JOURNAL OF VIROLOGY, Dec. 1990, p. 5783–5796
0022-538X/90/125783-14$02.00/0

Vol. 64, No. 12

Transduction of Cellular *neo* mRNA by Retrovirus-Mediated Recombination

HEIDI STUHLMANN,* MARIANNE DIECKMANN, AND PAUL BERG

Department of Biochemistry, Beckman Center, Stanford University School of Medicine, Stanford, California 94305-5307

Received 29 May 1990/Accepted 23 August 1990

Transduction of cellular oncogenes by retroviruses is thought to be a multistep process, involving transcriptional activation of a cellular gene by upstream proviral integration and joining of cellular DNA to retroviral transcriptional signals, followed by copackaging and recombination with a helper virus genome during reverse transcription. To examine the molecular mechanism of the reverse transcriptase-mediated recombination, we introduced into mouse fibroblast cells a variety of constructs in which the *neo* selectable marker was joined to flanking retroviruslike or cell-like sequences. After superinfection and copackaging with a replication-competent Mo-MuLV^{supF} virus, the formation of recombinant *neo* transducing viruses was assessed in a second round of virus infection by the ability to confer G418 resistance to infected cells. Our results showed that recombinant *neo* proviruses were generated from *neo* RNA containing either a 5′ or 3′ retroviral end, implying that one recombination event with helper virus RNA was sufficient to incorporate the *neo* gene into proviral DNA. Recombination occurred with an apparent frequency of 10^{-4} to 10^{-5} per replication cycle in the absence of homology between the two recombining partners. This frequency, however, increased at least 100-fold if homology was provided at the site of recombination. Our results support the hypothesis that *neo*-transducing viruses arise via reverse transcriptase-mediated recombination of RNA rather than by recombination proceeding through DNA intermediates. Unexpectedly, removal of the retroviral packaging site Ψ reduced the number of *neo* recombinants only slightly. Our data indicated that although RNAs lacking the Ψ site are poorly packaged into virions, those RNAs that are included in the virions undergo frequent recombination, even if there is no selection for recombination. Many of the *neo* recombinants formed with the Ψ$^-$ constructs had undergone additional recombinations and often incorporated the Ψ site from the helper RNA.

It is widely accepted that the transforming genes of retroviruses, v-*onc*s, are derived from cellular counterparts, c-*onc*s. Ordinarily, the acquisition of a cellular gene by a retrovirus is an extremely rare event, but oncogenic retroviruses can be identified by their ability to transform cells and to form tumors in vivo. Most likely, transduction is not restricted to cellular oncogenes. Thus, a portion of the actin gene was transduced with a v-*fgr* oncogene (29) and the T-cell-receptor β chain has been transduced by the feline leukemia virus (11).

Several models have been proposed to explain the origin of transducing retroviruses. These are based on comparison of the transducing viral genomes with the structures of the parental virus and the cellular oncogene (4, 19, 37, 38). In the prevailing model proposed by Swanstrom and colleagues (37, 38), a nontransforming retrovirus integrates upstream and in the same transcriptional orientation as the c-*onc* gene. DNA rearrangement then joins the 5′-proximal domain of the viral genome including the viral long terminal repeat (LTR) and the packaging site to c-*onc* sequences. Transcription of this unit results in a chimeric RNA which is copackaged as a heterodimer with the genome of a superinfecting or coresident retrovirus. In a subsequent infection, recombination between the heterodimeric RNA molecules during reverse transcription generates a recombinant provirus containing a portion of the c-*onc* gene and the 3′ end derived from the helper virus. Coffin (8) has proposed that the recombination occurs by a variation of the copy-choice model.

In this multistep model, two recombination events are required to form an oncogene-transducing virus, the first occurring at the DNA level, and the second at the RNA level. Recombination between Harvey murine sarcoma virus lacking the 3′ end of its genome and a helper virus (12) as well as the transduction of c-*fps* by avian sarcoma virus (17, 18) provide support for this mechanism. In a variation of this model, 5′ retroviral and c-*onc* sequences can be joined by readthrough transcription from the viral sequence into the downstream cellular gene followed by aberrant splicing (30, 34). Copackaging of such transcripts with helper virus RNA and the subsequent reverse transcriptase-mediated recombination could then produce functional transducing virus. The acquisition of c-*erbB* by avian leukosis virus (25, 30, 32) suggests that the latter mechanism can also account for the formation of defective transforming viruses. Alternative models proposed for the transduction of c-*onc*s involve recombination during reverse transcription of readthrough RNA (16), a variation of the displacement-assimilation mechanism during plus DNA strand synthesis (19), and recombination on the DNA level between c-*onc* and helper virus sequences (14). Furthermore, Linial (20) has recently demonstrated that cellular mRNAs can be packaged into virions. It is still a mystery how such packaged mRNAs are reverse transcribed after infection; nevertheless, recombination with a helper virus genome is required for the generation of a transducing virus.

In this report, we examine the mechanism by which

* Corresponding author.

transducing viruses arise during reverse transcription of copackaged RNAs. Variously modified retroviral constructs carrying the dominant selectable neomycin resistance gene, *neo*, were introduced into established mouse fibroblast cells. The formation of *neo*-transducing retrovirus after infection of such cells with wild-type virus allowed us to monitor the recombination events that occurred during reverse transcription in the next infection cycle.

MATERIALS AND METHODS

Vector constructions. (i) 5′ LTR vectors. The structures of the 5′ LTR vectors are diagrammed in Fig. 2. pMΨ$^+$*neo* was constructed by ligating a 2.4-kilobase (kb) *Eco*RI-*Xho*I fragment of pMov9 (6) containing the 5′ end of Moloney murine leukemia virus (Mo-MuLV) and 5′-flanking mouse sequences to a 1.5-kb *Xho*I-*Bam*HI *neo* fragment derived from pBR*neo* (36). pMΨ$^-$*neo* was made from a 2.1-kb *Eco*RI-*Xho*I fragment of pMovΨ$^-$ (24) joined to the 1.5-kb *Xho*I-*Bam*HI *neo* fragment.

pMΨ$^{+/-}$*neo*SV were obtained by ligating the corresponding *Eco*RI-*Bam*HI fragments from pMΨ$^+$*neo* or pMΨ$^-$*neo* with a 990-base-pair (bp) *Bcl*I-*Eco*RI fragment of simian virus 40 (SV40) (from nucleotides 2770 to 1782) containing the SV40 large T polyadenylation site.

pMΨ$^{+/-}$*neo*Δ*env*M were constructed by ligating the *Eco*RI-*Bam*HI fragments described above to the 2.9-kb *Bam*HI-*Pst*I fragment of pMov9 containing the 3′ end of Mo-MuLV and 3′-flanking mouse sequences and to the 3.6-kb *Pst*I-*Eco*RI fragment of pBR322.

pMΨ$^{+/-}$*neo*Δ*env*SV were made by ligating the *Eco*RI-*Cla*I fragment derived from either pMΨ$^+$*neo*Δ*env*M (5.0 kb) or pMΨ$^-$*neo*Δ*env*M (4.7 kb) to the 990-bp *Bcl*I-*Eco*RI fragment of SV40 (containing the large T polyadenylation site) and to pBR322 DNA.

pMΨ$^{+/-}$*neo*Δ*pBR*SV were obtained by ligating the 5.1- or 4.8-kb *Eco*RI-*Cla*I fragment from pMΨ$^+$*neo* or pMΨ$^-$*neo*, respectively, containing 1.3-kb pBR322 sequences at their 3′ ends to the 990-bp *Bcl*I-*Eco*RI fragment containing the SV40 large T polyadenylation site and to pBR322 DNA.

pΔRVMΨ$^+$*neo*SV was obtained by removing the *Eco*RV fragment from pMΨ$^+$*neo*SV; this removes the 5′-flanking mouse sequences and 215 bp of U3 LTR sequences up to nucleotide 8032 (40). pΔRVMΨ$^+$*neo* was made similarly from pMΨ$^+$*neo*.

(ii) 5′ SV40 vectors. The structures of the 5′ SV40 *neo* vectors are shown in Fig. 4. pSVΨ$^+$*neo* was constructed by ligating the *Bal*I-*Bam*HI fragment of pMΨ$^+$*neo* containing Mo-MuLV sequences from nucleotides 212 to 1560 and 1.5 kb of *neo* sequences to the 340-bp *Pvu*II-*Hin*dIII fragment from pSV2*neo* containing the SV40 promoter and the 2.7-kb fragment from pBR322. pSVΨ$^-$*neo* was obtained by ligating the 2.4-kb *Hin*dIII-*Bam*HI fragment from pMΨ$^-$*neo*, which extends from the Ψ$^-$ site to the end of the *neo* sequence, to the 340-bp fragment from pSV2*neo* containing the SV40 promoter region and to the 2.7-kb segment from pBR322.

pSVΨ$^{+/-}$*neo*SV were obtained by ligating the 4.0- or 3.6-kb *Eco*RI-*Xho*I fragment from the 5′ half of either pSVΨ$^+$*neo* or pSVΨ$^-$*neo*, respectively, to the 2.5-kb *Xho*I-*Eco*RI fragment containing the *neo* and SV40 large T polyadenylation sequence from the 3′ half of pMΨ$^+$*neo*SV.

The constructs pSVΨ$^{+/-}$*neo*Δ*env*M were derived by ligating the 5.5- or 5.1-kb *Eco*RI-*Bam*HI fragment containing the SV40 promoter and *neo* sequences from pSVΨ$^{+/-}$*neo* to the 2.9-kb fragment from pMΨ$^+$*neo*Δ*env*M containing Δ*env*, LTR, and 3′-flanking mouse sequences.

Cell culture and electroporation. NIH 3T3 and XC cells were maintained in Dulbecco modified Eagle medium supplemented with 10% calf serum. The Mo-MuLV^{supF} producer clone 3T3/supF was obtained from R. Jaenisch. NIH 3T3 cells were electroporated essentially as described by Chu et al. (5). Approximately 5×10^6 cells were suspended in 1 ml of HEPES-buffered saline (1× HBS is 20 mM HEPES [pH 7.05], 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 6 mM glucose) and mixed with 0.1 or 1 μg of linearized plasmid DNA and 500 μg of salmon sperm carrier DNA. Electroporation was performed at 270 V, and the cells were plated in culture medium. Cells were allowed to express the transformed phenotype for 48 h and then split 1:20 into Dulbecco modified Eagle medium supplemented with 10% calf serum and 0.5 mg of G418 per ml. Primary (I°) G418-resistant colonies were isolated 10 to 12 days later and grown up for further analysis.

