JOURNAL OF VIROLOGY, July 2005, p. 8410–8421 Vol. 79, No. 13
0022-538X/05/$08.00+0 doi:10.1128/JVI.79.13.8410–8421.2005

Mobilization and Mechanism of Transcription of Integrated Self-Inactivating Lentiviral Vectors

Hideki Hanawa, Derek A. Persons, and Arthur W. Nienhuis*

Division of Experimental Hematology, Department of Hematology/Oncology, St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, Tennessee 38105

Received 13 December 2004/Accepted 11 March 2005

Permanent genetic modification of replicating primitive hematopoietic cells by an integrated vector has many potential therapeutic applications. Both oncoretroviral and lentiviral vectors have a predilection for integration into transcriptionally active genes, creating the potential for promoter activation or gene disruption. The use of self-inactivating (SIN) vectors in which a deletion of the enhancer and promoter sequences from the 3′ long terminal repeat (LTR) is copied over into the 5′ LTR during vector integration is designed to improve safety by reducing the risk of mobilization of the vector genome and the influence of the LTR on nearby cellular promoters. Our results indicate that SIN vectors are mobilized in cells expressing lentiviral proteins, with the frequency of mobilization influenced by features of the vector design. The mechanism of transcription of integrated vector genomes was evaluated using a promoter trap design with a vector encoding tat but lacking an upstream promoter in a cell line in which drug resistance depended on tat expression. In six clones studied, all transcripts originated from cryptic promoters either upstream or within the vector genome. We estimate that approximately 1 in 3,000 integrated vector genomes is transcribed, leading to the inference that activation of cryptic promoters must depend on local features of chromatin structure and the constellation of nearby regulatory elements as well as the nature of the regulatory elements within the vector.

Retroviral vectors have been adapted to achieve stable, permanent modification of the genome of target cells, with the ultimate clinical goal of expressing gene products that achieve a therapeutic purpose. Beginning in the early 1980s, efforts focused on the development of therapeutic retroviral vectors based on murine oncoretroviruses. However, work in large animal models and clinical trials revealed the inefficiency of such vectors in transducing various target cells and therefore their ineffectiveness at achieving the desired therapeutic goal (44). Lentiviral vectors, because their preintegration complex traverses the nuclear membrane (10, 33), are potentially superior to oncoretroviral vectors at transducing nonmitotic cells, including hematopoietic stem cells ex vivo (13, 27) and liver (6, 19), muscle (19), brain (32, 49), and retina (45) in vivo. The relative stability of the lentiviral preintegration complex compared to that of oncoretroviral vectors also contributes to improved transduction efficiency (10). Lentiviral vector systems based on human immunodeficiency virus (HIV) (1), simian immunodeficiency virus (SIV) (13, 38), feline immunodeficiency virus (36), and equine infectious anemia virus (34, 35) are at various stages of preclinical development.

The clinical evaluation and use of lentiviral vectors for therapeutic applications will ultimately depend on their safety profile. Promising in this regard is the fact that the accessory proteins that contribute to lentiviral virulence have proved dispensable with respect to vector production and subsequent transduction of target cells (1). Among the potential adverse events associated with lentiviral vectors are the following: (i) activation of a proto-oncogene near the insertion site by transcriptional elements within the vector genome, (ii) interruption and inactivation of an important gene by vector insertion, and (iii) generation and propagation of replication-competent lentiviruses during vector production, a risk that has been diminished by separation of the coding sequences for the vector genome, structural proteins, regulatory proteins, and an envelope protein on four individual plasmids (1). The generation and transmission of recombinants that include the long terminal repeat (LTR) of the vector genome and the gag and pol coding sequences from the packaging plasmid during vector production are well documented (40, 51), but further separation of the gag and pol coding sequences on individual plasmids or development of stable packaging cell lines (52) substantially reduces this risk (22, 48). Finally, mobilization and spread of a vector genome by subsequent superinfection by wild-type HIV (4) must also be considered.

The risk of activation of a proto-oncogene is underscored by the development of leukemia in two young children who participated in a gene therapy trial for severe combined immune deficiency. Insertional activation of the LMO2 gene, a known proto-oncogene, by the LTR enhancer contributed to leukemogenesis in these patients (11). In future studies, characterization of vectors with respect to their capacity to activate nearby genes will become an important aspect of their preclinical evaluation. Interruption of a critical gene by vector integration will also remain as a finite risk, since both oncoretroviral and lentiviral vectors preferentially integrate into transcriptionally active genes (43, 50). Recently, in a collaboration between our group and Cynthia Dunbar's, a detailed comparison has been made of the pattern of oncoretroviral (Moloney leukemia virus) versus lentiviral (SIV) integration into hematopoietic stem cells of rhesus macaques (17). Both

* Corresponding author. Mailing address: St. Jude Children's Research Hospital, 332 N. Lauderdale, Mail Stop #272, Memphis, TN 38105. Phone: (901) 495-3301. Fax: (901) 525-2720. E-mail: arthur.nienhuis@stjude.org.

types of vectors exhibited a preference for integration into genes which are transcriptionally active in primitive hematopoietic cells. Moloney leukemia virus integration events also showed a predilection for the region of the transcriptional start site, whereas SIV integrations were more evenly distributed throughout genes. In each case, fewer than 5% of the integration events occurred within exons. These data suggest that oncoretroviral vectors may have a greater potential to activate genes, as was observed in the clinical trial (11), whereas both vectors have the potential to interrupt a gene leading to lowered expression. The risk of gene interruption is mitigated by the fact that most genes are recessive or exhibit a pathological effect only when expressed as a transdominant mutant.

The risk of mobilization of the vector genome by HIV superinfection has been reduced by the development of SIN vectors (30, 42, 53). Deletion of the enhancer and promoter sequences from the 3′ LTR, which is copied over into the 5′ LTR during the retroviral life cycle, inactivates both LTRs of the integrated proviral genome. Although this modification may reduce the risk that the integrated vector genome will enhance expression of surrounding genes, the internal enhancer and promoter required for transgene expression sustain this as a relevant risk. Available evidence also indicates that mobilization of an integrated vector genome is reduced but not totally eliminated by the SIN design of the vector (9).

Our studies were designed to evaluate the influence of vector design on the potential for mobilization of a SIN lentiviral vector genome. We found that integrated vector genomes were frequently transcribed to yield vector particles capable of transferring an intact mobilized genome into target cells. Vector design significantly influenced the likelihood of vector genome transcription. Mechanistic studies suggest that vector transcription most often occurs via activated cryptic promoter sites rather than from readthrough transcription from an upstream gene.

MATERIALS AND METHODS

Plasmid constructions. The following is a general outline of the derivation of the individual plasmids used in these studies. Further details are available on request.

pCL20c MpGFP. For pCL20c MpGFP, the vector genome of pCL20c MSCV GFP (13, 14) was modified to eliminate the U5 sequences from the murine stem cell virus (MSCV) LTR (16) by deleting the sequences downstream from the XmaI site through the AgeI site just upstream from the green fluorescent protein (GFP) coding sequences.

pCL20c S0 MpGFP. For pCL20c S0 MpGFP, the 3′ LTR of the base plasmid, pCL20c MpGFP, has a deletion which leaves only 35 bp from the 5′ end and 18 bp from the 3′ end of the LTR (see Fig. 2, below). To generate the derivative pCL20c S0 MpGFP, from which the 3′ 18 bp have been removed, the primers 5′-CCG GAA GAC AAG ATC GGG TCT CTC TGG TTA GAC CAG ATC-3′ and 5′-TGT TCA TGG CAG CCA GCA TA-3′ were used to amplify the sequences between the R region in the 3′ LTR and downstream sequences in the plasmid pCL20c MpGFP. A BbsI-XhoI fragment was recovered from the PCR product and ligated into the corresponding sites of pCL20c MpGFP.

pCL20c mAS0 MpGFP. For pCL20c mAS0 MpGFP, 20 of the 35 bp on the 5′ end of the 3′ LTR were eliminated by ligating annealed oligonucleotides 5′-GGC CGC GGT ACC TTT TTA AAA GAA AAG GGG GGA CTG GAA GGG CTA ATC GAT-3′ and 5′-GAT CAT CGA TTA GCC CTT CCA GTC CCC CCT TTT CTT TTA AAA AGG TAC CGC-3′ between the NotI and BbsI sites of pCL20c S0 MpGFP. This construction also eliminated the Δnef coding sequences from the parent plasmid.

pCL20c mAUS0 MpGFP. The pCL20c mAUS0 MpGFP plasmid was assembled by ligating oligo 5′-GGC CGC GGT ACC TTT TTA AAA GAA AAG GGG GGA CTG GAA GGG CTA ATC GAT TTG TGA AAT TTG TGA TAT TTG TAA C-3′ and 5′-GAT CGT TAC AAA TAT CAC AAA TTTT CAC AAA TCG ATT AGC CCT TCC AGT CCC CCC TTT TCT TTT AAA AAG GTA CCG C-3′ into the NotI-BbsI site of pCL20c S0 MpGFP. With this modification, three AUUUGURA sequences, which constitute the upstream efficiency element of the simian virus 40 (SV40) late polyadenylation signal (41), were inserted into ΔU3 of pCL20c mAS0 MpGFP.

pCL20c mMpGFP. For pCL20c mMpGFP, the enhancer sequences in the MSCV LTR (15, 28) were removed from the parent plasmid by digestion with HpaI, which cuts downstream from the Rev response element (RRE), and XbaI, which cuts within the MSCV LTR 3′ to the enhancer sequence. Following digestion, the overhang of the XbaI site was filled in with Klenow fragment and the plasmid was religated.

