JOURNAL OF VIROLOGY, Dec. 1997, p. 9466–9474 Vol. 71, No. 12
0022-538X/97/$04.00+0

Increased Probability of Expression from Modified Retroviral Vectors in Embryonal Stem Cells and Embryonal Carcinoma Cells

PAUL B. ROBBINS, XIAO-JIN YU, DIANNE M. SKELTON, KAREN A. PEPPER, RICHARD M. WASSERMAN, LI ZHU, AND DONALD B. KOHN*

Department of Molecular Microbiology and Immunology, University of Southern California School of Medicine, and Division of Research Immunology/Bone Marrow Transplantation, Childrens Hospital Los Angeles, Los Angeles, California

Received 9 April 1997/Accepted 4 September 1997

Gene expression from the Moloney murine leukemia retrovirus (Mo-MuLV) is highly restricted in embryonic carcinoma (EC) and embryonic stem (ES) cells. We compared levels of expression in PA317 fibroblasts, F9 (EC) cells, and CCE (ES) cells by Mo-MuLV-based vectors and vectors based on our previously reported MND backbone, which has alterations to address three viral elements implicated as repressors of expression by Mo-MuLV: the enhancer, the primer binding site, and the negative-control region. Expression was evaluated with three reporter genes, the chloramphenicol acetyltransferase (CAT) gene, whose expression was measured by enzymatic assay and by Northern blotting; a truncated nerve growth factor receptor (tNGFR), whose expression was measured by fluorescence-activated cell sorting (FACS) as a cell surface protein; and the enhanced green fluorescent protein (EGFP), whose expression was measured intracellularly by flow cytometry. We found significantly higher levels of CAT activity (5- to 300-fold) and greater quantities of vector-specific transcripts in ES and EC cells transduced with the modified MND-CAT-SN vector than in those transduced with L-CAT-SN. Northern blot analysis indicated that long terminal repeat transcripts from MND-CAT-SN are >80 times more abundant than the L-CAT-SN transcripts. FACS analysis of tNGFR expression from a pair of vectors, L-tNGFR-SN and MND-tNGFR-SN, indicated that only 1.04% of the CCE cells containing the L-tNGFR-SN vector expressed the cell surface reporter, while the MND-tNGFR-SN vector drove expression in 99.54% of the CCE cells. Of the F9 cells containing the L-tNGFR-SN vector, 13.32% expressed tNGFR, while 99.89% of the F9 cells transduced with MND-tNGFR-SN showed expression. Essentially identical results were produced with an analogous pair of vectors encoding EGFP. In unselected pools of F9 cells 48 h posttransduction, the L-EGFP-SN vector drove expression in only 5% of the population while the MND-EGFP-SN vector drove expression in 88% of the cells. After more than 3 weeks in culture without selection, the proportion of cells showing expression from L-EGFP-SN decreased slightly to 3% while expression from the MND-EGFP-SN vector persisted in 80% of the cells. Interestingly, in the few ES and EC cells which did show expression from the L-tNGFR-SN or L-EGFP-SN vectors, the magnitude of reporter expression was similar to that from the MND-tNGFR-SN or MND-EGFP-SN vector in nearly all cells, suggesting that the MND vectors are far less susceptible to position-dependent variegation of expression than are the Mo-MuLV-based vectors. Therefore, the modified retroviral vector, MND, achieves higher net levels of expression due to a greater frequency of expression, which may be useful for the expression of exogenous genes in EC and ES cells.

Moloney murine leukemia virus (Mo-MuLV)-based vectors are the predominant gene transfer vehicles used for stable transduction of mammalian cells in vitro, despite some recognized disadvantages (6a, 33). Gene expression from the Mo-MuLV long terminal repeat (LTR) has been disappointing in murine embryonic stem (ES), embryonic carcinoma (EC), and hematopoietic cells (8, 29) as well as in human hematopoietic cells (1).

Persistent high-level expression from retroviral vectors in ES and EC cells could be an extremely useful tool for the evaluation of the effects of exogenous genes on mammalian cell development. For example, ES cells can be induced to undergo differentiation into cells of the hematopoietic lineages; efficient retrovirus-mediated gene transfer and expression in ES cells would provide a model system with which to evaluate the effects of genes which regulate hematopoietic cell differentiation. Reliable expression of genes introduced into hematopoietic stem cells (HSC) is a central requisite for effective clinical gene therapy of a variety of genetic disorders (24). Therefore, understanding the mechanisms which control transcription from the LTRs of retroviral vectors may allow the development of modified vectors which achieve more effective expression in stem cells.

The enhancer/promoter of the Mo-MuLV has been shown to be inactive in EC and ES cells (4, 18, 20, 27). A related virus, the myeloproliferative sarcoma virus (MPSV) (25), arose as a spontaneous mutant of the Moloney murine sarcoma virus (MSV), which caused myeloproliferative disorders instead of sarcomas. Several groups have demonstrated that MPSV has greater transcriptional activity than Mo-MuLV in EC cells (20, 21, 50) and in hematopoietic cells (4, 39, 46). The LTR of MPSV has several base-pair changes compared to the LTR sequences of Mo-MuLV and MSV, with one single-base-pair alteration in the enhancer of MPSV creating a consensus binding site for the Sp1 transcription factor; this new Sp1 site was shown to be responsible for the altered pathogenic pattern of MPSV and increased transcriptional activity in EC cells (41).

By constructing and selecting for recombinant retroviruses which had enhanced infectivity and expressivity in EC cells, Barklis et al. (2), Loh et al. (28), and Weiher et al. (50) determined that sequences contained within the 5′-untrans-

* Corresponding author. Mailing address: 4650 Sunset Blvd., Mail stop no. 62, Los Angeles, CA 90027. Phone: (213) 669-4617. Fax: (213) 667-1021. E-mail: dkohn%smtpgate@chlais.usc.edu.

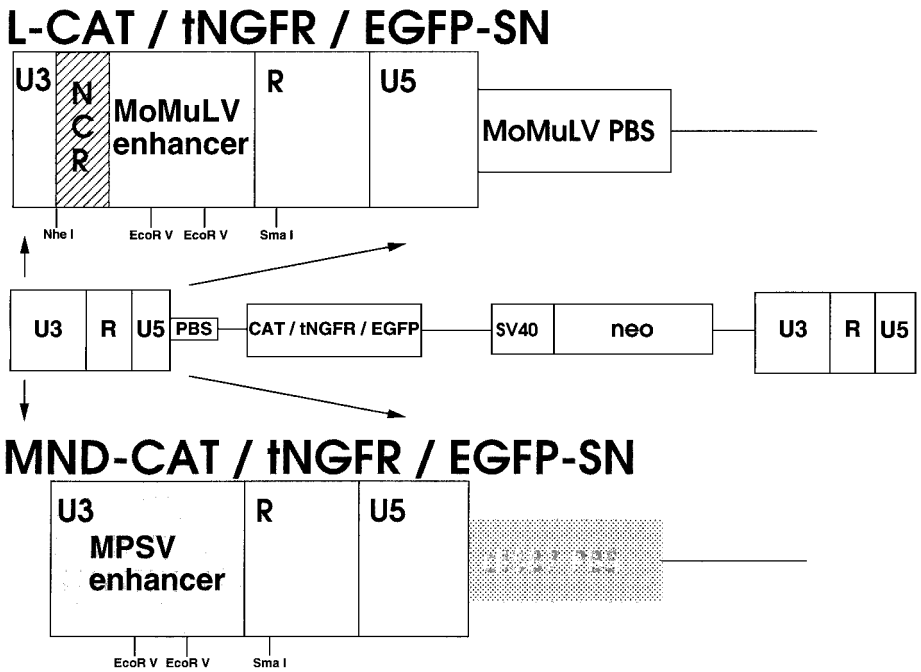

FIG. 1. Diagram of the retroviral vectors expressing either the CAT, tNGFR, or EGFP genes from the LTR. The relevant differences between the constructs are highlighted in the expanded views of the 5′ LTRs.

lated region immediately downstream of the 5′ LTR in or near the primer binding site (PBS) also inhibited LTR-directed expression. Barklis et al. (2) found that a single point mutation in the PBS of the B2 proviral recombinant was sufficient to permit expression from a Mo-MuLV template in EC cells. Subsequent research efforts identified a *trans*-acting factor termed the repressor binding protein which bound to the Mo-MuLV PBS and interfered with transcription from the LTR (14, 23, 40). An endogenous murine retrovirus, *dl*587rev, was shown to escape the PBS-associated repression by utilizing a variant PBS which has multiple base pairs divergent from Mo-MuLV that eliminate the PBS-associated repressor-binding site and result in the utilization of a $tRNA^{Glu}$ instead of a $tRNA^{Pro}$ for the priming of reverse transcription (11, 40, 47).