Virus infections and virus assays. Cells obtained from I° G418^r colonies were infected with Mo-MuLV^{supF} helper virus (33) at a multiplicity of infection of about 3 by overlaying 5×10^5 cells (seeded the day before in culture medium supplemented with 4 μg of Polybrene per ml) with 5 ml of culture supernatant from 3T3/supF which had been filtered through sterile filters (0.45-μm pore size; Syrfil-MF; Nuclepore Corp.). After 4 h, 5 ml of culture medium was added, and 3 days later, 5 ml of the infected culture supernatant was harvested, filtered, and used to infect fresh NIH 3T3 cells (split 1:20 the day before) in the presence of Polybrene (4 μg/ml). With some constructs, cells from I° clones were randomly subcloned in parallel 3 days after infection with Mo-MuLV^{supF}. Supernatant from these clones was used to infect NIH 3T3 cells as described above. The formation of recombinant virus genomes was assayed by trypsinizing the cells 2 days later and seeding them onto two large plates (14-cm diameter) in culture medium supplemented with 0.5 mg of G418 per ml. Secondary (II°) G418^r clones were isolated about 10 days later. Mo-MuLV^{supF} helper virus titer was determined by the XC plaque assay (35). For this, NIH 3T3 cells were plated at a density of 5×10^3 cells per well in 24-well dishes. Culture supernatants were serially diluted from 10^{-1} to 10^{-6}, and 0.2 ml of each dilution was used for infection. Six days later, cells were subjected to the XC assay. Supernatants from II° G418^r clones were analyzed for infectious and *neo*-containing virus by infecting NIH 3T3 cells with serial dilutions (10^{-1} to 10^{-6}) of filtered supernatants. Helper virus titers were determined by the XC plaque assay, and the titers of *neo*-containing virus were measured by selecting the infected cells in G418.

DNA preparation and Southern analysis. The preparation of high-molecular-weight DNA, agarose gel electrophoresis, and Southern hybridization were performed by standard procedures (22) with Zetabind membranes (AMF/Cuno). A 1.5-kb *Bam*HI-*Hin*dIII *neo* fragment from pBR*neo* (36), a 221-bp *Ava*II fragment from pMΨ$^+$*neo*Δ*env*M containing Mo-MuLV Ψ sequences from nucleotides 312 to 532, or a 220-bp *Eco*RI fragment from pVSUII (21) containing the bacterial *supF* tRNA suppressor gene was labeled with [α-^{32}P]dCTP by the hexamer method as described previously (10).

Slot-blot analysis of virion RNAs. Supernatants from I° G418^r clones (80 ml) were harvested and filtered before and after superinfection with Mo-MuLV^{supF} helper virus. Virion RNA was prepared by the following method (L. Donehower and H. Varmus, unpublished data). Supernatants were centrifuged for 1.5 h at 25,000 rpm and 4°C in an SW28 Beckman rotor. The pelleted virus was lysed in 150 μl of buffer A (10

mM Tris [pH 7.5], 200 mM NaCl, 10 mM EDTA, 0.5% sodium dodecyl sulfate) and extracted twice with phenol-chloroform (1:1) and once with ether. The ether was removed by warming the open tubes for 5 min to 60°C, and the solution was made 2.2 M with formaldehyde and incubated for 15 min at 60°C. Aliquots of virion RNA were adjusted to 10× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and spotted onto Nytran membranes (Schleicher & Schuell) with a slot-blot apparatus (Beckman Center research shop). The filters were hybridized with a ^{32}P-labeled 220-bp *Eco*RI *supF* fragment from pVSUII (21) or a 1.5-kb *Bam*HI-*Hin*dIII *neo* fragment from pBR*neo* (36).

Molecular cloning of recombinant retroviral fragments. High-molecular-weight DNA (60 μg) was digested to completion with *Xba*I and *Bam*HI endonucleases. DNA aliquots (10 μg) were analyzed on a 1% agarose gel, transferred to Zetabind filters, and hybridized with a ^{32}P-labeled *neo* probe as described above. The fragment sizes were determined by running in parallel λDNA digested with *Eco*RI and *Hin*dIII or with *Bst*EII endonucleases. To isolate restriction fragments for cloning, we electrophoresed 50 μg of digested DNA on a preparative gel, and 6-mm segments of the gel corresponding to the determined fragment sizes were excised. The DNA was purified by the glass powder method (39) and dissolved in 20 μl of TE (10 mM Tris [pH 7], 1 mM EDTA). A 10-μl sample was added to a large molar excess (0.5 μg) of *Xba*I-*Bam*HI-digested vector DNA [pBS KS(+); Stratagene] and ligated overnight in a total volume of 50 μl. After heat inactivation (10 min, 70°C), the total ligation mixture was used to transform 500 μl of competent *recA* mutant *E. coli* DH5α (15) [transformation efficiency of 5 × 10^7 to 10 × 10^7 colonies per μg of pBS KS(+) DNA]. The transformed bacteria were plated on Luria broth agar containing 50 μg of kanamycin per ml. Between 2 and 13 colonies were obtained from each ligation.

RESULTS

Experimental scheme. The aim of our experiments was to examine the mechanism of recombination between two copackaged RNA molecules during reverse transcription (Fig. 1). One of the recombination partners was provided by a series of mRNAs that contain the dominant selectable marker *neo*. The DNA templates for these mRNAs (see Fig. 2 and 4) were introduced into mouse NIH 3T3 cells under conditions that lead to a limited number of integrations. G418-resistant (G418^r) transformants were isolated, and the structure of the integrated sequences was confirmed by Southern analysis. To be certain that the primary (I°) G418^r clones selected for further analysis did not contain tandemly integrated copies of the transfected DNA, we analyzed the cellular DNA in addition by digestion with *Xho*I endonuclease, which cuts once at the 5′ border of the *neo* sequence (see Fig. 2 and 4). Three I° G418^r clones from each construct were selected which contained either a single copy (27 of 54 I° clones) or a few, non-tandemly integrated copies (27 of 54 I° clones). Cells from these clones were infected with Mo-MuLV^{supF} helper virus, a replication-competent Mo-MuLV containing the bacterial suppressor gene *supF* in the U3 region of the LTR (33). Virus produced by the infected I° clones was harvested 3 days later and used to infect fresh NIH 3T3 cells. G418^r colonies arising from this infection (secondary [II°] clones) score the recombinants formed between the superinfecting Mo-MuLV^{supF} viral RNA and the *neo* mRNA transcribed from the integrated *neo* genes in the I° G418^r cells. The recombinant proviral DNA contained

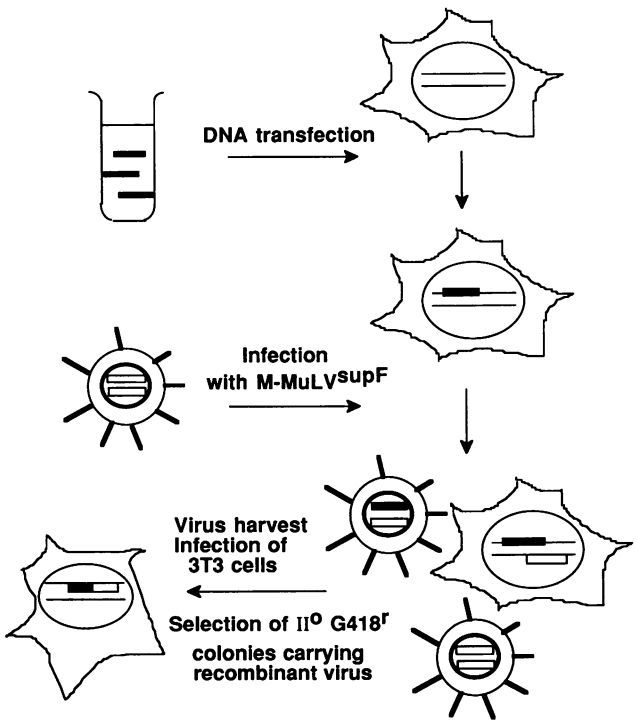

FIG. 1. Experimental scheme. NIH 3T3 cells were transfected by electroporation with DNA containing various constructs capable of expressing *neo*-containing RNA. Primary (I°) G418^r clones were selected and infected with Mo-MuLV^{supF} helper virus. Three days later, supernatant from these infected cultures was harvested and used to infect fresh NIH 3T3 cells. Virus particles in the supernatant are likely to contain either two copies of the helper virus RNA genome or, with a lower frequency, particles containing one helper virus and one *neo*-containing RNA. Viruses containing two different RNAs represent the precursors for the formation of a recombinant provirus. Recombinant proviruses containing the *neo* marker can be detected by selecting G418-resistant cells (secondary [II°] G418^r colonies).

in the II° clones was characterized by restriction enzyme digests and Southern blotting as well as by analysis of proviral transcripts. In addition, supernatants from the II° clones were assayed for production of Mo-MuLV^{supF} helper and *neo*-transducing viruses. In some cases, the structure of the recombinant proviruses was examined after molecular cloning.

***neo* mRNAs transcribed from viral LTR participate in recombination with helper virus.** First we examined recombination between Mo-MuLV^{supF} viral RNA and *neo* mRNA produced by the modified viral genomes shown in Fig. 2. These constructs contain the *neo* gene flanked at the 5′ end by the retroviral LTR sequences necessary for viral transcription, the primer binding site for initiation of minus DNA strand synthesis during reverse transcription (40), and about 1 kb of *gag* sequences. These constructs contain an LTR (MΨ$^{+/-}$*neo*Δ*env*M), no polyadenylation site (MΨ$^{+/-}$*neo* and ΔRVMΨ$^+$*neo*), or the SV40 early-region polyadenylation sequence (MΨ$^{+/-}$*neo*SV and ΔRVM$^+$*neo*SV) 3′ to the *neo* gene. The region Ψ which spans nucleotides 212 to 563 in the viral genomic RNA has been implicated in the encapsidation (24) and dimerization (40) of viral RNA. To examine whether packaging plays a role in the formation of *neo* recombinant proviruses, we compared constructs which

