

Improving Transcriptional Termination of Self-inactivating Gamma-retroviral and Lentiviral Vectors

Axel Schambach¹, Melanie Galla¹, Tobias Maetzig¹, Rainer Loew² and Christopher Baum^{1,3}

¹Department of Experimental Hematology, Hannover Medical School, Hannover, Germany; ²EUFETS AG, Idar-Oberstein, Germany;

³Division of Experimental Hematology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

Adverse events relating to insertional mutagenesis have reinforced the interest in self-inactivating (SIN) gamma-retroviral and lentiviral vectors without enhancer-promoter sequences in the U3 region of the long terminal repeats. However, SIN vectors suffer from leaky transcriptional termination, increasing the probability of read-through into cellular genes. To improve 3' end processing, we incorporated seven upstream polyadenylation enhancer elements (or upstream sequence elements, USEs) derived from viral or cellular genes into the 3' U3 region of gamma-retroviral and lentiviral SIN vectors. A 100-base-pair sequence representing a recombinant direct repeat of the USE derived from simian virus 40 (2xSV USE) gave the best results, improving both titer and gene expression. In both gamma-retroviral and lentiviral SIN vectors, the 2xSV USE partially substituted for effects provided by the much larger post-transcriptional regulatory element derived from woodchuck hepatitis virus (wPRE). By northern blot and reporter assays, we found that the 2xSV USE greatly improved proper messenger RNA (mRNA) processing at the retroviral termination signal. Importantly, the 2xSV USE was superior to the wPRE in suppressing transcriptional read-through, improving not only vector efficiency but potentially also biosafety.

Received 30 August 2006; accepted 19 February 2007; published online 3 April 2007. doi:10.1038/sj.mt.6300152

INTRODUCTION

Efficient 3' end processing is essential for eukaryotic messenger RNA (mRNA) processing. Correct termination of the nascent transcript involves cleavage at a predetermined site and subsequent polyadenylation.¹⁻³ The polyA motif is an AAUAAA sequence located 10–30 base pairs upstream of the cleavage site. Cleavage by cellular factors preferably occurs after a CA dinucleotide that is followed by a G/U-rich sequence. The 3' end is polyadenylated by the polyA polymerase, resulting in an average polyA tail of 200 nucleotides. Several molecules of the polyA binding protein II bind to the tail to enhance both mRNA stability and translation efficiency.^{2,4}

Retroviruses use mRNA not only as a translation template but also as the mobile form of genomic information. In both murine leukemia virus (a gamma-retrovirus) and human immunodeficiency virus (HIV, a lentivirus), the polyA motif is contained within the R region of the long terminal repeats (LTRs), implying its presence at both ends of the transcript. To prevent premature termination and polyadenylation in the 5' R region, retroviruses have evolved with relatively weak polyA sites and additional mechanisms of 5' polyA site suppression, as shown for HIV.⁵ Self-inactivating (SIN) vectors lack the majority of the U3 region.⁶⁻⁸ However, in the case of HIV this deletion increases the probability of read-through, suggesting that U3 contains termination enhancer motifs in addition to enhancer-promoter sequences.^{9,10}

Insufficient termination may cause read-through of randomly integrated retroviral sequences into cellular genes. This may have two problematic consequences. If the vector genome is stably integrated into the genome of retroviral packaging cells, read-through may promote the uptake of cellular proto-oncogenes into the vector genome, as exemplified by acutely transforming retroviruses. In transduced target cells, read-through into cellular genes may contribute to the up-regulation of cellular proto-oncogenes, potentially triggering malignant transformation. Although long-distance enhancer interactions are the most frequent form of retroviral insertional mutagenesis,^{11,12} we and others have previously demonstrated that read-through transcripts originating from retroviral vectors may induce leukemia.^{13,14} Read-through transcripts also represent a safety concern for lentiviral vectors, as these have an even higher propensity to insert within actively transcribed regions of the genome.¹⁵ Achieving improved mRNA termination and polyadenylation is thus of major importance for both gamma-retroviral and lentiviral vectors.

RESULTS

A duplicated element of the SV40 improves titer and expression of gamma-retroviral vectors

To improve mRNA termination of SIN vectors, we chose well-characterized upstream polyadenylation enhancer elements (or upstream sequence elements, USEs) derived from various viral or cellular genes: simian virus 40 (SV40) late mRNA,^{16,17} HIV type 1 (HIV-1),^{18,19} ground squirrel hepatitis virus (identical to

Correspondence: Christopher Baum, Experimental Hematology, Hannover Medical School, Carl-Neuberg-Strasse 1, D-30625 Hannover, Germany. E-mail: baum.christopher@mh-hannover.de

the sequence of woodchuck hepatitis virus, WHV),^{20–22} adenovirus,^{19,23,24} the human *THRB* gene (prothrombin, Thr),^{4,25} and the human C2 complement gene.^{26,27}

We inserted these auxiliary U-rich sequences (**Table 1**) into the 3' U3 deletion of a series of gamma-retroviral SIN vectors with three different internal promoters derived from the human cytomegalovirus (CMV), the human phosphoglycerate kinase gene (PGK), and the murine spleen focus-forming virus (SF) (**Figure 1a**). We produced replication-deficient vectors by co-transfection with packaging constructs. Depending on the configuration of the SIN vector, basic constructs lacking additional 3' RNA processing motifs generated titers in the order of $5 \times 10^5 - 5 \times 10^6$ infectious units per ml. The only USE that reproducibly enhanced retroviral titers (up to threefold) was that derived from SV40, especially when it was duplicated as a tandem repeat (2xSV) (**Supplementary Figure S1**). All subsequent expression analyses were normalized for equal transduction rates obtained under conditions of a low multiplicity of infection, as described earlier,⁸ and equivalent average copy number, as determined by real-time polymerase chain reaction (PCR) (data not shown). Depending on the promoter, the 2xSV USE increased gene expression by 45–100% ($P < 0.01$; **Figure 1b** and **c**). Increased expression was stable over time (**Figure 1c**). PCR performed in transduced cells revealed that the 2xSV USE was stably transmitted, without evidence of internal deletions (**Figure 1d**). Sequencing of the PCR products confirmed the correct transmission of the duplicated element (data not shown).