pCL20c INS1R MpGFP and pCL20c INS1L MpGFP. For pCL20c INS1R MpGFP and pCL20c INS1L MpGFP, the plasmid pCL20c MpGFP was digested with BbsI, which cuts in the ΔU3 region, and a 1.2-kb Ecl136II fragment containing hypersensitive site 4 from the chicken β-globin locus control region (cHS4) (37) was ligated into the plasmid. The INS1R version has the insulator in the tandem orientation relative to the internal promoter, whereas the INS1L version has the insulator reversed.

pCL20c EF1αGFP. In the pCL20c EF1αGF plasmid, the MSCV LTR promoter was replaced with the promoter for elongation factor Iα (EF1α) (31, 47). A blunt-ended HindIII-PmlI fragment from the plasmid pEF/myc/nuc (Invitrogen, Carlsbad, CA) was ligated into a blunt-ended HpaI-EcoRI fragment from the parent plasmid, pCL20c MSCV-GFP. The EF1α promoter fragment includes the first intron from that gene. The EF1α sequences included in this vector extend from −203 to +986 from the mRNA CAP site and correspond to sequences from nucleotide (nt) 373 to 1561 in the GenBank sequence (accession number JO4617).

pCL20c EF1αSAmGFP. For pCL20c EF1αSAmGFP, two PCR fragments were amplified using two primer sets, (i) 5′-TCGTGCTTGAGTTGAG-3′ and 5′-CA CGACACgTGAAATGGAAGAAA-3′ and (ii) 5′-TTCCATTTCAcGTGTCGT GAAC-3′ and 5′-GCTTGTCGGCCATGATATAG-3′, using pCL20c EF1αGFP as a template (lowercase letters indicate the mutation site). The two PCR products were connected by amplification using the first primer and fourth primers. Finally, the resulting PCR product was digested with FseI and SmaI and then ligated into the FseI-SmaI site of pCL20c EF1αGFP. Successful mutagenesis was confirmed by sequencing.

pCL20c sEF1αGFP. For pCL20c sEF1αGFP, a vector lacking the EF1α intron was derived by ligating a SmaI-BsrGI EF1α promoter–GFP cassette from vector HRST-IEF1αGFP-WS (provided by John Gray) into a HpaI-BsrGI fragment from pCL20c MpGFP. In this vector, the EF1α promoter sequences extend from the −197 to +34 from the mRNA CAP site and correspond to sequences from nt 374 to nt 614 in the GenBank record HUMEF1α.

Vectors encoding tat. pCL20c RT MpGFP, pCL20c RT INS1R MpGFP, pCL20c RT INS1L MpGFP, pCL20c RT mMpGFP, and pCL20c RT EF1aGFP were constructed by replacing the MfeI-BstBI RRE fragment of pCL20c MpGFP, pCL20c INS1R MpGFP, pCL20c INS1L MpGFP, pCL20c mMpGFP, or pCL20c EF1aGFP with the EcoRI-ClaI fragment from pCAG4-RTR2. The EcoRI-ClaI fragment includes the RRE fragment flanked by the rev and tat exons.

pCAG-KGP1.1R. The gag/pol expression plasmid pCAG-kGP1.1R is a modified version of pCAG-kGP1R (14). The sequence between BstEII and ClaI of pCAG-kGP1R was replaced with the annealed oligonucleotides 5′-GTT ACC ATG GGA GCA CGC GCC AGT GTC CTT TCA GGT GGC GAG CTC GAC-3′ and 5′-CGG TCG AGC TCG CCA CCT GAA AGG ACA CTG GCG CGT GCT CCC ATG-3′, thereby eliminating 44 bp of homology between the transfer vector and the gag/pol expression plasmid.

BSvRSC-PUR-BaGFP. The BSvRSC-PUR-BaGF plasmid was assembled by inserting a blunt-ended 812-bp HindIII-XbaI fragment containing the puromycin resistance gene cDNA from pPUR (Clontech) into the blunt-ended EcoRI site of BSvRSC BaGFP. BSvRSC BaGFP is an HIV type 1 (HIV-1)-based transfer vector which contains the wild-type LTR and the chicken β-actin promoter to express GFP. The EcoRI site in BSvRSC-BaGFP is between the RRE and the β-actin promoter, so that the puromycin gene is expressed at significant levels only when the wild-type HIV-1 LTR is transactivated by tat protein.

Vector production. Vesicular stomatitis virus G protein (VSV-G)-pseudotyped lentiviral vector particles were prepared using a four-plasmid system by transient transfection of 293T cells as previously described (13, 14). In brief, 293T cells were transfected with a mixture of plasmid DNAs consisting of 6 μg pCAGkGP1.1R (gag/pol), 2 μg pCAG4-RTR2 (rev/tat), 2 μg pCAG-VSVG (VSV-G envelope), and 10 μg of a gene transfer vector plasmid per 10-cm dish, using the calcium phosphate precipitation technique. Eighteen hours after transfection, the cells were washed twice with phosphate-buffered saline (PBS), and

Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 50 U/ml penicillin G, and 50 μg/ml streptomycin (D10) was added to each plate of cells. Twenty-four hours later, the medium containing vector particles was harvested, cleared by low-speed centrifugation, and filtered through a 0.22-μm-pore-size cellulose acetate filter. The GFP titer was determined on HeLa cells as described previously (13).

Mobilization assay. Twenty-four hours before transduction, 5×10^4 293T cells were seeded into a six-well plate in 2 ml of D10 (day zero). The 293T cells were transduced with 2.5×10^6 transducing units (TU) of a lentiviral vector medium containing 7.5 μg/ml of Polybrene three times at intervals of 24 h (days 1, 2, and 3). Following the final transduction, the 293T cells were transferred into a 10-cm dish (day 4). After an additional 3 days of culture, the cells were split 1:7 into 10-cm dishes (day 7). On the following day, the cells were transfected with the helper plasmids (6 μg pCAGkGP1.1R, 2 μg pCAG4-RTR2, and 2 μg pCAG-VSVG/10-cm dish). After 18 h, the transfection medium was removed, the cells were washed twice with PBS, D10 was added, and culture continued for another 24 h before the medium was harvested and titrated on HeLa cells for transfer of the GFP marker.

Characterization of mobilized vector genomes. Medium conditioned by 293T cells having multiple vector copies after transfection with helper plasmids was used to derive subclones of 293T cells containing a single copy of a mobilized vector genome. Briefly, 5×10^4 cells were seeded into six-well plates and transduced with 2 ml of a conditioned medium for 24 h. Polybrene (8 μg/ml) was added to the culture medium. Approximately 5% of the cells expressed GFP, and single, positive cells were recovered by fluorescence-activated cell sorting and cultured to provide GFP-expressing clones. Serial dilutions of a primary vector preparation were used to derive clones of 293T cells having a single copy of the vector as a control using similar methodology. Southern blot analysis was performed by standard methods using BbsI, which cuts in each LTR, to evaluate the integrity of the integrated genome or EcoRV, which cuts once in the vector genome, to determine copy number. A radiolabeled 720-bp BamHI-NotI GFP fragment was used as the probe. Further characterization of the integrated 5′ LTR was performed by PCR amplification using a 5′ primer which corresponded to the 5′ end of the LTR (5′-TGGAAGGGCTAATTCACTC-3′) and a reverse primer from the 5′ end of the gag sequences (5′-CGCTTAATACTGACGCTC TC-3′) to give a predicted 417-bp PCR product from a proviral genome having an intact LTR.

tat expression assay. The tat expression indicator cell line, HeLa-PUR, was made by subcloning HeLa cells transduced with the HIV-1 vector BScRSC-PUR-BaGFP. This cell line is sensitive to puromycin (≥1 μg/ml) but becomes resistant when transduced with a tat-expressing vector. To detect tat expression, confluent HeLa-PUR cells were split 1:8 24 h before transduction. For titration of each pCL20c RT vector, 1 ml, 0.1 ml, or 0.01 ml of medium containing vector particles was added to 10 ml of D10 containing 6 μg/ml of Polybrene. Two days after initiation of transduction, the medium was replaced with fresh D10 containing 1 μg/ml of puromycin, and 2 days later the medium was replaced with fresh D10 with the same concentration of puromycin. Two weeks after transduction, the cells were washed with PBS, fixed, and stained using crystal violet–10% formaldehyde–1× PBS, and puromycin-resistant (tat^+) colonies were counted. Clones were also isolated in puromycin-containing medium for mapping of integration and transcriptional start sites.

Mapping of integration sites. The determination of integration sites was carried out using linear amplification-mediated PCR (LAM-PCR). Briefly, each integration site sequence was linearly amplified (100 cycles) using biotinylated primers (HIV3-I for the 3′ end junction, HIV5-I for the 5′ end junction) and HotStarTaq master mix kit (QIAGEN, Valencia, CA) or the TaKaRa Ex *Taq* Hot Start version (TaKaRa, Japan) using 100 ng of genomic DNA as template. Amplified single-strand DNA fragments were conjugated to Dynabeads (Dynal Biotech Inc., Lake Success, NY) and then rendered double stranded using random primers. After digestion with restriction endonucleases (Tsp509I for the 3′ end junction and ApoI for the 5′ end junction), an asymmetric linker (annealed oligos LCa and LCb) was ligated to the digested double-strand DNA. Biotin-free, single-strand DNA was obtained by adding 5 μl of 0.1 N NaOH followed by 5 μl of 0.1 N HCl to the Dynabeads. The junctional sequences were amplified by nested PCR using primer sets specific for LTR and linker sequence. To amplify the 3′ LTR and junctional genomic sequence, outer primer set HIV3-II and LC1 and inner primer set HIV3-III and LC2 were used. For the amplification of the 5′ LTR and junctional genomic sequence, HIV5-I was used for linear amplification and then the outer primer set, HIV5-II and LC1, and inner primer set, HIV5-III and LC2, were used for exponential amplification. The PCR products were size fractioned by electrophoresis on an agarose gel and directly sequenced after purification. The junctional sequence was considered valid when the PCR fragment contained an intact linker and LTR sequences. The integration sites were mapped using the National Center for Biotechnology Information BLAST build 35.1 of the human genome and the "search for short nearly exact matches" tool. The oligos which were used in this assay were the following:

LCa, 5′-AAT TCT CTA GTA TGC TAC TCG CAC CGA TTA TCT CCG CTG TCA GT-3′;

LCb, 5′-ACT GAC AGC GGA GAT AAT CGG TGC GAG TAG CAT ACT AGA G-3′;

LC1, 5′-ACT GAC AGC GGA GAT AAT CG-3′;

LC2, 5′-GTG CGA GTA GCA TAC TAG AG-3′,

HIV3-I, 5′-TTT TGC CTG TAC TGG GTC TCT CTG-3′,

HIV3-II, 5′-TCT CTG GCT AAC TAG GGA AC-3′;

HIV3-III, 5′-GCC TTG AGT GCT TCA AGT AGT G-3′;

HIV5-I, 5′-AGG GTC TGA GGG ATC TCT AGT TAC-3′;

HIV5-II, 5′-CAG TGG GTT CCC TAG TTA GC-3′; and

HIV5-III, 5′-GCA AAA AGC AGA TCT TGT CTT C-3′.