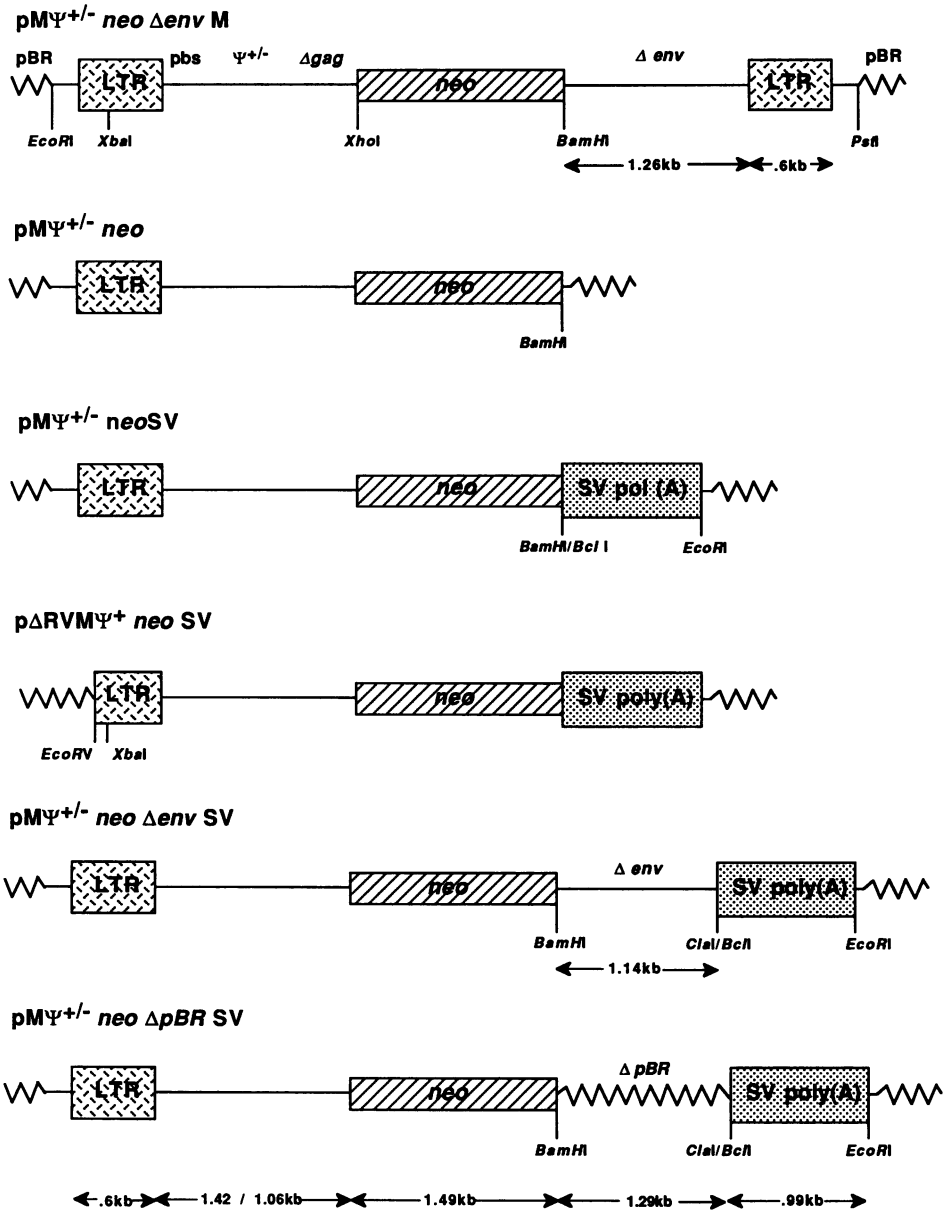

FIG. 2. Structure and derivation of recombinant 5′ LTR *neo* vectors. These vectors contain 5′ Mo-MuLV sequences up to the *Xho*I site at nucleotide 1560 (40) which includes the LTR, the primer-binding site (pbs), and 995 bp of *gag* sequences. In addition, they carry a 1.5-kb fragment containing the bacterial *neo* gene. Most constructs were generated in two forms: with (Ψ^+) or without (Ψ^-) the Ψ site. The restriction sites important for constructing the *neo* vectors as well as the length of fragments are indicated. Symbols: ▨, LTR; ▨, *neo* gene; ——, Mo-MuLV DNA sequences; ∧∧∧∧∧, pBR322 sequences; ▩, SV40 sequences.

differ only by the presence (Ψ^+) or absence (Ψ^-) of the Ψ site. Some constructs contain about 1.1 kb of *env* sequence (M$\Psi^{+/-}$*neo*Δ*env*SV) or 1.3 kb of pBR322 sequence (M$\Psi^{+/-}$*neo*Δ*pBR*SV) between the *neo* coding sequence and the 3′ end. These were used to test whether a region of homology between the defective *neo* virus and the helper virus genome would affect recombination. The constructs ΔRVMΨ^+*neo* (without a polyadenylation site; not shown in Fig. 2) and ΔRVMΨ^+*neo*SV contain a 215-bp deletion of the 5′-most U3 sequences (up to the *Eco*RV site at nucleotide 8032) in the viral LTR. This deletion includes sequences necessary for integration, but the TATAA and CCAAT sequences of the viral promoter are retained. In each of the constructs mentioned above, a single recombination event between the superinfecting Mo-MuLV^{supF} genomic RNA and the *neo* mRNA is necessary and sufficient to produce a G418^r integration-competent virus. However, with constructs lacking the packaging site Ψ, a second recombination event is required to generate high-titered *neo*-transducing virus. The construct which contains an authentic Mo-MuLV 3′ end in addition to the sequences necessary for plus DNA strand synthesis, polyadenylation, and proviral integration (M$\Psi^{+/-}$*neo*Δ*env*M) serves as a positive control. In this case, recombination is not required because packaging of the genomic viral RNA and reverse transcription are sufficient to generate II° G418^r clones.

Each plasmid DNA was linearized and transfected into mouse NIH 3T3 cells, and I° G418^r clones were selected (see Materials and Methods). The transformation efficiencies were similar for all constructs, with the exception of pMΨ$^{+/-}$*neo*, which lack a polyadenylation site and yielded 5- to 10-fold fewer transformants (data not shown). This is possibly due to a requirement for integration of the construct upstream from a cellular polyadenylation site. Three G418^r I° clones from each transfection, containing single or low copy numbers of the integrated constructs in nontandem array as determined by Southern analysis (data not shown), were infected with Mo-MuLV^{supF} helper virus at a multiplicity of infection of about 3. Three days later, culture supernatants were harvested and used to infect fresh NIH 3T3 cells.

The titers of helper virus and *neo*-transducing virus (see Materials and Methods) from three I° clones of each type are summarized in Table 1. Without infection by helper virus, the culture supernatants from each of the I° clone cells were negative in the XC plaque assay and unable to transduce G418^r. By contrast, after infection with Mo-MuLV^{supF}, the culture supernatants generated both XC plaque-forming virus and G418^r II° clones. Although the titers of the helper virus were the same in all cases and similar to that of the Mo-MuLV^{supF} producer clone, there were striking differences in the frequency of G418^r clones (Table 1). Thus, between 0.5×10^6 and 1×10^6 G418^r colonies per ml were obtained from I° clones containing the transfected control plasmid pMΨ$^+$*neo*Δ*env*M. The titer of *neo*-transducing virus produced by I° clones with the construct lacking the packaging site (pMΨ$^-$*neo*Δ*env*M) is about 10^{-4}-fold that of the control. This result is expected due to the inefficiency of packaging of the Ψ$^-$ viral RNA (23).

I° clones expressing *neo* RNAs with nonretroviral 3′ ends and therefore requiring recombination with the helper virus also produced G418^r clones, but at about 10^{-3} to 10^{-5} the frequency obtained with the positive control pMΨ$^+$*neo*Δ*env*M. Continuous passage of cells from superinfected I° clones did not change the number of *neo*-transducing virus in the supernatant (data not shown). About the same number of G418^r II° clones were obtained regardless of whether the constructs contain or lack a polyadenylation site (compare pMΨ$^+$*neo* and pMΨ$^+$*neo*SV), possibly due to usage of an endogenous polyadenylation site downstream from the integrated MΨ$^+$*neo*. When Δ*env* sequences were included 3′ to the *neo* gene, the number of II° G418^r clones increased about 100-fold (compare pMΨ$^+$*neo*Δ*env*SV with pMΨ$^+$*neo*SV, Table 1). We surmise that the homology between the *env* regions in the helper virus and *neo* RNAs contributes to the increased frequency of *neo* transduction. However, the inclusion of Δ*pBR* sequences of about the same length at this position also increased the number of II° G418^r clones about 10-fold (compare pMΨ$^+$*neo*Δ*pBR*SV with pMΨ$^+$*neo*SV), indicating that an increase in the length of RNA within which a productive recombination can occur accounts for some of the effect.

Recently, DNA recombination rather than reverse transcriptase-mediated RNA recombination has been proposed as a possible mechanism for the transduction of c-*onc*s (13, 14). Experimental evidence showed that transforming Harvey murine sarcoma virus was not, or was very inefficiently, recovered with constructs lacking a 3′ LTR and in which the U3 region of the 5′ LTR was truncated, whereas Harvey murine sarcoma virus was efficiently recovered from a construct containing an intact 5′ LTR. The U3 deletion removed the inverted repeat sequence required for integration of retroviral DNA (J. Murphy and S. Goff, personal communication) but did not affect transcription from the promoter of the LTR. Therefore, tandemly integrated DNA could be transcribed but could not generate transduced

TABLE 1. Virus production from primary G418^r clones[a]

Construct	Virus production			
	−MuLV^{supF}		+MuLV^{supF}	
	XC titer	G418^r titer	XC titer (10^6)	G418^r titer
MΨ$^+$*neo*Δ*env*M	0	0	1	5×10^5 5×10^5 1×10^6
MΨ$^-$*neo*Δ*env*M	0	0	2	17 34 29
MΨ$^+$*neo*	0	0	3	11 16 6
MΨ$^-$*neo*	0	0	2	1 0 2
MΨ$^+$*neo*SV	0	0	3	6 5 7
MΨ$^-$*neo*SV	0	0	2	1 1 1
ΔRVMΨ$^+$*neo*	0	0	1	0 12 4
ΔRVMΨ$^+$*neo*SV	0	0	1	2 5 7
MΨ$^+$*neo*Δ*env*SV	0	0	1	1,320 508 38
MΨ$^-$*neo*Δ*env*SV	0	0	1	33 29 6
MΨ$^+$*neo*Δ*pBR*SV	0	0	1	47 36 69
MΨ$^-$*neo*Δ*pBR*SV	0	0	1	1 2 1
Control (Mo-MuLV^{supF} producer)			2	0

[a] Primary G418^r clones (three from each construct) were superinfected with Mo-MuLV^{supF} helper virus as described in Materials and Methods. Supernatants were harvested 3 days later and titered for the presence of helper virus by the XC plaque assay (35) and for recombinant virus by selection in Dulbecco modified Eagle medium supplemented with 10% calf serum and 0.5 mg of G418 per ml (see Materials and Methods). The titers are given as numbers of XC plaques or G418^r colonies per ml. The XC titers are shown as an average number from three different clones. Undiluted supernatants from I° clones before superinfection with Mo-MuLV^{supF} were assayed for the presence of helper and *neo* virus.

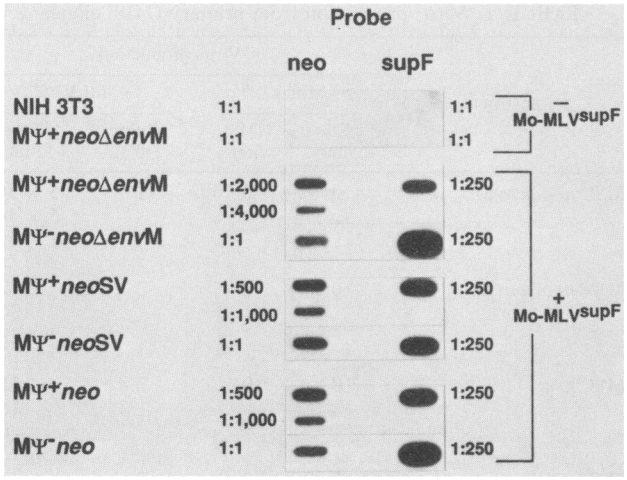

FIG. 3. RNA slot-blot analysis from packaged RNA. Virion RNA from I° G418^r clones was prepared before (− Mo-MuLV^{supF}) or after (+ Mo-MuLV^{supF}) superinfection with helper virus as described in Materials and Methods. Either undiluted or diluted samples were spotted onto Nytran membranes. The filters were hybridized with 10^7 cpm of a ^{32}P-labeled 1.5-kb *neo* fragment or with 10^7 cpm of a ^{32}P-labeled 220-bp *supF* fragment.

integrated Harvey murine sarcoma virus without additional recombination with helper virus RNA during reverse transcription. In contrast, when we examined similar constructs in which 215 bp was deleted from the 5′ end of the LTR U3 region (pΔRVMΨ$^+$*neo* and pΔRVMΨ$^+$*neo*SV), the frequency of II° G418^r clones was essentially the same as that obtained with pMΨ$^+$*neo*SV (Table 1). Therefore, this control experiment provides further support that transduction of *neo* in our experimental design results from recombination between the copackaged *neo* and helper virus RNA.

