goff, manuscript in preparation). Second, the LTR-LTR junction formed by circularization might be tagged, either covalently or nonco-

valently, perhaps by proteins, during the act of circularization; the internal junction might not be so tagged. Third, the context of the pIS5 and pVNBJ2 genomes might in some special way preclude the use of the LTR-LTR junction. Despite these reservations, we favor the notion that the LTR edges of the linear DNA are the major precursor form. Some evidence for this idea has recently come from the in vitro integration system of Brown and colleagues (3), in which the most active fractions contain linear precursor forms. Even stronger evidence that the linear DNA is favored, at least in the in vitro reactions, has appeared in a detailed analysis of the structure of intermediates formed in such reactions (8; Brown, personal communication). This work showed that two bases (AA) from the tips of the linear viral DNA remain unpaired and protruding at the site of joining to the target DNA; such structures are only consistent with a linear precursor.

These notions are in contrast to the results of a similar analysis of the spleen necrosis virus system (26), in which such a junction was reported to constitute a functional and efficient integration site. The explanation for this difference is unclear; it is possible that the two viruses are truly distinctive in the relative efficiency of utilization of precursor forms, although they are evolutionarily and structurally very similar. Analogous work in the avian sarcoma virus system aimed at detecting integration at an ectopic junction sequence suggests that here too the LTR-LTR junction is at best poorly used (J. Leis; A. Skalka, unpublished observations).

We found only one clone, the pB2J2 provirus, suggesting that the LTR-LTR junction can be used in integrative recombination. The reaction was essentially a deletion between the junction and a distant site in the provirus, similar to autointegration reactions described previously (33, 34). The deletion could have arisen in the unintegrated proviral DNA, or even after its integration into the host. While clearly a rare event, the clone suggests that at some frequency the junction can be recognized and used. The reaction occurring at the junction was not quite normal in that the cleavage was three bases rather than two away from the U5 edge, or 1 bp displaced from its correct site. It is interesting that in the avian system, cleavage of the LTR-LTR junction in vitro by the IN protein, $pp32^{pol}$, can occur either two or three bases from the LTR edge, depending on experimental conditions (7, 13, 14, 20).

The pB2J2 provirus, like another recovered without a deletion, is arranged in tandem with a helper virus genome. It is not clear whether the involvement of the helper genome was important in the use of the LTR-LTR junction. The presence of tandem arrays of proviral DNAs is very rare in integration-proficient viruses but was found to be the rule for insertion of viral constructs in the absence of a functional integrase protein (15). The blot analysis (Fig. 6) suggests that with helper genomes present and when the multiplicity of infection is high, tandem arrays are also readily formed in the absence of two functional LTR edges. Perhaps whenever proviral DNAs persist in the infected cell without integrating, whether by virtue of the absence of the *trans*- or the *cis*-acting functions required for the reaction, homologous recombination can occur and give rise to tandem arrays. Plausible models for the formation of the array include simple homologous recombination between the LTRs (Fig. 8) or more complex reactions during reverse transcription. The oligomerization could be similar to the routes by which DNAs introduced by nonviral means often form large linked units (27). The eventual insertion of the array into the chromosome would be mediated by leakiness of the inactivating mutations. In this scheme, the oligomerization of the genomes would serve no selective advantage for integration but would be a gratuitous by-product of the persistence of the DNA. Indeed, in our system it is difficult to imagine an advantage to the involvement of the helper genome, since it carried the same inactivating mutation, *dl*587, as the vector genome.