In several cases, the LAM-PCR procedure yielded a genome-vector junction fragment with an apparent deletion in the LTR sequences which was judged to be artifactual. This ambiguity was resolved in each case by utilizing a 5′ primer derived from adjacent genomic sequences and a 3′ primer within the LTR to generate a fragment for sequencing across the LTR genomic DNA junction.

Mapping of 5′ ends of tat-encoding RNAs. RNA ligase-mediated rapid amplification of 5′ cDNA ends (5′-RACE; GeneRacer kit; Invitrogen, Carlsbad, CA) was used to determine the start 5′ sites of tat-encoding, capped RNAs extracted from puromycin-resistant HeLa-PUR cells. Uncapped RNAs were initially dephosphorylated with calf intestinal phosphatase to prevent their subsequent participation in the reaction. Capped RNAs were decapped with tobacco acid pyrophosphatase and an RNA oligo which contained the outer and inner nested PCR primer sequences, and the sequence GAAA on the 3′ end was ligated to the 5′ end using T4 RNA ligase. The product was reversed transcribed using random primers, and nested PCR was performed using primers on the RNA oligo and vector sequences. If the PCR product retained the intact 3′-end RNA oligo sequence including the GAAA sequence, false priming was judged to be unlikely. The gene-specific primers are complementary to the tat coding sequence (outer, 5′-CGG GAT TGG GAG GTG GGT TGC TTT G-3′; inner, 5′-GCT GTC TCC GCT TCT TCC TGC CAT AG-3′) or to sequences between the splice acceptor for *vif* (which can be used for tat expression in this vector) and the splice acceptor for *rev* (outer, 5′-AGT GGC TCG AAT TGT CCC TCA TAT CTC C-3′; inner, 5′-CCT CAT CCT GTC TAC TTG CCA CAC AAT C-3′). The PCR products were shotgun cloned into a plasmid vector for DNA sequencing. In some cases, higher-molecular-weight products or those present in higher concentration were purified on an agarose gel prior to the cloning step.

RESULTS

Vector mobilization. Experiments were designed to determine whether our SIN lentiviral vector could be mobilized after integration into a target cell genome. The initial experiments were performed with an HIV-based vector (MpGFP) containing the MSCV LTR, which had been modified to eliminate the U5 region, to drive the GFP coding sequences. Polyclonal populations of 293T cells with multiple vector integrants were derived by exposing cells three times to vector particles at a high multiplicity of infection (MOI) of up to 50. After passage, the cells were transfected with a mixture of plasmids encoding the structural, regulatory (rev and tat), and envelope proteins necessary for vector particle formation. Conditioned media from these cells as well as from control, untransduced but transfected cells and transduced but not transfected cells were added to naive HeLa cells. Vector particles were mobilized into the experimental conditioned medium at a titer of 4.1×10^3 TU/ml (Fig. 1A), but no vector particles were detected in either of the controls (Fig. 1B and C). We speculated that the ability to mobilize the SIN lentiviral vector genome might reflect residual promoter activity in the 55 bp of the U3 region remaining in the vector genome. However, additional deletions did not reduce the titer of mobilized vector particles

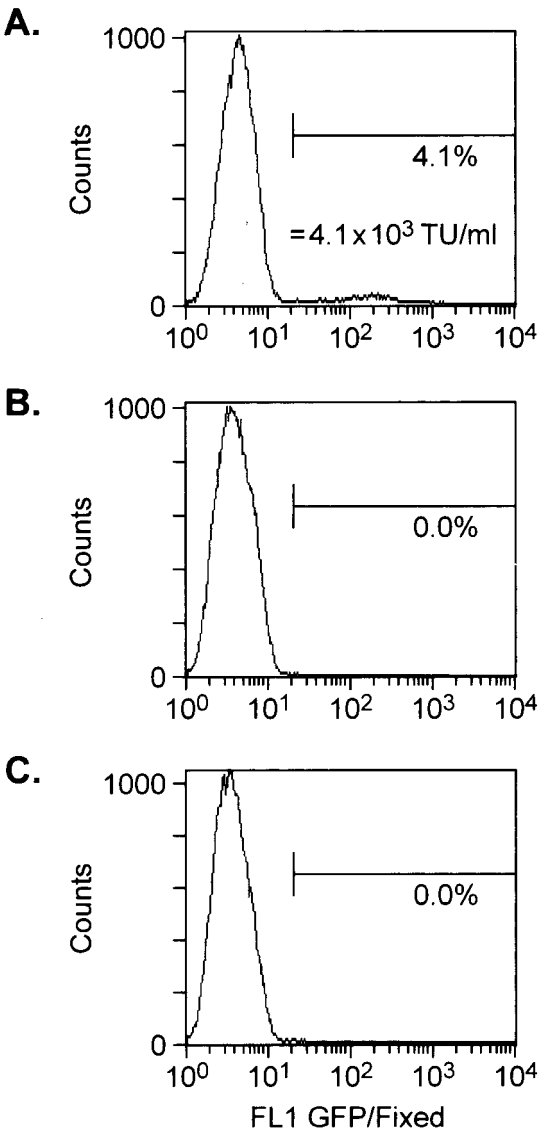

FIG. 1. Mobilization of a SIN lentiviral vector genome. (A.) Medium conditioned by 293T cells which had been multiply transduced with MpGFP vector particles at a high multiplicity of infection after transfection with the helper packaging plasmids. (B.) Medium conditioned by control cells which had not been transduced with the vector particles. (C.) Medium conditioned by transduced cells which were not transfected with the plasmids was titrated on HeLa cells. Vector mobilization was accessed by flow cytometry 5 days or 14 days after exposure of the HeLa cells to the conditioned medium.

(Fig. 2), whereas addition of the upstream efficiency element from the SV40 late polyadenylation region was associated with a more than twofold increase in the mobilized titer (3,700 ± 340 versus 9,600 ± 1,500 [mean ± standard error]; $P = 0.0001$). Preparations of each of these vectors had nearly equivalent titers ranging from 5×10^7 to 7.5×10^7.

Preliminary experiments suggested that the nature of the transcriptional regulatory sequences within the vector influenced the potential for mobilization. Substitution of the oncoretroviral vector LTR promoter (MSCV-U3) with the cellular EF1α promoter along with the first intron from that gene, which is retained in the vector genome during particle formation in the presence of Rev (data not shown), reduced the mobilized titer by more than 10-fold. To evaluate this phenomenon further, we derived a vector (mMpGFP) in which the U3 enhancer was deleted from the MSCV LTR, a vector in which the splice acceptor site at the 3′ end of the EF1 intron was mutated (EF1αSAmGFP), and a third vector in which the EF1α intron was eliminated (sEF1αGFP). The titers of preparations of the MpGFP, mMpGFP, and sEF1αGFP vectors were equivalent at 3×10^7, whereas the titers of the intron-containing vectors EF1αGFP and EF1αSAmGFP were lower, at 6.5×10^6 and 1×10^7, respectively. Two aliquots of 293T cells were transduced three times at an MOI of 50 with each of the five vectors depicted in Fig. 3. Following passage of the cells, DNA was extracted and vector copy number was estimated by real-time PCR with minor modifications from the original protocol (39). The copy numbers of the two replicates were very similar in each case, and overall the average copy number ranged from 80 to 112 for the five different vectors. Mobilized titers were estimated by characterizing conditioned medium from each of the 10 populations of cells by titration on HeLa cells for transfer of the GFP marker (Fig. 1 and 3). All mobilized titers were corrected for the copy number of the respective vector in the HeLa cell populations and expressed as transducing units/ml/100 copies. Passage of the transduced 293T cells for two additional weeks prior to transfection with the packaging plasmids did not result in a change in the mobilized titers (data not shown), confirming that the mobilized vector particles were derived from integrated proviral genomes.

As shown in Fig. 3, elimination of the enhancer from the LTR reduced the mobilized titer approximately twofold (1,170 ± 169 versus 670 ± 70; $P = 0.03$). Substitution of the EF1α promoter with intron for the MSCV LTR resulted in a ninefold reduction in apparent titer (1,170 ± 169 versus 148 ± 31; $P = 0.0003$). However, this effect seems predominantly due to the addition of the splice acceptor site into the vector, since a point mutation in the splice acceptor site restored the mobilized titer to that observed with the LTR-containing vector. Presumably, a significant proportion of transcripts from the unmodified EF1αGFP vector are being spliced during vector particle production using the downstream splice acceptor site of the EF1α intron, thereby eliminating the EF1α promoter and preventing detection and quantitation of mobilized vector particles in the GFP transfer assay. Indeed, substitution of the EF1α promoter without the intron for the LTR increased the mobilized titer approximately twofold (1,170 ± 169 versus 2,498 ± 105; $P = 0.0001$).