With some constructs, I° clone cells were randomly subcloned 3 days after superinfection with Mo-MuLV^{supF} helper virus. Southern analysis of DNA from 32 subclones (8 from each I° clone) showed that the structure of the integrated constructs remained unchanged; moreover, the clones did not contain any additional or new *neo* copies. Furthermore, the majority of these clones produced high titers of helper virus (32 of 32), as well as *neo*-transducing virus (25 of 32) with similar frequencies as observed for the corresponding I° clones before subcloning (data not shown). This result indicates that recombinant *neo* viruses did not arise by rare, illegitimate recombination between the integrated *neo* construct and the superinfecting helper virus on the DNA level or by recombination with helper virus during reinfection of I° clone cells.

Deletion of the packaging signal in the various constructs lowered the number of *neo* recombinant proviruses (Table 1). However, the frequencies were only 5- to 50-fold lower than those obtained with the respective Ψ$^+$ constructs. This was unexpected because of the marked difference in *neo* transductions found with the control constructs containing or lacking the Ψ site (pMΨ$^+$*neo*Δ*env*M and pMΨ$^-$*neo*Δ*env*M, Table 1). This surprising result might be explained if the mRNA transcribed from the Ψ$^-$ constructs is copackaged with the helper RNA genomes at a much higher than expected frequency. This could increase the likelihood of forming *neo* proviral DNAs in the subsequent infection. However, measurements of the amount of *neo* and helper RNA contained in virus particles in the supernatants of infected I° clones did not support that possibility (Fig. 3). As expected from the similar helper virus titers in such supernatants (Table 1), the amount of Mo-MuLV^{supF} RNA is about the same for each of the infected I° clones. More importantly, the virion particles in the supernatant from each of the infected clones carrying the Ψ$^-$ constructs contained about 10^3-fold-lower levels of *neo* RNA compared with their Ψ$^+$ counterparts, a result similar to that observed with the control constructs MΨ$^{+/-}$*neo*Δ*env*M (Fig. 3). This result indicates that the unexpectedly high frequency of *neo* transduction with the Ψ$^-$ constructs is not due to their more efficient packaging. Moreover, this finding indicates that Ψ$^+$*neo* recombinants are not formed in the I° clones because such recombinants would have been efficiently packaged in the released virus.

***neo* mRNAs transcribed from the SV40 promoter serve as substrates for recombination.** We next examined whether *neo* transcripts lacking 5′ LTR sequences recombine with helper virus RNAs to produce *neo*-transducing virus. The constructs used to create the I° G418^r clones contained the SV40 early region promoter in place of the Mo-MuLV LTR (Fig. 4). The SV40 promoter sequences were joined to Mo-MuLV sequences just beyond the primer-binding sequence and, therefore, included the Ψ site, about 1 kb of the viral *gag* sequences, and the *neo* gene; comparable constructs lacking the Ψ site were made as mentioned in Materials and Methods. The 3′ ends of these constructs contained the Mo-MuLV LTR, the SV40 early region polyadenylation site, or no polyadenylation sequences.

Table 2 shows that none of the G418^r I° clones produced *neo*-transducing virus without infection with the helper virus. Even after infection, the *neo* RNAs that are unable to serve as templates for reverse transcription because they lack the primer-binding sequence and, in addition, sequences necessary for plus DNA strand initiation (SVΨ$^+$*neo*, SVΨ$^+$*neo*SV) failed to produce G418^r II° clones. By contrast, *neo* RNAs lacking the 5′ site for initiating minus DNA synthesis but containing the characteristic retroviral 3′ end (SVΨ$^+$*neo*Δ*env*M) underwent recombination with the infecting viral RNA. Here too, *neo* RNAs lacking the Ψ site still underwent recombination with the helper viral RNA, albeit at about one-fourth the efficiency of the comparable Ψ$^+$ construct.

We rationalize these results as follows. Although SVΨ$^+$*neo*Δ*env*M may be copackaged with the helper viral RNA, only the helper RNA can serve for initiation of minus DNA strand synthesis. However, once initiated on the helper viral RNA, reverse transcription can continue at the 3′ terminus of the *neo* RNA after intermolecular strand transfer (31). Under these circumstances, only a single recombination event is needed to incorporate the *neo* sequence between two LTRs in the proviral DNA. By contrast, the *neo* RNAs transcribed from the two constructs lacking the 3′ LTR are unable either to prime the minus-strand DNAs or to accept the minus DNA strand initiated on the helper RNA. Accordingly, minus DNA strand synthesis could only continue at the 3′ end of the helper RNA, and two recombinations would be needed to acquire the *neo* gene. We conclude that the failure to detect recombinants with the I° clones containing SVΨ$^+$*neo* and SVΨ$^+$*neo*SV constructs stems from the low probability of two recombinations needed for the acquisition of the *neo* sequence into a proviral DNA. In addition, the inefficiency with which minus-strand DNA initiated at the 5′ end of the helper RNA can be transferred to the 3′ end of that RNA may contribute to the inability to detect recombinant *neo* viruses (31).

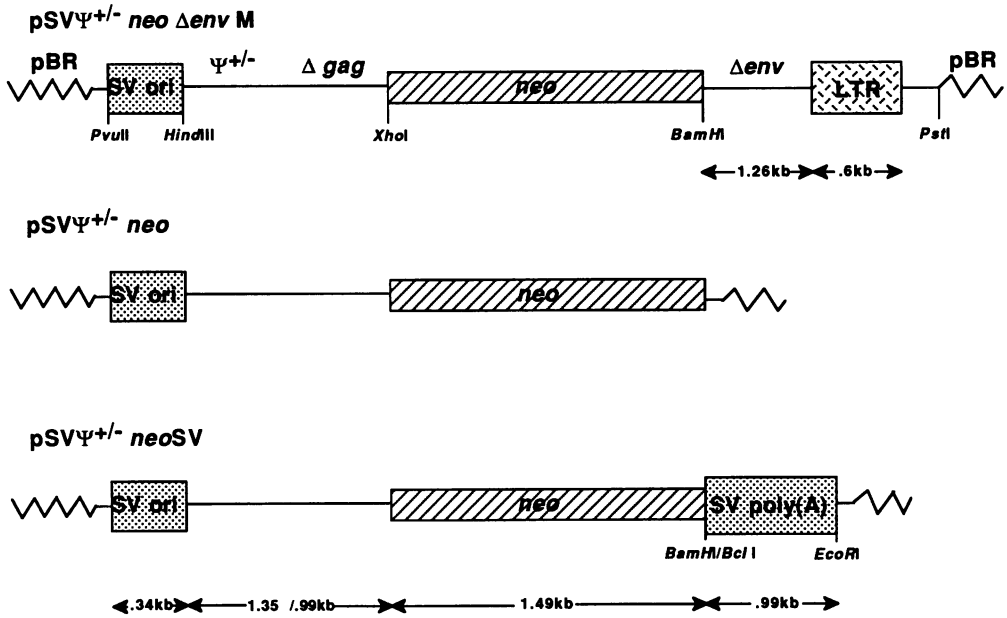

FIG. 4. Structure and deviation of recombinant 5′ SV40 *neo* vectors. All vectors contain the SV40 early promoter at their 5′ end. The Ψ^+ constructs contain Mo-MuLV-derived sequences from nucleotides 212 to 1560 including the packaging site and 995 bp of *gag* sequences (40). The Ψ^- constructs carry Mo-MuLV-derived sequences from nucleotide 563, the 3′ boundary of the packaging deletion (24), to nucleotide 1560. Restriction sites that were used in the construction of the *neo* vectors and the fragment sizes are indicated. Symbols: ▦, SV40 sequences; ▨, *neo*; , LTR; ——, Mo-MuLV DNA; ∧∧∧∧∧, pBR322 sequences. ori, Origin.

Structures of recombinant proviruses produced from Ψ^+ genomes. To provide an estimate of the number of *neo*-containing genomes per G418^r II° clone, we digested cellular DNAs with *Xho*I endonuclease, which cleaves the recombinant genomes once at the 5′ junction between the viral and *neo* sequences (Fig. 2). The *Xho*I digests of three G418^r II° clones from each of three independent I° clones obtained by transformation with the pMΨ^+*neo*Δ*env*M construct showed that each contained one or at most two copies of the *neo* provirus; moreover, these were integrated at different chromosomal locations (data not shown).

Information about the recombinant proviral structures was obtained by *Xba*I endonuclease digests of the G418^r II° clone DNAs. *Xba*I endonuclease cleaves proviral DNA in each LTR; consequently, the length of the *neo*-containing fragment is diagnostic of the recombination events that gave rise to each G418^r II° clone. As expected, the *Xba*I digests of DNA from the MΨ^+*neo*Δ*env*M-derived I° and II° clones contained the same-sized *neo*-containing fragment (Fig. 5). This follows because copackaging of the MΨ^+*neo*Δ*env*M RNA with the helper virus is all that is required to transfer G418^r to the II° clones; i.e., recombination is not necessary to produce *neo*-transducing virus.

The same analysis was done with G418^r II° clones resulting from recombination between the helper virus RNA and RNA transcribed from the integrated MΨ^+*neo*SV sequences. Judging from the number of *neo*-containing fragments produced by *Xho*I endonuclease cleavage of 24 G418^r II° clone DNAs derived from three independent I° clones (data not shown), we surmise that each contains only one or at most two recombinant proviruses per genome. Recombinant G418^r II° clones referred to above should contain the *neo* sequence flanked by two LTRs, and therefore, *Xba*I endonuclease digests should yield *neo* fragments that differ in size from one another depending on where the recombination occurred. Moreover, the *neo* fragments from the II°

TABLE 2. Virus production from primary G418^r clones[a]

Construct	Virus production			
	−MuLV^{supF}		+MuLV^{supF}	
	XC titer	G418^r titer	XC titer (10^6)	G418^r titer
SVΨ$^+$*neo*Δ*env*M	0	0	9	30 16 13
SVΨ$^-$*neo*Δ*env*M	0	0	4	4 6 6
SVΨ$^+$*neo*	0	0	5	0 1 0
SVΨ$^-$*neo*	0	0	12	0 0 0
SVΨ$^+$*neo*SV	0	0	6	0 0 0
SVΨ$^-$*neo*SV	0	0	7	0 0 0
Control (Mo-MuLV^{supF} producer)			2	0

[a] Primary G418^r clones (three from each construct) were analyzed for virus production before and after superinfection with Mo-MuLV^{supF} helper virus as described in Materials and Methods. Production of helper virus was determined by the XC plaque assay (35), and recombinant *neo* virus was detected by selection in Dulbecco modified Eagle medium supplemented with 10% calf serum and 0.5 mg of G418 per ml. The titers are given as numbers of XC plaques or G418^r colonies per ml. The XC titers shown represent the average number from three different clones.