We next tested whether the 2xSV USE synergizes with the 3' end processing mediated by the post-transcriptional regulatory element of WHV (wPRE). This element increases titer and often gene expression of retroviral and lentiviral vectors, potentially by affecting RNA export, stability, and polyadenylation.²⁸ As shown in **Figure 1b** and **c**, the relatively small 2xSV USE (100 base pairs) was nearly as potent as the much larger wPRE (600 base pairs) in elevating titer and transgene expression. Combining the 2xSV USE and the wPRE did not significantly improve titer in comparison to the wPRE alone. In lineage-negative primary murine hematopoietic cells, both the 2xSV USE and the wPRE increased gene expression from SIN vectors with the internal SF promoter ($P < 0.01$; **Figure 2**, equal average copy number in transduced cells confirmed by real-time PCR). In this context, the 2xSV USE even tended to be stronger than the wPRE ($P = 0.08$).

The 2xSV USE thus enhanced titer and gene expression of different internal promoters and in various cell types. In terms of titer and expression, the effects were similar although not always

as potent as those achieved with the wPRE. The level of the effects depended on promoter and target cell type, as previously observed for the wPRE and other 3' processing motifs.²⁹

The USE improves transcriptional termination of gamma-retroviral SIN vectors

We next addressed how the USE improved 3' mRNA processing. We designed read-through reporter plasmids in which the 3' LTR of the retroviral construct is followed by a cassette consisting of an internal ribosomal entry site, the DsRedExpress fluorescent protein and a polyA motif derived from the SV40 early mRNA (**Figure 3a**). For northern blot analyses only cell pools with comparable transfection efficiency were used. Read-through transcripts from the unmodified SIN vector were easily detected by northern blotting (**Figure 3b**). The wPRE and the USE derived from HIV-1, Thr, or WHV partially attenuated transcriptional read-through. In contrast, the SV40 USE, especially when duplicated (2xSV), almost completely suppressed read-through. Combining the 2xSV USE with the retroviral polyA motif was almost as efficient as the introduction of the cellular termination and polyA motif of the bovine growth hormone (bGH) into the 3' R region. The latter strategy improved vector titer but not expression in target cells because the 3' R region is not preserved during reverse transcription.

Flow cytometry confirmed these results at the level of single cells. Presence of the 2xSV USE reduced the population that expressed the *IL2RG* complementary DNA (cDNA) encoded within the vector and the 3'-located DsRed allele at least tenfold ($P < 0.01$), similar to the strong polyA signal from bGH (**Figure 3c**). Again, the wPRE and WHV USE were much less efficient (~50% inhibition of read-through, **Figure 3c**), and the combination of the 2xSV USE and the wPRE was not superior to the 2xSV USE alone.

We next explored the effect of the 2xSV USE in the context of a gamma-retroviral reporter cassette stably integrated into the genome of target cells by lentiviral SIN vectors based on HIV-1.⁷ To avoid premature termination, we introduced the reporter cassette in antisense orientation (**Figure 3a**). To overcome low titers relating to the formation of antisense transcripts in packaging cells, we placed the reporter cassette under the control of a tetracycline-inducible promoter (scheme in **Figure 3a**). The target cells were co-transduced with an independent vector expressing the tetracycline-inducible transactivator. The data were not only consistent with the transient transfection studies, revealing that the 2xSV USE was sufficient to repress read-through even in the

Table 1 Upstream sequence element (USE) oligonucleotides (T highlighted in bold)

USE	Sequence	Length (bp)	T (%)
Adenovirus (L3)	ACTTCTTTT GT CACTTGAAC	21	48
C2	GACTTGACTCATGCTT GT TTCACTTTCACAT GGAATTTCCAGTTATGAAAT	53	40
HIV	AGCTGCTTTT GT CCGTACTGGGTCTCTCTGGTTA	35	43
SV40 (SV)	TTTATTT GT GAAATTT GT GATGCTATTGCTTTATTTGTAACCAT	44	52
<i>THRB</i> (Thr)	AGAATTATTTT GT GTTTCTA	21	57
WHV	TCATGTATCTTTT TC ACTGTGCCTT GT TTTTGCCTGTG	39	51

Abbreviations: bp, base pair; HIV, human immunodeficiency virus; SV40, simian virus 40; WHV, woodchuck hepatitis virus.