Recently, we developed a lentiviral vector system based on a nonpathogenic variant of SIV for use in transducing primitive hematopoietic cells from rhesus macaques (13). The SIV vector genome is efficiently packaged by proteins derived from HIV packaging plasmids (data not shown). This vector genome, after integration into 293T cells, was also mobilized when these cells were transduced with the packaging plasmids from the HIV-based vector system (data not shown). The titers of the HIV and SIV vectors were nearly equivalent at $3.3 \times 10^7 \pm 0.3 \times 10^7$ and $3.0 \times 10^7 \pm 0.2 \times 10^7$, respectively. The mobilized titers from multiply transduced 293T cells were 2.0×10^3 and 2.9×10^3, respectively. Duplicate mobilization assays were done for each vector.

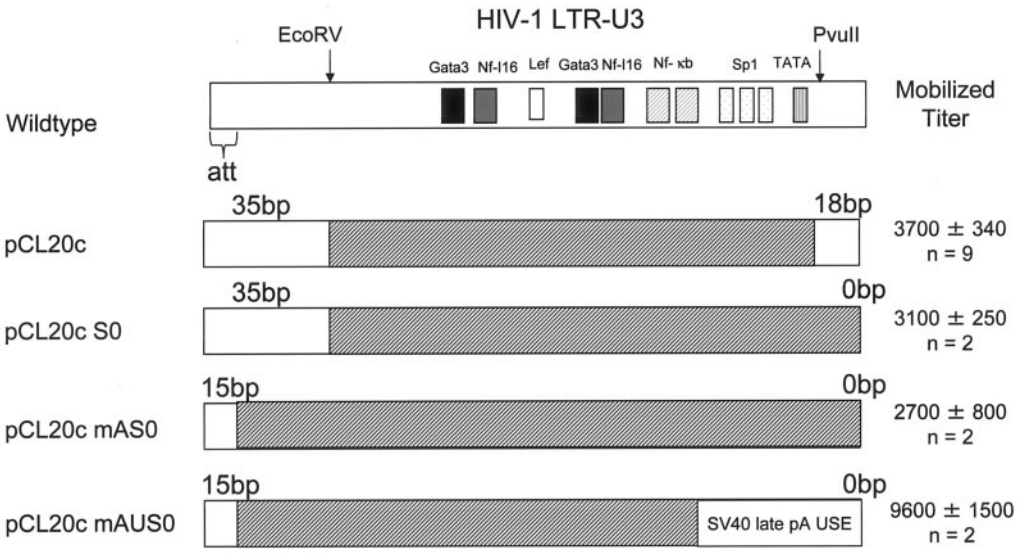

FIG. 2. Impact of modifications of the HIV-1 LTR-U3 region of the SIN vector on the apparent mobilized titer. Shown above is the diagram of the wild-type U3 region showing the protein binding sites that have been defined (5). Additional deletions of the residual sequences of the U3 region in the SIN vector were performed as described in Materials and Methods. A 75-bp fragment from the SV40 late polyadenylation region containing the upstream efficiency enhancer for polyadenylation was added to the lower vector as described in the text. *n* is the number of replicates for each determination.

Mobilized vector genomes are intact on reintegration. Several GFP-expressing clones were isolated from a 293T cell population that had been transduced with mobilized vector particles. Southern blot analysis demonstrated an intact proviral genome in each, since digestion with BbsI, which cuts once in each LTR, yielded the predicted 3.3-kb fragment (Fig. 4). Each clone was shown to have a single copy of the vector genome by Southern blot analysis using EcoRV, which cuts

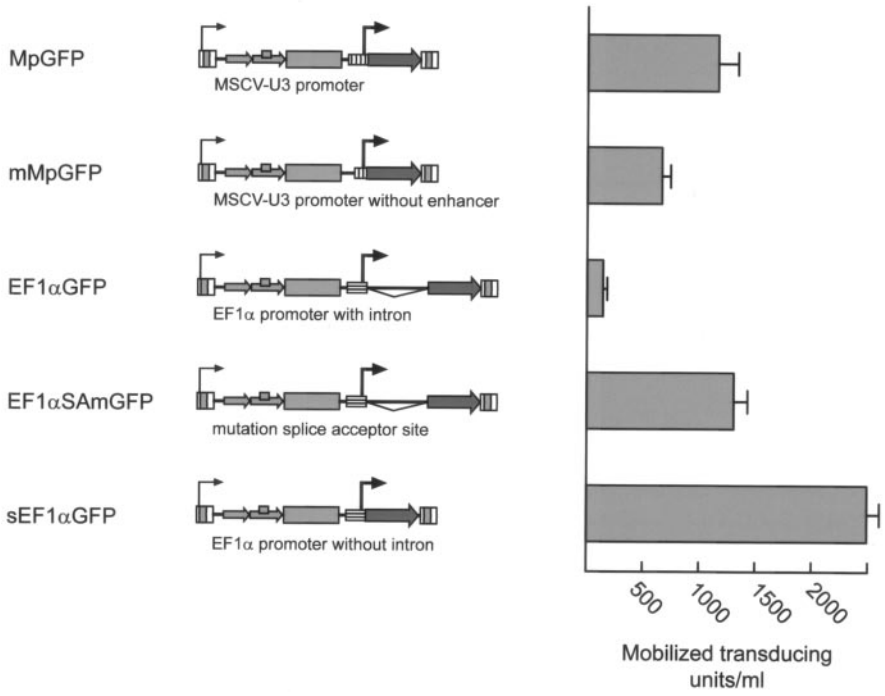

FIG. 3. Influence of vector design on mobilization of an integrated SIN lentiviral vector. Aliquots of 293T cells were transduced multiple times with each of the vectors individually at high multiplicities of infection to develop subpopulations containing 80 to 112 copies of the vector genome. After passage, these cell populations were transfected with the viral packaging plasmids and conditioned medium was titrated on HeLa cells. In each case, the mobilized titer was corrected for copy number in the 293T cells to yield the data displayed in the right panel. Duplicate populations of 293T cells were derived for each of the vectors, and the transfection experiment and evaluation of the mobilized titer were performed on each three separate times. Thus, the data displayed represent the mean and standard deviation of six independent determinations of the mobilized titer for each vector. The data are expressed as the titer of conditioned medium in transducing units/ml/100 copies of the vector genome.

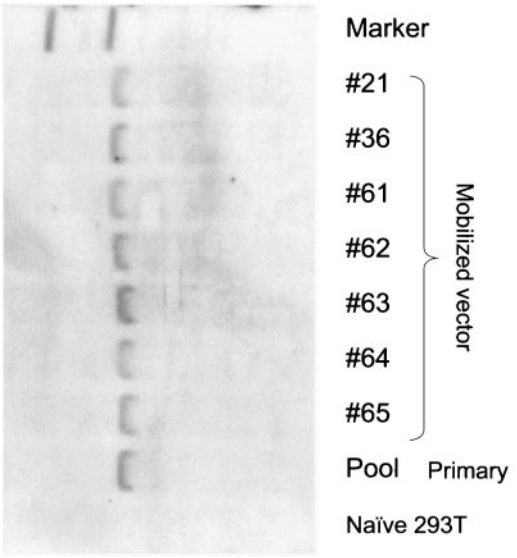

FIG. 4. Mobilized vector particles integrate an intact proviral genome in transduced cells. Southern blot analysis was performed on DNA samples from individual clones. The DNA samples were digested with BbsI, which cuts once in each proviral LTR, yielding a 3.3-kb fragment.

once in the vector genome (data not shown). Further characterization of the integrated provirus in each single-copy clone was achieved by PCR amplification of the 5′ LTR (Fig. 5). In each case, the expected 417-bp band was apparent on gel analysis, and sequencing of each PCR product confirmed that the LTR was intact. Characterization of the genomic vector junctions by LAM-PCR showed that the vector-derived sequences of junction fragments were intact in all cases (Table 1).

The clones derived with mobilized vector particles were capable of generating secondary, mobilized vector particles upon transfection with the helper plasmids as reflected by the generation of particles capable of transferring the GFP marker into naïve cells. The titers of media conditioned by seven subclones derived from primary vector particles were 102 ± 29 TU/ml (mean and standard error), and the titers of media conditioned by eight clones derived with mobilized vector particles averaged 78 ± 34 TU/ml ($P = 0.472$). There was considerable variation in the titer of individual clones, with the range for clones derived with primary vector particles being 10 to 230 TU/ml and that of the clones derived with mobilized vector particles being 10 to 150 TU/ml. Production of mobilized vector particles by these clones occurred regardless of whether the genome was inserted into an intergenic or intragenic position or, if within an intron, in the reverse or forward orientation (Table 1 and data not shown).