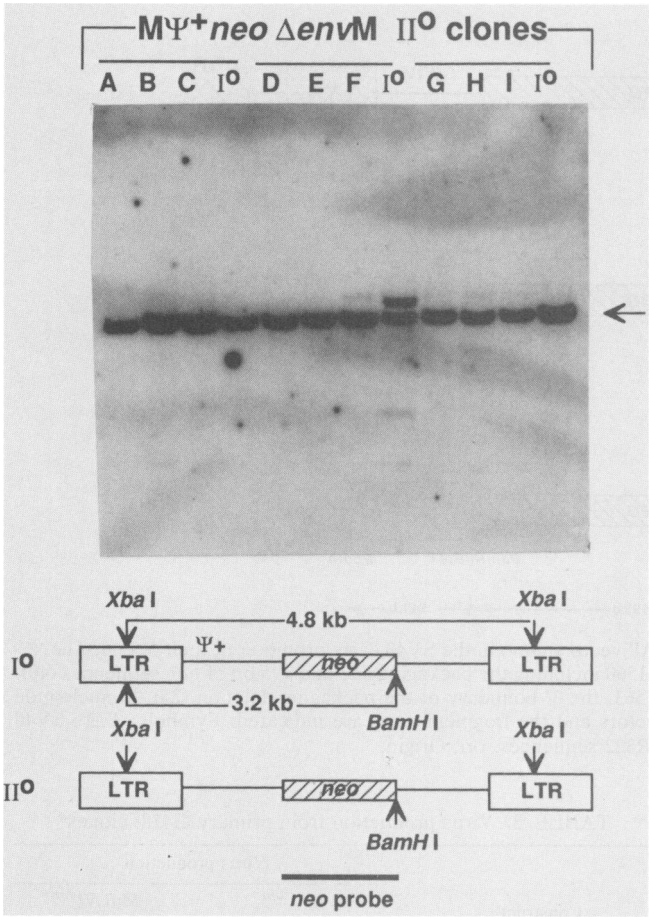

FIG. 5. Southern blot analysis of DNA from II° clones derived from MΨ+*neo*Δ*env*M. High-molecular-weight DNA (10 μg) was digested to completion with *Xba*I endonuclease, separated on 1% agarose gel, and transferred to Zetabind filters. λ DNA, digested with *Eco*RI and *Hin*dIII endonucleases, was run as a size marker in parallel. The filter was hybridized with 10^7 cpm of a ^{32}P-labeled 1.5-kb *neo* fragment. Top: Lanes labeled A to I represent *Xba*I digests of 10 μg of DNA from three sets of II° clones. The lanes marked I° associated with each set show the *Xba*I digests of the I° clone DNA from which each set of II° clones was obtained. The arrow indicates the position of the expected *Xba*I fragments. Bottom: Schematic structure of MΨ+*neo*Δ*env*M proviral DNA. The arrows indicate the locations of the *Xba*I and *Bam*HI cleavage sites and the sizes of the expected fragments in the I° clone DNA.

clones should differ from those of their I° progenitors. Figure 6 shows the digests of three sets of eight II° clones, each set being derived from an independent I° clone. Clearly, the sizes of the recombinant viral sequences differ from one another and from the I° clones from which they derive. This finding indicates that the II° clones have independent origins and that the recombinations do not occur at hot spots.

Similar analyses were done on multiple G418^r II° clones originating from I° clones transformed with pMΨ+*neo*, pMΨ+*neo*Δ*pBR*SV, ΔRVMΨ+*neo*, and ΔRVMΨ+*neo*SV (data not shown). In each case, the *neo*-containing fragments generated by *Xba*I endonuclease digestion differed in size from one another and from those in similar digests of I° clone DNA, indicating that here too each II° clone had an independent origin, probably due to recombination at different locations. In addition, Northern (RNA) analysis of total cellular RNA from several I° and II° G418^r clones supported these conclusions. Distinct *neo* RNA species, which differed in size among different II° clones and from the I° clones, were found, indicating that there were multiple independent recombination events during the formation of the G418^r II° clones (data not shown).

We next tested whether the recombinant *neo* proviruses in II° clones had acquired the 3′ LTR from the Mo-MuLV^{supF} helper virus genome. The blots shown in Fig. 6 were rehybridized with a ^{32}P-labeled *supF* fragment derived from the U3 region of the Mo-MuLV^{supF} LTRs (33). Most of the II° clone digests contained a complex array of *supF*-positive bands, indicating that they contained many (10 to 15) helper provirus copies per genome. However, almost all the *neo*-containing *Xba*I fragments cross-hybridized with the *supF* probe (data not shown). This result is consistent with the notion that the recombination between the helper virus RNA and the *neo* transcripts occurred during reverse transcription.

To determine whether recombination at homologous sites was favored, we analyzed the *Xba*I fragments from II° clones derived from the constructs MΨ+*neo*Δ*env*SV and SVΨ+*neo*Δ*env*M. Most of the recombinant proviruses appeared to result from recombinations within the homologous region. Thus, 22 of 24 of the *neo*-containing *Xba*I fragments from the II° clones derived from the I° clones containing the MΨ+*neo*Δ*env*SV construct were the same size, although they differed from the fragment cleaved from the I° clone DNA (data not shown). Similarly, 20 of 24 *neo* fragments from the II° clones generated from the SVΨ+*neo*Δ*env*M construct were the same size but different from the ones obtained from their corresponding I° clone DNA. This suggests that most of the recombinations occurred within the sequences flanking the *neo* gene, that is, within Δ*env*, or within sequences upstream of *neo*, respectively (see Fig. 2 and 4).

Because the recombinant *neo* proviruses in II° G418^r clones have acquired a retroviral 3′ end from Mo-MuLV^{supF} virus, they should produce recombinant *neo* virus at higher titers than did the I° clones. We found that the II° clones produced about the same titers of helper virus as the I° clones but that there was considerable variation in the amounts of *neo* virus (between 10^1 and 10^6 G418^r colonies per ml). We presume that this variation in the *neo* virus titer is due to the variation in expression of the different proviral integration sites and possibly to the variations in the structure of each recombinant.

Role of Ψ sequence in retroviral recombination. We have already noted that I° clones containing the MΨ+*neo*Δ*env*M construct yielded high levels of *neo*-transducing virus after infection with the helper Mo-MuLV^{supF} (Table 1). Moreover, there was a nearly 10^4- to 10^5-fold reduction in the production of *neo*-transducing virus when the Ψ site was removed (Table 1). The explanation for this seems quite straightforward. The *neo* RNA transcribed from the MΨ+*neo*Δ*env*M construct can be packaged into virions for transduction of the II° clones. Furthermore, recombination is not needed to form a *neo*-containing proviral DNA; the dispensability of recombination is supported by the finding that *Xba*I endonuclease digestion of these I° and II° clone DNAs generates the same *neo*-containing fragment (Fig. 5). By this reasoning, the markedly reduced frequency of *neo*-transducing viruses in I° clones containing the corresponding Ψ$^-$ construct stems from the very inefficient packaging of Ψ$^-$ RNA into virions (Fig. 3). Because recombination is not needed to produce MΨ$^-$*neo*Δ*env*M proviruses, we expected

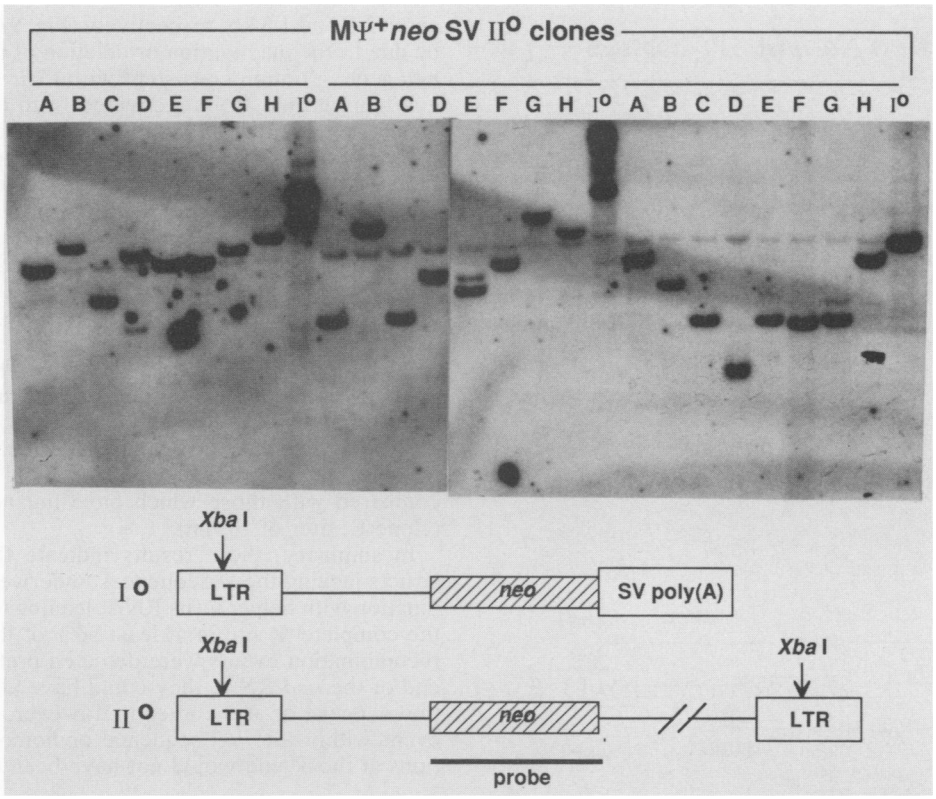

FIG. 6. Southern blot analysis of DNA from II° clones derived from MΨ+*neo*SV. Digestions, electrophoretic separations, and hybridizations were performed as described in the legend to Fig. 5. Top: *Xba*I digests of 10 μg of DNA from three sets of II° clones (lanes A to H), each set being derived from different I° clones. Bottom: Schematic structure of integrated MΨ+*neo*SV DNA in I° clones and predicted structure of the recombinant provirus in II° clones, indicating the positions of the *Xba*I restriction sites and the probe used for hybridization.

to find the same size *neo*-containing fragments in the corresponding II° and I° G418^r clones. But this is not the case, as the *Xba*I endonuclease digests of six of nine II° clones derived from Mo-MuLV^{supF}-infected I° clones containing the MΨ$^-$*neo*Δ*env*M construct contained either smaller or larger *neo* fragments than those obtained from their I° clone source (Fig. 7). In a second experiment (data not shown), in which 20 II° G418^r clones were obtained after Mo-MuLV^{supF} infection of two different MΨ$^-$*neo*Δ*env*M I° clones, 12 yielded *neo* fragments after *Xba*I endonuclease digestion that were different in size from their I° clone progenitors. These results suggest that retroviral RNAs lacking a Ψ sequence undergo frequent recombination during reverse transcription and that such events are detectable even in the absence of selection.