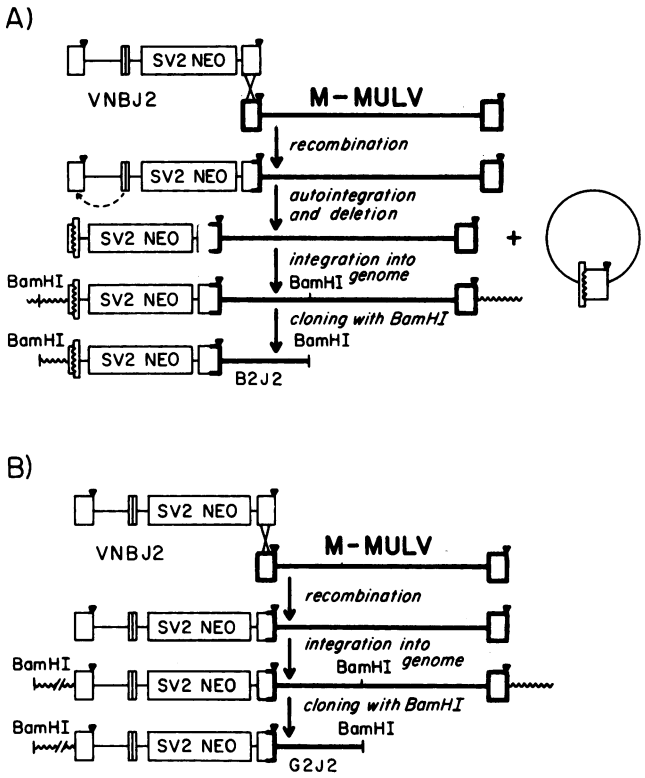

FIG. 8. Models of the formation of the pB2J2 and pG2J2 proviruses. In each case, a cell is infected with a mixed population of pVNBJ2 and helper genomes, forming a mixture of linear proviral DNAs, each bearing the *dl*587 deletion mutation in both LTRs. The DNAs oligomerize by homologous recombination in the LTRs. The oligomeric DNA is finally integrated into the host, utilizing the normal U3 edge of the 5′ LTR of pVNBJ2 and, in one possibility shown here, utilizing the mutant U5 edge of a 3′ LTR of the helper. (A) In forming the pB2J2 provirus, an autointegrative deletion results in excision of a region from the 5′ LTR to the LTR-LTR junction as a circle. (B) In forming pG2J2, no such loss occurs.

ACKNOWLEDGMENTS

We thank John Colicelli for helpful discussions.

This work was supported by Public Health Service grant CA 30488 from the National Cancer Institute.

LITERATURE CITED

1. **Battula, N., and L. A. Loeb.** 1974. The infidelity of avian myeloblastosis virus deoxyribonucleic acid polymerase in polynucleotide replication. J. Biol. Chem. **249**:4086–4093.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening lambda gt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
3. **Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop.** 1987. Correct integration of retroviral DNA in vitro. Cell **49**:347–356.
4. **Cepko, C. L., B. E. Roberts, and R. C. Mulligan.** 1984. Con-