Vector transcription. To further evaluate the mechanism by which the SIN vector genome was transcribed following integration, the HIV tat coding sequences were inserted downstream from several splice acceptor sites within the vector genomes containing the MSCV-U3 or EF1α promoter with intron but upstream from the internal promoter (Fig. 6). The puromycin resistance gene (Pur^r) under the control of the wild-type HIV promoter was introduced into HeLa cells to create the HeLa Pur^r cell line which, in the absence of tat, remains sensitive to puromycin. Vector preparations for the genomes containing the tat coding sequences were generated by standard techniques along with relevant controls (Fig. 6). The titer of these vector preparations with respect to the trans-

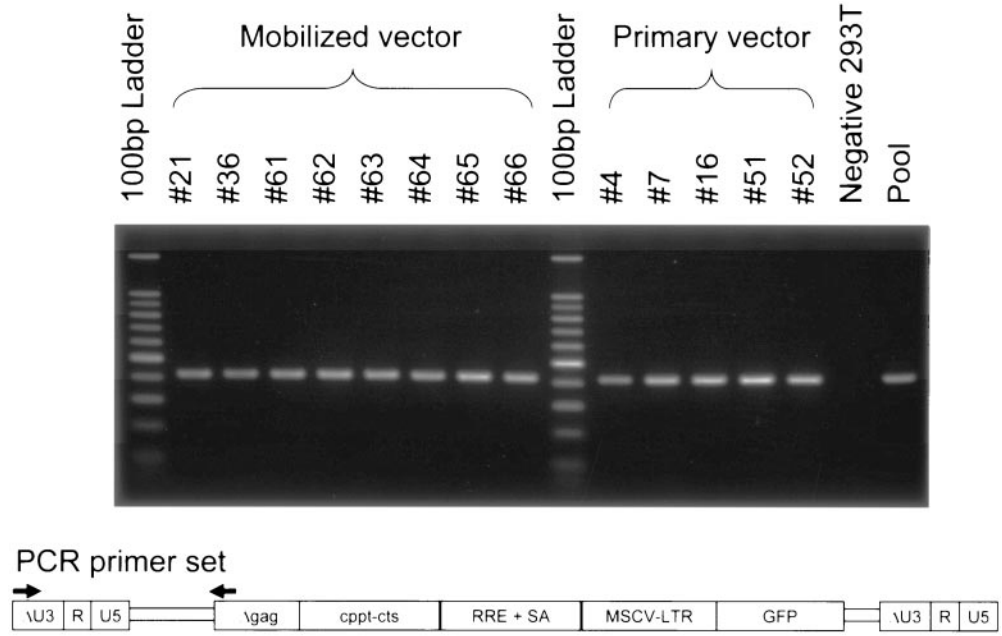

FIG. 5. The 5′ LTRs of integrated proviral genomes derived from mobilized vector particles are intact. PCR analysis was performed on DNA extracted from individual clones using primers positioned as shown on the diagram on the bottom of the figure. In each case, a 417-bp band predicted by the primer set was obtained. The control DNA was derived from a pool of 293T cells which had been transduced with vector particles at an MOI of 1.

TABLE 1. Characterization of integrated proviral genomes derived from mobilized or primary vector particles

Vector types and clone no.	Integrity				Integration site					
	5′ end of LTR at junction	LTR-leader sequence	LTR-LTR length at junction	3′ end of LTR at junction	Chromosome	Gene symbol	Gene ID	Distance[c]	Position	Orientation
Mobilized vector										
21	Intact	Intact	Intact	ND	20q13.33	DAFT1	11083	+15	Intron	Reverse
36	Intact	Intact	Intact	Intact	17q21.31	Intergenic[b]				
61	Intact	C643T[a]	Intact	Intact	22q11.21	DGCR8	54487	−7	Upstream	Forward
62	Intact	Intact	Intact	Intact	10q24	ADD3	120	+94	Intron	Reverse
63	Intact	Intact	Intact	Intact	17q12	LOC440434	440434	+44	Intron	Reverse
64	Intact	Intact	Intact	Intact	12q23.3	HCFC2	29915	+16	Intron	Reverse
65	Intact	Intact	Intact	Intact	22q12	KIAA1671	85379	+113	Intron	Forward
66	Intact	Intact	Intact	Intact	12q24.2	COX6A1	1337	−26	Upstream	Forward
Primary vector										
4	ND	C643T[a]	Intact	ND						
7	Intact	C643T[a]	Intact	ND	11q15	SWAP70	23075	+73	Intron	Forward
16	Intact	Intact	Intact	ND	19p13	PRM1	5619	−14	Upstream	Reverse
51	ND	Intact	Intact	ND						
52	ND	Intact	Intact	ND						
53	Intact	Intact	Intact	ND	7q22	RELN	5649	+378	Intron	Forward
54	ND	Intact	Intact	ND						
55	Intact	Intact	Intact	ND	6p21.1	POLH	5429	+8	Intron	Reverse

[a] Position corresponding to HXB2 (accession no. K03455).
[b] No RefSeq within 50 kb on either side of the integration site.
[c] Distance (kb) upstream (−) or downstream (+) from the transcriptional start site of a RefSeq gene.

fer of the GFP marker into HeLa cells (GFP titer) was compared to their titer with respect to their ability to generate Pur-resistant colonies after transduction of the HeLa Pur^r cell line (tat titer). A control vector in which the tat and GFP coding sequences were under the control of the wild-type HIV LTR gave equivalent GFP and tat titers, whereas the parent vectors lacking the tat coding sequences gave only rare Pur^r colonies (Fig. 6). The GFP titer for the vector containing the MSCV-U3 promoter was 7.2×10^6 with a tat titer of 2,600, whereas the GFP titer for the vector containing the EF1α

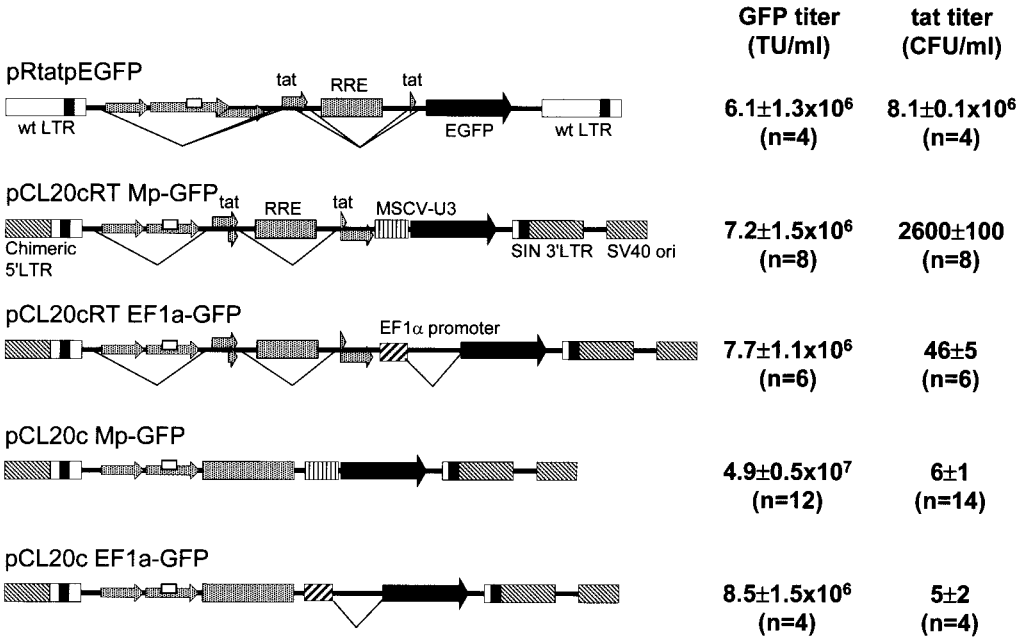

FIG. 6. Transcription of lentiviral self-inactivating genomes. Segments of the HIV genome encoding tat were introduced into several vectors as shown upstream from the internal promoter. Conventional titers of preparations of the individual vectors were determined by transfer of the GFP marker into naïve HeLa cells, whereas the tat titers of these preparations were determined by scoring puromycin-resistant colonies after transduction of a HeLa cell line in which the puromycin expression cassette was under the control of the wild-type HIV LTR. *n* is the number of independent replicators of each assay.

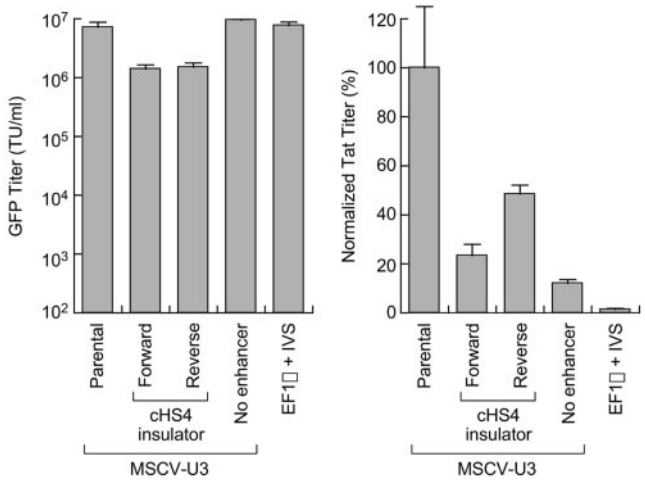

FIG. 7. Influence of vector design on genome transcription as reflected by the production of tat. The panel on the left shows the mean and standard deviation of the titers of preparations of each of the vectors ($n = 6$). The parental vector is MpGFP and the enhancer-deleted vector is mMpGFP, whereas the insulator vectors containing the chicken hypersensitive site 4 (cHS4) from the β-globin locus are designated INS1R MpGFP or INS1L MpGFP for the forward and reverse orientations, respectively. The vector containing the EF1α promoter plus intron is designated EF1αGFP. Shown on the right is the capacity of vector preparations of each to generate puromycin-resistant colonies when used to transduce the tat expression indicator cell line, HeLa-PUR. In each case the normalized tat titer reflects the number of puromycin-resistant colonies observed with serial dilutions of each vector preparation normalized, based on the GFP titer, to yield the data shown in the right panel. Six replicates were performed to derive the tat titers.

promoter (plus intron) was 7.7×10^7 with a tat titer of only 74 (Fig. 6). Thus, approximately 1 in 3,000 integration events of the vector containing the MSCV-U3 promoter was associated with transcription of the tat coding sequences. Modification of the vector genome by changing the promoter and adding the intron markedly reduced the proportion of vector genomes that generated functional tat mRNA (Fig. 6).

