## Title

Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter



## Journal

Nucleic Acids Research, 43(1)

## ISSN

0305-1048

## Authors

Cooper, Aaron R
Lill, Georgia R
Gschweng, Eric H
et al.

## Publication Date

2015-01-09

## DOI

10.1093/nar/gku1312

682–690 Nucleic Acids Research, 2015, Vol. 43, No. 1 Published online 17 December 2014
doi: 10.1093/nar/gku1312

# Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter

**Aaron R. Cooper[1], Georgia R. Lill[2], Eric H. Gschweng[2] and Donald B. Kohn[2,3,4,*]**

[1]Molecular Biology Interdepartmental PhD Program, University of California, Los Angeles, Los Angeles, CA 90095, USA, [2]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA, [3]Department of Pediatrics, University of California, Los Angeles, Los Angeles, CA 90095, USA and [4]Molecular Biology Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA

Received July 31, 2014; Revised December 03, 2014; Accepted December 04, 2014

## ABSTRACT

**Lentiviral vectors almost universally use heterologous internal promoters to express transgenes. One of the most commonly used promoter fragments is a 1.2-kb sequence from the human ubiquitin C (*UBC*) gene, encompassing the promoter, some enhancers, first exon, first intron and a small part of the second exon of *UBC*. Because splicing can occur after transcription of the vector genome during vector production, we investigated whether the intron within the UBC promoter fragment is faithfully transmitted to target cells. Genetic analysis revealed that more than 80% of proviral forms lack the intron of the UBC promoter. The human elongation factor 1 alpha (*EEF1A1*) promoter fragment intron was not lost during lentiviral packaging, and this difference between the UBC and EEF1A1 promoter introns was conferred by promoter exonic sequences. UBC promoter intron loss caused a 4-fold reduction in transgene expression. Movement of the expression cassette to the opposite strand prevented intron loss and restored full expression. This increase in expression was mostly due to non-classical enhancer activity within the intron, and movement of putative intronic enhancer sequences to multiple promoter-proximal sites actually repressed expression. Reversal of the UBC promoter also prevented intron loss and restored full expression in bidirectional lentiviral vectors.**

## INTRODUCTION

The HIV-1-based lentiviral vector (LV) is one of the most common tools used for genetic modifications in biological experiments and in gene therapy. Most LVs used are self-inactivating, meaning that the region within the long terminal repeat containing the promoter and enhancers has been removed (1). In order to express a transgene within such a vector, a promoter must therefore be placed within the vector payload along with the transgene. Typically, in order to express a protein-coding gene, a heterologous RNA Pol II viral or cellular promoter will be used, and common examples are viral promoters from cytomegalovirus, murine leukaemia virus, and spleen focus-forming virus, and cellular promoters from human genes such as elongation factor 1 alpha (*EEF1A1*), ubiquitin C (*UBC*) and phosphoglycerate kinase (*PGK1*) (2,3).

During the viral production process, RNA Pol II transcribes the vector genome, typically from a transfer plasmid that has been transfected into the producer cells. Virtually all systems incorporate the Rev protein from HIV-1, which binds to the Rev response element (RRE) within the HIV-1 genome and mediates splicing-independent nuclear export of the viral genome. Despite the incorporation of the RRE sequence into LV constructs, however, introns within the vector payload can be lost during packaging if the splicing event retains the packaging signal (Psi) in the transcript. With some expression cassettes, though, such as one including the intron-containing promoter of EEF1A1 and one containing the hybrid CAG promoter, intron loss has not been observed during lentiviral packaging (4,5). From these observations, it has sometimes been inferred that lentiviral gene transfer allows for the transmission of introns (6).

We set out to investigate whether the intron contained by the human UBC promoter is faithfully transmitted from a transfer plasmid through to proviral forms in stably transduced cells. We hypothesized that a loss of the UBC intron would result in a significant reduction in transgene expression, as the UBC intron has been reported to possess strong enhancer activity (7). In contrast to previous findings with the EEF1A1 intron, the UBC intron was found to be missing in the majority of proviral forms in cells transduced with vector produced from intron-containing plasmids. The lack

*To whom correspondence should be addressed. Tel: +1 310 794 1964; Fax: +1 310 206 0356; Email: dkohn1@mednet.ucla.edu

of the UBC intron resulted in a roughly 2-fold decrease in expression in both transient transfection and stable transduction experiments in cell lines, and a 4-fold decrease in transduction experiments in primary cells. This contrasted strikingly with experiments with the EEF1A1 promoter, in which the majority of proviral forms maintained the intron. Reversal of the UBC expression cassette prevented this splicing-mediated intron loss and maximized expression in uni- and bidirectional LVs. The difference in intron maintenance between the UBC and EEF1A1 promoters is caused by promoter exonic sequences, rather than the intronic sequences themselves.

