View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

VIROLOGY

Virology 312 (2003) 281–294

www.elsevier.com/locate/yviro

Frequent dual initiation of reverse transcription in murine leukemia virus-based vectors containing two primer-binding sites

Yegor A. Voronin[a,b] and Vinay K. Pathak[a,*]

[a] *HIV Drug Resistance Program, National Cancer Institute-Frederick, Frederick, MD 21702, USA*
[b] *Department of Biochemistry, West Virginia University, Morgantown, WV 25606, USA*

Received 11 December 2002; returned to author for revision 31 January 2003; accepted 28 March 2003

Abstract

Retroviruses package two copies of viral RNA into each virion. Although each RNA contains a primer-binding site for initiation of DNA synthesis, it is unknown whether reverse transcription is initiated on both RNAs. To determine whether a single virion is capable of initiating reverse transcription more than once, we constructed a murine leukemia virus-based vector containing a second primer-binding site (PBS) derived from spleen necrosis virus and inserted the green fluorescent protein gene (*GFP*) between the two PBSs. Initiation of reverse transcription at either PBS results in a provirus that expresses *GFP*. However, initiation at both PBSs can result in the deletion of *GFP*, which can be detected by flow cytometry and Southern blotting analysis. Approximately 22–29% of the proviruses formed deleted the *GFP* in a single replication cycle, indicating the minimum proportion of virions that initiated reverse transcription on both PBSs. These results show that a significant proportion of MLV-based vectors containing two PBSs have the capacity to initiate reverse transcription more than once.
© 2003 Elsevier Science (USA). All rights reserved.

Keywords: MLV; SNV; PBS; Initiation; tRNA; Reverse transcription

Introduction

Retroviral particles contain two copies of viral genomic RNA. The RNA is converted into double-stranded genomic DNA by the virally encoded reverse transcriptase (RT) shortly after the virus enters the host cell (Baltimore, 1970; Temin and Mizutani, 1970). All retroviruses use a cellular tRNA as a primer for initiation of reverse transcription and each virus uses a particular tRNA, which is specifically packaged into virion (Harada et al., 1975; Leis et al., 1993; Mak and Kleiman, 1997; Peters et al., 1977; Ratner et al., 1985). The tRNA used to initiate DNA synthesis is partially melted and its 3′ end is annealed by virtue of its complementarity to 18 nucleotides of a viral sequence called the primer-binding site (PBS) shortly after virion formation (Cen et al., 2000; Fu et al., 1997; Huang et al., 1998).

The components necessary for reverse transcription appear to be available in viral particles in excess of the amounts necessary to initiate and complete DNA synthesis. Each virion contains approximately 75–150 molecules of RT (Bauer and Temin, 1980; Levin et al., 1993; Panet and Kra-Oz, 1978; Vogt and Simon, 1999) and several specific primer tRNAs (Huang et al., 1994; Levin and Seidman, 1979; Mak et al., 1994; Peters et al., 1977). Murine leukemia viruses (MLVs) use $tRNA^{Pro}$ to initiate DNA synthesis; analysis of MLV particles indicates that each viral particle contains approximately six to eight $tRNA^{Pro}$ molecules (Levin and Seidman, 1979; Peters et al., 1977). Thus, it should be possible for a virus to initiate reverse transcription more than once. However, it is currently unknown whether dual initiation of reverse transcription can occur in a single virion. If a single virion cannot initiate reverse transcription more than once, then it can only form one genomic DNA molecule. However, if a single virion has the capacity to

* Corresponding author. HIV Drug Resistance Program, National Cancer Institute-Frederick, Building 535, Room 334, Frederick, MD 21702. Fax: +1-301-846-6013.
E-mail address: VPATHAK@ncifcrf.gov (V.K. Pathak).

0042-6822/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0042-6822(03)00300-3

initiate reverse transcription twice, then it could potentially form two genomic DNAs.

It is generally thought that only one provirus is formed as a result of one infection event. In studies of recombinant viruses, it was shown that only one provirus is present in most infected cells (Hu and Temin, 1990a). However, the fact that only one provirus is found in cells infected with a single virus cannot serve as a measure of the initiation capacity of a single virion. Even if two initiations occur during reverse transcription, inefficient minus-strand DNA transfer and intermolecular template switching events during minus-strand DNA synthesis could prevent the formation of two genomic DNAs (Anderson et al., 1998; Hu and Temin, 1990b). In addition, degradation of viral DNA, formation of one and two long terminal repeat (LTR) circles (Ju and Skalka, 1980; Shank and Varmus, 1978; Swanstrom et al., 1981), and autointegration (Lee and Craigie, 1994; Shoemaker et al., 1981) could reduce the efficiency of provirus formation (Butler et al., 2001, 2002; Follenzi et al., 2000; Sirven et al., 2000).

To date, it is not known whether one viral particle is capable of forming multiple functional initiation complexes. To determine whether a single virion has the capacity to initiate reverse transcription more than once, we introduced a second site of initiation in the viral genome and analyzed reverse transcription products and structures of proviruses that were generated. The results of these studies indicate that at least 22–29% of MLV vectors containing two initiation sites are able to initiate reverse transcription more than once.

Results

Construction of vectors

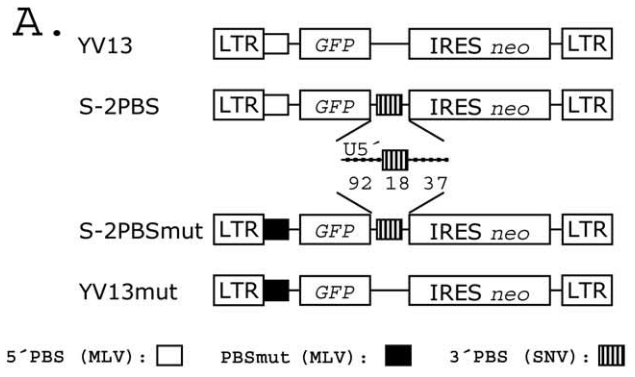
To determine whether two initiation events can occur in the same virion during reverse transcription, we constructed a series of retroviral vectors (Fig. 1). Vector YV13 contains all of the *cis*-acting elements required for vector propagation, including 5′ and 3′ LTRs, PBS, packaging signal (ψ), and polypurine tract. The PBS that is located at its natural position, just downstream of the 5′ LTR, is referred to as the 5′ PBS throughout the text. YV13 also expresses green fluorescent protein gene (*GFP*) and the neomycin phosphotransferase gene (*neo*) (Jorgensen et al., 1979) from a single bicistronic transcript expressed from the 5′ LTR. Translation of *neo* is facilitated by the presence of an internal ribosomal entry site (IRES) from encephalomyocarditis virus (Jang et al., 1988, 1989).

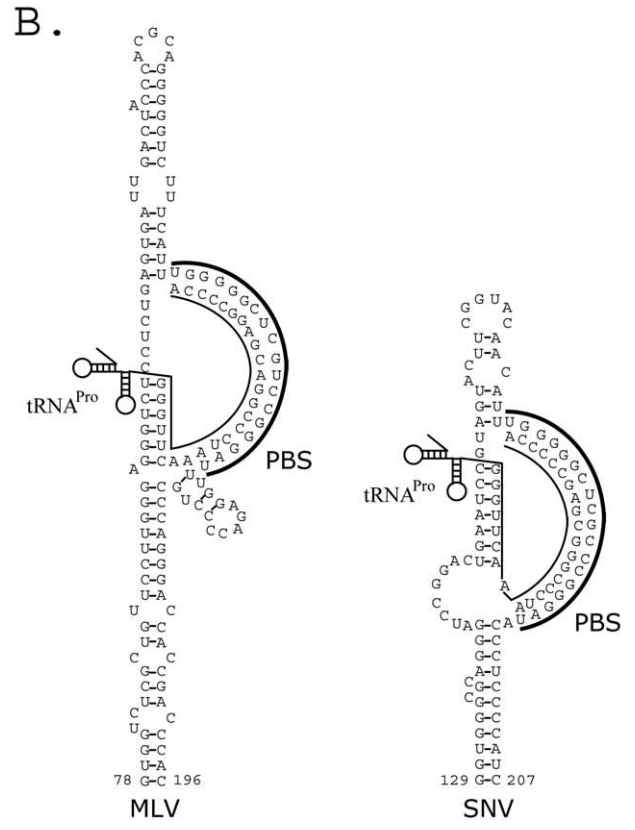
Next, we constructed vector S-2PBS by insertion of a 147-bp sequence derived from spleen necrosis virus (SNV) between *GFP* and IRES (Fig. 1). A predicted structure of the MLV and SNV PBS, sequences surrounding the PBS, and their interactions with the tRNA primer are shown in Fig. 1B. SNV, similar to MLV, uses $tRNA^{Pro}$ as a primer and the 18 nucleotides of the SNV PBS are identical to those of the MLV (Fig. 1C). However, the sequences surrounding the PBS in MLV and SNV have very little homology to each other. The inserted sequence contains the complete SNV U5 (92 nt), including the SNV attachment site (*att* site), the PBS (18 nt), and a portion of the 5′ untranslated leader sequences (37 nt). The SNV-derived PBS region was used to create the 3′ PBS because insertion of the MLV PBS resulted in the formation of directly repeated sequences in the vector that underwent deletion at a high frequency through RT switching templates during the process of reverse transcription (data not shown). Because the inserted SNV sequence has only 49% homology to the MLV sequence, and the longest stretch of sequence identity is 22 nucleotides, direct-repeat deletions are not expected to occur at a high frequency (Hwang et al., 2001; Svarovskaia et al., 2000).