Transcriptional repression of Mo-MuLV in ES and EC cells is mediated, in part, by the binding of the transcription factor YY1 (also known as UCRBP or NCRBP) to a highly conserved sequence (5′-TAACGCCATTT-3′) in the 5′ LTR termed the negative-control region (NCR) (3, 15, 16, 48).

Repression of expression of endogenous and exogenous gene sequences in mammalian cells is frequently associated with methylation of cytosine residues in DNA (5, 7). Although it is not clear whether cytosine methylation is a primary cause of transcriptional silencing or merely an associated result, the correlation between the two events suggests an active role for methylation in the down regulation or maintenance of transcriptional repression of many genes. Mo-MuLVs are prone to extensive de novo methylation in EC cells (38, 45), as well as in murine fibroblast cells (22). Challita and Kohn (8) and, recently, Lengauer et al. (26) demonstrated a strong correlation between methylation of the Mo-MuLV LTR and the lack of LTR-directed transcription in murine HSC and human colorectal cancer cells, respectively.

We previously described a novel retroviral vector, Mp-ncr-dl-neo (herein called MND-neo) which contains modifications of the LTR enhancer, the NCR, and the PBS (9). MND-neo was shown to have greatly increased expression in a titer assay with F9 EC cells compared to a parallel vector containing the standard elements from Mo-MuLV. To further examine the triply modified vector MND (MPSV LTR, NCR deleted, *dl*587 PBS), we constructed the MND-X-SN backbone, which allowed convenient insertion of different reporter genes. We analyzed chloramphenicol acetyltransferase (CAT) enzyme activity, truncated nerve growth factor receptor (tNGFR) surface reporter expression, and enhanced green fluorescent protein (EGFP) expression from the Mo-MuLV and MND vector backbones in ES, EC, and fibroblast cells.

MATERIALS AND METHODS

Cell lines. PA317 and F9 cell lines were obtained from the American Type Culture Collection. The PA317 amphotropic packaging cells (35), derived from NIH 3T3 fibroblasts, were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS), penicillin-streptomycin (50 U/ml), and L-glutamine (2 mM). The F9 EC cell line was maintained on gelatin-coated tissue culture plates in DMEM supplemented with 15% FCS, penicillin-streptomycin (50 U/ml), L-glutamine (2 mM), and 0.1 mM (final concentration) β-mercaptoethanol. The CCE ES cell line (13) was generously provided by Jeffrey Mann (City of Hope Medical Center, Duarte, Calif.). The CCE cells were propagated in the same medium as the F9 cells, with the addition of recombinant leukemia inhibitory factor (0.1 ng/ml; R+D Systems, Minneapolis, Minn.) to preserve their undifferentiated state. The GP+E-86 ecotropic packaging cells (31), a generous gift from Arthur Bank, Columbia University, New York, N.Y., were kept under selection in HXM medium (DMEM supplemented with 10% newborn calf serum, hypoxanthine [15 μg/ml], xanthine [250 μg/ml], and mycophenolic acid [25 μg/ml]) until viral supernatant was needed, at which point the medium was changed to DMEM supplemented with 10% FCS, penicillin-streptomycin, and L-glutamine.

Vectors. The LN vector was constructed in the laboratory of A. Dusty Miller (34). The L-CAT-SN vector was constructed by insertion of a fragment from pSV2-CAT encoding the bacterial CAT gene into the *Hpa*I site of the LXSN vector plasmid (Fig. 1) (34).

The construction of MND-neo was previously described by Challita et al. (9). To construct a multipurpose MND vector backbone, the multicloning site and the simian virus 40 (SV40) promoter from LXSN were cloned in front of *neo* in MND-neo to generate the vector MND-X-SN, where "X" represents unique *Eco*RI, *Hpa*I, and *Xho*I cloning sites. MND-CAT-SN was assembled by cleaving

the MND-neo vector upstream of the *neo* gene at *Bcl*I and inserting a *Hin*dIII fragment containing the CAT gene and SV40 promoter from L-CAT-SN (Fig. 1).

Two pairs of vectors which incorporate one of two fluorescence-activated cell sorting (FACS)-detectable reporter genes, either that a truncated form of rat nerve growth factor receptor (tNGFR) or that for the enhanced green fluorescent protein (EGFP), were constructed by using the vector backbones described above. The truncation of the rat nerve growth factor receptor and the assembly of the Mo-MuLV-driven vector, L-tNGFR-SN, are described in Malik et al. (30). The modified version, MND-tNGFR-SN, was constructed by inserting a 940-bp *Eco*RI-*Xho*I fragment containing the tNGFR gene from L-tNGFR-SN into the same sites in the MND-X-SN backbone (Fig. 1). The L-EGFP-SN and MND-EGFP-SN vectors were constructed by inserting *Bgl*II-*Not*I fragments containing the EGFP gene (Clontech Laboratories, Palo Alto, Calif.) into the *Hpa*I sites in the polylinkers of LXSN and MND-X-SN, respectively (Fig. 1).

Packaging of retroviral vectors. A. Dusty Miller supplied a clone of PA317 packaging cells producing the LN vector at a high titer (10^7/ml) (35). The L-CAT-SN, MND-CAT-SN, L-tNGFR-SN, MND-tNGFR-SN, L-EGFP-SN, and MND-EGFP-SN vectors were packaged by transfecting their vector plasmids into GP+E-86 cells (31). Ecotropic viral supernatant was used to generate G418-resistant PA317 pools from which amphotropic virus supernatant was collected as described by Challita et al. (9).

The amphotropic supernatants were used to transduce the target F9 and CCE cells under the same conditions as those for the fibroblasts. F9 and CCE cell pools were selected in G418 (500 μg/ml) for approximately 14 days, after which they were cultured in DMEM supplemented with 15% FCS, penicillin-streptomycin (50 U/ml), L-glutamine (2 mM), and 0.1 mM (final concentration) β-mercaptoethanol. The CCE cells were maintained in the presence of leukemia inhibitory factor (0.1 ng/ml). Cells were assayed for CAT activity, tNGFR, or EGFP expression while in log phase of growth after at least 24 h without G418, which was observed to interfere with the CAT assay.

Protein extraction and quantitation. Cell extracts were made from cells harvested by trypsinization during log phase of growth. Cell pellets were washed twice in Hanks balanced salt solution, and 10^6 cells were resuspended in 0.1 ml of 0.1 M Tris-HCl (pH 7.8) and lysed with three freeze/thaw cycles in a dry ice-ethanol bath. The lysates were centrifuged at 3,000 × *g* for 10 min to remove the membrane debris. The remaining liquid was transferred to an Eppendorf tube and incubated at 70°C for 15 min to destroy a cellular protein that interferes with the CAT assay (37). After heat treatment, the lysate was centrifuged again at 12,000 × *g* for 10 min to remove additional proteins. The second centrifugation step was crucial for reproducibility and maximal CAT activity.

To standardize the amount of cellular protein in the lysates for the CAT assay, we utilized the BCA Protein Assay quantitation kit (Pierce, Rockford, Ill.). Two microliters of each sample was checked against a bovine serum albumin standard curve prepared for each experiment. Samples for analysis were standardized at 900 μg of total protein and assayed in duplicate or triplicate.

CAT diffusion assay. CAT assays were performed by a modified version of the method published by Neumann et al. (37). Alterations to the protocol include the addition of a centrifugation step after the heat inactivation of the lysate, an increase in the sample volume and total aqueous phase volume from 50 to 100 μl, and the use of 5-ml plastic scintillation vials instead of 7-ml glass vials. The positive control for the assay was a dilution (0.028 μl/sample) of recombinant CAT enzyme (Pharmacia, Uppsala, Sweden). The negative control consisted of 900 μg of protein extracted from cells containing the LN vector that does not encode CAT. The levels of CAT activity expressed by the cells with the LN vector were considered background and were subtracted from the test sample results.