To determine whether there was a hot spot for this Ψ$^-$-related hyper-recombination, we compared the structures of the *neo*-containing proviral DNA in the II° clones derived from the MΨ$^+$*neo*Δ*env*M and MΨ$^-$*neo*Δ*env*M constructs (Fig. 8). After digestion with *Xba*I and *Bcl*I endonucleases, the Ψ$^+$ construct yielded *neo*-containing fragments of 2.1 and 2.75 kb from the 5′ and 3′ halves, respectively, while the Ψ$^-$ version generated 1.7- and 2.75-kb fragments from the same regions (Fig. 8A, bottom). Three II° clones and their I° clone progenitor derived from MΨ$^+$*neo*Δ*env*M produced the expected size fragments, indicating that all retained the structure of the MΨ$^+$*neo*Δ*env*M DNA. However, three of the five II° clones derived from the MΨ$^-$*neo*Δ*env*M construct yielded *neo*-containing fragments from the 5′ half that differed in size from the expected 1.7 kb, from each other, and from the fragment produced by the I° clone DNA (Fig. 8A). By contrast, all the II° clones and their I° progenitor produced the expected 3′-specific 2.75-kb fragment. To further identify the recombination site, we digested the same II° and I° clone DNAs with *Hin*dIII and *Bcl*I (Fig. 8B). Here we focused specifically on the region between the *Hin*dIII restriction site created by the removal of the Ψ sequence (24) and the *Bcl*I restriction site at the 5′ end of the *neo* coding sequence (Fig. 8B, bottom). Instead of the 1.3-kb fragment characteristic of the MΨ$^-$*neo*Δ*env*M DNA in the I° clone, each II° clone yielded different size fragments. These results suggest that both homologous and nonhomologous recombinations took place at the 5′ end of the MΨ$^-$*neo*Δ*env*M between the Ψ site and the *neo* gene.

When DNA from MΨ$^-$*neo*- and MΨ$^-$*neo*SV-derived II° clones was digested with *Xba*I and *Bcl*I endonucleases, five of seven of the recombinant proviruses produced *neo*-containing fragments that differed in length from the 1.7 kb characteristic of the 5′ end of the constructs (data not shown). Thus, in these cases, in which recombination at the 3′ end is needed for the provirus to acquire the *neo* sequence, recombination often occurs at the 5′ end as well.

To locate the recombination hot spot at the 5′ end more precisely, we cloned the 5′ segment of several proviral genomes from II° clones derived from the MΨ$^-$*neo*Δ*env*M constructs. Cellular DNA was digested with *Xba*I and *Bam*HI endonucleases, and the appropriate size *neo*-containing fragment was identified by Southern blot hybridization with a ^{32}P-labeled *neo* fragment (Fig. 9A). DNA enriched for the 5′ *Xba*I-*Bam*HI fragment was purified from a

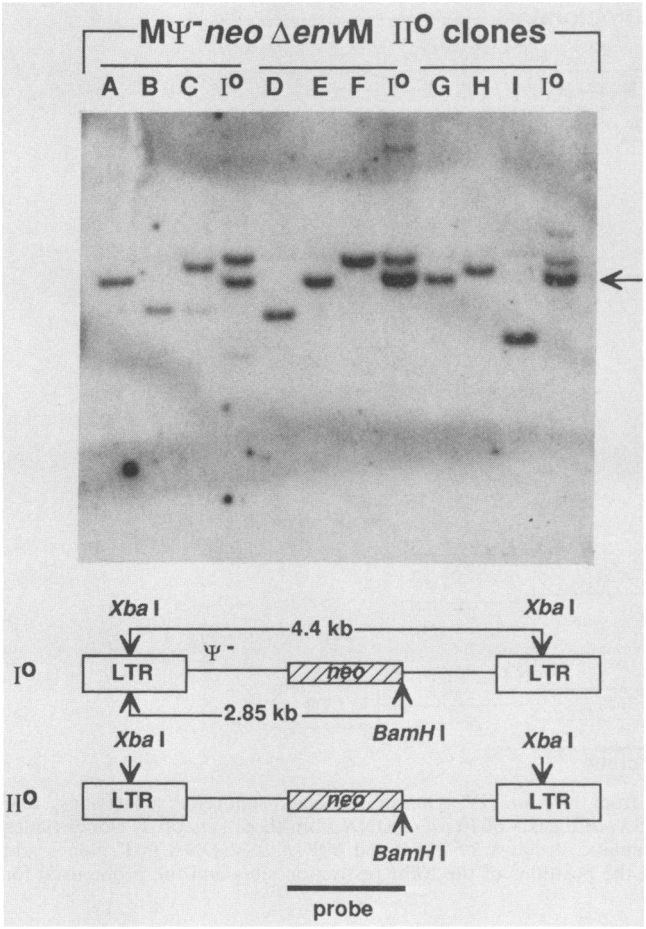

FIG. 7. Southern blot analysis of DNA from II° clones derived from MΨ⁻*neo*Δ*env*M. Digestions, electrophoretic separations, and hybridizations were performed as described in the legend to Fig. 5. Top: Lanes labeled A to I show the *Xba*I endonuclease digests of DNA from three sets of II° clones. The lanes marked I° associated with each set contain the digests of the I° clone DNA from which each set of II° clones was derived. The arrow indicates the position of the expected *Xba*I fragment. Bottom: Schematic structure of MΨ⁻*neo*Δ*env*M proviral DNA, the relevant restriction sites, and the expected sizes of fragments in I° and II° clone DNAs.

preparative gel, ligated to vector DNA [Bluescript KS(+)] cut with *Xba*I and *Bam*HI, and transfected into highly competent *recA* mutant *E. coli* (strain DH5α), and *neo* clones were selected for kanamycin resistance. We readily obtained bacterial clones from II° clones C, E, L, and H (Fig. 9A). Restriction enzyme analysis showed that the plasmid DNA recovered from the bacterial clones contained the same size inserts as those identified in the II° clone proviral DNA digests (Fig. 9A). Furthermore, digestion of the cloned proviral DNA inserts with *Dde*I or *Ava*II (generating 10 or 8 fragments, respectively) showed that the insert derived from II° clone E was identical to the 5′ end of MΨ⁻*neo*Δ*env*M and that the inserts from II° clones C, L, and H were identical to the 5′ end of MΨ⁺*neo*Δ*env*M (data not shown). In each case, the size of the cloned *Xba*I-*Bam*HI fragment was in agreement with the restriction analysis of the proviral DNA in the corresponding II° clones. We did not recover any cloned plasmids containing inserts from II° clones D and I, each of which yielded altered 5′ fragments of about 1.9 and 1.3 kb, respectively (Fig. 9A). This failure may be due to the inactivation or deletion of the promoter of the *neo* gene, a notion consistent with their smaller sizes.

When plasmid DNA recovered from the bacterial clones was digested with *Xba*I-*Bam*HI endonucleases and probed with a ^{32}P-labeled fragment containing the Ψ sequence, only the 3.2-kb fragments from clones C, L, and H and from the pMΨ⁺*neo*Δ*env*M control DNA hybridized strongly (Fig. 9B), confirming that these recombinants have acquired the packaging site (see above). DNA from clone E showed the same size 2.85-kb band of very low intensity as pMΨ⁻*neo*Δ*env*M, indicating that clone E contains a *neo* provirus which still lacks the packaging site (Fig. 9B). The weak hybridization is due to weak cross-hybridization of the probe with sequences flanking Ψ. Further support for the acquisition of the Ψ site by clones C, L, and H stems from the finding that these II° clones produce consistently higher titers of *neo* virus (10^4 to 10^5 G418^r colonies per ml) compared with those which have not acquired the Ψ site (clone E, titer of 10^1/ml).

In summary, these results indicate that retroviral constructs lacking the Ψ sequences underwent frequent recombination with helper virus RNA, leading to the acquisition of the complete Ψ site in at least 50% of the cases. Although recombination events were detected preferentially at the 5′ end of the *neo* RNA, they could have also occurred within the *neo* gene or at the 3′ end. However, any recombination event within the *neo* sequence or homologous recombinations at the 3′ end would not have been detected.

DISCUSSION

Several mechanisms have been proposed for the genesis of transducing retroviruses. The most widely accepted model, proposed by Swanstrom and colleagues (37, 38) assumes that the initial step involves rearrangement of cellular sequences with those of an upstream integrated provirus, resulting in a hybrid transcriptional unit in which 5′ retroviral sequences are joined to 3′ cellular sequences. The incorporation of cellular sequences into proviral DNA is presumed to occur only after infection with wild-type virus when viral and hybrid RNAs are copackaged into virions. Ostensibly, recombinant proviral DNAs are formed during the next round of infection by reverse transcriptase as it copies sequences from each RNA. The experiments reported here examine several of the parameters that influence the efficiency of this reverse transcriptase-mediated recombination.

Cells transformed with a variety of different DNA constructs designed to produce retroviruslike or cell-like RNAs containing the selectable *neo* gene provided potential partners for recombination with a superinfecting helper virus. One set of constructs produced RNAs with 5′ retroviral termini and either a retroviral or cellular 3′ terminus; such RNAs either included or lacked a viral sequence, Ψ, that has been implicated in highly efficient packaging of viral RNA into virions (24). Additionally, cell lines were established whose *neo* transcripts lacked retroviruslike 5′ ends but contained either a retroviral or cellular 3′ terminus; here too, the RNAs included or lacked the Ψ sequence. None of these cell lines can transduce the *neo* gene to other cells spontaneously, but several generate *neo*-transducing virus after infection with wild-type Mo-MuLV^{supF}. Our data indicate that only those cells producing *neo* RNAs with either a viral 5′ or 3′ end are able to generate transducing virus. This would be expected if reverse transcriptase is to jump once between the two templates and thus incorporate the *neo*