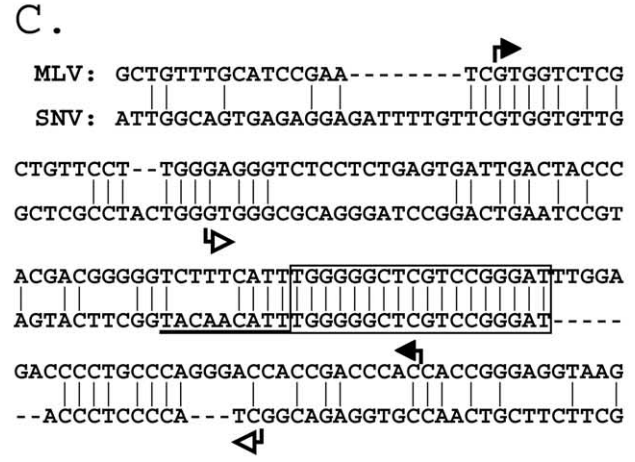
To determine whether the 3′ PBS (derived from SNV) was functional, we constructed vector S-2PBSmut (Fig. 1A). This vector is identical to S-2PBS, except that the 5′ PBS was mutated by substitution of the 3-bp sequence TGG at the 5′ end of the PBS with ACC. This 3-bp substitution is expected to prevent the basepairing of the 3′-end CCA nucleotides of the primer tRNA to the PBS and thereby interfere with initiation of reverse transcription. As a result, the replication of S-2PBSmut should be dependent on initiation of reverse transcription at the 3′ PBS. Finally, vector YV13mut was generated to confirm that the 3-bp substitution introduced in vector S-2PBSmut severely reduced initiation of reverse transcription from the 5′ PBS.

Experimental design

The structures of proviruses generated after one round of replication with vectors containing two functional PBS regions can be analyzed to determine whether initiation of reverse transcription occurred once or more than once during infection. Each viral RNA containing two PBS regions can initiate reverse transcription at either the 5′ PBS (Fig. 2A), the 3′ PBS (Fig. 2B), or at both PBSs (Fig. 2C). Even though each virion contains two copies of viral genomic RNA, only one RNA is shown for the sake of simplicity (the possible outcomes of initiating reverse transcription on both RNAs are considered under Discussion). The first possibility is that initiation will occur only at the 5′ PBS (Fig. 2A). Reverse transcription will proceed in a normal fashion and the resulting provirus will have a structure that is identical to the original vector. We refer to this provirus structure as 1GFP. The second possibility is that initiation will occur only at the 3′ PBS (Fig. 2B). During reverse transcription, an unusually long minus-strand strong-stop DNA containing R, U5, PBS, ψ, and *GFP* will form. This minus-strand strong-stop DNA will be transferred to the 3′ end of viral RNA. Plus-strand DNA synthesis will start at PPT and continue through LTR, PBS, ψ, *GFP*, and stop after copying tRNA sequences complementary to PBS. Since the se-







quence of the PBS is the same for SNV and MLV, the plus-strand strong-stop DNA will be transferred to the 5′ end of minus-strand DNA. The resulting provirus will contain a second *GFP* downstream of the 3′ LTR and is referred to as the 2GFP structure. After completion of reverse transcription, the DNA will have heterologous attachment sites; the 5′-*att* site will be derived from U3 of MLV and 3′-*att* site will be derived from U5 of SNV. Even though the four bases at the 3′ ends of both attachment sites are the same (Fig. 1C), this could potentially reduce the efficiency of integration.

Finally, the third possibility is that reverse transcription will be initiated at both the 5′ PBS and the 3′ PBS (Fig. 2C). Initiation of reverse transcription at the 5′ PBS will result in the formation of a normal minus-strand strong-stop DNA, which can be transferred to the 3′ end of the viral RNA. Initiation of reverse transcription at the 3′ PBS will result in the formation of a minus-stand DNA coding for the *GFP*, *ψ*, and PBS; because initiation of reverse transcription at the 5′ PBS will result in degradation of the R and U5 regions by the RNase H activity of RT, DNA synthesis initiating at the 3′ PBS will not be extended past the 5′ PBS (Artzi et al., 1996). This reverse transcription product will lack any homology to the R region and therefore will not be used for minus-strand DNA transfer; consequently, it will not be involved in subsequent steps during reverse transcription. DNA synthesis that is initiated at the 5′ PBS will be used for minus-strand DNA transfer and will be extended to the 3′ PBS but will not be extended further because the *GFP*, *ψ*, and 5′ PBS will have been degraded during DNA synthesis initiating from the 3′ PBS. Plus-strand DNA synthesis will

Fig. 1. Vector design. (A) Structures of MLV-based vectors. All vectors contain MLV LTRs, other *cis*-acting elements needed for vector propagation, *GFP* that is expressed from the MLV LTR promoter, and a *neo* that is expressed from IRES. Vector YV13 has only a 5′ PBS derived from MLV (open box between 5′LTR and *GFP*). Vector S-2PBS has an additional 3′PBS derived from SNV (striped box between *GFP* and IRES *neo*). The 3′PBS fragment consists of 92 nt from the SNV U5 (dotted line labeled U5′), 18 nt of SNV PBS (striped box), and 37 nt of SNV 5′ untranslated leader (dotted line). Vector S-2PBSmut is identical to S-2PBS, except that the 5′ PBS has been inactivated by mutation (black box between 5′ LTR and *GFP*). Vector YV13mut possesses only the 5′ PBS that has been inactivated by mutation (black box between 5′ LTR and *GFP*). (B) Comparison of predicted secondary structures of the 5′PBS region derived from MLV and 3′PBS region derived from SNV in complex with $tRNA^{Pro}$ (modified from Leis et al., 1993). The $tRNA^{Pro}$ sequences are shown in part as a schematic of a double hairpin structure (thin line). The 18 nt of 5′ and 3′ PBS sequences are indicated by a thick line. The numbers at the base of each predicted structure refer to nucleotide positions in the MLV and SNV genomes, starting at the beginning of the 5′R regions. (C) Homology comparison of MLV (top line) and SNV (bottom line) nucleotide sequences in the vicinity of the PBS regions. The SNV sequence shown is of the fragment used to generate S-2PBS and S-2PBSmut vectors and consists of the U5 region, PBS, and 37 nt of 5′ untranslated leader. The homologous MLV sequence shown indicates that 49% of the nucleotides are identical (vertical lines). The boxed sequences indicate 18 nt of PBS and the underlined sequence represents the *att* site. The black and white arrowheads indicate the sequences that form the predicted secondary structures of MLV and SNV, respectively, as shown in B.

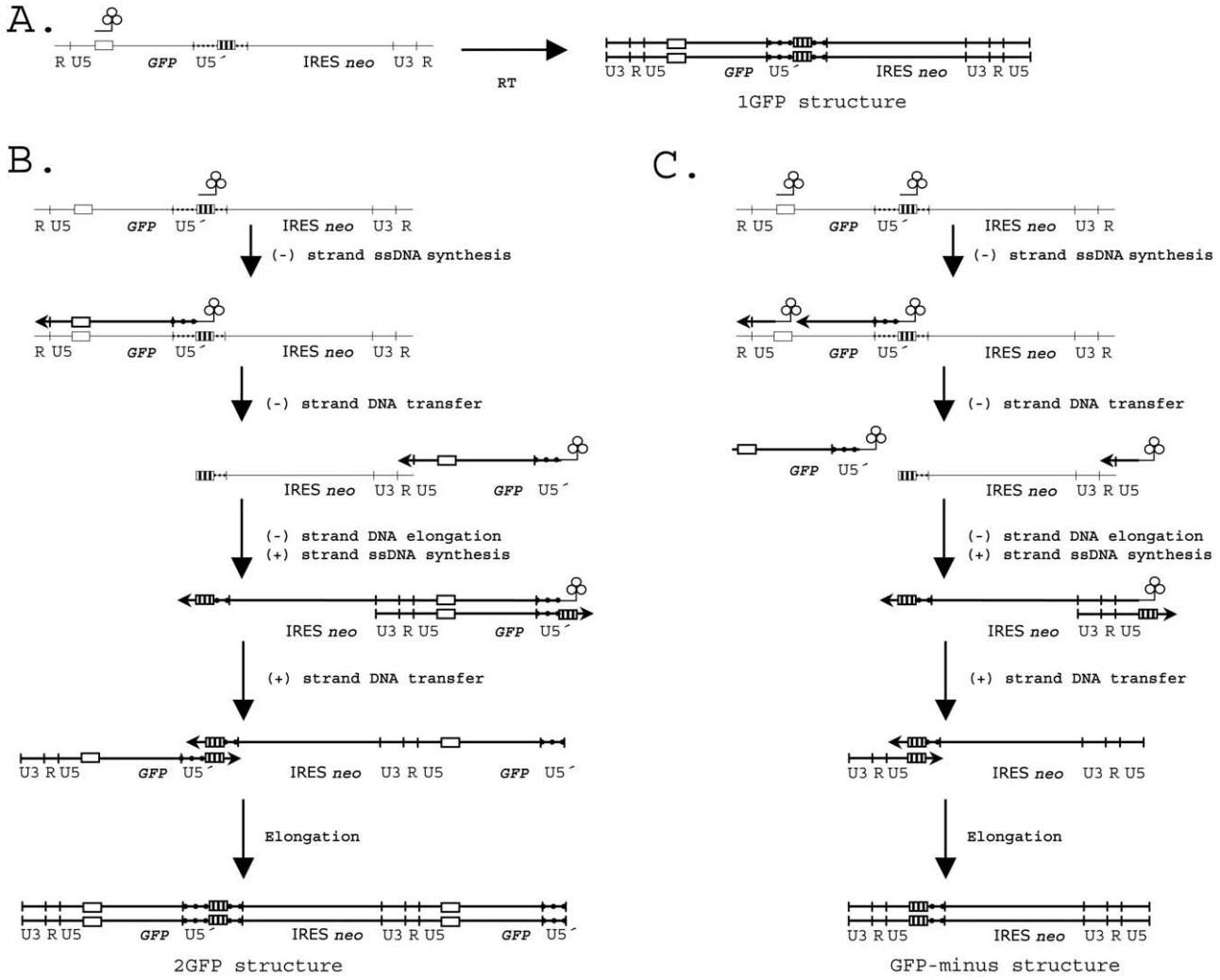


Fig. 2. Reverse transcription of vectors containing two PBSs. Thin lines indicate genomic RNA; thick lines indicate newly synthesized DNA; dotted lines indicate SNV-derived sequences; cloverleaf symbol represents tRNA and site(s) of DNA synthesis initiation; open boxes represent 5′ PBS; and striped boxes represent 3′ PBS. Arrows show the direction of DNA synthesis. Unique 5′ region (U5), unique 3′ region (U3), and repeat region (R) form the MLV LTR. Unique 5′ region of the SNV LTR is indicated as U5′. (A) Initiation of reverse transcription at the 5′ PBS, followed by completion of normal reverse transcription (Gilboa et al., 1979; Temin, 1981), results in the formation of a 1GFP proviral structure. (B) Initiation of reverse transcription at the 3′ PBS results in the formation of a 2GFP proviral structure, with *GFP* being duplicated at the 3′ end (see Results for detailed description). (C) Initiation of reverse transcription at both 5′ and 3′ PBS regions results in the formation of a GFP-minus proviral structure (see Results for detailed description).

be initiated at the PPT, extended through the first 18 nt of the primer tRNA, and transferred to the 3′ PBS. Consequently, the provirus that is formed will not contain *GFP* and ψ sequences and is referred to as the GFP-minus structure.