Samples were prepared by diluting 900 μg of total cell lysate to a volume of 100 μl with 0.1 M Tris-HCl (pH 7.8), which was pipetted directly into a scintillation vial. The reaction cocktail (200 μl/sample) was preprepared with 25 μl of 1 M Tris (pH 7.8)–25 μl of 10 mM chloramphenicol (Sigma, St. Louis, Mo.)–0.2 μl per sample of 0.5-μCi/μl [^{3}H]acetyl coenzyme A (Dupont, NEN, Boston, Mass.)–149.8 μl of double-distilled H_2O and added to the lysates in the scintillation vials. The samples were vortexed, and 4.7 ml of Econofluor scintillation fluid (Dupont, NEN) was overlaid. The samples were placed in a shaking water bath (37°C) for up to 18 h. The numbers of counts per minute were monitored periodically, starting at 1 h, for activity within the linear range of the assay. The monitoring was performed by scintillation counting with an LS 3800 (Beckman Instruments, Palo Alto, Calif.). The most consistent results were achieved at approximately 12 h.

Antibody staining and flow cytometry. For analysis of cell surface expression of the tNGFR reporter, CCE, F9, and PA317 cells were harvested with enzyme-free cell dissociation buffer (Gibco-BRL, Bethesda, Md.) in log phase of growth, washed twice in RPMI 1640, resuspended at 10^7 cells/ml in phosphate-buffered saline, and placed on ice. One million cells were stained for each sample, and all subsequent steps were performed on ice. Nonspecific antibody binding was blocked by adding 5 μl of murine immunoglobulin G (IgG) (10 mg/ml) and incubating for 30 min. Then, 5 μl of anti-rat NGFR monoclonal antibody MC192 (100 μg/ml; Oncogene Science, Uniondale, N.Y.) or the negative control (murine IgG) were added, and the cells were incubated for 30 min. Excess antibody was removed by two washes with phosphate-buffered saline. The fluorescent second antibody, phycoerythrin-labeled goat anti-mouse antibody (5 μl/10^6 cells of a 1:20 dilution of the stock [Caltag, San Francisco, Calif.]), was added. After 30 min, the cells were washed twice with phosphate-buffered saline and fixed in 1% paraformaldehyde. Stained cells were evaluated with a FACScan flow cytometer (Becton Dickinson, San Jose, Calif.). To establish the background for the staining procedure and to set the gates for data acquisition, sham-transduced cells were prepared in parallel and analyzed for each experiment.

Preparation for the analysis of green fluorescent protein (GFP) expression required unselected F9 cells to be trypsinized either 48 h, 9 days, or 22 days posttransduction, washed twice in phosphate-buffered saline, and resuspended at a concentration of 10^7/ml in phosphate-buffered saline. Detection of GFP expression was accomplished on a FACScan cytometer equipped with a 488-nm argon laser for excitation of the reporter protein and a 530/30-nm bandpass filter for monitoring the fluorescent emissions. Experimental background was established by including non-EGFP (LN)-transduced cells in the analysis.

Northern and Southern blot analyses. Genomic DNA and total cellular RNA were extracted from cell pellets of retrovirally transduced PA317, CCE, and F9 cells for Southern and Northern blot analysis. Genomic DNA was extracted by sodium dodecyl sulfate (SDS)-proteinase K digestion overnight at 37°C in proteinase K buffer containing 0.01 M Tris-HCl (pH 7.4), 0.15 M NaCl, 0.02 M EDTA (pH 8.0), 0.08% SDS, 50 μl of proteinase K (10 mg/ml) (Gibco-BRL), and 5 μl of RNAse (Gibco-BRL) per ml. Protein was removed by phenol-chloroform extraction, and the DNA was precipitated in ethanol and resuspended in Tris-EDTA buffer (43).

Quantitative Southern blots were performed to determine proviral copy number as described by Challita et al. (9). The intensities of the proviral/β-actin bands were compared with a U.S. Biochemicals (Cleveland, Ohio) Sci-Scan 5000 densitometer.

Southern blots (also described in reference 9) were used to determine the extent of proviral methylation at the *Sma*I site in the R region of the LTR (Fig. 1). The extent of proviral methylation was assessed with densitometry by comparing the intensity of the unique band created by *Sma*I digestion to the intensity of the band created by the larger *Sma*I-resistant fragment.

RNA was extracted from the cells with guanadinium isothiocyanate and phenol-chloroform as described by Chomczynski and Sacchi (10). RNA (15 to 20 μg) was electrophoresed on a 1.2% formaldehyde gel, denatured, neutralized, and transferred to a nylon membrane via capillary action. Nucleic acids were cross-linked to the membrane with a UV Stratalinker (Stratagene, La Jolla, Calif.). The nylon membrane was then hybridized to either a *neo* or β-actin cDNA probe labeled with [^{32}P]dCTP incorporated with a Prime-It random primer kit from Stratagene. The membrane was washed as outlined above and then used to expose Kodak X-Omat X-ray film.

TABLE 1. CAT activity in cell lysates

Cell type (*n*)	Avg CAT activity (cpm)[a] ± SEM (% control[b]) from:	
	L-CAT-SN	MND-CAT-SN
PA317 (10)	99,182 ± 11,586 (100)	112,232 ± 10,666 (100)
CCE (7)	11,876 ± 2,960 (12)	54,240 ± 12,349 (48)
F9 (3)	122 ± 62 (0.1)	38,122 ± 3,439 (34)

[a] Activity is expressed as the number of counts per minute of ^{3}H-labeled acetylated chloramphenicol produced by lysates of CCE, F9, or PA317 cells containing the indicated vector minus the background counts per minute in cells transduced with the LN vector from the same experiment.

[b] % control, number of counts per minute of ^{3}H-labeled acetylated chloramphenicol produced by lysates of CCE or F9 cells divided by the number of counts per minute of ^{3}H-labeled acetylated chloramphenicol produced by lysates of transduced PA317 cells times 100%.

RESULTS

Expression by Mo-MuLV is restricted in EC and ES cells. Observations in our laboratory by Challita et al. (9) indicated that the MND vector increased the expression of the neomycin resistance gene in F9 cells. Therefore, MND was selected as the backbone for the novel vectors in this study. To accurately quantitate relative gene expression levels from the standard, Mo-MuLV-based vectors (L-CAT-SN, L-tNGFR-SN, and L-EGFP-SN) and our novel, modified constructs (MND-CAT-SN, MND-tNGFR-SN, and MND-EGFP-SN) (Fig. 1) in F9 EC, CCE ES, and PA317 cells, we compared CAT enzyme activities and levels of tNGFR protein or EGFP production.

Measurement of CAT activity. Table 1 shows the average CAT assay results measured for each vector in seven independent experiments on cell pools from five separate transductions of CCE cells and three experiments from two transductions of

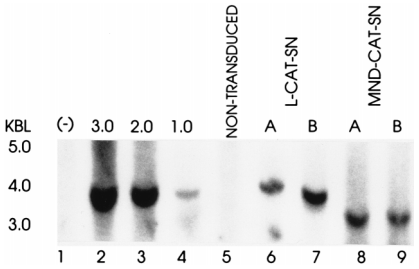

FIG. 2. Determination by Southern blotting of the copy numbers of CAT vectors in F9 cells. Total genomic DNA from two sets of separately transduced F9 cells (labeled A and B) was digested with restriction enzymes to release discrete vector fragments and was electrophoresed on a 1.2% agarose gel. Standard dilution series representing 3, 2, and 1 vector copies are shown in lanes 2, 3, and 4, respectively. β-actin probing of the gel was used as a DNA loading control (data not shown; see the text for an explanation). Densitometry results are included in the text. Lane (−), negative control containing DNA from nontransduced cells. KBL, kilobase ladder.

F9 cells. The background in the CAT assay was established as the level of CAT activity measured in the CCE or F9 target cells transduced with the LN vector. The percent of control was defined as the level of CAT expression (in counts per minute with the background subtracted) from each vector in CCE or F9 cells divided by the level of expression for the same vector in PA317 fibroblast cells in the same experiment. The numbers of counts per minute (with the background subtracted) are shown for each vector in all three cell types.