struction and applications of a highly transmissible murine retrovirus shuttle vector. Cell **37**:1053–1062.
5. **Colicelli, J., and S. P. Goff.** 1985. Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. Cell **42**:573–580.
6. **Dhar, R., W. L. McClements, L. W. Enquist, and G. F. Vande Woude.** 1980. Terminally repeated sequences (TRS) of integrated Moloney sarcoma provirus: nucleotide sequence of TRS and its host and viral junctions. Proc. Natl. Acad. Sci. USA **77**:3937–3941.
7. **Duyk, G., J. Leis, M. Longiaru, and A. M. Skalka.** 1983. Selective cleavage in the avian retroviral long terminal repeat sequence by the endonuclease associated with the alpha-beta form of avian reverse transcriptase. Proc. Natl. Acad. Sci. USA **80**:6745–6749.
8. **Fujiwara, T., and K. Mizuuchi.** 1988. Retroviral DNA integration: structure of an integration intermediate. Cell **54**:497–504.
9. **Gilboa, E., S. W. Mitra, S. Goff, and D. Baltimore.** 1979. A detailed model of reverse transcription and tests of crucial aspects. Cell **18**:93–100.
10. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
11. **Goff, S. P., E. Gilboa, O. N. Witte, and D. Baltimore.** 1980. Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. Cell **22**:777–785.
12. **Goff, S. P., P. Traktman, and D. Baltimore.** 1981. Isolation and properties of Moloney murine leukemia virus mutants: use of a rapid assay for release of virion reverse transcriptase. J. Virol. **38**:239–248.
13. **Golomb, M., and D. P. Grandgenett.** 1979. Endonuclease activity of purified RNA-directed DNA polymerase from avian myeloblastosis virus. J. Biol. Chem. **254**:1606–1613.
14. **Grandgenett, D. P., A. C. Vora, and R. D. Schiff.** 1978. A 32,000-dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. Virology **89**:119–132.
15. **Hagino-Yamagishi, K., L. A. Donehower, and H. E. Varmus.** 1987. Retroviral DNA integrated during infection by an integration-deficient mutant of murine leukemia virus is oligomeric. J. Virol. **61**:1964–1971.
16. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–371.
17. **Hughes, S. H., A. Mutschler, J. M. Bishop, and H. E. Varmus.** 1981. A Rous sarcoma virus provirus is flanked by short direct repeats of a cellular DNA sequence present in only one copy prior to integration. Proc. Natl. Acad. Sci. USA **78**:4299–4303.
18. **Hughes, S. H., P. R. Shank, D. H. Spector, H.-J. Kung, J. M. Bishop, H. E. Varmus, P. K. Vogt, and M. L. Breitman.** 1978. Proviruses of avian sarcoma viruses are terminally redundant, coextensive with unintegrated linear DNA and integrated at many sites. Cell **15**:1397–1410.
19. **Kirschmeier, P. T., G. M. Housey, M. D. Johnson, A. S. Perkins, and I. B. Weinstein.** 1988. Construction and characterization of a retroviral vector demonstrating efficient expression of cloned cDNA sequences. DNA **7**:219–225.
20. **Leis, J., G. Duyk, S. Johnson, M. Longiaru, and A. Skalka.** 1983. Mechanism of action of the endonuclease associated with the alphabeta and betabeta forms of avian RNA tumor virus reverse transcriptase. J. Virol. **45**:727–739.
21. **Lerner, N., S. Brigham, S. P. Goff, and A. Bank.** 1987. Human β-globin gene expression after gene transfer using retroviral vectors. DNA **6**:573–582.
22. **Mann, R., R. C. Mulligan, and D. Baltimore.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell **33**:153–159.
23. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 499–599.
24. **McCutchan, J. H., and J. S. Pagano.** 1968. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. J. Natl. Cancer Inst. **41**:351–357.
25. **Panganiban, A. T., and H. M. Temin.** 1983. The terminal nucleotides of retrovirus DNA are required for integration but not virus production. Nature (London) **306**:155–160.
26. **Panganiban, A. T., and H. M. Temin.** 1984. Circles with two LTRs are precursors to integrated retrovirus DNA. Cell **36**: 673–679.
27. **Perucho, M., D. Hanahan, and M. Wigler.** 1980. Genetic and physical linkage of exogenous sequences in transformed cells. Cell **22**:309–317.
28. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labelling of DNA to high specific activity by nick translation. J. Mol. Biol. **113**:237–258.
29. **Shank, P. R., S. H. Hughes, H. Kung, J. E. Majors, N. Quintrell, R. V. Guntaka, J. M. Bishop, and H. E. Varmus.** 1978. Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell **15**:1383–1395.
30. **Shank, P. R., and H. E. Varmus.** 1978. Virus-specific DNA in the cytoplasm of avian sarcoma virus-infected cells is a precursor to covalently closed circular viral DNA in the nucleus. J. Virol. **25**:104–114.
31. **Shimotohno, K., S. Mizutani, and H. M. Temin.** 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. Nature (London) **285**:550–554.
32. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
33. **Shoemaker, C. S., S. Goff, E. Gilboa, M. Paskind, S. W. Mitra, and D. Baltimore.** 1980. Structure of a cloned circular Moloney murine leukemia virus molecule containing an inverted segment: implications for retrovirus integration. Proc. Natl. Acad. Sci. USA **77**:3932–3936.
34. **Shoemaker, C., J. Hoffmann, S. Goff, and D. Baltimore.** 1981. Intramolecular integration within Moloney murine leukemia virus DNA. J. Virol. **40**:164–172.
35. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
36. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. **1**:327–341.
37. **Van Beveren, C., E. Rands, S. K. Chattopadhyay, D. R. Lowy, and I. M. Verma.** 1982. Long terminal repeat of murine retroviral DNAs: sequence analysis, host-proviral junctions, and preintegration site. J. Virol. **41**:542–556.
38. **Varmus, H. E.** 1983. Retroviruses, p. 411–503. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
39. **Wahl, G. M., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization using dextran sulphate. Proc. Natl. Acad. Sci. USA **76**:3683–3687.
40. **Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel.** 1979. Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell **16**:777–785.
41. **Yoshimura, F. K., and R. A. Weinberg.** 1979. Restriction endonuclease cleavage of linear and closed circular murine leukemia virus DNAs: discovery of a smaller circular form. Cell **16**:323–332.