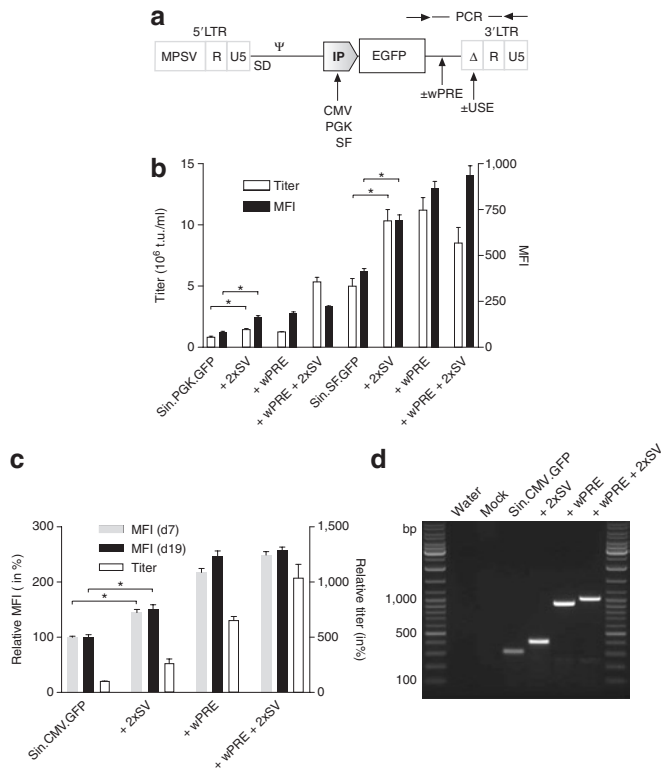


Figure 1 Performance of different upstream sequence elements (USEs) in gamma-retroviral vectors. **(a)** Modular vector design with the long terminal repeats (LTRs) (U3, R, U5), splice donor (SD), packaging signal (Ψ), internal promoter (IP), and transgene (enhanced green fluorescence protein, EGFP). In the proviral plasmid, the 5' LTR is driven by the myeloproliferative sarcoma virus (MPSV) enhancer–promoter. In self-inactivating (SIN) vectors the enhancer–promoter of the 3' U3 region was deleted and replaced with different upstream polyadenylation enhancer element (USE) sequences (see **Table 1**, Results, and **Supplementary Figure 1**). Optionally, the post-transcriptional regulatory element of woodchuck hepatitis virus (wPRE) was inserted into the 3' untranslated region (UTR). The transgene is expressed under the control of different internal promoters (SF; cytomegalovirus, CMV; phosphoglycerate kinase, PGK; see Results). **(b)** Effects of basic vector Sin.PGK/SF.GFP alone or with USE derived from simian virus 40 (2xSV USE), with wPRE, and with both USE and wPRE on titer (white columns) and gene expression as determined by mean fluorescence intensity (MFI, black columns) in transduced SC1 cells. The basic construct Sin.CMV.GFP harbors an internal CMV promoter and no wPRE. Analysis 5 days after transduction. Error bars indicate standard deviations of four independent experiments. **(c)** Relative effects of vectors encoding 2xSV, wPRE, and 2xSV plus wPRE in comparison with basic vector Sin.CMV.GFP (100%) on MFI (filled columns) and titer (white columns). MFI was analyzed 7 days (d7, gray columns) and 19 days (d19, black columns) after transduction ($n = 8$). **(d)** Polymerase chain reaction (PCR) from genomic DNA samples transduced with basic vector Sin.CMV.GFP, vector plus 2xSV, vector plus wPRE, and vector plus wPRE plus 2xSV (expected sizes 305, 405, 936, and 1036 base pairs, respectively) or untransduced (mock). Primers are directed against EGFP and the U5 region to confirm the presence of indicated elements (see **a**). DNA markers on the left and right. $*P < 0.01$. t.u., transducing unit.

presence of relatively high expression levels (**Figure 3d**); they also suggested that the amount of read-through from integrated transgenes depends on the expression level, which may be position dependent (**Figure 3d**).

The vectors used in these studies expressed the *IL2RG* cDNA (CD132), which is of interest for gene therapy for X-linked severe combined immunodeficiency.¹⁴ These data suggest that the 2xSV

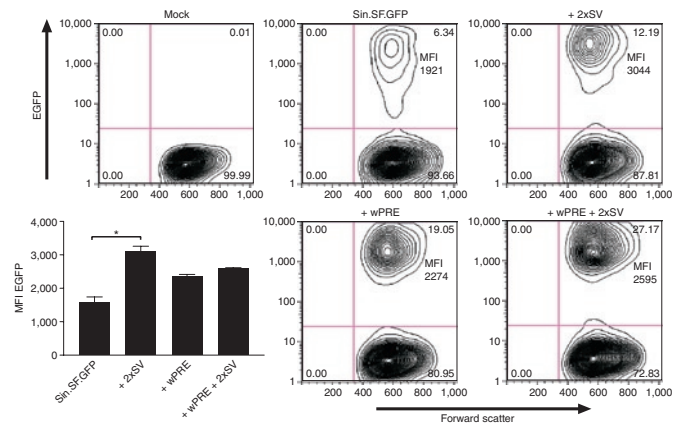


Figure 2 Effect of upstream polyadenylation enhancer elements derived from simian virus 40 (2xSV USE) and/or post-transcriptional regulatory element of woodchuck hepatitis virus (wPRE) in transduced primary murine hematopoietic cells. The lineage marker–depleted cell populations were analyzed 8 days after transduction by flow cytometry. Representative dot plots (x -axis, forward scatter; y -axis, enhanced green fluorescence protein, EGFP) are shown for untransduced cells (mock) and cells transduced with the vector Sin.SF.GFP, vector plus 2xSV, vector plus wPRE, or vector with both elements combined (wPRE and 2xSV). The upper-right quadrant shows the mean fluorescence intensity (MFI). Results of three experiments are summarized in the lower-left panel. Quantitative polymerase chain reaction revealed comparable vector copy number in all samples (data not shown). $*P < 0.01$.

USE suppresses read-through transcription in the context of clinically relevant vectors.

The USE improves transcriptional termination and polyadenylation of lentiviral SIN vectors

We next tested whether the 2xSV USE also enhanced titer and expression of HIV-based SIN vectors⁷ if present in their LTRs. These vectors had the same internal cassettes (internal promoters CMV, PGK, SF) as the gamma-retroviral counterparts (**Figure 4a**; cf. **Figure 1a**) and were co-transfected with HIV gag/pol, Rev, and murine leukemia virus ecotropic env expression plasmids. Again, the 2xSV USE increased titer (up to 2.8-fold, $P < 0.01$) and gene expression (1.6–2.2-fold, $P < 0.01$) in transduced target cells. Interestingly, in the lentiviral context, the effects of the 2xSV USE on titer and expression were even closer to those of the wPRE (**Figure 4b**).

Of note, the lentiviral vectors contained the same plasmid sequences downstream of the 3' LTR as used in the gamma-retroviral SIN vectors. Initial attempts to utilize the 2xSV USE in the context of the third-generation lentiviral vector plasmids as described by Dull *et al.*⁷ resulted in an unexpected loss of infectious titer. This may have been due to the interaction of the 2xSV USE with a second polyA motif located downstream of the 3' LTR in these constructs, which is not found in the gamma-retroviral plasmids.