In a subsequent experiment three additional vectors were compared to the RT MpGFP and RT EF1αGFP vectors with respect to their relative GFP versus tat titers (Fig. 7). In two, the insulator element from the locus control region of the chicken β-globin locus (cHS4) was inserted either in the forward or reverse orientation into the 3′ LTR within the recombinant vector plasmid. Integration of this vector genome results in transfer of the cHS4 insulator to the 5′ LTR so that the vector transcriptional unit is flanked by insulator elements. In the forward orientation, the insulator resulted in a reduction in the tat titer relative to the GFP titer to 23% ± 4.6% compared to the control ($P < 0.05$). The insulator element, when in the reverse orientation, resulted in a reduction to 48% ± 3.6% relative to the control, which was not statistically significant ($P = 0.1$). Deletion of the enhancer element from the MSCV-U3 region did not decrease the GFP titer, but it did reduce the tat titer to 12% of the control ($P < 0.01$). The change from the MSCV-U3 promoter to the EF1α promoter with the intron reduced the tat titer to 1.4% ± 0.4% of the control ($P < 0.005$). Based on the results obtained with the various vectors in the mobilization assay (Fig. 3), we infer that all or a part of this reduction is likely to be due to introduction of an alternative downstream splice acceptor site into the vector along with the EF1α promoter.

Mechanism of transcription. Fundamental to the interpretation of these data is the identification of the transcriptional start sites for the transcripts that encode tat. To that end, we isolated seven Pur^r colonies derived from HeLa Pur^r cells transduced with vector particles encoding the tat sequences with no upstream promoter and GFP sequences under the control of the MSCV promoter. One colony lacked an integrated vector genome, whereas the other six colonies contained a single copy of the vector genome (Table 2). The junction of the vector genome and cellular DNA was recovered from each by LAM-PCR and sequenced. Of these six colonies, three contained the vector genome in a reverse or forward orientation downstream from the transcriptional start site within a RefSeq gene transcriptional unit and the other three contained the vector genome in an intergenic position (no RefSeq gene 50 kb upstream or downstream from the integration site).

The 5′ ends of transcripts encoding tat were defined using 5′-RACE. The results suggested that there were multiple transcriptional start sites for tat-encoding mRNAs in each of the clones rendered Pur^r following transduction (Fig. 8). Sequencing of the transcripts after recovery by plasmid subcloning defined the exact transcriptional start sites for several tat-encoding RNAs for each clone (Fig. 9). Elimination of the decapping reaction in the analysis of RNA from clone 2 as a control, to exclude the possibility that the tat-encoding RNAs arose by degradation of larger transcripts, resulted in disappearance of the specific bands and the emergence of a single smaller band which on cloning and sequencing was found to be a priming artifact (data not shown). From these data, we concluded that generation of tat-encoding RNA sequences most often occurred by virtue of the activation of cryptic promoter sites either within the vector or upstream rather than from

TABLE 2. Integration position of tat-expressing SIN vector genomes

Clone no.	Chromosome	Gene symbol	Gene ID	Distance[b]	Position	Orientation
2	8q24	Intergenic[a]				
3	8q24	Intergenic[a]				
4	8q24	Intergenic[a]				
5	19q13	PSG2	5670	+19	Downstream	Reverse
6	17q23	VMP1	81671	+91	Upstream	Forward
7	19q13	LOC126052	126052	+4.4	Downstream	Forward

[a] No RefSeq genes within 50 kb on either side of the integration site.
[b] Distance (in kb) upstream (−) or downstream (+) from the transcriptional start site of a RefSeq gene.

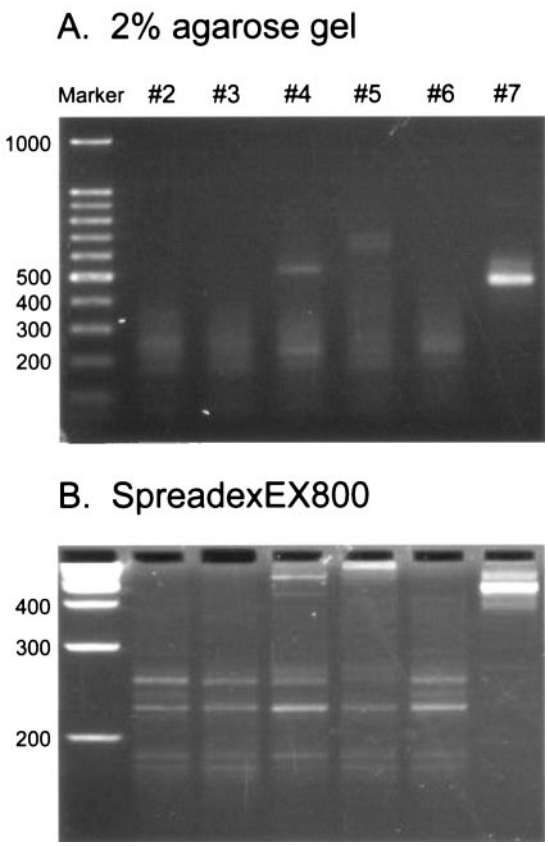

FIG. 8. Mapping of the 5′ ends of tat-encoding transcripts. (A.) Analysis on the 2% agarose gel detected longer transcripts in clones 4, 5, and 7 as well as shorter species in clones 2 through 6. (B.) The SpreadexEX800 gel further defined multiple shorter species in all clones.

readthrough transcription from the native promoter of a gene into which the vector genome had integrated.

DISCUSSION

Our results confirm that SIN lentiviral vector genomes can be mobilized at a readily detectable frequency by expression of viral proteins in cells in which the vector genome is integrated. Vector design significantly influenced mobilization frequency. The mobilized vector particles yielded an intact, unrearranged proviral genome upon reintegration into target cells. The mechanism of transcription of the integrated vector genome was evaluated using a promoter trap design with a vector encoding tat but lacking an upstream promoter in a cell line in which drug resistance depended on tat expression. The location of transcribed integrants in intergenic regions or in a reverse orientation within a gene suggested a transcriptional mechanism other than readthrough from an endogenous upstream promoter. Indeed, we found that in all cases, transcripts encoding tat arose from cryptic promoters either within or upstream of the integrated vector genome.

We demonstrated that vector particles containing a mobilized genome were capable of transferring an intact unrearranged proviral genome into naïve target cells. In all cases studied, the LTRs were fully intact, as determined by sequencing of PCR-amplified products, and Southern blot analysis demonstrated the genome to be intact. Clones of 293T cells containing a single copy of the proviral genome, whether derived from primary or mobilized vector particles, gave rise to vector particles de novo when transduced with helper plasmids. These data suggest that virtually every integrated proviral genome is transcribed, albeit often at low frequency, with considerable variation in the frequency of transcription depending on the integration position. However, we estimate that only approximately 1 in 3,000 integrated vector genomes containing the MSCV LTR was transcribed at a level sufficient to generate tat in amounts adequate to active the wild-type HIV LTR. This estimate is derived from the ratio of the tat titer of 2,600 divided by the GFP titer of 7.2×10^6 (Fig. 6). Thus, the activity of cryptic promoters must depend on local features of chromatin structure and the constellation of nearby regulatory elements and regulatory elements within the vector that facilitate transcription.

Based on the demonstration that relatively high-level transcription of tat-encoding proviral genomes occurs via cryptic promoters, we infer that most or all of the transcripts which result in vector mobilization also arise from cryptic promoters. Because the puromycin-resistant clones derived by virtue of tat expression also contain an integrated, wild-type LTR driving the Pur^r gene, it is not possible to evaluate mobilization of tat-encoding proviral vectors directly, since the genomes that include the wild-type LTR are likely to be far more efficiently mobilized. Furthermore, this genome also contains the GFP marker which, although expressed at a low level, would further confound efforts to evaluate mobilization of the tat-encoding genome by transfer of GFP expression. Any reverse transcript derived from a cryptic promoter that includes the R region of the 5′ LTR may yield an intact DNA proviral genome which is a substrate for reintegration, since first-strand transfer may occur when all, or a portion of the R region, has been transcribed (46). We have shown that the titer of mobilized vector particles from individual clones containing the GFP-encoding proviral genome ranges from 10 to 230 IU/ml, indicating that the relative likelihood of mobilization of any integrant is highly variable. SIN LTRs are expressed at about 15% of the level of a wild-type LTR in the absence of tat (20). This difference is likely to be greater when tat is expressed, since the wild-type but not the SIN LTR is activated by tat. Although a vector containing a wild-type LTR can be mobilized from cells by HIV infection (4), mobilization of a vector with a SIN LTR, although feasible as demonstrated in our experiments, is likely to be less efficient.

The existence of cryptic promoters in eukaryotic DNA is well described. For example, upstream transcription initiation sites for globin genes, which account for a small fraction of globin-encoding RNAs, have been defined within a few hundred base pairs of the major CAP sites (2, 8, 25). In addition, there are long RNA species that may span segments of the locus which appeared to be derived from specific promoter structures (7). Cryptic promoters within the 5′ untranslated region of cellular genes, e.g., the gene for the translation initiation factor eIF4G, may account for only a small portion of the eIF4G-encoding mRNAs but may be the dominant translated species because of the extensive secondary structure in