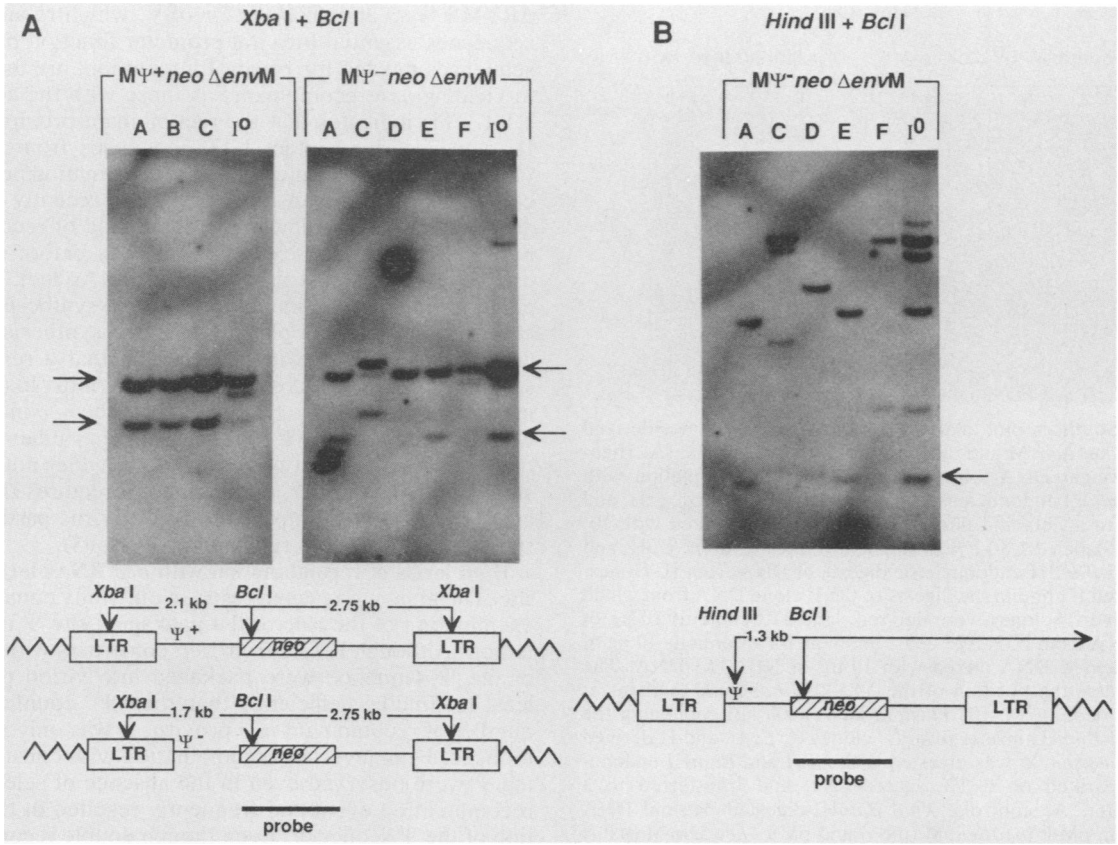

FIG. 8. Characterization of proviral DNA structure in II° clones derived from MΨ⁻*neo*Δ*env*M. Cellular DNA (10 μg) from II° clones, digested to completion with *Xba*I and *Bcl*I endonucleases (A) or *Hin*dIII and *Bcl*I endonucleases (B), was separated on 0.8% agarose gels and transferred to Zetabind filters. The filters were hybridized with 10^7 cpm of a ^{32}P-labeled 1.5-kb *neo* fragment. (A) The left panel contains the *Xba*I-*Bcl*I endonuclease digests of three II° clones (lanes A, B, and C) and a I° clone derived from MΨ⁺*neo*Δ*env*M. The right panel shows the digests of five II° clones (lanes A, C, D, E, and F) and a I° clone derived from MΨ⁻*neo*Δ*env*M. (B) DNA from five II° clones digested with *Hin*dIII and *Bcl*I endonucleases (lanes A, C, D, E, and F) and a I° clone derived from MΨ⁻*neo*Δ*env*M. The arrows beside each panel indicate the positions of the fragments expected from the Ψ⁺ and Ψ⁻ constructs. Bottom: Schematic structure of integrated MΨ⁺*neo*Δ*env*M and MΨ⁻*neo*Δ*env*M proviral DNA, showing the cleavage sites for *Xba*I, *Bcl*I, and *Hin*dIII endonucleases, the expected fragment sizes, and the probe used in the hybridization.

gene into a proviral DNA. Unexpectedly, the viral packaging sequence Ψ can be deleted from the *neo* RNA without drastically altering the frequency with which the *neo* sequence is incorporated into recombinant proviruses. Possible reasons for this finding are discussed later in this section.

Recombination between *neo* and helper virus RNA is not dependent on homology. *neo*-transducing viruses arise with an apparent frequency between 10^{-5} and 10^{-3} when only one recombination is needed between copackageable *neo* and helper virus RNAs. These transduction frequencies are minimal estimates because only those recombination events that include the complete *neo* sequence in the proviral DNA are scored. Any recombinations that occur within the *neo* gene, or in regions required for its expression, are not detected in our assay. It is possible that, due to reinfections, the number of G418^r clones overestimates the frequency of recombination. However, we believe this to be unlikely, since no new *neo* proviruses can be detected in randomly subcloned I° clone cells, and since the structures of recombinant *neo* viruses in independent II° clones are different from each other (see next section).

Our findings indicate that the recombinations leading to the incorporation of the *neo* gene into proviral structures can occur in regions where there is no apparent sequence homology between the two recombining partners. For example, there is no sequence homology between MΨ⁺*neo* or MΨ⁺*neo*SV and Mo-MuLV at the 3′ ends where the recombination must occur. Nevertheless, sequence homology at the *env* region of MΨ⁺*neo*Δ*env*SV increases the frequency of *neo* transduction about 100-fold over that found with MΨ⁺*neo*SV. Part of this effect seems to be due to the increased distance separating the *neo* and SV40 polyadenylation sequence because substituting an equal length of pBR322 sequence produces a 10-fold increase in *neo* transduction. This seems reasonable, as the recombination must occur between *neo* and the SV40 polyadenylation site to form a functional proviral *neo* sequence.

***neo*-transducing viruses arise by reverse transcriptase-mediated recombination.** Several lines of evidence in this report support the view that *neo*-transducing viruses arise by reverse transcriptase-mediated recombination of RNAs rather than by recombination at the DNA level. (i) The structure of the one, or few, nontandemly integrated copies of *neo* DNA in the I° clones we selected for study was the same as that of the transfected construct, thus excluding the possibility of recombination during or after transfection of the input DNA. Moreover, the *neo* mRNA produced by these I° clones was the same size as that expected from the

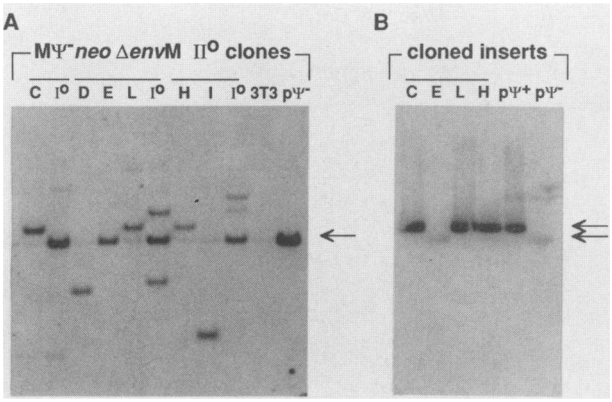

FIG. 9. Southern blot analysis of DNA from II° clones derived from pMΨ⁻*neo*Δ*env*M and cloned proviral DNA inserts. (A) High-molecular-weight DNA (10 μg) was digested to completion with *Xba*I and *Bam*HI endonucleases, separated on 1% agarose gels, and transferred to a Zetabind filter. The filter was hybridized with 10^7 cpm of a ^{32}P-labeled 1.5-kb *neo* fragment. Lanes C, D, E, L, H, and I show *Xba*I-*Bam*HI endonuclease digests of DNA from II° clones. Lanes marked I° contain the digests of the I° clone DNA from which the respective II° clones were derived, lane 3T3 contains 10 μg of digested DNA from NIH 3T3 cells, and lane pΨ⁻ contains 10 pg of pMΨ⁻*neo*Δ*env*M DNA mixed with 10 μg of NIH 3T3 DNA. The arrow indicates the position of the 2.85-kb *Xba*I-*Bam*HI fragment from pMΨ⁻*neo*Δ*env*M. (B) Plasmid DNA (100 ng) containing the cloned *Xba*I-*Bam*HI inserts from II° clones C, E, L, and H derived from MΨ⁻*neo*Δ*env*M was digested with *Xba*I and *Bam*HI endonucleases, separated on a 1% agarose gel, and transferred to a Zetabind filter. As controls, *Xba*I-*Bam*HI-digested plasmid DNA (100 ng) from pMΨ⁺*neo*Δ*env*M (pΨ⁺) and pMΨ⁻*neo*Δ*env*M (pΨ⁻) was run in parallel. The filter was hybridized with 5×10^6 cpm of a ^{32}P-labeled 220-bp *Ava*II fragment of pMΨ⁺*neo*, containing Ψ sequences of Mo-MuLV from nucleotides 312 to 532 (40). The arrows indicate the positions of the 3.2-kb *Xba*I-*Bam*HI fragment from pMΨ⁺*neo*Δ*env*M (upper) and the 2.85-kb *Xba*I-*Bam*HI fragment from pMΨ⁻*neo*Δ*env*M (lower). Compare Fig. 5 and 7 (bottom) for expected fragment sizes and probes used for Southern hybridization.

structure of the transfected *neo* DNA. Thus, with the exception of the MΨ⁺*neo*Δ*env*M construct, the integrated *neo* DNA in the I° clones does not exist in a proviruslike form. (ii) No evidence could be found for the possibility that the transducing *neo* virus resulted from rare recombinations in the I° clones between the integrated *neo* DNA and the superinfecting helper virus or from recombination with helper virus during reinfection of the I° clone cells. Thus, after random subcloning of I° clone cells 3 days after infection with helper virus, the structure of the integrated *neo* constructs remained unchanged; moreover, the clones did not contain any additional or new *neo* copies. Furthermore, the majority of these clones produced high titers of helper virus, as well as *neo*-transducing virus with frequencies similar to those observed for the corresponding I° clones before subcloning. (iii) The structures of the recombinant *neo* DNA in independent II° clones differ from one another and from the I° clone from which they originate, indicating multiple independent recombination events during virus replication rather than a rare preexisting recombinant present in the I° clone. Moreover, the *neo* sequence in the II° clones is flanked by viral LTRs. And, the fact that recombinant proviruses contain the *supF* marker in their flanking LTR shows that these derived from the infecting helper virus. (iv) RNAs transcribed from the constructs ΔRVMΨ⁺*neo* and ΔRVMΨ⁺*neo*SV, which retain the LTR sequences essential for viral promoter function but lack the sequences needed for proviral integration, are as proficient in yielding *neo* recombinants as those with the complete 5′ LTR. This indicates that the recombinant proviruses in the II° clones derived intact LTR sequences from the helper virus, rather than via illegitimate recombination between *neo* constructs during transfection, as has recently been suggested (13, 14). (v) Finally, the likelihood of recombination between two reverse-transcribed DNAs before integration can be excluded, because the *neo* RNAs lack sequences either for initiation of minus DNA strand synthesis at their 5′ end or for initiation of plus DNA strand synthesis at their 3′ end. These lines of evidence indicate that a reverse transcriptase-mediated recombination is the most likely mechanism to account for the transduction of the *neo* marker. This conclusion concurs with earlier genetic studies indicating that recombination between retroviruses does not occur in a single infection; rather, recombination requires the copackaging of two RNAs into heterozygous virus particles and a second round of virus replication (7, 41–43).