## MATERIALS AND METHODS

### Plasmid construction

All plasmid sequences used in these studies are included as Supplementary files, available at NAR online.

The human ubiquitin C promoter was amplified via polymerase chain reaction (PCR) from FUGW (8), phosphorylated with T4 polynucleotide kinase and ligated into linearized and blunted pCafe (Cassette for expression) to generate pCafe-UBC. The woodchuck hepatitis virus post-transcriptional regulatory element sequence (herein 'PRE,' referred to as 'LPRE' in Schambach *et al.*) was PCR amplified and cloned into pCafe-UBC linearized with KpnI using In-Fusion (Clontech Laboratories, Mountain View, CA, USA, Cat. No. 639645). The Emerald variant of EGFP was PCR amplified from pRSET-EmGFP (Life Technologies, Carlsbad, CA, USA, Cat. No. V353-20) and cloned into HpaI-linearized pCafe-UBC-PRE using In-Fusion to generate pCafe-UBC-EmGFP-PRE. pCafe-UBCs-EmGFP-PRE was generated in a similar fashion, with UBC cloning primers designed to omit the UBC intron sequence.

For the expression cassettes in the reverse orientation (ro) plasmids, pCafe-roUBC-EmGFP-bGHpA and pCafe-roUBCs-EmGFP-bGHpA, the bovine growth hormone polyadenylation signal (bGHpA) was PCR amplified from pcDNA4/HisMax A (Life Technologies, Cat. No. V864-20) and inserted after the transgene.

For constructs with the UBC intron repositioned (i), pCafe-iUBC-EmGFP-PRE, pCafe-roiUBC-EmGFP-PRE and pCafe-rofiUBC-EmGFP-PRE, UBC intronic sequences were PCR amplified from pCafe-UBC-PRE and cloned into EcoRV-linearized pCafe-UBCs-EmGFP-PRE using In-Fusion.

For a construct with the UBC enhancer deleted (dEnh), pCafe-dEnhUBC-EmGFP-PRE, pCafe-UBC-EmGFP-PRE was PCR amplified using overlapping, outward-facing primers flanking the putative intronic enhancer region and recircularized with In-Fusion after DpnI treatment.

For all pCCLc (3) LVs, expression cassettes were removed from pCafe plasmids with EcoRV/KpnI digestion and ligated into EcoRV/KpnI-linearized pCCLc with the NEB Quick Ligase Kit (New England Biolabs, Ipswitch, MA, USA).

The bidirectional (BD) vector was constructed by assembly of PCR amplicons of the human growth hormone polyadenylation signal (hGHpA) and bidirectional mCMV/UBC promoter (9) and EGFP and WPRE from FUGW, and mCherry from EFS-single-IDLV (10) designed with overlapping homology with the pCCLc backbone using In-Fusion. roBD vector was constructed by restriction digest of BD to invert the mCherry-bidirectional promoter-EGFP cassette between inverse hGHpA and WPRE, and ligated with the NEB Quick Ligase Kit.

### Cell culture

D10 medium was prepared by adding 50 ml heat-inactivated foetal bovine serum (Gemini Bio-Products, West Sacramento, CA, USA, Cat. No. 900-208) and 5.5 ml 100X L-Glutamine:Penicillin:Streptomycin solution (Gemini Bio-Products Cat. No. 400-110) to 500 ml Dulbecco's modified Eagle's medium without L-glutamine (Mediatech, Herndon, VA, USA, Cat. No. 15-013-CV). R10 medium was prepared by adding the same two components to 500 ml RPMI 1640 medium without L-glutamine (Mediatech Cat. No. 15-040). HEK293T cells, hereafter '293T,' (ATCC, Manassas, VA, USA, Cat. No. CRL-1268) were maintained in D10 medium, and K562 (ATCC Cat. No. CCL-243) cells were maintained in R10 medium.

### Vector production

LV supernatant was produced by transfection of $1 \times 10^7$ 293T cells with 10 μg pCMVΔR8.91 (11), 10 μg of the appropriate pCCLc vector plasmid and 2 μg pCAG-VSV-G (12). Transfection mixtures were prepared in 1.5 ml DPBS by adding the plasmids and 66 μl 1 mg/ml branched PEI solution (Sigma-Aldrich, St. Louis, MO, USA, Cat. No. 408727-100ML), and then vortexing for several seconds. After incubation at room temperature for 5–10 min, transfection mixes were added dropwise to 293T cells plated 24 h earlier in 10 cm dishes. After ∼16 h, the medium was changed to UltraCULTURE medium (Lonza, Basel, Switzerland, Cat. No. 12-725F) supplemented with 50 U/ml penicillin, 50 μg/ml streptomycin, 2 mM L-glutamine and 20 mM HEPES. Viral supernatant was harvested 24–48 h after this medium change.