In the permissive PA317 cells, the L-CAT-SN vector showed an average CAT activity (± standard error of the mean [SEM]) of 99,182 (±11,586) cpm. MND-CAT-SN gave an average of 112,232 (±10,666) cpm in the PA317 cells, which is similar to the value from L-CAT-SN (Student's *t* test result: $P = 0.44$). However, when the vectors were tested in CCE ES cells, which are restrictive to expression from the Mo-MuLV LTR, we found that L-CAT-SN produced an average of only 11,876 (±2,960) cpm. In contrast, the MND-CAT-SN vector produced an average of 54,240 (±12,349) cpm in the CCE cells, which is significantly greater than the value for the L-CAT-SN vector ($P = 0.0031$). Vector MND-CAT-SN achieved a level of expression in the CCE ES cells that was 48% of that in PA317 cells, while vector L-CAT-SN produced only 12% of the activity in CCE cells as in PA317 cells.

In F9 cells, repression of the Mo-MuLV LTR appears to be more extreme than in CCE cells, with an average CAT activity of only 122 (±62) cpm from L-CAT-SN. In contrast, MND-CAT-SN produced a significantly greater average CAT activity of 38,122 (±3,439) cpm than did L-CAT-SN ($P = 0.00038$). While the relative expression pattern between the Mo-MuLV-based vector and the MND construct is maintained in the F9 EC cells compared with the results for CCE ES cells, the absolute levels of expression are decreased. In the F9 EC cells, the MND-CAT-SN vector produced 34% of the activity seen in the PA317 cells, while the L-CAT-SN vector produced only 0.1% of the CAT activity seen in the PA317 cells. Therefore, MND-CAT-SN produced 4 and 340 times more CAT activity than the L-CAT-SN vector in CCE and F9 cells, respectively.

Quantitation of proviral integrants. In order to allow the comparison of gene expression on a copy-per-cell basis, it was necessary to determine the number of proviral integrants in the target cells. Quantitative Southern blotting was performed on DNA from pools of transduced CCE and F9 cells. To prepare the samples for Southern blotting, phenol-extracted DNA was cut with restriction enzymes that recognize sites within the vectors, thereby releasing a discrete fragment regardless of integration site or copy number. The standard dilution series on the left of Fig. 2 (lanes 2 to 4) are derived from a cell line known to contain a single proviral integrant. The dilutions in lanes 2 to 4 represent the equivalent of three vector copies, two vector copies, and a single proviral integrant, respectively. After correction for the DNA loading on the gel by re-hybridization of the blot with a β-actin probe (data not shown) densitometric analysis of the fragments hybridized to the labeled *neo* gene probe revealed that the L-CAT-SN (lanes 6 and 7) and MND-CAT-SN (lanes 8 and 9) vectors were present at one or two copies per cell in two separately transduced pools of F9 cells. The β-actin hybridization results are not shown because the different restriction enzymes required to release discrete vector fragments produced complicated digestion patterns upon reprobing. Therefore, the cumulative density measured for all of the β-actin complementary bands in a given sample was used as the denominator to extrapolate the number of vector integrants. Southern blot analysis of two sets of transduced CCE cells similarly showed that the L-CAT-SN and MND-CAT-SN vectors were present at one or two copies per cell (data not shown).

Northern blot analyses of vector-derived transcripts. Total cellular RNA from transduced CCE cells was analyzed by Northern blotting to assess vector expression in the target cells (Fig. 3). Three transcripts may potentially be produced from these vectors: a full-length genomic transcript initiated in the 5′ LTR of approximately 3.4 kb, a spliced form of the genomic transcript approximately 3.1 kb in length, and a transcript initiated at the SV40 promoter approximately 1.4 kb in length, with all transcripts terminating at the polyadenylation signal of the 3′ LTR.

Table 2 shows the numerical values for the relative transcript abundances from two separate Northern blots. Figure 3 (lanes 1 and 2) shows strong expression of the full-length genomic transcript and the SV40-promoted *neo* message from the L-CAT-SN and MND-CAT-SN vectors in PA317 cells. In the F9 cells transduced with the L-CAT-SN vector (lane 3), no measurable signal was produced for either the full-length or spliced LTR transcripts, although the SV40-promoted *neo* message was visible. The MND-CAT-SN vector produced easily detectable signals for the singly spliced LTR transcript and the SV40 *neo* transcript in the F9 cells (lane 4). The modified vector produced an average of 27% as much LTR-promoted RNA in

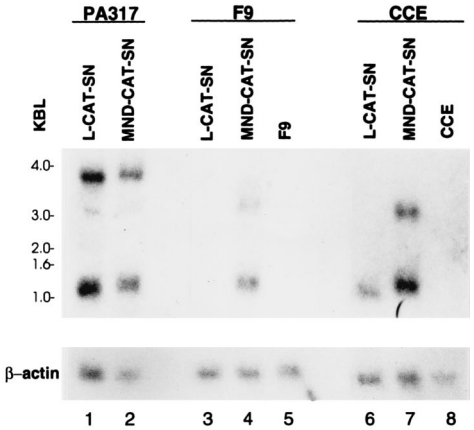

FIG. 3. Northern blot analysis of vector-derived RNA from transduced PA317, F9, and CCE cells. The membrane was hybridized with a ^{32}P-labeled *neo* gene probe, stripped, and reprobed with labeled β-actin cDNA to allow for correction in loading efficiency. Densitometry results are included in the text. KBL, kilobase ladder.

TABLE 2. Relative abundance of vector-derived LTR transcripts

Vector(s)	Relative abundance[a] of LTR transcripts in:			
	CCE cells		F9 cells	
	Blot 1	Blot 2	Blot 1	Blot 2
LN	0.026	0.04	0.049	0.14
L-CAT-SN	0.0028	0	0	0
MND-CAT-SN	0.23	0.81	0.072	0.47
MND-CAT-SN/L-CAT-SN	81	ND[b]	ND	ND

[a] Values were normalized to the abundance of transcripts from L-CAT-SN vector LTR in PA317 cells, which was defined as 1.0.
[b] ND, not mathematically defined.

the F9 cells as it did in the PA317 cells and >1,000 times more RNA than the L-CAT-SN vector did in the F9 cells. As expected, no evidence of *neo* expression is observed in the nontransduced F9 cells (lane 5) or the CCE cells (lane 8).

In the CCE cells, the L-CAT-SN vector failed to produce a measurable amount of full-length genomic or spliced RNA, although a modest signal from the SV40 *neo* transcript is visible (Fig. 3, lane 6). However, in a second blot (data not shown), the L-CAT-SN vector produced an average of 0.14% of the amount of LTR-promoted RNA in the CCE cells as it did in the PA317 packaging cells (Table 2). Remarkably, CCE cells containing the MND-CAT-SN vector (lane 7) produced 52% of the amount of LTR-driven RNA as did PA317 cells containing the same vector and an average of 370 times more RNA than did CCE cells containing the Mo-MuLV-based L-CAT-SN vector. Interestingly, increased levels of the SV40-directed transcripts from the MND-CAT-SN vector were seen in both the CCE and F9 cells compared to those from L-CAT-SN.

Southern blot analysis of proviral methylation. Challita and Kohn (8) demonstrated that methylation at the *Sma*I site inside the R region of the 5′ LTR (Fig. 1) is inversely correlated with expression from the provirus; therefore, we performed the same analysis on these samples. Densitometry of the blots confirms that the L-CAT-SN vector (Fig. 4, lane 2) and the MND-CAT-SN vector (data not shown) are completely unmethylated in PA317 cells. The Mo-MuLV LTR within the L-CAT-SN vector was 61% methylated in the F9 cells (Fig. 4, lane 4) and 92% methylated in the CCE cells (Fig. 4, lane 6). In contrast, the MND LTR is only 32% methylated in the F9 EC cells (Fig. 4, lane 8) and 14% methylated in transduced CCE cells (Fig. 4, lane 10). Therefore, the extent of methylation of each vector showed that there was an inverse relation between methylation of the provirus and LTR-driven expression in all three cell models.