To summarize, initiation of reverse transcription at either the 5′ or the 3′ PBS will result in the formation of provirus structures that can express *GFP* (the 1GFP or 2GFP structures, respectively), but initiation of reverse transcription at both the 5′ PBS and the 3′ PBS will result in the formation of a provirus structure that cannot express *GFP* (the GFP-minus structure). Because the expression of *GFP* can be easily detected by flow cytometry and the structures of the proviruses can be determined by Southern blotting analysis, this system can be used to detect events in which reverse transcription initiated more than once in a single infectious viral particle. It is important to note that reverse transcription could initiate at both the 5′ and the 3′ PBS but not result in deletion of *GFP;* for example, if minus-strand DNA transfer occurs intermolecularly, *GFP* will not be deleted. Thus, the frequency of proviruses that have deleted GFP represents a minimum proportion of the virion that initiated reverse transcription more than once.

Replication of vectors containing two PBSs

We determined the ability of vectors YV13, YV13mut, S-2PBS, and S-2PBSmut to be propagated in MLV and SNV helper cells (Table 1). The vectors were transfected into the helper cell lines and pools of transfected cells were

Table 1
Titers of viruses produced by MLV and SNV packaging cell lines

Vector	Helper cells[a]	No. of expts.	Titer (CFU/ml $\times 10^4$) (mean ± SE[b])	Relative titer[c]
YV13	PG13	4	16 ± 10	100
YV13mut	PG13	5	0.007 ± 0.003	0.04
S-2PBS	PG13	4	3.65 ± 0.54	23
S-2PBSmut	PG13	4	0.25 ± 0.03	1.6
YV13	DSH134G	5	31.2 ± 7.4	100
YV13mut	DSH134G	4	0.004 ± 0.002	0.013
S-2PBS	DSH134G	4	30.5 ± 8.1	100
S-2PBSmut	DSH134G	4	14.1 ± 5.1	45

[a] PG13 is an MLV-based helper cell line; DSH134G is an SNV-based helper cell line.
[b] SE is standard error.
[c] For each helper cell line the titer of YV13 was set as 100%.

selected and expanded. The titers of virus produced from four to five independent transfected pools were determined by infection of D17 target cells and selection of G418-resistant colonies. The MLV helper cell line, PG13, produced an average YV13 titer of 16×10^4 CFU/ml; the SNV helper cell line, DSH134G, produced a similar average YV13 titer of 31×10^4 CFU/ml. In contrast, both PG13 and DSH134G helper cell lines produced average YV13mut titers that were at least 2500-fold lower than the titers generated by the control YV13 vector, indicating that the 3-nucleotide substitution in the PBS severely reduced the ability of the YV13mut vector to initiate DNA synthesis and complete viral replication. Both PG13 and DSH134G helper cell lines produced average S-2PBS titers that were similar to the YV13 control vector, indicating that the presence of both 5′ and 3′ PBSs did not have a deleterious effect on viral replication. The DSH134G helper cell line produced average S-2PBSmut titers that were only twofold lower than the titers of the S-2PBS vector. Since S-2PBSmut lacked a functional 5′ PBS and was dependent on the 3′ PBS for initiation of DNA synthesis, the result indicated that the 3′ PBS was used efficiently for initiation of DNA synthesis in the SNV helper cell line; the result also indicated that the SNV proteins could efficiently use the heterologous MLV *att* sites for integration (see Fig. 2B). In contrast, the PG13 helper cell line produced average S-2PBSmut titers that were 14-fold lower than the S-2PBS vector and 62-fold lower than the YV13 vector. The result suggested that the MLV proteins did not efficiently use the SNV PBS and/or the SNV *att* site.

GFP expression in vectors containing two PBSs

As described earlier (Fig. 2C), the initiation of DNA synthesis at both the 5′ and the 3′ PBS should result in the deletion of the *GFP* coding sequences and loss of *GFP* expression. To determine whether the presence of a functional 3′ PBS affected *GFP* expression, we performed flow cytometry analysis of pools of cells infected with YV13, S-2PBS, and S-2PBSmut virus derived from the DSH134G helper cell line (Fig. 3). At least three independent experiments were performed. As expected, 99.9% of the uninfected D17 cells were negative for *GFP* expression. Analysis of cells infected with YV13 virus indicated that 4.9 ± 0.5% of the infected cells were GFP-negative. This suggested that approximately 5% of the GFP genes were inactivated by mutations that occurred during transfection of the DSH134G helper cells or during reverse transcription. Similarly, flow cytometry analysis of cells infected with S-2PBSmut virus indicated that the frequency of GFP-negative cells was 6.7 ± 0.8%. In contrast to the cells infected with YV13 or S-2PBSmut, 31.3 ± 2% of cells infected with S-2PBS virus were negative for *GFP* expression. Since the S-2PBS virus is expected to have two functional PBSs and the control vectors YV13 and S-2PBSmut are expected to have only one functional PBS, the higher proportion of GFP-negative cells was correlated with the presence of two functional PBSs. The result suggested that in 26% of the cells infected with S-2PBS virus derived from the DSH134G helper cell line, the GFP gene was deleted through initiation of reverse transcription at both the 5′ and the 3′ PBSs (31% of GFP inactivation for S-2PBS minus 5% background of GFP inactivation for YV13).

The proportion of GFP-negative cells was also determined for cells infected with YV13, S-2PBS, and S-2PBSmut viruses produced from the PG13 helper cell line (data not shown). The results obtained from three independent experiments were similar to the results obtained for cells infected with virus produced from the DSH134G cells; 3.3 ± 0.6% of cells infected with YV13 virus were GFP-

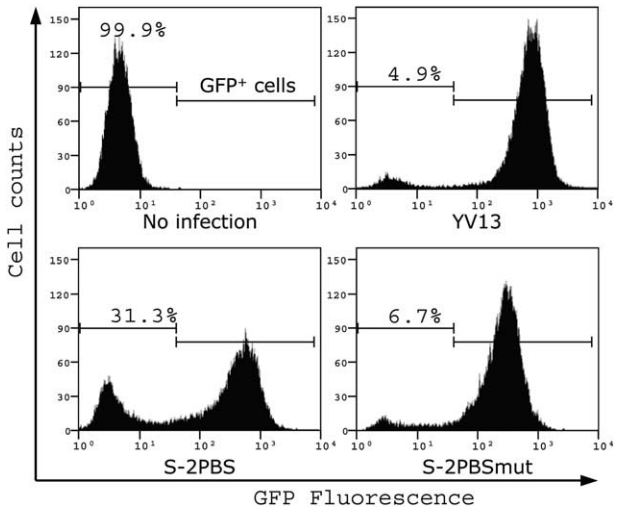


Fig. 3. Flow cytometry analysis. Representative flow cytometric analysis of D17 cells infected with YV13, S-2PBS, and S-2PBSmut virus produced from SNV-based helper cell line DSH134G. Infected cells were selected for resistance to G418, and pools of resistant cells were analyzed for expression of *GFP*. The average proportions of GFP-negative cells determined in three to four independent experiments are shown. At least 5000 events were analyzed in each experiment.

negative, 5.1 ± 1.3% of cells infected with S-2PBSmut virus were GFP-negative, and 32 ± 7% of cells infected with S-2PBS virus were GFP-negative. The result suggested that in cells infected with S-2PBS virus derived from PG13 cells, GFP was deleted at a similar frequency of 29% through initiation of reverse transcription at both the wild type and the internal PBS (32% GFP-negative cells for S-2PBS minus 3.3% GFP-negative cells for YV13; data not shown).

Quantitative PCR analysis of reverse transcription initiation in vectors containing two PBSs

To determine the kinetics and relative efficiencies of reverse transcription initiation at the wild-type and the internal PBS, we analyzed the products of reverse transcription 1, 2, 3, and 24 h after infection of target cells with virus derived from DSH134G cells. The real-time PCR primer and probe sets are shown in Fig. 4A. The primer sets R-U5 and GFP were designed to detect early reverse transcription products that initiated at the wild-type and internal PBS, respectively. The Neo primer set was designed to detect reverse transcription products after minus-strand DNA transfer and the U5-ψ primer set was designed to detect late reverse transcription products after plus-strand DNA transfer. Two independent infection experiments were performed and the copy number of each target sequence was determined twice for approximately 10^5 infected cells from each experiment. The threshold cycle values and the estimated copy number of each DNA based on a standard curve generated from a control plasmid are shown for a representative experiment in Table 2. The amounts of cellular DNAs analyzed were normalized by determining the copy number of a hygromycin phosphotransferase gene (*hygro*) that was previously introduced into the target cells by transfection (see Materials and methods).

To determine the extent of reverse transcription initiation that occurred 1, 2, and 3 h after infection, we compared the amounts of reverse transcription products detected by the R-U5 primer set at these time points with the amount of product detected 24 h after infection. The amounts of reverse transcription products detected by the R-U5 primer set 1, 2, and 3 h after infection with YV13 virus derived from DSH134G helper cell line were 2–5, 25–35, and 50–90% of the products detected 24 h after infection, respectively (data not shown). This result indicated that the majority of reverse transcription initiation events occurred within the first 3 h after infection.