For roUBC vectors, 5 μg pcDNA3-NovB2 was included to prevent a drop in titers caused by the presence of transcripts antisense to the vector genomic RNA (13). An additional 15 μl 1 μg/ml PEI solution was added to compensate for the increased plasmid DNA.

Vector was concentrated ∼150-fold by ultracentrifugation at 26,000 rpm for 90 min at 4°C in a Beckman Coulter SW-32Ti rotor for transduction of human CD34+ HSPCs.

### Transfection

293T cells were seeded at $8 \times 10^5$ cells/well in 6-well plates (Corning, Corning, NY, USA, Cat. No. 3516) in D10 medium. Twenty-four hours later, 1.5 μg of plasmid was prepared for transfection in 200 μl Opti-MEM I medium (Life Technologies, Carlsbad, CA, USA, Cat. No. 31985-062) in 1.5 ml microcentrifuge tubes, 4.5 μl of TransIT-293 transfection reagent (Mirus Bio, Madison, WI, USA, Cat. No. MIR 2700) was added, and the mixtures were vortexed briefly and incubated at room temperature for 5 min before being added dropwise to the cells. Cells were collected 48

h after transfection by brief trypsinization and analyzed for green fluorescent protein reporter expression on a BD LSR-Fortessa flow cytometer.

### Transduction

For lentiviral expression analysis, K562 cells were plated in 24-well plates at $5 \times 10^4$ cells/well and treated with a range of vector doses to obtain populations with 10% transduction or lower, thus ensuring that the majority of cells received only single integrations. Cells were cultured for 1–2 weeks before flow cytometric analysis to dilute out non-integrated vector and to allow fluorescent protein levels to reach steady state.

For expression analysis in primary human CD34+ HSPCs from mobilized peripheral blood, cryopreserved cells were thawed and prestimulated overnight in X-VIVO 15 medium (Lonza) supplemented with 50 ng/ml human *fms*-like tyrosine kinase 3 (FLT-3) ligand, 50 ng/ml human stem cell factor and 50 ng/ml human thrombopoietin (PeproTech, Rocky Hill, NJ, USA). Viral vector was then added in an equal volume of the same medium to achieve a final vector concentration of $3 \times 10^5$ transducing units/ml, as determined by transduction of K562 cells. This vector dose yielded ∼10% transduction. Twenty-four hours after vector addition, 2 ml of myeloid differentiation medium was added, composed of IMDM supplemented with 20% foetal bovine serum, 0.5% bovine serum albumin, 5 ng/ml human interleukin-3, 10 ng/ml human interleukin-6 and 25 ng/ml human stem cell factor (PeproTech).

### PCR analysis of splicing

Genomic DNA from transduced K562 cells was analyzed via PCR using KAPA HiFi Hot Start polymerase and primers UBC intron F (AAGTAGTCCCTTCTCGGCGAT), UBC intron R (GGTCAGCTTGCCGTAGGT), EEF1A1 intron F (GTTCTTTTTCGCAACGGGTTTG) and EEF1A1 intron R (TGTGGCCGTTTACGTCGC).

Quantitative droplet digital PCR (ddPCR) was carried out by analysis of genomic DNA from UBC vector-transduced K562 cells using primers UBCint F (GGCGAGTGTGTTTTGTGAAGTTT) and EmGFP R (TACGTCGCCGTCCAGCTC), and probe FAM-EmGFP (FAM-CACCACCCCGGTGAACAGCTCCTCG). For EEF1A1 vector analysis, the UBCint F primer was substituted with EEF1A1int F (TCTCAAGCCTCAGACAGTGGT). The spliced form of UBC was quantified using UBCs F (GCTGTGATCGTCACTTGACA) instead of UBCint F. ddPCR was carried out according to the manufacturer's instructions, using 100 ng of template gDNA. One unit of DraI enzyme (New England Biolabs) was added to the ddPCR master mix containing ddPCR Supermix for Probes (Bio-Rad, Hercules, CA, USA), and predigestion was carried out in the PCR reaction mixes for 1–2 h at 37°C before droplet generation and thermal cycling.