High levels of recombination with *neo* RNAs lacking the Ψ site. The most unexpected result in our study comes from the examination of the role of the packaging site Ψ in recombination. Although RNAs from *neo* constructs with a deletion of the Ψ sequence were packaged into virion particles at least 10^3-fold less efficiently than their Ψ⁺ counterparts, the number of recombinant *neo* proviruses was only severalfold reduced. High levels of recombination with constructs lacking Ψ were observed even in the absence of selection for a recombination event and frequently resulted in the acquisition of the Ψ sequence. Even though double recombinations appear to be extremely rare with RNAs containing the Ψ site, they do occur frequently in the Ψ⁻ counterparts. Perhaps, the high recombination frequency we observed with Ψ⁻ constructs explains the high frequency with which wild-type virus appears from retroviral vector-producing packaging cell lines which contain either ecotropic or amphotropic Ψ⁻ genomes (9, 26).

The packaging site Ψ spans a 350-bp region in the 5′ leader sequence of Mo-MuLV (24). There is considerable evidence indicating that the Ψ region contains *cis*-acting sequences necessary for efficient retroviral encapsidation (24, 40). In addition to Ψ, retroviral sequences extending into the *gag* coding region, as well as sequences from the U5 LTR, have been shown to be important for efficient packaging of genomic RNA (1, 2, 27). The *gag* region is present in all our constructs, and the U5 sequence is present in those constructs in which *neo* RNA is transcribed from a retroviral LTR. Recent experimental evidence suggests that the Ψ region also has a role in dimerization of genomic RNA (40; C. Roy and J. Darlix, personal communication; R. Alford and J. Belmont, personal communication). The dimer linkage structure functions to hold two genomic RNA strands together at or near their 5′ termini in mature virion particles (3, 28). Although both the encapsidation and the dimer linkage structure have been localized to the Ψ region, the precise nature of the sequences involved in each function is unknown. For example, it remains to be established whether the sequences responsible for packaging and dimer formation are distinctive, alike, or overlapping.

The finding that the lack of the Ψ site in one of the copackaged virion RNAs results in an increased frequency of recombination during reverse transcription could be explained if orderly retroviral reverse transcription requires proper packaging of genomic RNA and specific interactions

of the dimer linkage structure with nucleocapsid proteins. Disruption of the RNA dimer linkage structure (by deletion of Ψ) could cause an aberrant organization of the two RNA templates and might cause reverse transcriptase to switch templates inappropriately more frequently. Recent studies on the replication of spleen necrosis virus indicate that reverse transcription involves both genomic RNA strands in a highly orderly fashion (31). Additional information on the role of the Ψ sequence could be obtained by analyzing the effects of specific mutational alterations of the Ψ site on recombination.

Mechanism of transduction by retroviruses. The results presented in this study are consistent with the multistep model for the transduction of cellular oncogenes as originally proposed by Swanstrom and colleagues (37, 38). Our experiments support the proposed second step, namely, a recombination during reverse transcription of a copackaged retroviral genomic RNA with a hybrid RNA containing an LTR at either the 5′ or 3′ end. Thus, RNAs in which the *neo* gene is joined to 5′ retroviral sequences participate in recombination with an infecting helper virus during reverse transcription. Moreover, RNAs that are transcribed from a cell-like promoter and have retroviral 3′ ends also participate in recombination with a copackaged viral RNA. Because the absence of a packaging sequence does not entirely preclude the packaging of such cellular RNAs, retrovirus-mediated transduction of cellular oncogenes might also occur if the provirus resides downstream from a cellular oncogene.

ACKNOWLEDGMENTS

We are thankful to Rudolf Jaenisch for the cell line 3T3/supF. We thank Ned Landau, Ian Chambers, Mark Walberg, and David Schultz for comments on the manuscript.

H.S. is a fellow of the Cystic Fibrosis Foundation. This work was supported by Public Health Service grant GM-13235-23 from the National Institutes of Health to P.B.

LITERATURE CITED

1. **Armentano, D., S.-F. Yu, P. W. Kantoff, T. von Ruden, W. F. Anderson, and E. Gilboa.** 1987. Effect of internal viral sequences on the utility of retroviral vectors. J. Virol. **61**:1647–1650.
2. **Bender, M. A., T. D. Palmer, R. E. Gelinas, and A. D. Miller.** 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. J. Virol. **61**:1639–1646.
3. **Bender, W., Y.-H. Chien, S. Chattopadhyay, P. K. Vogt, M. B. Gardner, and N. Davidson.** 1978. High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. J. Virol. **25**:888–896.
4. **Bishop, J. M., and H. Varmus.** 1984. Functions and origins of retroviral transforming genes, p. 999–1108. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Chu, G., H. Hayakawa, and P. Berg.** 1987. Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res. **15**:1311–1326.
6. **Chumakov, I., H. Stuhlmann, K. Harbers, and R. Jaenisch.** 1982. Cloning of two genetically transmitted Moloney leukemia proviral genomes: correlation between biological activity of the cloned DNA and viral genome activation in the animal. J. Virol. **42**:1088–1098.
7. **Clavel, F., M. D. Hoggan, R. L. Willey, K. Strebel, M. A. Martin, and R. Repaske.** 1989. Genetic recombination of human immunodeficiency virus. J. Virol. **63**:1455–1459.
8. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. J. Gen. Virol. **42**:1–26.
9. **Danos, O., and R. C. Mulligan.** 1988. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. Proc. Natl. Acad. Sci. USA **85**:6460–6464.
10. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
11. **Fulton, R., D. Forrest, R. McFarlane, D. Onions, and J. C. Neil.** 1987. Retroviral transduction of T-cell antigen receptor β-chain and *myc* genes. Nature (London) **326**:190–194.
12. **Goldfarb, M. P., and R. A. Weinberg.** 1981. Generation of novel, biologically active Harvey sarcoma viruses via apparent illegitimate recombination. J. Virol. **38**:136–150.
13. **Goodrich, D. W., and P. H. Duesberg.** 1988. Retroviral transduction of oncogenic sequences involves viral DNA instead of RNA. Proc. Natl. Acad. Sci. USA **85**:3733–3737.
14. **Goodrich, D. W., and P. H. Duesberg.** 1990. Evidence that retroviral transduction is mediated by DNA, not by RNA. Proc. Natl. Acad. Sci. USA **87**:3604–3608.
15. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 109–136. *In* D. M. Glover (ed.), DNA cloning. A practical approach. IRL Press, Oxford.
16. **Herman, S. A., and J. M. Coffin.** 1987. Efficient packaging of readthrough RNA in ALV: implications for oncogene transduction. Science **236**:845–848.
17. **Huang, C.-C., C. Hammond, and J. M. Bishop.** 1985. Nucleotide sequence and topography of chicken c-*fps*. Genesis of a retroviral oncogene encoding a tyrosine-specific protein kinase. J. Mol. Biol. **181**:175–186.
18. **Huang, C.-C., N. Hay, and J. M. Bishop.** 1986. The role of RNA molecules in transduction of the proto-oncogene c-*fps*. Cell **44**:935–940.
19. **Junghans, R. P., L. R. Boone, and A. M. Skalka.** 1982. Retroviral DNA H structures: displacement-assimilation model of recombination. Cell **30**:53–62.
20. **Linial, M.** 1987. Creation of a processed pseudogene by retroviral infection. Cell **49**:93–102.
21. **Lobel, L., and S. P. Goff.** 1984. Construction of mutants of Moloney murine leukemia virus by suppressor-linker insertional mutagenesis: positions of viable insertion mutations. Proc. Natl. Acad. Sci. USA **81**:4149–4153.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Mann, R., and D. Baltimore.** 1985. Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. J. Virol. **54**:401–407.
24. **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.
25. **Miles, B. D., and H. L. Robinson.** 1985. High-frequency transduction of c-*erb*B in avian leukosis virus-induced erythroblastosis. J. Virol. **54**:295–303.
26. **Miller, A. D., and C. Buttimore.** 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. Mol. Cell. Biol. **6**:2895–2902.
27. **Murphy, J. E., and S. P. Goff.** 1989. Construction and analysis of deletion mutations in the U5 region of Moloney murine leukemia virus: effects on RNA packaging and reverse transcription. J. Virol. **63**:319–327.
28. **Murti, K. G., M. Bondurant, and A. Tereba.** 1981. Secondary structural features in the 70S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. J. Virol. **37**:411–419.
29. **Naharro, G., K. C. Robbins, and E. P. Reddy.** 1984. Gene product of v-*fgr onc*: hybrid protein containing a portion of actin and a tyrosine-specific protein kinase. Science **223**:63–66.
30. **Nilsen, T. W., P. A. Maroney, R. G. Goodwin, F. M. Rottman, L. B. Crittenden, M. A. Raines, and H.-J. Kung.** 1985. c-*erb*B activation in ALV-induced erythroblastosis: novel RNA processing and promoter insertion result in expression of an amino-truncated EGF receptor. Cell **41**:719–726.
31. **Panganiban, A. T., and D. Fiore.** 1988. Ordered interstrand and intrastrand DNA transfer during reverse transcription. Science **241**:1064–1069.

32. **Raines, M. A., N. J. Maihle, C. Moscovici, L. Crittenden, and H.-J. Kung.** 1988. Mechanism of c-*erbB* transduction: newly released transducing viruses retain poly(A) tracts of *erbB* transcripts and encode C-terminally intact *erbB* proteins. J. Virol. **62**:2437–2443.
33. **Reik, W., H. Weiher, and R. Jaenisch.** 1985. Replication-competent Moloney murine leukemia virus carrying a bacterial suppressor tRNA gene: selective cloning of proviral and flanking host sequences. Proc. Natl. Acad. Sci. USA **82**:1141–1145.
34. **Rosson, D., D. Dugan, and E. P. Reddy.** 1987. Aberrant splicing events that are induced by proviral integration: implications for *myb* oncogene activation. Proc. Natl. Acad. Sci. USA **84**:3171–3175.
35. **Rowe, W. P., W. E. Pugh, and J. W. Hartley.** 1970. Plaque assay techniques for murine leukemia viruses. Virology **42**:1136–1139.
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. **Swanstrom, R., R. C. Parker, H. E. Varmus, and J. M. Bishop.** 1983. Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. Proc. Natl. Acad. Sci. USA **80**:2519–2523.
38. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216**:812–820.
39. **Vogelstein, B., and D. Gillespie.** 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA **76**:615–619.
40. **Weiss, R., N. Teich, H. Varmus, and J. Coffin.** 1985. RNA tumor viruses: appendixes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Weiss, R. A., W. S. Mason, and P. K. Vogt.** 1973. Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumor viruses. Virology **52**:535–552.
42. **Wyke, J. A., and J. A. Beamand.** 1979. Genetic recombination in Rous sarcoma virus: the genesis of recombinants and lack of evidence for linkage between *pol*, *env* and *src* genes in three factor crosses. J. Gen. Virol. **43**:349–364.
43. **Wyke, J. A., J. G. Bell, and J. A. Beamand.** 1975. Genetic recombination among temperature-sensitive mutants of Rous sarcoma virus. Cold Spring Harbor Symp. Quant. Biol. **39**:897–905.