Results achieved by transient transfection of reporter constructs suggested that the 2xSV USE also suppressed transcriptional read-through of lentiviral SIN vectors (**Figure 5**). Within the population expressing the vector-encoded cDNA (*IL2RG*), the frequency of cells co-expressing DsRed was 20% in the unmodified vector, 8% in the presence of the wPRE, and 3% in the presence of the 2xSV USE. Again, the combination of the wPRE and the 2xSV USE further suppressed read-through (levels $< 1.5\%$;

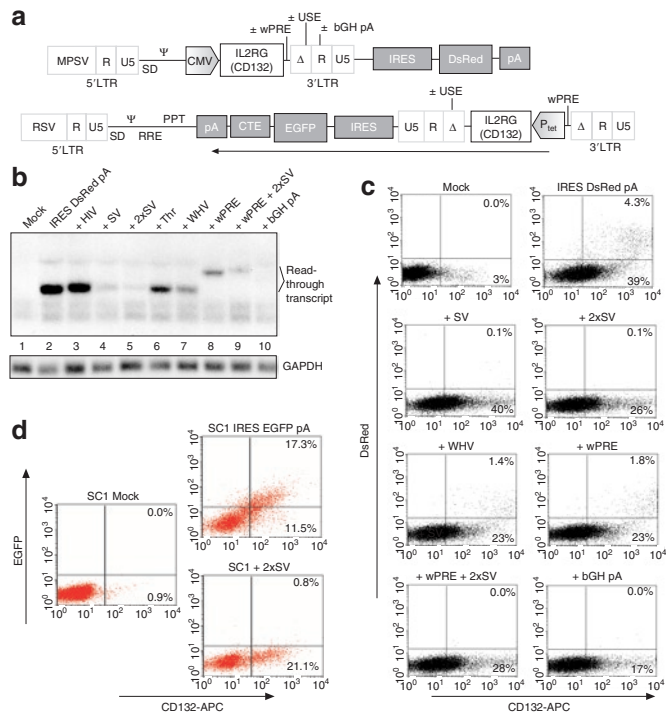


Figure 3 Detection and prevention of read-through transcription. **(a)** Vector diagrams for read-through reporter constructs for transient transfection (upper) and target cell transduction (lower). The upper panel shows a gamma-retroviral self-inactivating (SIN) vector with the internal cytomegalovirus (CMV) promoter driving the interleukin-2 receptor gamma chain (IL2RG, CD132). Behind the 3' LTR an expression cassette consisting of an internal ribosomal entry site (IRES), DsRed Express complementary DNA, and an early simian virus 40 (SV40) polyA (pA) signal was introduced. The lower panel shows a lentiviral vector with a tet-inducible (P_{tet}) antisense cassette consisting of IL2RG, 3' gamma-retroviral SIN LTR, IRES, enhanced green fluorescence protein (EGFP), a constitutive transport element, and SV40 pA. Here the complete detection cassette is maintained in transduced cells. **(b)** Northern blot of total RNA prepared from transfected 293T cells. The basic vector construct from **a**, termed IRES Δ SRedpA, lacks upstream polyadenylation enhancer elements (upstream sequence elements, USEs) and the post-transcriptional regulatory element of woodchuck hepatitis virus (WHV) (wPRE). Different USEs (HIV, SV, 2xSV, Thr, WHV) and/or the wPRE were added as indicated. As a control the strong pA signal from bovine growth hormone (bGH pA) was introduced into the R region instead of the original retroviral pA. To detect the read-through transcript (marked on the right), a radioactive probe was directed against DsRed. GAPDH served as a loading control. **(c)** Transfected 293T cells (mRNA shown in **Figure 3b**; constructs shown in **Figure 3a**) were analyzed 3 days after transfection. The cells were stained for CD132 (IL2RG) using a biotinylated anti-CD132 antibody and a streptavidin- α -allophycocyanin (APC) antibody. The dot blots indicate cells with read-through transcripts in the upper-right quadrant (with percentage of upper-right and lower-right quadrants). **(d)** Read-through detection in transduced SC1 cells measured 4 days after transduction. Analysis is performed as in **c** but using EGFP as the read-through marker.

mean values of two independent experiments). **Figure 5** shows results from one of these experiments.

Polyadenylation in retroviral vectors containing the 2xSV USE

To determine whether the USE increased the polyA length, we analyzed vector RNA on high-resolution gels. We transfected 293T cells with equal amounts of plasmids designed as shown in

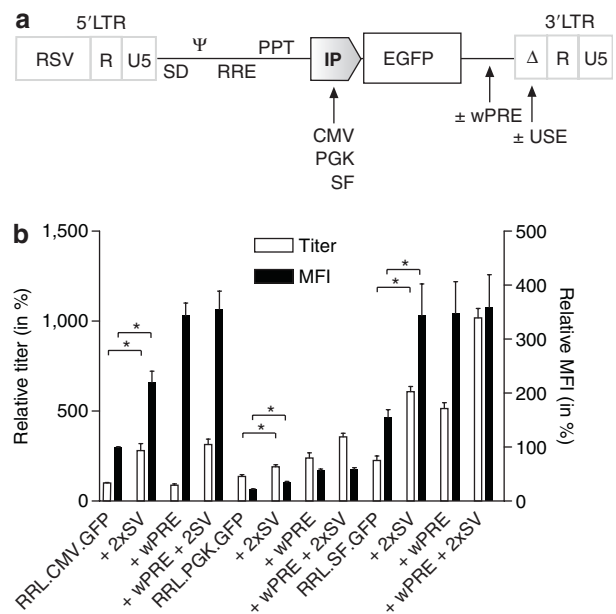


Figure 4 Upstream polyadenylation enhancer elements (upstream sequence elements, USEs) and post-transcriptional regulatory element of woodchuck hepatitis virus (wPRE) in lentiviral vectors. **(a)** Scheme of the lentiviral vectors used with 5' long terminal repeat (LTR) (RSV U3, R, U5), splice donor (SD), packaging signal (Ψ), rev-responsive element (RRE), central polypurine tract (PPT), internal promoter (IP: cytomegalovirus, CMV; phosphoglycerate kinase gene, PGK; or SF), transgene (enhanced green fluorescence protein, EGFP), and 3' LTR (Δ U3, R, U5). The wPRE and/or USEs were incorporated into the 3' UTR and the Δ U3 deletion, respectively. **(b)** Relative titer and mean fluorescence (percentage) in transduced SC1 cells 5 days after transduction, tested for three IPs (CMV, PGK, SF) ($n = 6$). The basic construct RRL.CMV.GFP was set to 1. Titer is indicated in white columns; mean fluorescence intensity (MFI) in black columns. * $P < 0.01$. SV, simian virus.