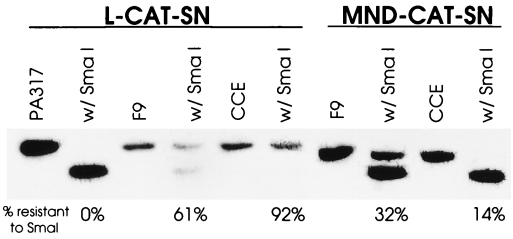

FIG. 4. Southern blot to determine the extent of proviral methylation. Total genomic DNA from PA317, CCE, and F9 cells transduced with either the L-CAT-SN or MND-CAT-SN vectors was digested with *Nhe*I (lanes 1, 3, and 5), *Nhe*I and *Sma*I (lanes 2, 4, and 6), *Eco*RV (lanes 7 and 9), or *Eco*RV and *Sma*I (lanes 8 and 10) to release discrete vector fragments. Digested samples were electrophoresed and transferred to nylon membranes. The membranes were hybridized to a ^{32}P-labeled *neo* gene and used to expose X-ray film. Densitometry was performed to quantitate the intensities of the *Sma*I-resistant fragments (upper bands) and *Sma*I-sensitive fragments (lower bands) in lanes 2, 4, 6, 8, and 10. The proportion of the total density measured from both bands that was contributed by the *Sma*I-resistant fragment represents the extent of proviral DNA methylation.

Quantitation of surface antigen and intracellular reporter expression in individual cells. (i) tNGFR expression. Because the CAT assay and Northern blot analyses assess average gene expression from the bulk population of cells, the contribution of individual cells is obscured. FACS was employed to assess the level of LTR-promoted expression of a surface antigen reporter in individual CCE or F9 cells transduced with either a Mo-MuLV LTR-driven vector, L-tNGFR-SN, or the modified version, MND-tNGFR-SN. Both vectors encode rat tNGFR, which is expressed on the cell surface and which can be detected with monoclonal antibody MC192 (30). Since the fluorescence intensity is proportional to the amount of protein expressed on the surface of each cell, FACS permits discrimination between high-level gene expression in a minority of cells and uniform expression in a majority of cells within a population. Because the cells are G418 selected, each contains at least one copy of one of the vectors and, therefore, has equal potential for tNGFR reporter gene expression.

In the PA317 fibroblasts, the MND-tNGFR-SN vector showed strong expression in essentially every cell, as did the L-tNGFR-SN vector (Fig. 5A). Mo-MuLV-based vector L-tNGFR-SN performed poorly in the CCE ES cells, producing detectable tNGFR protein in only 1.04% of the cells (Fig. 5B). In contrast, the MND-tNGFR-SN vector drove expression of tNGFR in 99.54% of the CCE cells. In the F9 cells, the dichotomy is also extreme. The L-tNGFR-SN vector expressed the cell surface reporter in only 13.32% of the cells, but the modified MND-tNGFR-SN vector expressed the reporter in 99.89% of the F9 cells (Fig. 5C). Interestingly, the level of expression (fluorescence intensity) from the fraction of CCE or F9 cells which did show expression from L-tNGFR-SN was equivalent to the expression levels from MND-tNGFR-SN in the same cells.

(ii) EGFP expression. To explore the possibility that G418 selection of cell pools in the previous experiments biased our observations on the frequency of expression or that the results were construct specific, we compared expression levels of the GFP from the Mo-MuLV LTR and the MND LTR in pools of transduced F9 cells without exposure to G418 (Fig. 1). In the packaging cells, GFP expression from both the L-EGFP-SN vector and the MND-EGFP-SN vector was strong (data not shown). Forty-eight hours after transduction, the L-EGFP-SN vector drove expression in only 5% of the transduced F9 cells (Fig. 6). As observed with the tNGFR reporter, the MND-EGFP-SN vector drove expression in a large proportion of the cells (88%) (Fig. 6). Also in agreement with the tNGFR results, the few cells that were expressing EGFP from the Mo-MuLV LTR were expressing it at the same intensity as those that were expressing it from the MND LTR. The titers of the viral supernatants used to produce these populations were similar ($\sim 10^6$/ml), and quantitative Southern blot analyses (performed as described above) indicated that the L-EGFP-SN and MND-EGFP-SN vectors were present in similar quantities ($\sim$1 copy per cell) in the nonselected F9 cells (data not shown).

To determine if the differences in the frequency of vector expression were due to variations in the kinetics of vector activation or inactivation, we analyzed EGFP expression 48 h after infection and then cultured the cells in the absence of

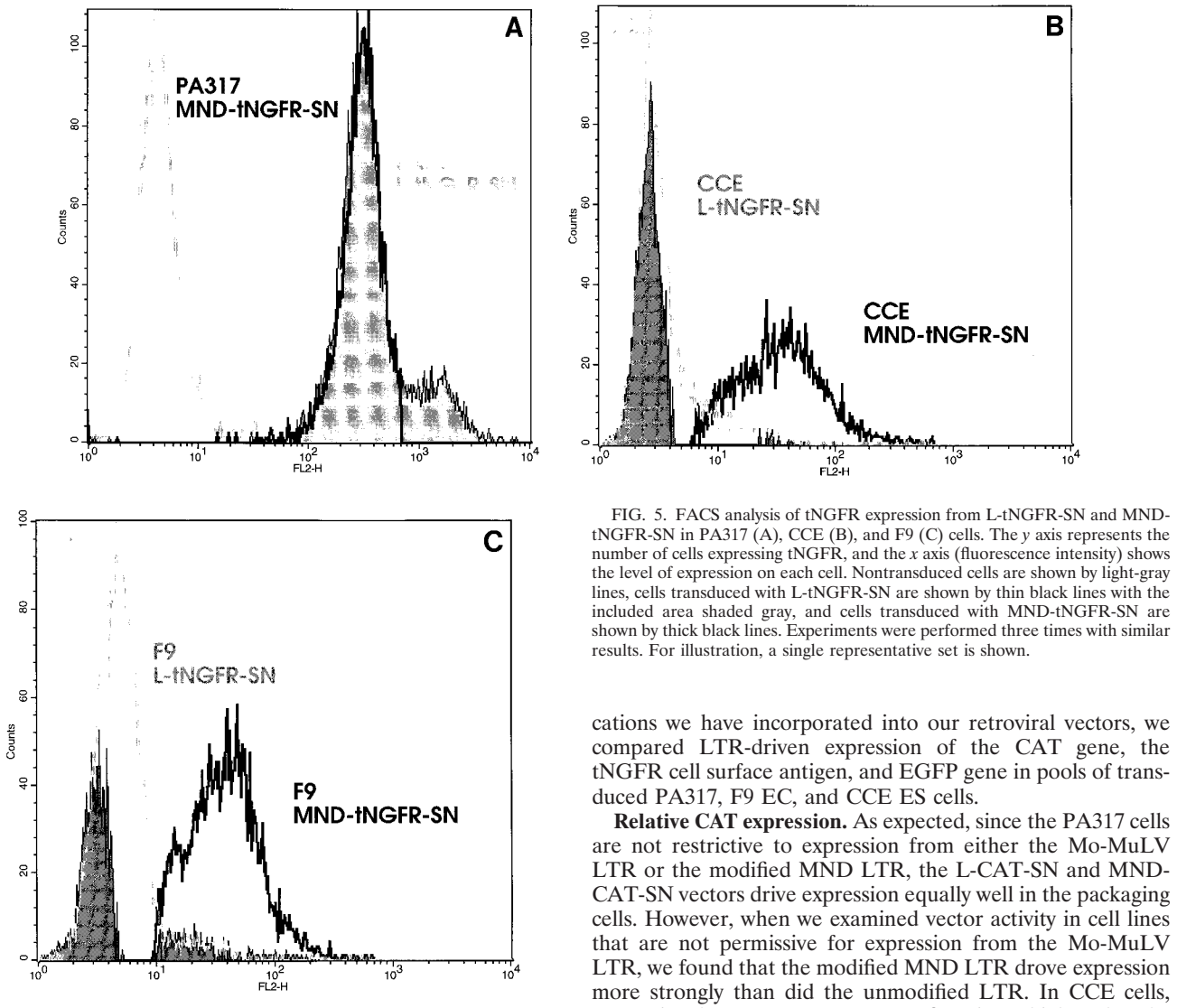

FIG. 5. FACS analysis of tNGFR expression from L-tNGFR-SN and MND-tNGFR-SN in PA317 (A), CCE (B), and F9 (C) cells. The *y* axis represents the number of cells expressing tNGFR, and the *x* axis (fluorescence intensity) shows the level of expression on each cell. Nontransduced cells are shown by light-gray lines, cells transduced with L-tNGFR-SN are shown by thin black lines with the included area shaded gray, and cells transduced with MND-tNGFR-SN are shown by thick black lines. Experiments were performed three times with similar results. For illustration, a single representative set is shown.