For analysis of vector genomes in vector supernatant, RNA was purified from 500 μl of raw vector supernatant using the PureLink RNA Mini Kit liquid sample procedure (Life Technologies). Reverse transcription was carried out before PCR using iScript cDNA Synthesis Kit (Bio-Rad).

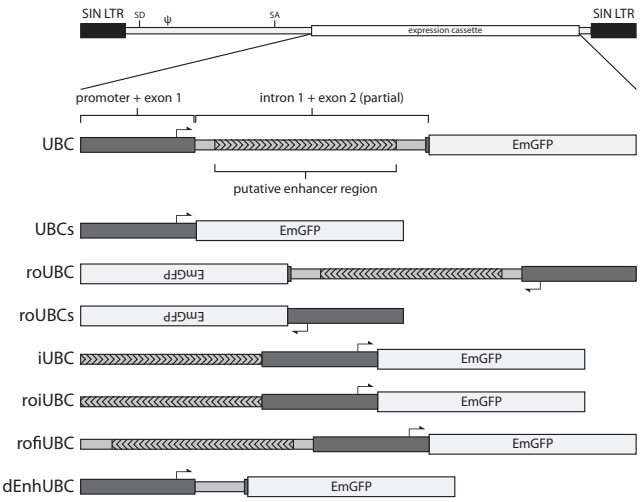


**Figure 1.** Expression vectors used for studies. Lentiviral diagram depicts location in CCLc vectors. roUBC and roUBCs vectors contain a bovine growth hormone polyadenylation signal (not depicted), in the proper reverse orientation, after the end of the EmGFP reading frame. pCafe expression plasmids contained identical cassettes upstream of an SV40 polyadenylation signal.

### Luciferase assay

pGL4.25 vector (Promega, Madison, WI, USA) containing an optimized luciferase open reading frame driven by a minimal TATA-box promoter was used to assay for enhancer activity of the UBC and EEF1A1 introns. A promoterless enhancer sequence from the CMV promoter was used as a positive control. All inserts were cloned via PCR and Gibson assembly into pGL4.25 linearized with EcoRV and KpnI. All plasmid sequences are provided as Supplementary data. Luciferase assays were performed in 293T cells plated on 96-well tissue culture-treated plates. $5 \times 10^4$ cells per well were plated in D10 medium, and 18 h later, transfection mixes were prepared in OPTI-MEM with 100 ng reporter plasmid and 0.3 μl TransIT-293 per well. Samples were prepared 48 h after transfection with the Dual-Luciferase Reporter Assay System (Promega) and luminescence readings were taken with a Tecan Infinite M1000 PRO plate reader (Tecan, Männedorf, Switzerland).

## RESULTS

### UBC intron is missing from proviral forms, and expression cassette reversal prevents loss

To assess whether UBC intron 1 is maintained during packaging, pCCLc LV DNA constructs and simpler pCafe expression plasmid constructs for transient transfection were created with various modifications of the UBC promoter (Figure 1). All constructs contained the Emerald variant of green fluorescent protein (EmGFP), which allowed for expression analysis via flow cytometry (14). UBC constructs contained the full UBC promoter fragment, as it exists in the human genome, whereas shorter UBCs constructs were designed with a full deletion of UBC intron 1, which would be the expected proviral form if canonical splicing occurred during packaging. To test whether movement of the expres-

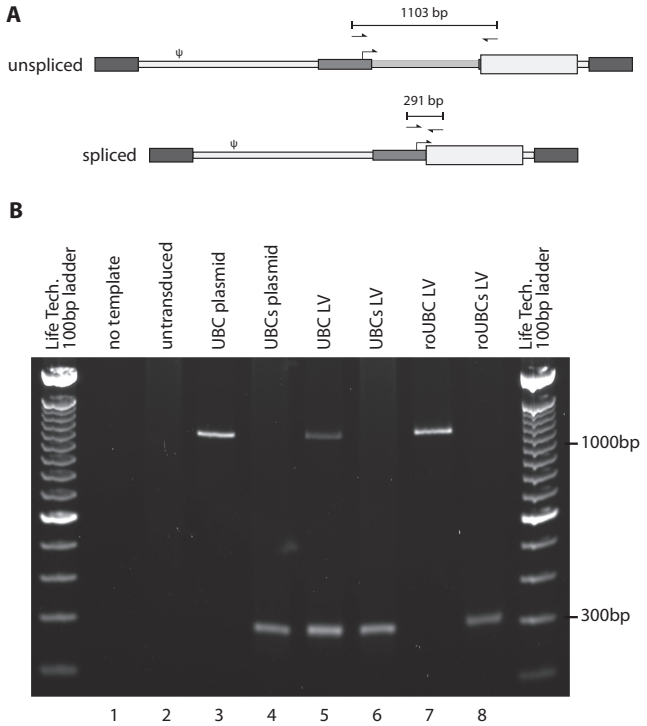


**Figure 2.** Genetic analysis of UBC splicing. (**A**) PCR strategy with primer locations and expected product sizes. (**B**) Electrophoresis of PCR products from controls and gDNA from cells transduced with lentiviral vectors bearing UBC promoter variants.