Figure 1a (encoding enhanced green fluorescent protein, EGFP, under the control of the internal CMV promoter). The results shown in **Figure 6** suggest that the USE increases the amount of polyadenylated vector transcript and the proportion of mRNAs with longer polyA tails (average length of at least 200 nucleotides). Similar results were obtained by northern blot analyses of mRNA harvested from retrovirally transduced unselected mass cultures (data not shown). The latter analyses also showed that the 2xSV USE increased the levels of mRNA originating from the internal promoter (after correction for transduction efficiency, $n = 4$; **Figure 6b**). We thus found that vector mRNA levels and mean expression levels of EGFP were closely correlated (**Figure 6b**).

Together, these data reveal that the 2xSV USE acts as a termination enhancer (**Figures 4** and **5**), may improve the recognition of the retroviral polyA signal (**Figure 6a**), and increases the amount of mRNA originating from the internal promoter of SIN vectors (**Figure 6b**). The latter effect may be sufficient to explain the gain of protein expression obtained with this element.

DISCUSSION

This study shows that USE motifs improve the efficiency of 3' end mRNA processing in gamma-retroviral and lentiviral SIN vectors. As a practical consequence, the resulting vectors showed

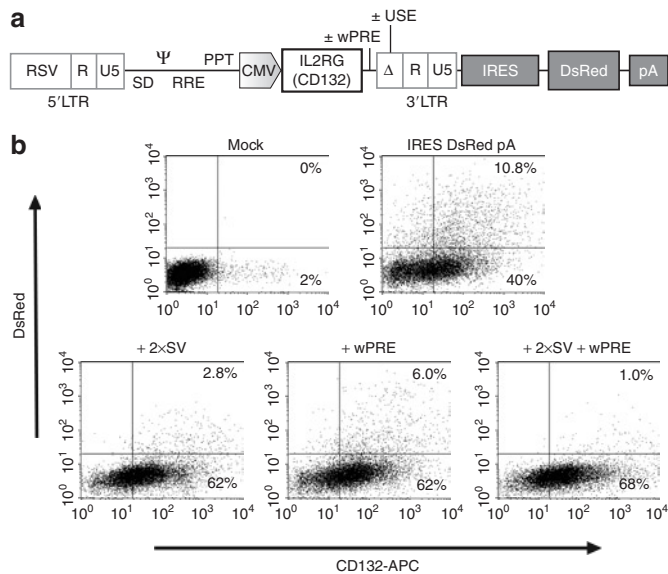


Figure 5 Read-through suppression in lentiviral vectors. **(a)** Scheme of the read-through reporter construct, as in **Figure 3a**, but using a third-generation lentiviral vector backbone. **(b)** Fluorescence-activated cell sorting analysis of lentiviral vectors (as in **Figure 3c**). The upper-right quadrant is indicative of read-through transcripts (with percentage of populations in upper-right and lower-right quadrants). APC, allophycocyanin; CMV, cytomegalovirus; IRES, internal ribosomal entry site; LTR, long terminal repeat; pA, polyA; PPT, polypurine tract; RRE, rev-responsive element; SD, splice donor; SV, simian virus; USEs, upstream polyadenylation enhancer elements; wPRE, post-transcriptional regulatory element of woodchuck hepatitis virus.

improved titer and transgene expression in target cells. The mechanism underlying the increase in titer remains to be elucidated. It could be explained at the level of mRNA stability in packaging cells, uptake of mRNA in particles (packaging), or reverse transcription, the late steps of which are target cell dependent. Increased transgene expression in target cells was closely correlated with increased amounts of mRNA. This could be explained by improved termination, which we found to occur both after transient transfection and after stable, semi-random transgene integration. In addition, we obtained results that suggest that the incorporation of USE motifs into SIN LTRs increases the length of the polyA tail, which may further support translation.¹

The importance of USE motifs in regulating termination in the presence of a suboptimal cleavage site is reminiscent of corresponding signals in the human *THRB* gene, mutations of which may give rise to hereditary forms of thrombophilia.^{4,25} The fact that the SV40 USE and especially its duplicated version mediated the strongest enhancement of retroviral termination suggests that these motifs facilitate the binding of crucial cellular factors such as cleavage and polyadenylation specificity factor,³⁰ its subunit Fip1,²⁴ the U1 small nuclear ribonucleoprotein A,^{31,32} the polypyrimidine tract-binding protein, and cleavage stimulatory factor.²⁷ Interestingly, the SV40 USE core sequence AUUUGURA is also found in the upstream polyadenylation regions of ground squirrel hepatitis virus²² and cauliflower mosaic virus,³³ arguing for a conserved mechanism.

Efficient 3' processing of HIV-1 depends on an RNA stem-loop structure (TAR) that juxtaposes the USE and the core polyA

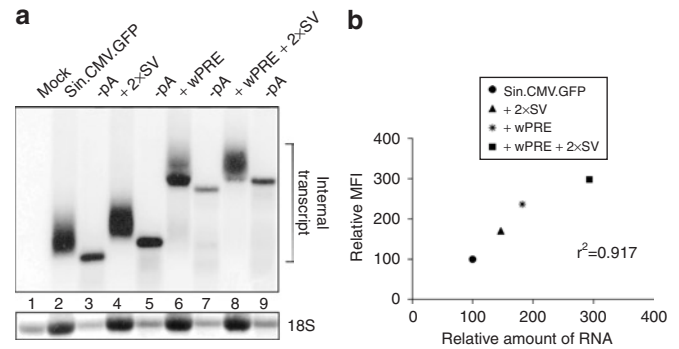


Figure 6 (a) Detection of polyA (pA) tail length after transfecting 293T cells with the vector derivatives of Sin.CMV.GFP. An aliquot of the RNA was treated with oligo(dT) primers and RNase H to digest the pA tail, so that adenylated and de-adenylated RNA could be viewed side by side in a northern blot, specific probe directed against enhanced green fluorescence protein (EGFP). As a loading control, 18S RNA was used. **(b)** SC1 cells were transduced with the indicated vectors, and RNA and protein expression (mean fluorescence intensity, MFI) were analyzed at the same time point ($n = 4$). For RNA detection northern blots were probed for EGFP, re-probed for 18S (loading control), and analyzed using a phosphorimager. EGFP expression was monitored by flow cytometry as previously mentioned. The basic vector Sin.CMV.GFP was set to 100% and the relative amounts of RNA and MFI are displayed as a mean of four independent experiments. CMV, cytomegalovirus; SV, simian virus; wPRE, post-transcriptional regulatory element of woodchuck hepatitis virus.