G418 for an additional 20 days and reanalyzed GFP expression. After 9 days posttransduction, the percentages of cells expressing EGFP from the Mo-MuLV LTR and the modified MND LTR were essentially unchanged from that observed at 48 h: 6 and 84%, respectively (Fig. 6). After 22 days in culture in the absence of G418, the L-EGFP-SN vector was active in only 3% of the population while expression from the modified MND-EGFP-SN vector persisted at the same magnitude in 80% of the cells.

DISCUSSION

In this report, we compared expression levels of Mo-MuLV LTR-driven retroviral vectors and vectors with modifications in *cis* elements in restrictive ES and EC cells. The modifications to the MND retroviral vectors target three obstacles to gene expression in EC and ES cells. They consist of replacement of the inactive Mo-MuLV enhancer with the enhancer from MPSV, which is active in EC and ES cells, the removal of the NCR to prevent binding of YY1, and substitution of the PBS with the *dl*587 PBS to prevent binding of the inhibitory repressor-binding protein. To gauge the effectiveness of the modifications we have incorporated into our retroviral vectors, we compared LTR-driven expression of the CAT gene, the tNGFR cell surface antigen, and EGFP gene in pools of transduced PA317, F9 EC, and CCE ES cells.

Relative CAT expression. As expected, since the PA317 cells are not restrictive to expression from either the Mo-MuLV LTR or the modified MND LTR, the L-CAT-SN and MND-CAT-SN vectors drive expression equally well in the packaging cells. However, when we examined vector activity in cell lines that are not permissive for expression from the Mo-MuLV LTR, we found that the modified MND LTR drove expression more strongly than did the unmodified LTR. In CCE cells, expression by MND-CAT-SN was five times higher than by L-CAT-SN as measured by the CAT assay. In the highly repressive F9 cells, the modified MND-CAT-SN vector produced ~300 times more CAT activity than did the L-CAT-SN vector. The densitometric analyses of Northern blots from transduced CCE and F9 cells showed >80-fold increases in the quantity of RNA from MND-CAT-SN, which suggested that the enhanced CAT activity is the result of an increased steady-state level of vector-derived RNA.

To rule out the possibility that the large increases in expression by MND-CAT-SN and MND-EGFP-SN were due to the presence of increased numbers of proviral templates, copy number Southern blots were performed. The blots indicated that the numbers of vector integrants ranged between 1 and 2 copies per cell for all of the vectors. Therefore, differences in gene transfer do not explain the lower level of expression from L-CAT-SN compared to the level from MND-CAT-SN.

Proviral methylation. It is clear that the modifications that we have incorporated into the MND vector backbone allow the provirus to be more resistant to the process(es) responsible for gene silencing than the standard Mo-MuLV LTR. One potential epigenetic mechanism of gene silencing is DNA methylation at cytosine residues. Gene silencing and de novo methylation frequently occur in transcriptionally quiescent DNA

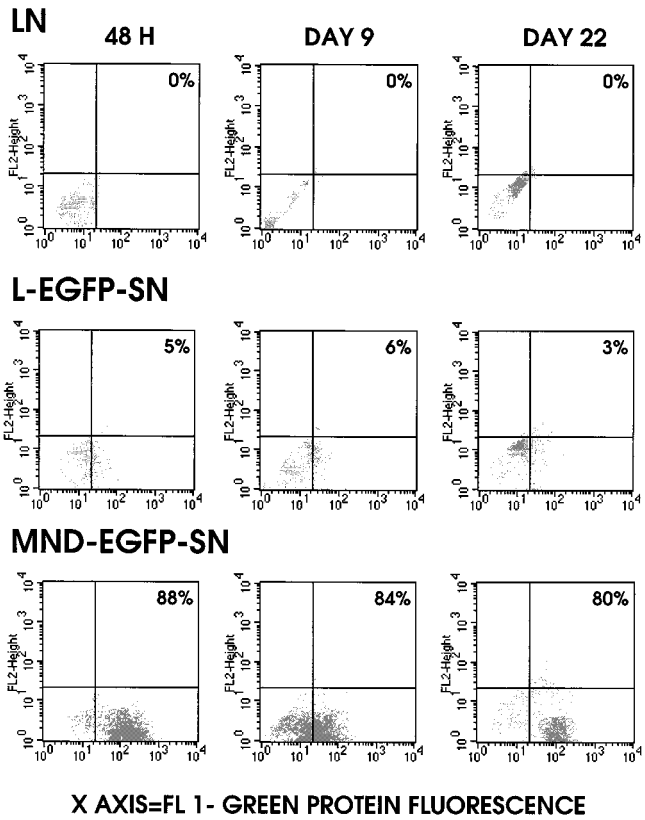

FIG. 6. Temporal expression of EGFP in F9 cells without selection. EGFP expression was analyzed by flow cytometry in pools of transduced F9 cells at 48 h, 9 days, and 22 days posttransduction. The *x* axis (fluorescence intensity) represents the level of EGFP expression in individual cells. The top row of panels (LN), representing the negative control, shows the lack of EGFP expression from F9 cells transduced with the LN vector, which does not encode EGFP. The middle row (L-EGFP-SN) shows low-level EGFP expression driven by the L-EGFP-SN vector at the three time points. The bottom row (MND-EGFP-SN) shows high-level EGFP expression from the modified MND-EGFP-SN vector at all three time points. The proportions of cells driving expression are given in the upper right corners of the dot plots.

sequences, while transcriptionally active DNA tends to remain unmethylated (7, 42). Analysis by Southern blotting indicated that in ES and EC cells the unmodified Mo-MuLV LTR in the L-CAT-SN vector was more extensively methylated at the *Sma*I site than was the modified MND-CAT-SN LTR.

It is possible that the Sp1 site created during the divergence of MPSV and Mo-MuLV might play a role in reducing the extent of methylation in the MND LTR. Silke et al. (44) showed that the incorporation of Sp1 transcription factor binding sites could induce substantial demethylation of target sequences in F9 cells. Brandeis et al. (6) and Macleod et al. (29a) demonstrated the involvement of Sp1 sites in the demethylation of sequences with transgenic mice. However, in vitro and in vivo, the presence of Sp1 sites alone was not sufficient to ensure complete demethylation of surrounding sequences. The results presented here support an association between methylation and vector silencing but do not illuminate any potential cause-and-effect relationship between these events.

Probability of expression. Although our studies with the CAT reporter vectors have demonstrated increased expression from the MND-CAT-SN vector in pools of transduced cells, the analysis of CAT expression in bulk populations obscures the contribution from individual cells. Walters et al. (49) suggested that promoters function as binary controls which are either "on" or "off" and that their relative strengths, measured in a pool of cells, are a result of the frequencies with which they are "on" and not the rate of transcription. In their model, enhancers act to increase the probability of forming an active transcription complex and, subsequently, the frequency of transcription from a given promoter determines the quantities of RNA produced in an individual cell or in a bulk population.

To examine the effects of the MND vector modifications on the frequency of vector expression, we analyzed cell surface expression of rat tNGFR from a Mo-MuLV-driven construct, L-tNGFR-SN, and a modified vector, MND-tNGFR-SN, in pools of CCE ES cells and F9 EC cells. Flow cytometry allows the quantitation of reporter gene expression on individual cells; therefore, analysis of retrovirally transduced cell pools can provide an unbiased assessment of a vector's ability to overcome the effects of chromosomal position and gene-silencing mechanisms in a given cell type.

The modified MND-tNGFR vector drove expression in essentially all of the CCE and F9 cells, while the Mo-MuLV LTR-driven L-tNGFR-SN vector drove expression in only a small percentage of the CCE and F9 cells. However, the magnitudes of expression in the fraction of CCE or F9 cells which did stain positive for tNGFR were similar for the two vectors. Presumably, the L-tNGFR-SN proviruses that achieve expression of tNGFR in ES and EC cells found infrequent, favorable integration sites.