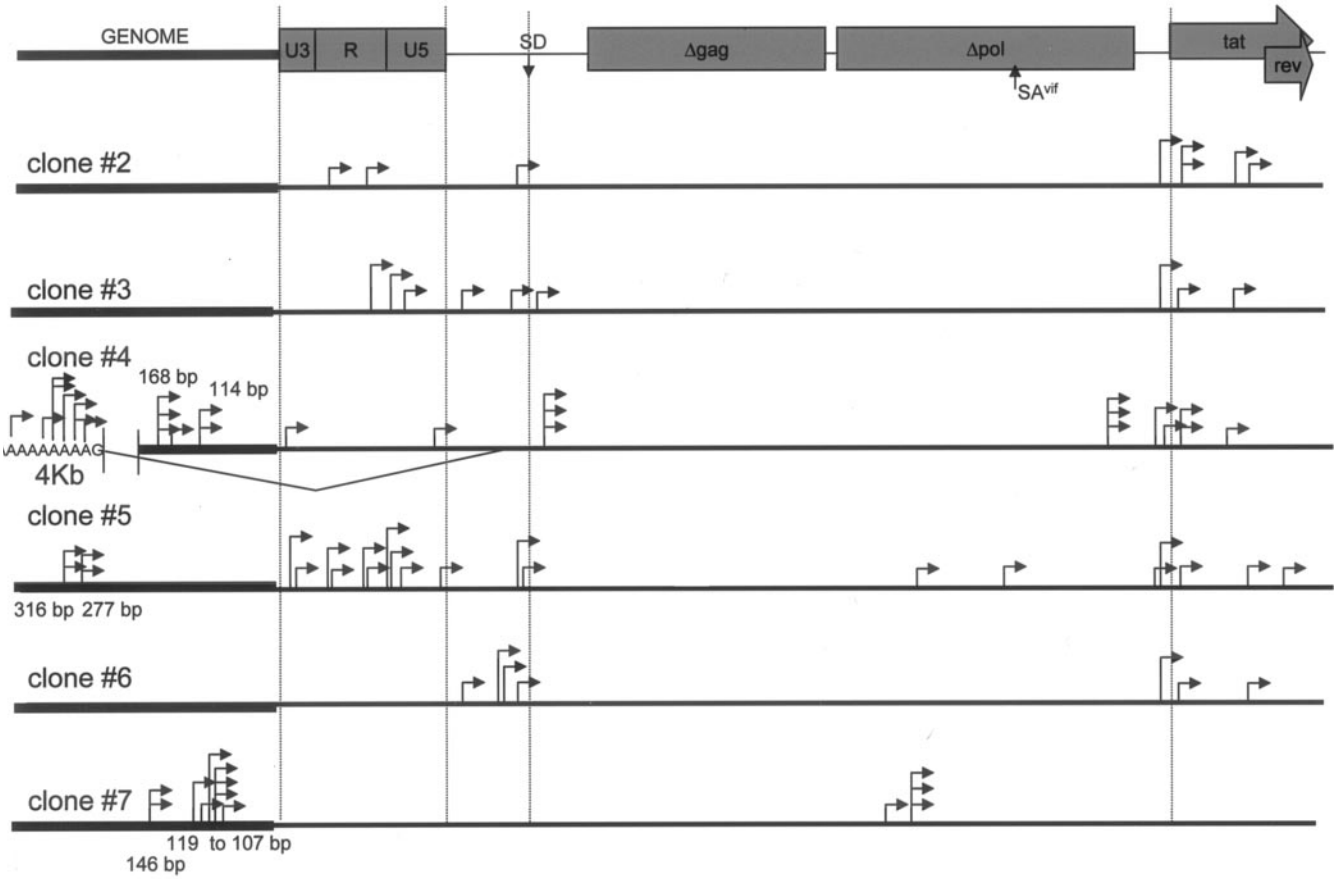

FIG. 9. Localization of the 5′ start sites for tat-encoding transcripts. Multiple start sites were identified for each clone by shotgun cloning of the 5′-RACE products followed by sequencing of the junctions. The sizes shown for clones 4, 5, and 7 are the distances from the transcriptional start site to the 5′ end of the LTR.

the 5′ untranslated region of the most abundant mRNA species which inhibits translation beginning at its CAP structure (12). Cryptic promoters apparently reflect coincident location of binding sites for one or more transcription factors which attract the RNA polymerase II (pol II) complex with a frequency sufficient to generate transcripts of downstream sequences (12). The ability of cryptic promoters to generate functional transcripts adds an element of caution to promoter trap experiments (3, 24, 29) and suggests that inferences regarding promoter inducibility should be validated by defining the 5′ ends of the induced transcripts.

DNA methylation and histone acetylation are two features of chromatin structure that are likely to influence the probability of cryptic promoter function (18, 23). Vector genomes integrated into or near methylated CpG islands may be less likely to be expressed than those which are in regions of undermethylated DNA. The process of transcription may expose cryptic promoter sites and allow their activation. For example, a mutation in the transcription elongation factor sPt6 in yeast results in altered chromatin structure of transcribed genes, permitting aberrant cryptic promoter function in coding regions (21). sPt6 is thought to participate in the restoration of chromatin structure following gene transcription. In the globin locus, intergenic transcription is thought to remodel chromatin to permit transcription from the individual globin gene promoters (7). In addition, enhancer and locus control region elements near the integrated genome may influence cryptic promoter activity. In summary, those vector genomes which integrate within or near transcribed genes are more likely to be expressed through activation of cryptic promoters.

Our data indicate that vector genome mobilization remains a risk despite the use of SIN vectors. Specific modifications, e.g., addition of insulator elements to the LTRs, appear to reduce the probability of transcription. Splicing events affecting the vector genome transcript may influence the ability to detect the mobilized vector, as revealed by our studies of the EF1α promoter with or without the downstream intron, and should be considered when comparing two vector designs with respect to the potential for mobilization. The nature of the regulatory elements in the vector and the cellular environment may also influence cryptic promoter activation. For example, globin locus control regions may enhance vector transcription from cryptic promoter sites in erythroid cells but not in lymphoid cells. Experiments are in progress in our laboratory to test this hypothesis.

We found that HIV proteins efficiently package vector genomes based on SIV (data not shown), leading to the prediction, supported by our data, that SIV genomes could be mo-

bilized by transfection of plasmids encoding HIV proteins. An alternative strategy explored by others (9), namely, the use of an artificial tRNA binding site to initiate reverse transcription of the vector genome when complemented by a modified tRNA during vector particle production, may diminish the probability of secondary transduction of host cells lacking the complementary tRNA by a mobilized vector.

Our results are consistent with the recently published studies by Logan et al. (26) in which integrated self-inactivating lentiviral vectors were shown to produce full-length genomic transcripts competent for encapsidation and integration. Their work focused on the identification of sequences in the SIN lentiviral vector which are responsible for transcriptional activation. Primers positioned within the encoded transgene and a second set that amplified the 5′ LTR confirmed that a significant proportion of the transcripts, perhaps the majority, extended to the R region of the LTR, but the actual transcriptional start sites were not mapped. The binding sites for two transcriptional activators, DBF1 and SP1, within the leader region of the proviral genome were identified as influencing the level of proviral gene transcription. Transcripts beginning at cryptic promoters, such as those we demonstrated, could indeed be packaged and give rise to an intact proviral genome upon transduction of target cells provided that the R region is included in the transcript. Undoubtedly, the transcriptional factor binding sites identified in the studies by Logan et al. (26) could influence the frequency of upstream transcription from cryptic promoters. Although we agree that most integrated genomes are transcribed at variable frequency, our work indicates that only rare integrants are transcribed with sufficient frequency to generate tat in quantities adequate to activate the wild-type HIV LTR promoter.

Our data indicate that careful attention to vector design and screening of various vectors in assays designed to detect vector transcription may reduce the risk of vector mobilization. In parallel with these studies, we are evaluating the effect of vector integration on expression of nearby genes using microarray and real-time PCR analysis. The predilection of oncoretroviral vectors to integrate near the promoter region has not been observed in primary hematopoietic stem cells with an SIV-based lentiviral vector system (17). Rather, SIV vector integrants are distributed throughout genes, potentially reducing the risk for promoter activation. Our current studies are focused on looking for evidence of gene activation by lentiviral globin gene vectors in erythroid cells (15). Such studies, when combined with the evaluation of vectors in tumor-prone mouse models (20), should give us a better appreciation of the risk of stem cell-targeted gene transfer and help determine whether this promising approach to the correction of gene defects is sufficiently safe to allow widespread clinical use.

ACKNOWLEDGMENTS

This work was supported by NHBLI program project grant P01 HL 53749, Cancer Center Support CORE grant P30 CA 21765, and American Lebanese Syrian Associated Charities.

We are grateful to Jean Johnson and Pat Streich for their help in the preparation of the manuscript. We also thank Richard Ashmun and the staff of the Flow Cytometry Lab for sharing their expertise.