site.³⁰ Similar stable stem-loops may be formed upstream of the polyA site in the late SV40 mRNA¹⁶ and by the R region of murine leukemia virus,³⁴ although the relevance for termination remains to be shown. Similarly, in hepatitis B viruses (HBVs, ground squirrel hepatitis virus, WHV) the rather weak polyA signal UAU AAA (usually AAUAAA) is functional only in the presence of the USE²² and, again, a stem-loop structure is predicted by computer algorithms as an intervening sequence between the USE and the polyA site. These findings set the stage for rational approaches to further improve 3' end processing of vector mRNA.

As shown in this study, the wPRE further increased titer and gene expression even in the presence of the USE (**Figure 1b** and **c** for gamma-retroviral vectors; **Figure 4** for lentiviral vectors). Our results suggest that the major effect of the wPRE is to improve mRNA levels (**Figure 6b**), which can hardly be explained by improved utilization of the retroviral termination and polyadenylation signal (**Figures 3, 5** and **6a**). In line with this observation, a distinct USE isolated from the same viral mRNA from which the wPRE is derived increased the rate of correct termination (WHV USE, **Figure 3c**). This suggests that WHV carries at least two distinct elements to improve 3' mRNA processing, namely the wPRE and the WHV USE, which, when used alone, are not sufficient to prevent read-through. The exact mechanisms and underlying sequence motifs by which the wPRE improves mRNA processing (e.g., mRNA export or stability) remain to be determined. Their determination could lead to designer modules for post-transcriptional enhancement of retroviral titer and gene expression.

A final question is whether the suppression of transcriptional read-through into downstream alleles may improve vector biosafety during both production and target cell modification.

During vector production, proper termination prevents the uptake of cellular sequences into the retroviral particle, thus improving product identity. In transduced cells used for therapy, the increase in expression by improved 3' processing of vector mRNA may allow the use of weaker transcriptional enhancer motifs, thus reducing the danger of insertional mutagenesis. In the not-too-rare cases where vector insertion may give rise to fusion transcripts,^{12,13,35} proper termination may prevent proto-oncogene up-regulation. There is only one hypothetical scenario in which improved termination may reduce vector biosafety: integration into the intron of a transcription unit whose residual expression of full-length transcript is monoallelic and has a tumor-suppressive role. However, database searches suggest that this scenario is unlikely.^{11,12} It remains to be determined whether the impact of the USE on insertional mutagenesis can be revealed by recently introduced assay systems that report insertional side effects of integrating vectors.^{36,37}

A clear advantage of the USE sequences is their small size, leaving the packaging capacity of retroviral vectors largely unaffected. The USE can be used alone or in combination with the wPRE sequence, which may synergize with the USE to enhance titer and transgene expression further. Tuning the size, sequence, and location of the USE may result in even stronger effects on termination, titer, and expression, potentially achieved by even more compact elements. USEs thus represent an important addition to the toolbox of vector design.

MATERIALS AND METHODS

Vector design. The gamma-retroviral SIN vector pSin.CMV.GFP⁸ contains an internal CMV promoter and the EGFP cDNA, and an *EcoRI* site in the residual U3 region of the 3' LTR. The USEs originating from adenovirus (L3, adeno), C2 complement gene (C2), HIV-1 (HIV), SV40 late mRNA (SV), THRβ gene (Thr), and WHV were generated using oligonucleotide self-assembly with flanking *EcoRI* sites and sequenced after introduction into the SIN vector. For the USEs of adenovirus and SV40, duplicated elements (2 × adeno, 2xSV) were constructed (Table 1).

To detect read-through transcripts, a cassette consisting of an encephalomyocarditis virus internal ribosomal entry site, DsRed Express cDNA, and early SV40 mRNA polyA signal (both from Clontech, Mountain View, CA) was inserted into the *XhoI* site behind the 3' LTR. In addition, we replaced EGFP with the human IL2RG cDNA using *AgeI* and *SaII*. The basic lentiviral construct pRRL.PPT.PGK.GFPpre was kindly provided by L. Naldini (San Raffaele Telethon Institute for Gene Therapy, Milano, Italy).⁷ In pRRL.PPT.CMV the PGK promoter was substituted for the CMV promoter, and variants with or without the wPRE were cloned. By overlap PCR, a multiple cloning site harboring restriction sites *BamHI*, *XhoI*, *EcoRI*, and *ClaI* was introduced into the 3' LTR ΔU3, and the 2xSV USE was cloned into the *EcoRI* site. To exclude possible interference with the SV40 polyA/ORI element in the plasmid backbone (see Results), this was replaced with the pUC19 backbone after introducing *AflIII* and *XhoI* sites before the 5' LTR and after the 3' LTR, respectively. In analogy to the gamma-retroviral vectors, an internal ribosomal entry site.DsRedExpress.polyA cassette was inserted into the *XhoI* site behind the 3' LTR and EGFP was replaced with IL2RG for the read-through assay. As a positive control for an optimal polyA signal, the bGH polyA was amplified using primers 5'bGHpAKpn (5'-GCATGTACCCTAGAGCTCGCTGATCAGCCT-3') and 3'bGHpAKpn (5'-CGATGGTACCTCCCAGCATGCCTGCTATT-3') and cloned into the *KpnI* site upstream of the retroviral polyA motif.

For detection of read-through in transduced cells, the DNA cassette consisting of IL2RG cDNA, 3' LTR (±2xSV), internal ribosomal entry site, DsRed, and SV40 polyA was inserted as an antisense cassette into the lentiviral vector backbone pRRL.PPT.CMV.GFPpre in the *HpaI* and *SaII* sites. The DsRed sequence was replaced with EGFP and constitutive transport element. Details of the construct shown in Figure 3a are available on request.