The relative amounts of reverse transcription products after infection with YV13 and S-2PBS vectors are shown in Fig. 4B, and the average amount of R-U5 products was set to 100%. One hour after infection, the primary reverse transcription product was detected in cells infected with YV13 was R-U5; relative to the R-U5 primer set, the reverse transcription products detected by the Neo primer set were less than 20%, indicating that most of the DNA synthesis initiating from the 5′ PBS did not undergo minus-strand DNA transfer. Similarly, reverse transcription products detected by the GFP and U5-ψ primer sets were less than 10% of the products detected by the R-U5 primer set, indicating that even fewer reverse transcription products had progressed beyond minus-strand DNA transfer. In sharp contrast, the profile of reverse transcription products detected 1 h after infection with S-2PBS differed from that of YV13. While the reverse transcription products detected by the Neo and U5-ψ primers remained below 20% of the R-U5 primers, the products detected by the GFP primers were 120% relative to the R-U5 primer set, indicating that these products were present at similar levels. The similar amounts of reverse transcription products detected by the R-U5 and GFP primers early after infection with S-2PBS

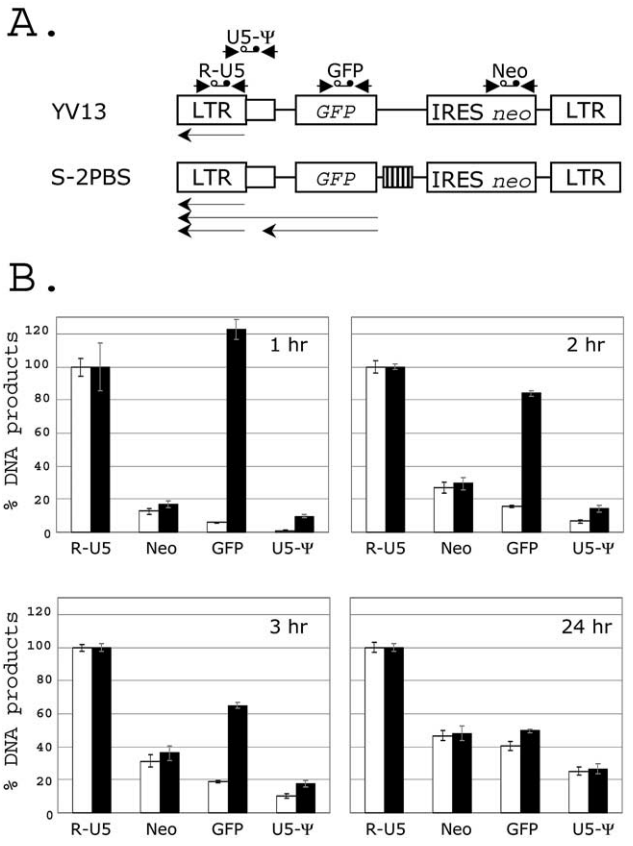


Fig. 4. Quantitative real-time PCR analysis. (A) Locations of real-time PCR primer-probe sets are shown above the structures of YV13 and S-2PBS vectors. The early products of reverse transcription that could form before minus-strand DNA transfer are represented as arrows below the structures of YV13 and S-2PBS vectors. (B) Comparison of relative amounts of reverse transcription products determined 1, 2, 3, and 24 h after infection. Target D17 cells were infected with virus obtained from SNV-based helper cell-line DSH134G. The amounts of DNA products determined after YV13 infection (white bars) and S-2PBS infection (black bars) are shown. Two independent experiments were analyzed and each DNA sample was analyzed by real-time PCR four times. The average copy numbers of various DNA products determined were expressed as a percentage of the DNA products determined by the R-U5 primer-probe set. The error bars represent the standard error of the mean.

Table 2
Real-time PCR-based quantitation of viral DNA obtained 1, 2, 3, and 24 h after infection[a]

Sample	R-U5		U5-ψ		GFP		Neo		Hygro	
	Ct[b]	Copies[c]	Ct	Copies	Ct	Copies	Ct	Copies	Ct	Copies
Control DNA Copies[d]										
1,000,000	16.95	—	18.86	—	17.08	—	21.03	—	18.76	—
333,333	18.07	—	20.46	—	18.45	—	22.31	—	20.29	—
111,111	20.13	—	22.11	—	19.61	—	24.11	—	21.97	—
37,037	21.67	—	24.02	—	21.37	—	25.41	—	23.56	—
12,346	23.27	—	24.91	—	23.24	—	27.22	—	25.22	—
4115	25.00	—	27.22	—	24.68	—	28.63	—	26.78	—
1372	26.24	—	29.26	—	26.54	—	30.29	—	28.47	—
457	27.54	—	30.84	—	27.76	—	32.36	—	29.93	—
152	30.03	—	32.67	—	29.48	—	34.15	—	31.65	—
51	31.35	—	34.16	—	31.08	—	35.05	—	33.29	—
YV13-1 h	24.03	7082	33.03	108	28.08	394	30.62	1207	20.93	220702
YV13-2 h	22.93	15,036	29.33	1142	25.79	1940	29.01	3647	22.41	80864
YV13-3 h	21.84	31,717	27.28	4210	24.18	5970	27.14	13158	22.47	77748
YV13-24 h	22.34	22,521	26.42	7258	23.81	7709	27.25	12173	23.40	41530
S-2PBS-1 h	26.99	937	32.69	134	26.29	1366	32.72	286	21.42	157,594
S-2PBS-2 h	24.09	6787	29.09	1327	24.45	4934	29.68	2309	22.34	84,948
S-2PBS-3 h	22.50	20,123	26.93	5243	23.04	13215	27.37	11277	22.58	72,017
S-2PBS-24 h	23.58	9639	27.42	3836	24.46	4919	28.26	6088	22.84	60,488
S-2PBSmut-1 h	26.11	1716	28.38	2092	24.43	5010	32.20	407	20.79	241,386
S-2PBSmut-2 h	24.89	3928	27.22	4366	23.97	6913	29.86	2027	22.44	79,491
S-2PBSmut-3 h	23.35	11,281	25.45	13,444	22.77	15,916	27.71	8902	22.99	54,755
S-2PBSmut-24 h	23.16	12,880	25.27	15,163	23.05	13,133	27.44	10,686	22.53	74,563

[a] Representative results from one of four experiments performed on two independent infections are shown; R-U5, U5-ψ, GFP, Neo, and Hygro are primer-probe sets as shown in Fig. 4.
[b] Ct is the cycle at which the amplification curve reaches a defined threshold.
[c] Number of DNA copies determined on the basis of standard curve generated using various quantities of control plasmid.
[d] Dilutions of control DNA plasmid used to generate a standard curve; the numbers represent an estimate of the number of plasmid molecules in each reaction.

indicated that DNA synthesis was initiating at the 5′ and 3′ PBS sites with similar efficiencies.

Analyses of the reverse transcription products 2 and 3 h after infection with YV13 and S-2PBS indicated that the products detected by the GFP primer set were consistently higher for the S-2PBS virus than for the YV13 virus. This result confirmed the view that DNA synthesis was being initiated at both the 5′ and the 3′ PBS sites for the S-2PBS vector.

Analyses of the reverse transcription products 24 h after infection with YV13 and S-2PBS indicated that the products detected by the Neo, GFP, and U5-ψ primer sets were between 30 and 50% relative to the R-U5 primer set. This result was expected since reverse transcription leads to duplication of the LTRs, which include the R-U5 regions. For each completely synthesized YV13 viral DNA, there are two copies that are detected by the R-U5 primer set, and one copy that is detected by the Neo, GFP, and U5-ψ primer sets.

Analysis of proviral structures generated after infection with vectors containing two PBSs

As described in Fig. 2, reverse transcription of the vectors containing a 5′ and 3′ PBS could result in the formation of the 1GFP, 2GFP, or GFP-minus structures. The structures of proviruses generated after infection with YV13, S-2PBS, and S-2PBSmut vectors are outlined in Fig. 5A. When genomic DNA of YV13-infected cells is digested with *Xba*I and hybridized to a *neo* probe, a 4.1-kb band is generated. When genomic DNA of S-2PBS-infected cells is digested with *Xba*I and hybridized to the *neo* probe, the GFP-positive proviral structures (1GFP and 2GFP) generate a 4.3-kb band (Fig. 5A); in contrast, the GFP-minus proviral structure generates a smaller 2.0-kb band. We first analyzed genomic DNAs isolated from clones of D17 cells infected with the S-2PBS virus (Fig. 5B). A total of 27 single cell clones was isolated; visualization under a microscope indicated that 16 clones were fluorescent (GFP-positive) and 11 were nonfluorescent (GFP-negative). Genomic DNAs from each of the cell clones were digested with *Xba*I and analyzed by Southern blotting; the results obtained from eight GFP-positive and eight GFP-negative cell clones are shown in Fig. 5B. As expected, all GFP-positive cell clones generated a 4.3-kb band, whereas all GFP-negative cell clones generated a smaller 2.0-kb band. The presence of the 2.0-kb band in the GFP-negative cell clones provided strong evidence that reverse transcription was initiated at both the 5′ and the 3′ PBS in the same virion and resulted in the