To further examine the effects of our vector alterations and allay concerns that our observations on the frequency of expression were specific to the reporter construct or were influenced by G418 selection of the cell pools used in the analyses, we compared levels of EGFP expression in transduced but unselected F9 cells. As with the tNGFR reporter, expression of EGFP from the Mo-MuLV LTR was detectable in only a small percentage of the cells. In contrast, expression of EGFP from the modified MND LTR could be detected in nearly all of the cells. In accordance with the tNGFR results and the model of Walters et al. (49), the magnitude of EGFP expression measured from the minority of cells which did show expression from the L-EGFP-SN vector was similar to the magnitude seen in the cells transduced with the MND-based vector.

To explore the possibility that vector expression might vary with time posttransduction, we analyzed EGFP expression from L-EGFP-SN and MND-EGFP-SN shortly after transduction and after 16 and 22 days in culture without selection. We found that the expression profile of the Mo-MuLV-based vector was consistently poor (~4%). The probability of EGFP expression from the modified MND vector was also stable for the duration of the experiment, occurring in a large proportion of the cells (~80%).

Therefore, neither the use of different reporters nor the time of analysis posttransduction nor selection of the target cells with G418 altered the outcome or interpretation of these experiments. The ability of the modified MND LTR to increase the probability of gene expression is transferable to different reporter constructs, is not dependent upon G418 selection, and does not appear to have a temporal expression window.

The Mo-MuLV LTR appears to be sensitive to chromosomal position effects, is prone to de novo methylation (which was correlated with vector inactivity), and has a low probability of expression in ES and EC cells. In contrast, the MND vector is relatively resistant to position effects and other repressive forces acting in ES and EC cells. Incorporation of other elements, such as locus control regions or scaffold attachment regions may further insulate vectors from inhibitory positional effects.

Franz et al. (17) and Hilberg et al. (21) isolated the PCC4 cell-passaged myeloproliferative sarcoma virus (PCMV), which was active in EC but not ES cells, by applying selection pressure for mutants of MPSV that grew in PCC4 EC cells. One of the major alterations of the MPSV LTR which characterizes the PCMV LTR is the loss of one of the 75-bp tandem enhancer repeats present in most viruses of this family. Grez et al. (19) demonstrated an expanded host cell range, including EC and ES cells, with the construction of their murine embryonic stem cell virus (MESV) vector. The MESV vector contained the LTR from the PCMV and the altered PBS from the *dl*587rev virus. The combined use of both modifications was shown to be essential for improved vector expression.

The MND vector backbone described here is similar, in some respects, to that of the MESV in that the MND and MESV vectors have the same substitution of the PBS and repressor protein-binding site from the *dl*587rev endogenous retrovirus (11). The LTRs of MND and MESV come from, respectively, MPSV (25) and PCMV (17, 21). In the LTR of the MND vector, we deleted a 63-bp fragment ($Nhe\mathrm{I}_{33}$-$Sau3\mathrm{a}_{97}$) which spanned the NCR region as defined by Flanagan et al. (16), completely eliminating the YY1 binding site. In the wild-type MPSV LTR, the YY1 binding site is disrupted by a point mutation; however, during the selection process that created the PCMV LTR, the YY1 site was restored to the consensus sequence and subsequently persists in the MESV vector. The relative effects of these different LTR alterations on expression are not known.

Mo-MuLV silencers. The observation that greater quantities of the SV40 transcript were produced in cells containing the MND-CAT-SN vector than in those with L-CAT-SN suggested that negative elements within the L-CAT-SN vector may interfere with expression from the internal SV40 promoter as well as from the LTR. Experiments incorporating Mo-MuLV sequences (including but not restricted to the LTR) in gene cassettes used to produce transgenic mice showed that multiple sequence elements from Mo-MuLV (such as the PBS) are capable of interfering with or shutting down expression of functional transgenes when incorporated upstream or downstream in transgene cassettes (12, 32). These results suggest that the MND backbone may be advantageous for vectors intended to express genes from internal promoters, such as those from housekeeping genes, e.g., phosphoglycerate kinase, or lineage-specific enhancers, such as the β-globin locus.

Since retroviral vectors integrate randomly and chromosomal context can have profound positive or negative influences on gene expression constructs including retroviral vectors, effective vectors must be refractory to positional silencing effects in vitro and in vivo (36, 51). The resistance of the MND provirus to methylation may improve the long-term performance of the construct. The dramatic increase in the probability of expression with the MND vector could be particularly useful in cells which are difficult to transduce, such as HSC, or where the cumulative expression increase would be most evident, e.g., for cell surface or secreted proteins and genes which provide no selective advantage. Augmented expression from the modified MND LTR might make the difference between detectable and undetectable experimental results in vitro or between therapeutic and nontherapeutic levels of gene expression in vivo. Therefore, these findings are potentially important for many research or clinical applications using retroviral vectors.

ACKNOWLEDGMENTS

These studies were supported by grants from the National Institute of Diabetes and Digestive and Kidney Diseases (DK4900) and the National Cancer Institute (CA59318) of the National Institutes of Health. Donald B. Kohn is the recipient of an Elizabeth Glaser Scientist Award from the Pediatric AIDS Foundation.

REFERENCES

1. **Agrawal, Y. P., R. S. Agrawal, A. M. Sinclair, D. Young, M. Maruyama, F. Levine, and A. D. Ho.** 1996. Cell-cycle kinetics and VSV-G pseudotyped retrovirus-mediated gene transfer in blood-derived CD34+ cells. Exp. Hematol. **24:**738–747.
2. **Barklis, E., R. Mulligan, and R. Jaenisch.** 1986. Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell **47:**391–399.
3. **Becker, K. G., P. Jedlicka, N. S. Templeton, L. Liotta, and K. Ozato.** 1994. Characterization of hUCRBP (YY1,NF-E1, δ): a transcription factor that binds to the regulatory regions of many viral and cellular genes. Gene **150:**259–266.
4. **Bowtell, D. D. L., S. Cory, G. R. Johnson, and T. J. Gonda.** 1988. Comparison of expression in hemopoietic cells by retroviral vectors carrying two genes. J. Virol. **62:**2464–2473.
5. **Boyes, J., and A. Bird.** 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell **64:**1123–1134.
6. **Brandeis, M., D. Frank, I. Keshet, Z. Siegfried, M. Mendelsohn, A. Nemes, V. Temper, A. Razin, and H. Cedar.** 1994. Sp1 elements protect a CpG island from de novo methylation. Nature (London) **371:**435–438.

6a. **Byun, J., J. M. Kim, S. Yu, P. D. Robbins, J. Yim, and S. Kim.** 1996. Analysis of the relative level of gene expression from different retroviral vectors used for gene therapy. Gene Ther. **3:**780–788.