Production of vector supernatants. Retroviral and lentiviral supernatant production was performed using Phoenix gp (kindly provided by G. Nolan, Stanford) or 293T cells as previously described.^{8,38} Phoenix gp, 293T, and SC1 (murine fibroblast) cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 100 U/ml penicillin/streptomycin, and 2 mmol/l glutamine.

Genomic PCR. To confirm stability of the 2xSV elements in integrated proviruses, we isolated the genomic DNA (DNA Blood Mini Kit; Qiagen, Hilden, Germany) and performed a PCR on the genomic DNA of transduced SC1 samples using primers 5'EGFP.Cterm (5'-CATGGTCCTGCTGGAGTTCGTG-3') and 3'U5 (5'-GGAGACCCTCCCAAGGAT-3'), *Taq* polymerase, and standard conditions (annealing temperature 59°C).

Quantitative PCR. Real-time PCR was performed on an Applied Biosystems 7300 Real Time PCR System (Foster City, CA) using the QuantiTect SYBR Green Kit (Qiagen, Hilden, Germany) as previously described.³⁶ To measure vector copy numbers, PCR was performed using EGFP-specific [present on vector; 5'EGFP-RT2 (5'-CTATATCATGGCCGACAA GCAGA-3') and 3'EGFP-RT2 (5'-GGACTGGGTGCTCAGGTAGTGG-3')] and *Flk1*-specific (fetal liver kinase gene 1, housekeeping gene) primers. Results were analyzed using the ΔΔCT method.

Northern blots. Standard procedures were used for northern blots.⁸ Specific probes were directed against EGFP, DsRed, GAPDH, or 18S. For the detection of polyA tail length, agarose gels were run 4 hours longer to resolve the polyA tail, using a modified protocol from Holtmann and colleagues.³⁹ De-adenylated transcripts were prepared *in vitro* by mixing total RNA with Oligo(dT) (0.5 μg for every 10 μg of total RNA) in buffer containing 200 mM KCl and 1 mM EDTA (pH 8.0). The samples were incubated at 90°C for 2 minutes and then annealed at 25°C for 10 minutes. Two volumes of a solution containing 100 mM KCl, 0.5 mM EDTA (pH = 8.0), 20 mM Tris-HCl (pH = 8.0), 28 mM MgCl₂, 20 U of RNase inhibitor, and 1 U of RNase H were added to de-adenylate the transcript. The sample was incubated at 37°C for 30 minutes and precipitated with ethanol.

Primary cells. Bone marrow was harvested from C57BL/6 mice obtained from Charles River Laboratories (Sulzfeld, Germany). Lineage-depleted bone marrow cells were cultured in Stem Span medium (Stem Cell Technologies, London, UK) supplemented with 200 U/ml penicillin/streptomycin, 50 ng/ml murine stem cell factor, 100 ng/ml hFlt3L, 100 ng/ml hIL11, and 20 ng/ml mL3 for 2 days.⁴⁰ Transduction of cultured cells was performed on non-tissue-culture-treated 12-well plates coated with Retronectin (Takara, Shiga, Otsu, Japan) using a multiplicity of infection of 3.

Flow cytometry. SC1 or whole bone marrow cells (1–5 × 10⁵) were harvested for analysis 5 days after transduction and washed with phosphate-buffered saline. IL2RG staining was performed using a biotinylated anti-CD132 antibody, followed by a streptavidin-allophycocyanin antibody (both from BD Pharmingen, Heidelberg, Germany). Samples were analyzed for CD132, EGFP, or DsRed in a FACSCalibur using Cell Quest software (Becton-Dickinson, Heidelberg, Germany). A gate was set on a homogeneous cell population, as determined by scatter characteristics, and 20,000 events were monitored. Marker gates were set to calculate the percentage and mean fluorescence intensity of positive cells.

Statistical analysis. Data from experiments are expressed as mean \pm SD. Student's paired *t*-test was used for comparison of differences between indicated groups. *P* < 0.05 was considered significant.

ACKNOWLEDGMENTS

This work was supported by grants from the European Union (CONCERT, LSHB-CT-2004-005242, to C.B.) and the Else-Kröner-Stiftung (to A.S.). T.M. received a stipend from the German National Merit Foundation (Studienstiftung des Deutschen Volkes). We thank Ute Modlich and Sabine Knoess for help with real-time PCR, Maimona Id for technical assistance, and Helmut Holtmann and Jens Bohne (all from Hannover Medical School, Germany) for helpful comments and suggestions.

SUPPLEMENTARY MATERIAL

Figure S1. Screening of different USEs for enhancement of titer and mean fluorescence intensity.