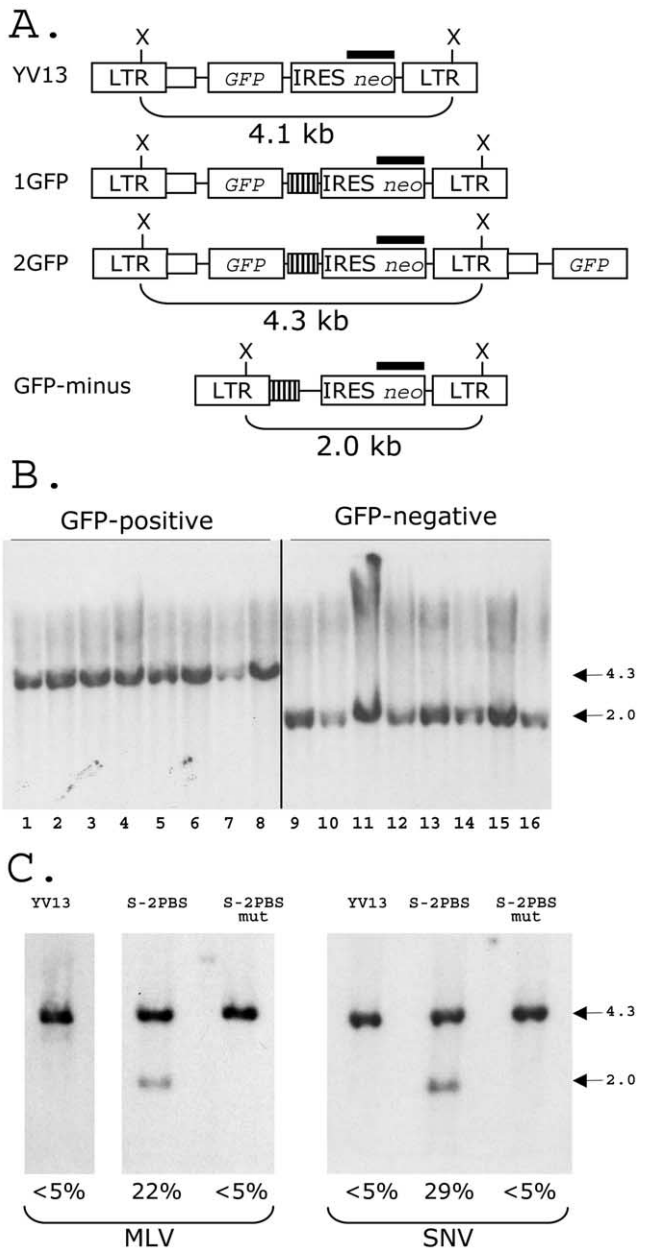


Fig. 5. Measurement of the frequency of GFP-minus provirus formation by Southern blotting analysis. (A) Potential structures of the proviruses, sizes of the corresponding bands after digestion with *Xba*I (X), and the location of the *neo* probe (black bar) are shown. Digestion of 1GFP and 2GFP proviral structures is expected to generate a 4.3-kb band and digestion of the GFP-minus proviral structure is expected to generate a 2.0-kb band. (B) Southern blotting analysis of eight single cell clones that were fluorescent (GFP-positive) and eight clones that were nonfluorescent (GFP-negative) is shown. (C) Southern blotting analysis of pools of cells infected with YV13, S-2PBS, and S-2PBSmut virus derived from MLV and SNV helper cell lines. The proportion of the GFP-minus proviruses was quantified by PhosphorImager analysis and is shown as the percentage of total number of proviruses for each vector.

formation of the GFP-minus proviral structure as shown in Fig. 2C. To determine the relative proportions of GFP-positive and GFP-minus proviruses, we analyzed genomic DNAs from pools of infected cells (Fig. 5C). The frequency of GFP-minus proviral structures was estimated by quantitative PhosphorImager analysis of the 4.3- and 2.0-kb bands. Genomic DNA from cells infected with YV13 was used as a control. As expected, analysis of YV13-infected cells revealed a single 4.2-kb band after *Xba*I digestion; this band is slightly smaller than the 4.3-kb band generated by S-2PBS and S-2PBSmut because it lacks the 3′ PBS fragment. Because only the 5′ PBS was present in this vector, deletion of GFP through initiation of reverse transcription at two PBS sites was not expected to occur. Southern blotting analyses of three independent sets of cells infected with S-2PBS virus derived from the PG13 helper cell line showed that approximately 22% of the cells contained the GFP-minus proviral structure. Similarly, analyses of three independent sets of cells infected with S-2PBS virus derived from the DSH134G helper cell line indicated that 29% of the proviruses had the GFP-minus structure. These GFP deletion frequencies were in agreement with the 26–29% frequency of GFP-negative cells observed after flow cytometry analysis of infected cells (Fig. 3). As expected, Southern blotting analyses of three independent sets of S-2PBSmut-infected cells did not produce a detectable 2.0-kb band; because only the 3′ PBS was functional in this vector, deletion of *GFP* through double initiation events was not expected to occur. The fact that the 2.0-kb bands were not observed for YV13 or S-2PBSmut vectors indicated that the GFP-minus proviral structure was formed only when the vector contained two functional PBSs. This result also strongly supported the view that reverse transcription was initiated at both the 5′ and the 3′ PBSs in the same virion and resulted in the formation of the GFP-minus proviral structures.

Frequencies of 1GFP and 2GFP proviruses generated by initiating reverse transcription once at the 5′ or the 3′ PBS of the S-2PBS vector

As described in Figs. 2A and B, the 1GFP and 2GFP proviral structures can be generated by initiating reverse transcription once at the 5′ PBS or the 3′ PBS, respectively. The strategy for estimating the relative proportions of 1GFP and 2GFP structures is outlined in Fig. 6A. Genomic DNAs isolated from pools of cells infected with YV13, S-2PBS, and S-2PBSmut were digested with *Msc*I, which cut once in *GFP* and once in *neo*. The probe was located in *neo,* which was cut once by *Msc*I, producing two *neo*-containing fragments. Each fragment bound the probe with approximately equal efficiency (data not shown). When genomic DNA of YV13-infected cells was digested with *Msc*I and hybridized to the *neo* probe, only a 1.4-kb band was expected. When genomic DNA of S-2PBS- and S-2PBSmut-infected cells was digested with *Msc*I and hybridized to the *neo* probe, only the 5′ portions of the 1GFP and 2GFP structures generated a specific 1.6-kb band. On the other hand, an additional *Msc*I site was present in the 2GFP proviral struc-

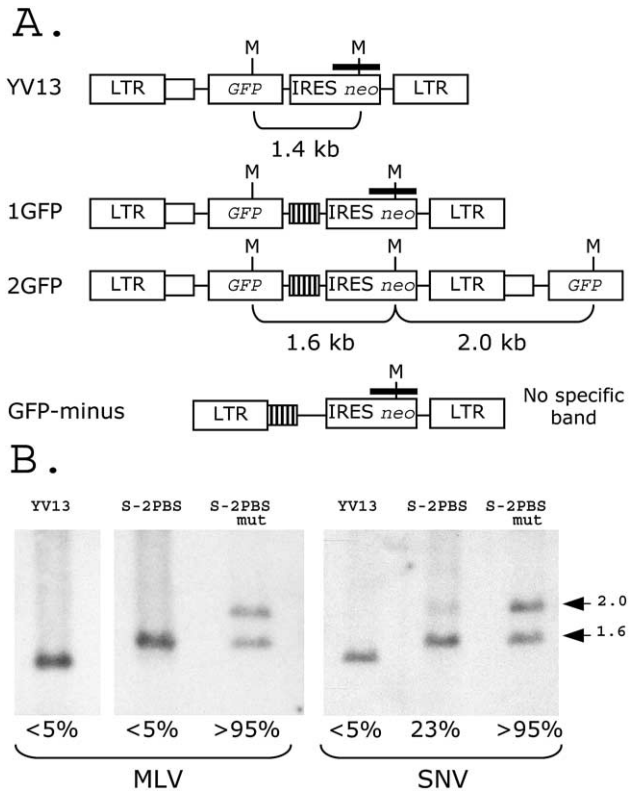


Fig. 6. Measurement of relative amounts of 1GFP and 2GFP structures among GFP-positive proviruses by Southern blotting analyses. (A) Potential structures of the proviruses, sizes of the corresponding bands after digestion with *Msc*I (M), and the location of the *neo* probe (black bar) are shown. Digestion of the 1GFP proviral structure is expected to generate a 1.6-kb band that hybridizes to the 5′ half of the *neo* probe; digestion of the 2GFP proviral structures is expected to generate 1.6- and 2.0-kb bands that hybridize to the 5′ and 3′ fragments of the neo probe, respectively; digestion of the GFP-minus proviral structure is not expected to generate a specific band. (B) Southern blotting analysis of pools of cells infected with YV13, S-2PBS, and S-2PBSmut virus derived from MLV and SNV helper cell lines. The proportion of the 2GFP proviruses was determined by quantitative PhosphorImager analysis and is shown as the percentage of total number of GFP-positive proviruses for each vector.

ture, generating a second specific band of 2 kb. GFP-minus proviruses have only one *Msc*I site and do not generate a specific band. Because both 1GFP and 2GFP structures generated the 1.6-kb band, quantitation of this band provided a measure of the total number of GFP-positive proviruses, whereas the 2-kb band provided a measure of the 2GFP structures.

Southern blotting analysis of genomic DNAs from pools of cells infected with the viruses produced by MLV and SNV packaging cells is shown in Fig. 6B. As expected, the genomic DNA from cells infected with YV13 generated a 1.4-kb band that was slightly smaller than the 1.6-kb band because it lacked the 3′ PBS. In the S-2PBSmut vector, the wild-type PBS was inactivated, and only the 2GFP proviral structures were expected. Indeed, Southern blotting analysis of cells infected with S-2PBSmut virus derived from both MLV and SNV helper cells generated 2.0- and 1.6-kb bands of similar intensity, indicating that most of the proviruses (>95%) had the 2GFP structure. Genomic DNAs from S-2PBS-infected cells produced different results for viruses derived from the PG13 and DSH134G helper cell lines. Analysis of cells infected with virus produced from PG13 cells revealed the presence of a single 1.6-kb band, indicating that very few of the proviruses (<5%) were generated by DNA synthesis initiating only at the 3′ PBS. Analysis of cells infected with virus derived from the DSH134G helper cell line indicated that the 2.0-kb band intensity was 23% the intensity of the 1.6-kb band, indicating that 77% of the GFP-positive proviruses had the 1GFP structure and 23% of the GFP-positive proviruses had the 2GFP structure. This result indicated that among the GFP-positive proviruses generated by S-2PBS virus produced from the SNV helper cell line, the proportion of proviruses that initiated DNA synthesis only once at the 5′ PBS was 77%, and the proportion of proviruses that initiated DNA synthesis only once at the 3′ PBS was 23%.

In contrast, most of the GFP-positive proviruses (>95%) generated by S-2PBS virus produced from the MLV helper cell line had the 1GFP structure. The viral DNAs that form the 1GFP and GFP-minus proviral structures are expected to possess the MLV *att* sites at both ends. However, the viral DNA that forms the 2GFP structure possesses an MLV *att* site at the 5′ end and an SNV *att* site at the 3′ end (Fig. 2B). The observation that the 2GFP proviral structure could not be detected in cells infected with virus produced from the MLV helper cell line suggested that the MLV IN proteins did not utilize the heterologous SNV *att* site efficiently.