REFERENCES

1. **Ailles, L. E., and L. Naldini.** 2002. HIV-1-derived lentiviral vectors. Curr. Top. Microbiol. Immunol. **261:**31–52.
2. **Allan, M., W. G. Lanyon, and J. Paul.** 1983. Multiple origins of transcription in the 4.5 kb upstream of the epsilon-globin gene. Cell **35:**187–197.
3. **Chen, W. V., J. Delrow, P. D. Corrin, J. P. Frazier, and P. Soriano.** 2004. Identification and validation of PDGF transcriptional targets by microarray-coupled gene-trap mutagenesis. Nat. Genet. **36:**304–312.
4. **Evans, J. T., and J. V. Garcia.** 2000. Lentivirus vector mobilization and spread by human immunodeficiency virus. Hum. Gene Ther. **11:**2331–2339.
5. **Flint, S. J., L. W. Enquist, R. M. Krug, V. R. Racaniello, and A. M. Skalka (ed.)** 2000. Principles of virology: molecular biology, pathogenesis, and control. ASM Press, Washington, D.C.
6. **Follenzi, A., M. Battaglia, A. Lombardo, A. Annoni, M. G. Roncarolo, and L. Naldini.** 2004. Targeting lentiviral vector expression to hepatocytes limits transgene-specific immune response and establishes long-term expression of human antihemophilic factor IX in mice. Blood **103:**3700–3709.
7. **Gribnau, J., K. Diderich, S. Pruzina, R. Calzolari, and P. Fraser.** 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. Mol. Cell **5:**377–386.
8. **Grindlay, G. J., W. G. Lanyon, M. Allan, and J. Paul.** 1984. Alternative sites of transcription initiation upstream of the canonical cap site in human gamma-globin and beta-globin genes. Nucleic Acids Res. **12:**1811–1820.
9. **Grunwald, T., F. S. Pedersen, R. Wagner, and K. Uberla.** 2004. Reducing mobilization of simian immunodeficiency virus based vectors by primer complementation. J. Gene Med. **6:**147–154.
10. **Haas, D. L., S. S. Case, G. M. Crooks, and D. B. Kohn.** 2000. Critical factors influencing stable transduction of human $CD34^+$ cells with HIV-1-derived lentiviral vectors. Mol. Ther. **2:**71–80.
11. **Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M. P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C. S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J. I. Cohen, G. de Saint Basile, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. Macintyre, F. Sigaux, J. Soulier, L. E. Leiva, M. Wissler, C. Prinz, T. H. Rabbitts, F. Le Deist, A. Fischer, and M. Cavazzana-Calvo.** 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy SCID-X1. Science **302:**415–419.
12. **Han, B., and J. T. Zhang.** 2002. Regulation of gene expression by internal ribosome entry sites or cryptic promoters: the eIF4G story. Mol. Cell. Biol. **22:**7372–7384.
13. **Hanawa, H., P. Hemmatti, K. Keyvanfar, M. E. Metzger, A. Krouse, R. E. Donahue, S. Kepes, J. Gray, C. E. Dunbar, D. A. Persons, and A. W. Nienhuis.** 2004. Efficient gene transfer into rhesus repopulating hematopoietic stem cells using a simian immunodeficiency virus-based lentiviral vector system. Blood **103:**4062–4069.
14. **Hanawa, H., P. F. Kelly, A. C. Nathwani, D. A. Persons, J. A. Vandergriff, P. Hargrove, E. F. Vanin, and A. W. Nienhuis.** 2002. Comparison of various envelope proteins for their ability to pseudotype lentiviral vectors and transduce primitive hematopoietic cells from human blood. Mol. Ther. **5:**242–251.
15. **Hargrove, P. W., H. Hanawa, S. Kepes, C. Cheng, G. Neale, A. W. Nienhuis, and D. A. Persons.** 2004. Assessment of changes in gene expression caused by insertions of a globin lentiviral vector containing globin regulatory elements or a lentiviral vector containing retroviral LTR elements. Blood. **104**(Suppl. I)**:**145a.
16. **Hawley, R. G., F. H. Lieu, A. Z. Fong, and T. S. Hawley.** 1994. Versatile retroviral vectors for potential use in gene therapy. Gene Ther. **1:**136–138.
17. **Hematti, P., B. K. Hong, C. Ferguson, R. Adler, H. Hanawa, S. Sellers, I. E. Holt, C. E. Eckfeldt, Y. Sharma, M. Schmidt, C. von Kalle, D. A. Persons, E. M. Billings, C. M. Verfaillie, A. W. Nienhuis, T. G. Wolfsberg, C. E. Dunbar, and B. Calmels.** 2004. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS Biol. **2:**2183–2190.
18. **Herman, J. G., and S. B. Baylin.** 2003. Gene silencing in cancer in association with promoter hypermethylation. N. Engl. J. Med. **349:**2042–2054.
19. **Kafri, T., U. Blomer, D. A. Peterson, F. H. Gage, and I. M. Verma.** 1997. Sustained expression of genes delivered directly into liver and muscle by lentiviral vectors. Nat. Genet. **17:**314–317.
20. **Kamijo, T., S. Bodner, E. van de Kamp, D. H. Randle, and C. J. Sherr.** 1999. Tumor spectrum in ARF-deficient mice. Cancer Res. **59:**2217–2222.
21. **Kaplan, C. D., L. Laprade, and F. Winston.** 2003. Transcription elongation factors repress transcription initiation from cryptic sites. Science **301:**1096–1099.
22. **Kappes, J. C., X. Wu, and J. K. Wakefield.** 2003. Production of transplentiviral vector with predictable safety. Methods Mol. Med. **76:**449–465.
23. **Khorasanizadeh, S.** 2004. The nucleosome: from genomic organization to genomic regulation. Cell **116:**259–272.
24. **Lai, Z., I. Han, M. Park, and R. O. Brady.** 2002. Design of an HIV-1 lentiviral-based gene-trap vector to detect developmentally regulated genes in mammalian cells. Proc. Natl. Acad. Sci. USA **99:**3651–3656.
25. **Ley, T. J., and A. W. Nienhuis.** 1983. A weak upstream promoter gives rise to long human beta-globin RNA molecules. Biochem. Biophys. Res. Commun. **112:**1041–1048.
26. **Logan, A. C., D. L. Haas, T. Kafri, and D. B. Kohn.** 2004. Integrated

self-inactivating lentiviral vectors produce full-length genomic transcripts competent for encapsidation and integration. J. Virol. **78:**8421–8436.
27. **Logan, A. C., C. Lutzko, and D. B. Kohn.** 2002. Advances in lentiviral vector designs for gene-modification of hematopoietic stem cells. Curr. Opin. Biotechnol. **13:**429–436.
28. **Manley, N. R., M. O'Connell, W. Sun, N. A. Speck, and N. Hopkins.** 1993. Two factors that bind to highly conserved sequences in mammalian type C retroviral enhancers. J. Virol. **67:**1967–1975.
29. **Medico, E., G. Gambarotta, A. Gentile, P. M. Comoglio, and P. Soriano.** 2001. A gene trap vector system for identifying transcriptionally responsive genes. Nat. Biotechnol. **19:**579–582.
30. **Miyoshi, H., U. Blomer, M. Takahashi, F. H. Gage, and I. M. Verma.** 1998. Development of a self-inactivating lentivirus vector. J. Virol. **72:**8150–8157.
31. **Mizushima, S., and S. Nagata.** 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. **18:**5322.
32. **Naldini, L., U. Blomer, F. H. Gage, D. Trono, and I. M. Verma.** 1996. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc. Natl. Acad. Sci. USA **93:**11382–11388.
33. **Naldini, L., U. Blomer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, and D. Trono.** 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science **272:**263–267.
34. **Olsen, J. C.** 2001. EIAV, CAEV and other lentivirus vector systems. Somat. Cell Mol. Genet. **26:**131–145.
35. **O'Rourke, J. P., G. C. Newbound, D. B. Kohn, J. C. Olsen, and B. A. Bunnell.** 2002. Comparison of gene transfer efficiencies and gene expression levels achieve with equine anemia virus- and human immunodeficiency virus type-1-derived lentivirus vectors. J. Virol. **76:**1510–1515.
36. **Poeschla, E. M.** 2003. Non-primate lentiviral vectors. Curr. Opin. Mol. Ther. **5:**529–540.
37. **Recillas-Targa, F., M. J. Pikkart, B. Burgess-Beusse, A. C. Bell, M. D. Litt, A. G. West, M. Gaszner, and G. Felsenfeld.** 2002. Position-effect protection and enhancer blocking by the chicken beta-globin are separable activities. Proc. Natl. Acad. Sci. USA **99:**6883–6888.
38. **Ruggieri, A., D. Negre, and F. L. Cosset.** 2003. SIV vectors. Methods Mol. Biol. **229:**233–249.
39. **Sastry, L., T. Johnson, M. J. Hobson, B. Smucker, and K. Cornetta.** 2002. Tittering lentiviral vectors: comparison of DNA, RNA and marker expression methods. Gene Ther. **9:**1155–1162.
40. **Sastry, L., Y. Xu, T. Johnson, K. Desai, D. Rissing, J. Marsh, and K. Cornetta.** 2003. Certification assays for HIV-1-based vectors: frequent passage of gag sequences without evidence of replication-competent viruses. Mol. Ther. **8:**830–839.
41. **Schek, N., C. Cooke, and J. C. Alwine.** 1992. Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. Mol. Cell. Biol. **12:**5386–5393.
42. **Schnell, T., P. Foley, M. Wirth, J. Munch, and K. Uberla.** 2000. Development of a self-inactivating, minimal lentivirus vector based on simian immunodeficiency virus. Hum. Gene Ther. **11:**439–447.
43. **Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110:**521–529.
44. **Sorrentino, B. P., and A. W. Nienhuis.** 2001. Gene therapy for hematopoietic diseases, p. 969–1003. *In* G. Stamatoyannopoulos, P. W. Majerus, R. M. Perlmutter, H. Varmus (ed.), The molecular basis of blood diseases, 3rd ed. W.B. Saunders Company, Philadelphia, Pa.
45. **Takahashi, M.** 2004. Delivery of genes to the eye using lentiviral vectors. Methods Mol. Biol. **246:**439–449.
46. **Telesnitsky, A., and S. P. Goff.** 1997. Reverse transcription and the generation of retroviral RNA, p. 121–160. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Press, Woodbury, N.Y.
47. **Uetsuki, T., A. Naito, S. Nagata, and Y. Kaziro.** 1989. Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1 alpha. J. Biol. Chem. **264:**5791–5798.
48. **Westerman, K. A., T. Morales, R. Pawliuk, E. Cohen, and P. LeBoulch.** 2003. A high titer, "supersplit" packaging system in combination with a self-inactivating (SIN) vector for the prevention of replication competent lentivirus (RCL) contamination. Mol. Ther. **7:**S3.
49. **Wong, L. F., M. Azzouz, L. E. Walmsley, Z. Askham, F. J. Wilkes, K. A. Mitrophanous, S. M. Kingsman, and N. D. Mazarakis.** 2004. Transduction patterns of pseudotyped lentiviral vectors in the nervous system. Mol. Ther. **9:**101–111.
50. **Wu, X., Y. Li, B. Crise, and S. M. Burgess.** 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science **13:**1749–1751.
51. **Wu, X., J. K. Wakefield, H. Liu, H. Xiao, R. Kralovics, J. T. Prchal, and J. C. Kappes.** 2000. Development of a novel trans-lentiviral vector that affords predictable safety. Mol. Ther. **2:**47–55.
52. **Xu, K., H. Ma, T. J. McCown, I. M. Verma, and T. Kafri.** 2001. Generation of a stable cell line producing high-titer self-inactivating lentiviral vectors. Mol. Ther. **3:**97–104.
53. **Zufferey, R., T. Dull, R. J. Mandel, A. Bukovsky, D. Quiroz, L. Naldini, and D. Trono.** 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J. Virol. **72:**9873–9880.