REFERENCES

- Colgan, DF and Manley, JL (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev* **11**: 2755–2766.
- Zhao, J, Hyman, L and Moore, C (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **63**: 405–445.
- Proudfoot, N and O'Sullivan, J (2002). Polyadenylation: a tail of two complexes. *Curr Biol* **12**: R855–R857.
- Danckwardt, S, Hartmann, K, Gehring, NH, Hentze, MW and Kulozik, AE (2006). 3' end processing of the prothrombin mRNA in thrombophilia. *Acta Haematol* **115**: 192–197.
- Furger, A, Monks, J and Proudfoot, NJ (2001). The retroviruses human immunodeficiency virus type 1 and Moloney murine leukemia virus adopt radically different strategies to regulate promoter-proximal polyadenylation. *J Virol* **75**: 11735–11746.
- Yu, SF, von Ruden, T, Kantoff, PW, Garber, C, Seiberg, M, Ruther, U *et al.* (1986). Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc Natl Acad Sci USA* **83**: 3194–3198.
- Dull, T, Zufferey, R, Kelly, M, Mandel, RJ, Nguyen, M, Trono, D *et al.* (1998). A third-generation lentivirus vector with a conditional packaging system. *J Virol* **72**: 8463–8471.
- Schambach, A, Bohne, J, Chandra, S, Will, E, Margison, GP, Williams, DA *et al.* (2006). Equal potency of gammaretroviral and lentiviral SIN vectors for expression of O6-methylguanine-DNA methyltransferase in hematopoietic cells. *Mol Ther* **13**: 391–400.
- Zaiss, AK, Son, S and Chang, LJ (2002). RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. *J Virol* **76**: 7209–7219.
- Gilmartin, GM, Fleming, ES and Oetjen, J (1992). Activation of HIV-1 pre-mRNA 3' processing *in vitro* requires both an upstream element and TAR. *EMBO J* **11**: 4419–4428.
- Uren, AG, Kool, J, Berns, A and van Lohuizen, M (2005). Retroviral insertional mutagenesis: past, present and future. *Oncogene* **24**: 7656–7672.
- Kustikova, O, Geiger, H, Li, Z, Brugman, MH, Chambers, SM, Shaw, CA *et al.* (2006). Retroviral vector insertion sites associated with dominant hematopoietic clones mark "stemness" pathways. *Blood* **109**: 1897–1907.
- Li, Z, Dullmann, J, Schiedlmeier, B, Schmidt, M, von Kalle, C, Meyer, J *et al.* (2002). Murine leukemia induced by retroviral gene marking. *Science* **296**: 497.
- Hacein-Bey-Abina, S, Von Kalle, C, Schmidt, M, McCormack, MP, Wulffraat, N, Leboulch, P *et al.* (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415–419.
- Wu, X, Li, Y, Crise, B and Burgess, SM (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**: 1749–1751.
- Schek, N, Cooke, C and Alwine, JC (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.
- Carswell, S and Alwine, JC (1989). Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. *Mol Cell Biol* **9**: 4248–4258.
- Valsamakis, A, Zeichner, S, Carswell, S and Alwine, JC (1991). The human immunodeficiency virus type 1 polyadenylation signal: a 3' long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. *Proc Natl Acad Sci U S A* **88**: 2108–2112.
- DeZazzo, JD, Falck-Pedersen, E and Imperiale, MJ (1991). Sequences regulating temporal poly(A) site switching in the adenovirus major late transcription unit. *Mol Cell Biol* **11**: 5977–5984.
- Cherrington, J, Russnak, R and Ganem, D (1992). Upstream sequences and cap proximity in the regulation of polyadenylation in ground squirrel hepatitis virus. *J Virol* **66**: 7589–7596.
- Russnak, R and Ganem, D (1990). Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes Dev* **4**: 764–776.
- Russnak, RH (1991). Regulation of polyadenylation in hepatitis B viruses: stimulation by the upstream activating signal PS1 is orientation-dependent, distance-independent, and additive. *Nucleic Acids Res* **19**: 6449–6456.
- Prescott, J and Falck-Pedersen, E (1994). Sequence elements upstream of the 3' cleavage site confer substrate strength to the adenovirus L1 and L3 polyadenylation sites. *Mol Cell Biol* **14**: 4682–4693.
- Kaufmann, I, Martin, G, Friedlein, A, Langen, H and Keller, W (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J* **23**: 616–626.
- Danckwardt, S, Gehring, NH, Neu-Yilik, G, Hundsdoerfer, P, Pforisch, M, Frede, U *et al.* (2004). The prothrombin 3' end formation signal reveals a unique architecture that is sensitive to thrombophilic gain-of-function mutations. *Blood* **104**: 428–435.
- Moreira, A, Wollerton, M, Monks, J and Proudfoot, NJ (1995). Upstream sequence elements enhance poly(A) site efficiency of the C2 complement gene and are phylogenetically conserved. *EMBO J* **14**: 3809–3819.
- Moreira, A, Takagaki, Y, Brackenridge, S, Wollerton, M, Manley, JL and Proudfoot, NJ (1998). The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms. *Genes Dev* **12**: 2522–2534.
- Hope, T (2002). Improving the post-transcriptional aspects of lentiviral vectors. *Curr Top Microbiol Immunol* **261**: 179–189.
- Schambach, A, Wodrich, H, Hildinger, M, Bohne, J, Krausslich, HG and Baum, C (2000). Context dependence of different modules for posttranscriptional enhancement of gene expression from retroviral vectors. *Mol Ther* **2**: 435–445.
- Gilmartin, GM, Fleming, ES, Oetjen, J and Graveley, BR (1995). CPSF recognition of an HIV-1 mRNA 3'-processing enhancer: multiple sequence contacts involved in poly(A) site definition. *Genes Dev* **9**: 72–83.
- Lutz, CS and Alwine, JC (1994). Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev* **8**: 576–586.
- Lutz, CS, Murthy, KG, Schek, N, O'Connor, JP, Manley, JL and Alwine, JC (1996). Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency *in vitro*. *Genes Dev* **10**: 325–337.
- Sanfacon, H, Brodmann, P and Hohn, T (1991). A dissection of the cauliflower mosaic virus polyadenylation signal. *Genes Dev* **5**: 141–149.
- Trubetskoy, AM, Okenquist, SA and Lenz, J (1999). R region sequences in the long terminal repeat of a murine retrovirus specifically increase expression of unspliced RNAs. *J Virol* **73**: 3477–3483.
- Kustikova, O, Fehse, B, Modlich, U, Yang, M, Dullmann, J, Kamino, K *et al.* (2005). Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science* **308**: 1171–1174.
- Modlich, U, Bohne, J, Schmidt, M, von Kalle, C, Knoss, S, Schambach, A *et al.* (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood* **108**: 2545–2553.
- Montini, E, Cesana, D, Schmidt, M, Sanvito, F, Ponzone, M, Bartholomae, C *et al.* (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol* **24**: 687–696.
- Schambach, A, Galla, M, Modlich, U, Will, E, Chandra, S, Reeves, L *et al.* (2006). Lentiviral vectors pseudotyped with murine ecotropic envelope: increased biosafety and convenience in preclinical research. *Exp Hematol* **34**: 588–592.
- Winzen, R, Gowrishankar, G, Bollig, F, Redich, N, Resch, K and Holtmann, H (2004). Distinct domains of AU-rich elements exert different functions in mRNA destabilization and stabilization by p38 mitogen-activated protein kinase or HuR. *Mol Cell Biol* **24**: 4835–4847.
- Li, Z, Schwieger, M, Lange, C, Kraunus, J, Sun, H, van den Akker, E *et al.* (2003). Predictable and efficient retroviral gene transfer into murine bone marrow repopulating cells using a defined vector dose. *Exp Hematol* **31**: 1206–1214.