GFP deletions do not occur through premature plus-strand DNA transfer

One possible explanation for the GFP-minus proviral structures observed after infection with S-2PBS is that reverse transcription is initiated at the 5′ PBS, but during plus-strand DNA synthesis the strong-stop DNA is transferred to the 3′ PBS. This may occur if the plus-strand strong-stop DNA is transferred to the first PBS that is copied into minus-strand DNA. To test this possibility, we constructed vectors YV20 and YV21. Vector YV20 contains a 3′ PBS region derived from SNV, which includes 10-nt of SNV U5, 18-nt of PBS, and 37-nt of the untranslated leader; because plus-strand DNA transfer is performed using the 18-nt of PBS, this PBS could theoretically serve as a substrate for plus-strand DNA transfer. However the 3′ PBS region of YV20 lacks sequences that form the U5-IR and U5-leader stems, which are critical for efficient initiation of DNA synthesis (Cobrinik et al., 1988, 1991). Thus, the YV20 vector contains a 3′ PBS that cannot be used to initiate DNA synthesis but could theoretically be used as a site for premature plus-strand DNA transfer. Any GFP-minus proviral structures that are formed after infection with YV20 should result from premature plus-strand DNA

transfer and not through initiation of DNA synthesis at both the 5′ and the 3′ PBS sites.

To verify that the 3′ PBS site in YV20 was inefficient in initiation of DNA synthesis, we constructed vector YV21, which contained the same 3-nucleotide substitution in the 5′ PBS that was present in vector YV13mut and was shown to inactivate the PBS. Thus, replication of YV21 is dependent on initiation of DNA synthesis at the 3′ PBS.

Vectors YV20 and YV21 were transfected into the SNV helper cell line and pools of G418-resistant cell clones were isolated. The titers of viruses produced from two independent pools of transfected cells were determined by infection of D17 target cells. As expected, the YV20-transfected cells produced average titers (2.5×10^5 CFU/ml) that were similar to those of YV13 (Table 1). However, the YV21-transfected cells produced average viral titers (8×10^3 CFU/ml) that were approximately 30-fold lower than the titers derived from YV13 transfected cells; this result indicated that the 3′ PBS region inserted in YV20 and YV21 could not be used to efficiently initiate DNA synthesis and complete viral replication.

To determine whether infection with YV20 resulted in the formation of GFP-minus proviral structures, we performed Southern blotting analysis as described in Fig. 5A. Genomic DNAs were isolated from pools of cells infected with YV20 and digested with *Xba*I. As shown in Fig. 7B, YV20- and YV21-infected cell pools generated only the 4.3-kb bands that indicated the presence of the 1GFP or 2GFP structures and did not generate a 2.0-kb band. The lack of the 2.0-kb band after infection with YV20 demonstrated that GFP-minus proviral structures were not generated. This result provided strong evidence that the 3′ PBS was not being used as a site for premature plus-strand DNA transfer.

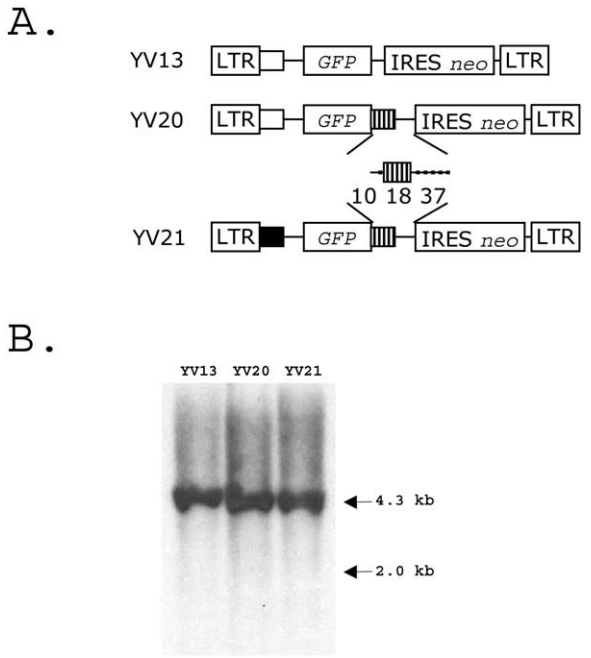


Fig. 7. GFP is not deleted through the transfer of plus-strand strong-stop DNA to the 3′ PBS. (A) Structures of vectors YV13, YV20, and YV21. All abbreviations and symbols describing the vector structures are identical to Fig. 1. YV20 and YV21 vectors contain a modified 3′ PBS fragment derived from SNV; most of the SNV U5 region, which is needed to form the predicted secondary structure shown in Fig. 1B, was deleted. (B) Southern blotting analysis of DNA isolated from pools of cells infected with each vector. The DNAs were digested with *Xba*I and analyzed as described in Fig. 5.

Discussion

The results of this study indicate for the first time that vectors containing two PBSs can initiate reverse transcription more than once in at least 22 and 29% of infectious virion derived from MLV or SNV helper cell lines, respectively. These frequencies of virions that initiated DNA synthesis multiple times are underestimates because it is possible for a virion to initiate DNA synthesis more than once without leading to the formation of a GFP-minus structure.

The analysis outlined in Fig. 8 presents possible scenarios in which up to 100% of the infectious viruses might initiate reverse transcription twice, but result in only a 25% frequency of *GFP* deletion. There are four PBSs in each virion, and dual initiation can result from four different combinations of tRNA placements. First, initiation of DNA synthesis could occur at the 5′ PBS of both genomic RNAs (Fig. 8A), resulting in the formation of the 1GFP structure (shown in Fig. 2). Second, initiation of DNA synthesis could occur at the 3′ PBS of both genomic RNAs (Fig. 8B), resulting in the formation of the 2GFP structure (shown in Fig. 2). Third, initiation of DNA synthesis could occur at the 5′ and 3′ PBS of the same RNA (Fig. 8C); intramolecular minus-strand DNA transfer would result in the formation of a GFP-minus proviral structure (shown in Fig. 2), whereas intermolecular minus-strand transfer would result in the formation of the 1GFP structure. Finally, initiation of DNA synthesis could occur at the 5′ PBS of one of the genomic RNAs and the 3′ PBS of the copackaged RNA (Fig. 8D); depending on which minus-strand strong-stop DNA is used for minus-strand DNA transfer, and whether minus-strand DNA transfer is intramolecular or intermolecular, the GFP-minus, 1GFP, 2GFP, or 3GFP structures could potentially form. The 3GFP structure (not shown) contains an internal duplication of *GFP* in addition to a third *GFP* at the 3′ end of the proviral structure as in the 2GFP structure. For simplicity, we first assume that the initiations of DNA synthesis occur randomly on the four PBSs; second, we assume that intramolecular and intermolecular minus-strand DNA transfers occur with similar efficiencies. Based on these two assumptions, we expect that 1/6 of the virions will initiate DNA synthesis at the 5′ or the 3′ PBS of both genomic RNAs (Figs. 8A and B, respectively), 1/3 of the virions will initiate DNA synthesis at the 5′ and 3′ PBS of the same genomic RNA (Fig. 8C), and 1/3 of the virions will initiate

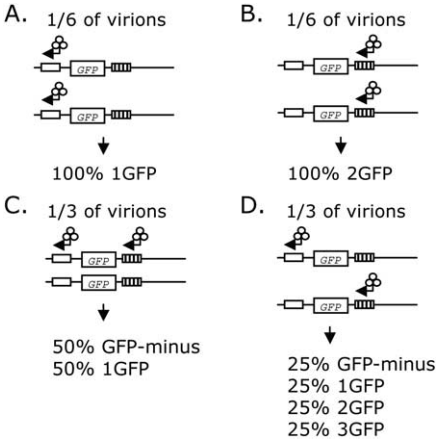


Fig. 8. Possible scenarios for dual initiation of DNA synthesis in vectors containing two PBSs. Four different combinations of tRNA placements (A–D) can lead to dual initiation. (A) Initiation of DNA synthesis at the 5′ PBS of both genomic RNAs. (B) Initiation of DNA synthesis at the 3′ PBS of both genomic RNAs. (C) Initiation of DNA synthesis at the 5′ and 3′ PBS of the same genomic RNA. (D) Initiation of DNA synthesis at the 3′ PBS of one genomic RNA and the 3′ PBS of the other genomic RNA. Assuming that initiation of DNA synthesis occurs randomly at the four PBSs, scenarios A and B are each expected to occur in 1/6 of the virion and scenarios C and D are each expected to occur in 1/3 of the virion. Assuming that intramolecular and intermolecular minus-strand DNA transfers occur with equal efficiencies, 50% of the proviruses generated through scenario C (1/6 of total) will be GFP-minus proviral structures and 25% of the proviruses generated through scenario D (1/12 of total) will be GFP-minus proviral structure. Thus, 25% of all proviruses will be expected to have the GFP-minus proviral structure (1/6 + 1/12 = 3/12 = 25%).

DNA synthesis at the 5′ PBS of one RNA and the 3′ PBS of the other RNA (Fig. 8D). The resulting predicted frequency of GFP-minus proviral structures is 25% if all of the virion initiated reverse transcription twice (see Fig. 8). This predicted frequency of GFP-minus proviral structures is close to the 22 and 29% observed frequencies of *GFP* deletion. Similar analysis of possible scenarios in which initiation of DNA synthesis occurs three times per virion indicates that the expected frequency of GFP-minus proviral structures is 50% (not shown). It is important to point out that the predicted frequency of *GFP* deletion is highly dependent on the assumptions made about the nature of reverse transcription; for example, the predicted frequency of GFP-minus proviral structures is 100% if it is assumed that both initiations always occur on the same RNA and minus-strand DNA transfer is always intramolecular.

In addition to initiating DNA synthesis twice in each S-2PBS virion, DNA synthesis could also initiate once, three times, or four times per virion. If DNA synthesis is initiated only once per virion, then the GFP-minus proviral structure cannot form. On the other hand, if DNA synthesis is initiated four times (at the 5′ and 3′ PBS of both genomic RNAs), then 100% of the proviruses formed should have the GFP-minus structure. The result that 22–29% of the proviruses have the GFP-minus structure strongly suggests that on average two or three initiations of DNA synthesis occurred per virion.

The results of this study indicate that the primer tRNAs, RT, and other components necessary for formation of two functional initiation complexes are present in the virion. Consequently, it is possible that two functional initiation complexes are formed in association with the two PBSs present in wild-type virion, and DNA synthesis may be frequently initiated on both genomic RNAs. Evidence for DNA synthesis on both genomic RNAs was previously obtained by electron microscopy and led to the proposal of a plus-strand model for retroviral recombination (Junghans et al., 1982). However, the frequency with which DNA synthesis occurs on both copies of genomic RNA in a single virion was not determined. Assuming that the results obtained with vectors containing two PBSs can be extrapolated to wild-type virion, at least 22–29% of the virion undergo initiation of reverse transcription on both RNAs.