7. **Cedar, H.** 1988. DNA methylation and gene activity: minireview. Cell **53:**3–4.
8. **Challita, P.-M., and D. B. Kohn.** 1994. Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation *in vivo*. Proc. Natl. Acad. Sci. USA **91:**2567–2571.
9. **Challita, P.-M., D. Skelton, A. El-Khoueiry, X.-J. Yu, K. Weinberg, and D. B. Kohn.** 1995. Multiple modifications in *cis* elements of the long terminal repeat of retroviral vectors lead to increased expression and decreased DNA methylation in embryonic carcinoma cells. J. Virol. **69:**748–755.
10. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
11. **Colicelli, J., and S. P. Goff.** 1987. Isolation of a recombinant murine leukemia virus utilizing a new primer tRNA. J. Virol. **57:**37–45.
12. **Ellis, J., X. Wu, and P. Pasceri.** 1996. Evaluation of β-globin gene therapy vectors in single copy transgenic mice. Presented at the Cold Spring Harbor Conference on Gene Therapy, Cold Spring Harbor, N.Y.
13. **Evans, M. J., and M. H. Kaufman.** 1981. Establishment in culture of pluripotent cells from mouse embryos. Nature **292:**154–156.
14. **Feuer, G., M. Taketo, R. C. Hanecak, and H. Fan.** 1989. Two blocks in Moloney murine leukemia virus expression in undifferentiated F9 embryonal carcinoma cells as determined by transient expression assays. J. Virol. **63:**2317–2324.
15. **Flanagan, J. R., K. G. Becker, D. L. Ennist, S. L. Gleason, P. H. Driggers, B.-Z. Levi, E. Appella, and K. Ozato.** 1992. Cloning of a transcription factor that binds to the upstream conserved region of Moloney murine leukemia virus. Mol. Cell. Biol. **12:**38–44.
16. **Flanagan, J. R., A. M. Krieg, E. E. Max, and A. S. Khan.** 1989. Negative control region at the 5′ end of murine leukemia virus long terminal repeat. Mol. Cell. Biol. **9:**739–746.
17. **Franz, T., F. Hilberg, B. Seliger, C. Stocking, and W. Ostertag.** 1986. Retroviral mutants efficiently expressed in embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA **83:**3292–3296.
18. **Gorman, C. M., P. W. J. Rigby, and D. P. Lane.** 1985. Negative regulation of viral enhancers in undifferentiated embryonic stem cells. Cell **42:**519–526.
19. **Grez, M., E. Akgun, F. Hilberg, and W. Ostertag.** 1990. Embryonic stem cell virus, a recombinant murine retrovirus with expression in embryonic stem cells. Proc. Natl. Acad. Sci. USA **87:**9202–9206.
20. **Grez, M., M. Zornig, J. Nowock, and M. Ziegler.** 1991. A single point mutation activates the Moloney murine leukemia virus long terminal repeat in embryonal stem cells. J. Virol. **65:**4691–4698.
21. **Hilberg, F., C. Stocking, W. Ostertag, and M. Grez.** 1987. Functional analysis of a retroviral host-range mutant: altered long terminal repeat sequences allow expression in embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA **84:**5232–5236.
22. **Hoeben, R. C., A. A. J. Mighielsen, R. C. M. van der Jagt, H. van Ormondt, and A. J. van der Eb.** 1991. Inactivation of the Moloney murine leukemia virus long terminal repeat in murine fibroblast cell lines is associated with methylation and dependent on its chromosomal position. J. Virol. **65:**904–912.
23. **Kempler, G., B. Freitag, B. Berwin, O. Nanassy, and E. Barklis.** 1993. Characterization of the Moloney murine leukemia virus stem cell-specific

repressor binding site. Virology **193:**590–599.
24. **Kohn, D. B.** 1995. The current status of gene therapy using hematopoietic stem cells. Curr. Opin. Pediatr. **7:**56–63.
25. **Kollek, R., C. Stocking, F. Smadja-Joffe, and W. Ostertag.** 1984. Molecular cloning and characterization of a leukemia-inducing myeloproliferative sarcoma virus and two of its temperature sensitive mutants. J. Virol. **50:**717–724.
26. **Lengauer, C., K. W. Kinzler, and B. Vogelstein.** 1997. DNA methylation and genetic instability in colorectal cancer cells. Proc. Natl. Acad. Sci. USA **94:**2545–2550.
27. **Linney, E., B. Davis, J. Overhauser, E. Chao, and H. Fan.** 1984. Nonfunction of a Moloney murine leukemia virus regulatory sequence in F9 embryonal carcinoma cells. Nature **308:**470–472.
28. **Loh, T. P., L. L. Sievert, and R. W. Scott.** 1987. Proviral sequences that restrict retroviral expression. Mol. Cell. Biol. **7:**3775–3784.
29. **Lund, A. H., M. Duch, and F. S. Pedersen.** 1996. Transcriptional silencing of retroviral vectors. J. Biomed. Sci. **3:**365–378.
29a. **Macleod, D., J. Charlton, J. Mullins, and A. P. Bird.** 1994. Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. Genes Dev. **8:**2282–2292.
30. **Malik, P., S. A. McQuiston, X.-J. Yu, K. A. Pepper, W. J. Krall, G. M. Podsakoff, G. J. Kurtzman, and D. B. Kohn.** 1997. Recombinant adeno-associated virus mediates a high level of gene transfer but less efficient integration in the K562 human hematopoietic cell line. J. Virol. **71:**1284–1296.
31. **Markowitz, D., S. Goff, and A. Bank.** 1988. A safe packaging line for gene transfer: separating viral genes on two different plasmids. J. Virol. **62:**1120–1124.
32. **McCune, S. L., and T. M. Townes.** 1994. Retroviral vector sequences inhibit human β-globin gene expression in transgenic mice. Nucleic Acids Res. **22:**4477–4481.
33. **Miller, A. D.** 1996. Cell-surface receptors for retroviruses and implications for gene transfer. Proc. Natl. Acad. Sci. USA **93:**11407–11413.
34. **Miller, A. D., and G. J. Rosman.** 1989. Improved retroviral vectors for gene transfer and expression. BioTechniques. **7:**980–990.
35. **Miller, A. D., and C. Buttimore.** 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. Mol. Cell. Biol. **6:**2895–2902.
36. **Milot, E., P. Fraser, and F. Grosfeld.** 1996. Position effects and genetic disease. Trends Genet. **12:**123–126.
37. **Neumann, J. R., C. A. Morency, and K. O. Russian.** 1987. A novel rapid assay for chloramphenicol acetyltransferase gene expression. BioTechniques. **5:**444–447.
38. **Niwa, O., Y. Yokota, H. Ishida, and T. Sugahara.** 1983. Independent mechanisms involved in suppression of Moloney murine leukemia virus genome during differentiation of murine teratocarcinoma cells. Cell **32:**1105–1113.
39. **Ostertag, W., K. Vehemeyer, B. Fagg, I. B. Pragnell, W. Paetz, M. C. Le Bousse, F. Smadja-Joffe, B. Klein, C. Jasmin, and H. Eisen.** 1980. Myeloproliferative virus, a cloned murine sarcoma virus with spleen focus-forming properties in adult mice. J. Virol. **33:**573–582.
40. **Petersen, R., G. Kempler, and E. Barklis.** 1991. A stem cell-specific silencer in the primer-binding site of a retrovirus. Mol. Cell. Biol. **11:**1214–1221.
41. **Prince, V. E., and P. W. J. Rigby.** 1991. Derivatives of Moloney murine sarcoma virus capable of being transcribed in embryonal carcinoma stem cells have gained a functional Sp1 site. J. Virol. **65:**1803–1811.
42. **Rhazin, A., and H. Cedar.** 1993. DNA methylation and embryogenesis, p. 343–357. *In* J. P. Jost and H. P. Saluz (ed.), DNA methylation: molecular biology and biological significance. Birkhauser Verlag, Basel, Switzerland.
43. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
44. **Silke, J., K. I. Rother, O. Georgiev, W. Schaffner, and K. Matsuo.** 1995. Complex demethylation patterns at Sp1 binding sites in F9 embryonal carcinoma cells. FEBS Lett. **370:**170–174.
45. **Stewart, C. L., H. Stuhlmann, D. Jahner, and R. Jaenisch.** 1982. De novo methylation, expression, and infectivity of retroviral genomes into embryonal carcinoma cells. Proc. Natl. Acad. Sci. **79:**4098–4102.
46. **Stocking, C., R. Kollek, U. Bergholz, and W. Ostertag.** 1985. Long terminal repeat sequences impart hematopoietic transformation properties to the myeloproliferative sarcoma virus. Proc. Natl. Acad. Sci. USA **82:**5746–5750.
47. **Taketo, M., and D. J. Shaffer.** 1989. Deletion in a recombinant retrovirus genome associated with its expression in embryonal carcinoma cells. J. Virol. **63:**4431–4433.
48. **Tsukiyama, T., O. Niwa, and K. Yokoro.** 1989. Mechanism of suppression of the long terminal repeat of Moloney leukemia virus in mouse embryonal carcinoma cells. Mol. Cell. Biol. **9:**4670–4676.
49. **Walters, M. C., S. Fiering, J. Eidemiller, W. Magis, M. Groudine, and D. I. K. Martin.** 1995. Enhancers increase the probability but not the level of gene expression. Proc. Natl. Acad. Sci. USA **92:**7125–7129.
50. **Weiher, H., E. Barklis, W. Ostertag, and R. Jaenisch.** 1987. Two distinct sequence elements mediate retroviral gene expression in embryonal carcinoma cells. J. Virol. **61:**2742–2746.
51. **Wilson, C., H. J. Bellen, and W. J. Gehring.** 1990. Position effects on eukaryotic gene expression. Annu. Rev. Cell Biol. **6:**679–714.