This result suggests the intriguing possibility that two complete genomic DNAs may frequently form during infection. The process of intermolecular template switching events leading to recombination between two copackaged RNAs may be affected by whether DNA synthesis is occurring on a single RNA template or in parallel on both genomic RNAs. This question was previously studied by using two vectors in which one contained a defective PBS and lacked the ability to initiate DNA synthesis (Anderson et al., 1998); however, it was not possible to measure the frequency of recombination in this study. On the other hand, an intermolecular template switch during minus-strand DNA synthesis may prevent the formation of two complete DNAs. Consistent with this prediction, it has been observed that only one provirus is formed in each infectious event in a recombining population of viruses (Hu and Temin, 1990a).

The assumption that initiation events on two different RNAs occur as frequently in wild-type viruses as in vectors containing two PBSs is supported by our observation that the location of the PBS did not affect the efficiency of its use in DNA synthesis initiation. On the other hand, the efficiency of placing the initiator tRNA on the PBS may influence how often DNA synthesis is initiated. For example, if 50% of PBSs are associated with tRNAs as reported for avian leukosis virus (Fu et al., 1997), then DNA synthesis will be initiated on average more than once in S-2PBS virion and once in wild-type virion. It is also conceivable that interactions between the two PBSs on copackaged RNAs and the adjacent dimer linkage sequences preclude initiation of reverse transcription on both RNAs.

In a previous study, HIV-1 genomes containing two PBSs were characterized (Li et al., 1997). A virus containing a second PBS appeared as a reversion mutation that allowed replication of a virus containing a mutated PBS. Similar to the results reported here, viruses containing two functional PBSs deleted one PBS and the spacer regions after multiple rounds of replication, whereas viruses containing a functional and a defective PBS were found to be stable.

The results of this study show that SNV integrase is able to carry out integration of viral DNAs that contain a 5′-*att* site derived from MLV and a 3′-*att* site derived from SNV. In contrast, the absence of 2GFP structures in cells infected with MLV-derived S-2PBS virus suggests that MLV integrase could not efficiently utilize the heterologous *att* sites. Real-time PCR analysis of infections carried out with virion derived from MLV helper cells also indicated that MLV RT could efficiently initiate DNA synthesis using SNV PBS (data not shown).

The system of vectors containing two PBSs described in this study could be used to determine the effects of *cis*- and *trans*-acting elements that influence initiation of reverse transcription. For example, the role of NC and other viral proteins in the placement of tRNA on the PBS could be analyzed and mutational analysis of the PBS and surrounding secondary structure could be performed to determine the effects of mutations on the efficiency of DNA synthesis initiation.

Materials and methods

Construction of vectors

Plasmid names begin with "p," while names of viruses derived from the plasmids do not. The MLV-based vector pYV13 was derived from pLAEN, which was a kind gift from A.D. Miller (Adam et al., 1991). The 5′ and 3′ LTR regions of the pLAEN vector were derived from Moloney murine sarcoma virus (MoMSV) and MLV, respectively (Miller and Rosman, 1989). All vectors contained *neo*, which was expressed from IRES.

To construct pYV13, a second PBS region that is present downstream of the 3′ LTR in pLAEN was deleted; in addition, a *Not*I restriction fragment derived from pGL1 (Gibco) that encoded *GFP* was inserted upstream of the IRES.

Vector pYV13mut was derived from pYV13 and contained substitution of the 5′ TGG nucleotides of the PBS with ACC. The mutagenesis was performed by two overlapping PCR reactions; the presence of the desired 3-bp substitution and the absence of undesired mutations were confirmed by DNA sequencing.

The SNV U5-PBS-leader region was PCR amplified from pJD214 (Dougherty and Temin, 1986). The amplicon included the entire SNV U5 region (92 bp), 18 bp of the PBS, and 37 bp of the leader sequence and was inserted into an *Eco*RI site downstream of *GFP* in pYV13 and pYV13mut to form pS-2PBS and pS-2PBSmut, respectively.

Vector pYV12, used to generate standard curves for real-time PCR analysis, was derived from pYV13 by deleting the 3′ LTR.

Cells, transfections, and virus propagation

D17 and PG13 cells were obtained from the American Type Culture Collection. DSH134G, an SNV-based packaging cell line, was a kind gift from Ralph Dornburg (Martinez and Dornburg, 1995). D17 and DSH134G are dog osteosarcoma cell lines that are permissive to infection by MLV and SNV. PG13 is a murine packaging cell line that expresses MLV *gag-pol* and gibbon ape leukemia virus *env* (Miller et al., 1991). Cells were grown in Dulbecco's modified Eagle's medium supplemented with 6% (D17 and DSH134G) or 10% (PG13) calf serum (HyClone Laboratories, Inc.). Penicillin (50 U/ml, Gibco) and streptomycin (50 μg/ml, Gibco) were also added to the medium. In addition, transfected DSH134G cells were grown in 3′-azido-3′-deoxythymidine (AZT) containing medium (1 μM final concentration) to prevent reinfection of the virus-producing cells. Cells were maintained at 37°C with 5% CO_2.

Transfections into PG13 cells were performed with the TransFast Transfection Kit (Promega), followed by G418 selection at 600 μg/ml. Transfections into DSH134G cells were performed by the dimethyl sulfoxide-Polybrene method (Kawai and Nishizawa, 1984). Because DSH134G cells are already G418 resistant, the vectors were cotransfected with pSV*hygro* (ratio 10:1) and selection for hygromycin resistance was performed at 120 μg/ml. Plasmid pSV*hygro* codes for the hygromycin phosphotransferase B gene (*hygro*), which confers resistance to hygromycin (Gritz and Davies, 1983).

To generate virus for infection experiments, the vectors were individually transfected into helper cell lines. Transfected cells were pooled, expanded, and plated at a density of 5×10^6 cells per 100-mm-diameter dish. Fresh AZT-free medium was added to virus producer cells 24 h before the virus was harvested. Virus-containing cell-culture medium was centrifuged at 6000 *g* for 10 min to pellet cellular debris. The cell-free supernatant was then used for infection of D17 cells. Flow cytometry analysis of virus-producing cells was performed to ensure that the majority (>80–90%) of transfected cells expressed *GFP*.

Tenfold serial dilutions of virus-containing supernatant were used to infect 2×10^5 D17 cells per 60-mm-diameter dish. Infections were performed in the presence of Polybrene (50 μg/ml) for 4 h. The infected cells were subjected to G418 selection 24 h postinfection. Viral titers were determined by quantitation of G418-resistant colonies approximately 2 weeks after infection. At least 10,000 G418-resistant colonies were pooled per experiment and analyzed by flow cytometry to obtain the percentage of cells that did not express *GFP*.

Detection of GFP expression by flow cytometry

The cells were infected at an m.o.i. <0.1. Approximately 10,000 G418-resistant infected colonies were pooled for analysis. The percentage of GFP-expressing cells was measured using flow cytometry (FACScan; Becton–Dickinson); 10,000 events were collected for each experiment and results were analyzed using CellQuest software (Becton–Dickinson).

Southern blotting analyses

Genomic DNA was isolated from infected cells with the AquaPure isolation kit (Bio-Rad) and proviral structures were analyzed by Southern blot hybridization as previously described (Delviks et al., 1997). A 1.2-kb DNA fragment containing *neo* was used to generate a probe. Quantitation of bands was performed with the Quantity One program (Bio-Rad).

Quantitative real-time PCR analysis

Each vector was transfected into the PG13 or DSH134G helper cell line and pools of G418-resistant cells were selected for at least 2 weeks to minimize contamination of viral preps with transfected plasmid DNA. Virus-containing medium was collected from the pools of helper cells and cleared from cellular debris by filtration (0.22 μm Millex-GS, Drummond). The media was incubated for 30 min at room temperature with DNase I (30 units/ml) and $MgCl_2$ (10 mM final concentration) to digest any DNA transferred from helper cells and was used to infect D17 cells. Infection was performed in the presence of 50 μg/ml Polybrene for 1, 2, 3 or 24 h, after which the cells were washed once with phosphate-buffered saline. For harvesting cells 24 h after infection, fresh media was added to cells 3 h postinfection. Total cellular DNA was extracted from infected cells with QIAmp DNA Blood Mini Kit (Qiagen) and dissolved in water. DNA from approximately 10^5 cells was used for each real-time PCR reaction. The following primer-probe sets were used: For the R-U5 region, the forward primer was 5′-TCCCAATAAAGCCTCTTGCTG-3′, the reverse primer was 5′-AGGAGACCCTCCCAAGGAAC-3′, and the probe was 5′-FAM-TTGCATCCGAATCGTGGTCTCGC-TAMRA-3′. For the U5-ψ region, the forward primer was 5′-GCCTCTTGCTGTTTGCATCC-3′, the reverse primer was 5′-GTCTCCAAATCCCGGACGA-3′, and the probe was 5′-FAM-ATCGTGGTCTCGCTGTTCCTTGGGAG-TAMRA-3′. For the GFP region, the forward primer was 5′-AATCGAGTTGAAGGGCATTGAC-3′, the reverse primer was 5′-TGTGGGAGTTATAGTTGTATTCCAGCT-3′, and the probe was 5′-FAM-TTAAGGAAGATGGAAACATTCTCGGCCAC-TAMRA-3′. For the *neo* region, the forward primer was 5′-GCGCCAGCCGAACTGTT-3′, the reverse primer was 5′-AGCCGGCCACAGTCGAT-3′, and the probe was 5′-FAM-CAAGGCGCGCATGCCCG-TAMRA-3′. The final concentrations of primers and probe were 600 and 75 nM, respectively.

pYV12 was derived from pYV13 by deletion of the 3′LTR and therefore contained a single copy of the R-U5, U5-ψ, GFP, and *neo* regions. Threefold serial dilutions of pYV12 were used to generate a standard curve ranging from 50 to 3,000,000 copies of DNA per PCR reaction. The same dilutions were used to generate a standard curve for each primer-probe set, which allowed accurate measurement of relative amounts of DNA products detected by different sets. The correlation coefficient for all standard curves was >0.99. The amount of each PCR product in the sample was determined from a standard curve generated with that particular primer set. At least two independent experiments were performed and each sample was analyzed at least twice with all four primer sets.

To normalize the amount of DNA analyzed in the real-time PCR experiments, D17 that stably expressed *hygro* were used as targets of infection. In addition to the primer sets used for monitoring viral infection, each sample was also analyzed with a primer set designed to detect the copy number of *hygro*. Because the number of *hygro* genes per cell was constant, the copy number of *hygro* was used to normalize the amount of input DNA in each infection. The following primer-probe set was used to determine the copy number of *hygro:* the forward primer was 5′-ACGAGGTCGCCAACATCTTC-3′, the reverse primer was 5′-CGCGTCTGCTGCTCCAT-3′, and the probe was 5′-FAM-CAAGCCAACCACGGCCTCCAGA-TAMRA-3′.

Acknowledgments

We especially thank Wei-Shau Hu, John Coffin, and Steve Hughes for intellectual input and valuable discussion of this manuscript, and Anne Arthur for expert editorial help and revisions. We also thank Que Dang, Jean Mbisa, Terence Rhodes, David Thomas, and Wen-Hui Zhang for critical reading of the manuscript. This work was supported by the HIV Drug Resistance Program, National Cancer Institute.

References

Adam, M.A., Ramesh, N., Miller, A.D., Osborne, W.R., 1991. Internal initiation of translation in retroviral vectors carrying picornavirus 5′ nontranslated regions. J. Virol. 65 (9), 4985–4990.

Anderson, J.A., Teufel 2nd, R.J., Yin, P.D., Hu, W.S., 1998. Correlated template-switching events during minus-strand DNA synthesis: a mechanism for high negative interference during retroviral recombination. J. Virol. 72 (2), 1186–1194.

Artzi, H.B., Shemesh, J., Zeelon, E., Amit, B., Kleiman, L., Gorecki, M., Panet, A., 1996. Ribonuclease H activity during initiation of reverse transcription using tRNA(lys)/RNA primer/template of human immunodeficiency virus. Arch. Biochem. Biophys. 325 (2), 209–216.

Baltimore, D., 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature 226 (252), 1209–1211.

Bauer, G., Temin, H.M., 1980. Radioimmunological comparison of the DNA polymerases of avian retroviruses. J. Virol. 33 (3), 1046–1057.

Butler, S.L., Hansen, M.S., Bushman, F.D., 2001. A quantitative assay for HIV DNA integration in vivo. Nat. Med. 7 (5), 631–634.

Butler, S.L., Johnson, E.P., Bushman, F.D., 2002. Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. J. Virol. 76 (8), 3739–3747.

Cen, S., Khorchid, A., Gabor, J., Rong, L., Wainberg, M.A., Kleiman, L., 2000. Roles of Pr55(gag) and NCp7 in tRNA(3)(Lys) genomic placement and the initiation step of reverse transcription in human immunodeficiency virus type 1. J. Virol. 74 (22), 10796–10800.

Cobrinik, D., Aiyar, A., Ge, Z., Katzman, M., Huang, H., Leis, J., 1991. Overlapping retrovirus U5 sequence elements are required for efficient integration and initiation of reverse transcription. J. Virol. 65 (7), 3864–3872.

Cobrinik, D., Soskey, L., Leis, J., 1988. A retroviral RNA secondary structure required for efficient initiation of reverse transcription. J. Virol. 62 (10), 3622–3630.

Delviks, K.A., Hu, W.S., Pathak, V.K., 1997. Psi-vectors: murine leukemia virus-based self-inactivating and self-activating retroviral vectors. J. Virol. 71 (8), 6218–6224.

Dougherty, J.P., Temin, H.M., 1986. High mutation rate of a spleen necrosis virus-based retrovirus vector. Mol. Cell. Biol. 6 (12), 4387–4395.

Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M., Naldini, L., 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. Nat. Genet. 25 (2), 217–222.

Fu, W., Ortiz-Conde, B.A., Gorelick, R.J., Hughes, S.H., Rein, A., 1997. Placement of tRNA primer on the primer-binding site requires pol gene expression in avian but not murine retroviruses. J. Virol. 71 (9), 6940–6946.

Gilboa, E., Mitra, S.W., Goff, S., Baltimore, D., 1979. A detailed model of reverse transcription and tests of crucial aspects. Cell 18 (1), 93–100.

Gritz, L., Davies, J., 1983. Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene 25 (2–3), 179–188.

Harada, F., Sawyer, R.C., Dahlberg, J.E., 1975. A primer ribonucleic acid for initiation of in vitro Rous sarcarcoma virus deoxyribonucleic acid synthesis. J. Biol. Chem. 250 (9), 3487–3497.

Hu, W.S., Temin, H.M., 1990a. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc. Natl. Acad. Sci. USA 87 (4), 1556–1560.

Hu, W.S., Temin, H.M., 1990b. Retroviral recombination and reverse transcription. Science 250 (4985), 1227–1233.

Huang, Y., Khorchid, A., Gabor, J., Wang, J., Li, X., Darlix, J.L., Wainberg, M.A., Kleiman, L., 1998. The role of nucleocapsid and U5 stem/A-rich loop sequences in tRNA(3Lys) genomic placement and initiation of reverse transcription in human immunodeficiency virus type 1. J. Virol. 72 (5), 3907–3915.

Huang, Y., Mak, J., Cao, Q., Li, Z., Wainberg, M.A., Kleiman, L., 1994. Incorporation of excess wild-type and mutant tRNA(3Lys) into human immunodeficiency virus type 1. J. Virol. 68 (12), 7676–7683.

Hwang, C.K., Svarovskaia, E.S., Pathak, V.K., 2001. Dynamic copy choice: steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of in vivo template switching. Proc. Natl. Acad. Sci. USA 98 (21), 12209–12214.

Jang, S.K., Davies, M.V., Kaufman, R.J., Wimmer, E., 1989. Initiation of protein synthesis by internal entry of ribosomes into the 5′ nontranslated region of encephalomyocarditis virus RNA in vivo. J. Virol. 63 (4), 1651–1660.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., Wimmer, E., 1988. A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. 62 (8), 2636–2643.

Jorgensen, R.A., Rothstein, S.J., Reznikoff, W.S., 1979. A restriction enzyme cleavage map of Tn5 and location of a region encoding neomycin resistance. Mol. Gen. Genet. 177 (1), 65–72.

Ju, G., Skalka, A.M., 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. Cell 22 (2 Pt. 2), 379–386.

Junghans, R.P., Boone, L.R., Skalka, A.M., 1982. Retroviral DNA H structures: displacement-assimilation model of recombination. Cell 30 (1), 53–62.

Kawai, S., Nishizawa, M., 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. Mol. Cell. Biol. 4 (6), 1172–1174.

Lee, M.S., Craigie, R., 1994. Protection of retroviral DNA from autointegration: involvement of a cellular factor. Proc. Natl. Acad. Sci. USA 91 (21), 9823–9827.

Leis, J., Aiyar, A., Cobrinik, D., 1993. Regulation of initiation of reverse transcription of retroviruses, in: Skalka, A.M., Goff, S.P. (Eds.), Reverse Transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 33–47.

Levin, J.G., Hatfield, D.L., Oroszlan, S., Rein, A., 1993. Mechanisms of translational suppression used in the biosynthesis of reverse transcriptase, in: Skalka, A.M., Goff, S.P. (Eds.), Reverse Transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 5–31.

Levin, J.G., Seidman, J.G., 1979. Selective packaging of host tRNA's by murine leukemia virus particles does not require genomic RNA. J. Virol. 29 (1), 328–335.

Li, Y., Kang, S.M., Morrow, C.D., 1997. Stability of HIV type 1 proviral genomes that contain two distinct primer-binding sites. AIDS Res Hum Retroviruses 13 (3), 253–262.

Mak, J., Jiang, M., Wainberg, M.A., Hammarskjold, M.L., Rekosh, D., Kleiman, L., 1994. Role of Pr160gag-pol in mediating the selective incorporation of tRNA(Lys) into human immunodeficiency virus type 1 particles. J. Virol. 68 (4), 2065–2072.

Mak, J., Kleiman, L., 1997. Primer tRNAs for reverse transcription. J. Virol. 71 (11), 8087–8095.

Martinez, I., Dornburg, R., 1995. Improved retroviral packaging lines derived from spleen necrosis virus. Virology 208 (1), 234–241.

Miller, A.D., Garcia, J.V., von Suhr, N., Lynch, C.M., Wilson, C., Eiden, M.V., 1991. Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. J. Virol. 65 (5), 2220–2224.

Miller, A.D., Rosman, G.J., 1989. Improved retroviral vectors for gene transfer and expression. Biotechniques 7 (9), 980–982, 984–986, 989–990.

Panet, A., Kra-Oz, Z., 1978. A competition immunoassay for characterizing the reverse transcriptase of mammalian RNA tumor viruses. Virology 89 (1), 95–101.

Peters, G., Harada, F., Dahlberg, J.E., Panet, A., Haseltine, W.A., Baltimore, D., 1977. Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. J. Virol. 21 (3), 1031–1041.

Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., et al., 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature 313 (6000), 277–284.

Shank, P.R., Varmus, H.E., 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 (1), 104–104.

Shoemaker, C., Hoffman, J., Goff, S.P., Baltimore, D., 1981. Intramolecular integration within Moloney murine leukemia virus DNA. J. Virol. 40 (1), 164–172.

Sirven, A., Pflumio, F., Zennou, V., Titeux, M., Vainchenker, W., Coulombel, L., Dubart-Kupperschmitt, A., Charneau, P., 2000. The human immunodeficiency virus type-1 central DNA flap is a crucial determinant for lentiviral vector nuclear import and gene transduction of human hematopoietic stem cells. Blood 96 (13), 4103–4110.

Svarovskaia, E.S., Delviks, K.A., Hwang, C.K., Pathak, V.K., 2000. Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. J. Virol. 74 (15), 7171–7178.

Swanstrom, R., DeLorbe, W.J., Bishop, J.M., Varmus, H.E., 1981. Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. Proc. Natl. Acad. Sci. USA 78 (1), 124–128.

Temin, H.M., 1981. Structure, variation and synthesis of retrovirus long terminal repeat. Cell 27 (1 Pt. 2), 1–3.

Temin, H.M., Mizutani, S., 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature 226 (252), 1211–1213.

Vogt, V.M., Simon, M.N., 1999. Mass determination of rous sarcoma virus virions by scanning transmission electron microscopy. J. Virol. 73 (8), 7050–7055.