sion cassette to the opposite strand would avoid splicing-mediated loss of the intron, reverse orientation (ro) constructs roUBC and roUBCs were created by reversing the promoter and transgene and inserting a polyadenylation signal after the transgene. Importantly, while the payloads of the pCCLc LVs pass through an RNA intermediate stage and are susceptible to splicing-mediated loss, payloads of the pCafe expression plasmids have no RNA intermediate and therefore cannot lose genetic elements due to splicing.

Viral vectors were produced in 293T cells and used to transduce K562 cells for PCR-based genetic analysis of proviral forms (Figure 2A). PCR analysis of gDNA two weeks post-transduction revealed that many CCLc-UBC-EmGFP-PRE proviral forms contained an amplicon consistent with intron loss, as indicated by analysis of UBCs proviral forms (Figure 2B, lanes 5 and 6). Sanger sequencing of the short product confirmed that the expected canonical splicing had occurred (data not shown). In contrast, roUBC proviral forms yielded no truncated PCR product, suggesting that reversal of the expression cassette fully prevented intron loss (Figure 2B, lane 7). Because of the significant difference in predicted PCR product size between the intron-containing templates and intron-lacking templates, there could be a substantial bias toward amplification of the intron-lacking templates and overestimation of the amount of intron loss from this result. Therefore, to quantify the frequency of intron loss, a duplex digital PCR assay was set up in which the signal from a primer and probe set spanning the intron and EmGFP transgene was normalized using a primer and probe set to the LV packaging signal (Figure 3A and B). This analysis showed that only 18% of UBC vector forms retained the UBC intron (Figure 3C), while roUBC vector forms fully retained the intron.

In order to assess whether events during transduction and reverse transcription influenced the proportion of proviral forms containing introns, we collected RNA from UBC viral supernatants and quantified the fraction of RNA genomes containing spliced UBC introns. We then compared this to the fraction of vector proviral forms containing spliced introns in K562 cells transduced with the same supernatants. These values agreed very closely, suggesting that the introns were already missing in vector particles and were therefore removed in the packaging cells (Supplementary Figure S1).

## Loss of intron lowers expression from UBC promoter

To assess the effect of intron loss on transgene expression, pCafe expression plasmids containing the full UBC promoter element or the truncated UBCs promoter with the intron region deleted were transiently transfected into 293T cells and analyzed at 48 h post-transfection via flow cytometry. The UBC promoter yielded significantly higher expression than the UBCs promoter, by a margin of ∼2-fold (Figure 4A). A similar 2-fold difference was observed between the roUBC and roUBCs constructs. Because these plasmids were transfected directly into cells, no intron loss was possible, and the UBC promoter plasmid molecules assayed therefore all contained the intron.

Having established that the presence of the intron confers higher expression in these transfection experiments where intron loss was not possible, we next examined expression from the various constructs packaged as LVs 2 weeks after transduction of K562 cells. As the genetic analysis revealed that the majority of UBC LV forms lack the intron, we reasoned that the UBC vector would express levels of EmGFP similar to the UBCs vector. Indeed, the fluorescence of EmGFP-expressing cells in populations transduced with the UBC vector was nearly equivalent to that in populations transduced with UBCs vector (Figure 4B). In contrast, the roUBC vector showed ∼2-fold higher fluorescence in cells than the UBC vector, consistent with the genetic analysis indicating that the roUBC vector retains the intron.

We also transduced human CD34+ hematopoietic stem and progenitor cells enriched from the peripheral blood of a healthy donor treated with granulocyte-colony stimulating factor to determine if the improved expression from the roUBC vector compared to the UBC vector would also be observed in a primary cell type relevant to lentiviral gene therapy. After 10 days of culture post-transduction in myeloid differentiation conditions, cells transduced with roUBC vector showed 4-fold higher expression than cells transduced with UBC (Supplementary Figure S2). Genetic analysis showed that intron loss was similar in the UBC-transduced cells to that observed in K562 cells and that the intron was fully maintained in roUBC-transduced cells (Supplementary Figure S3).

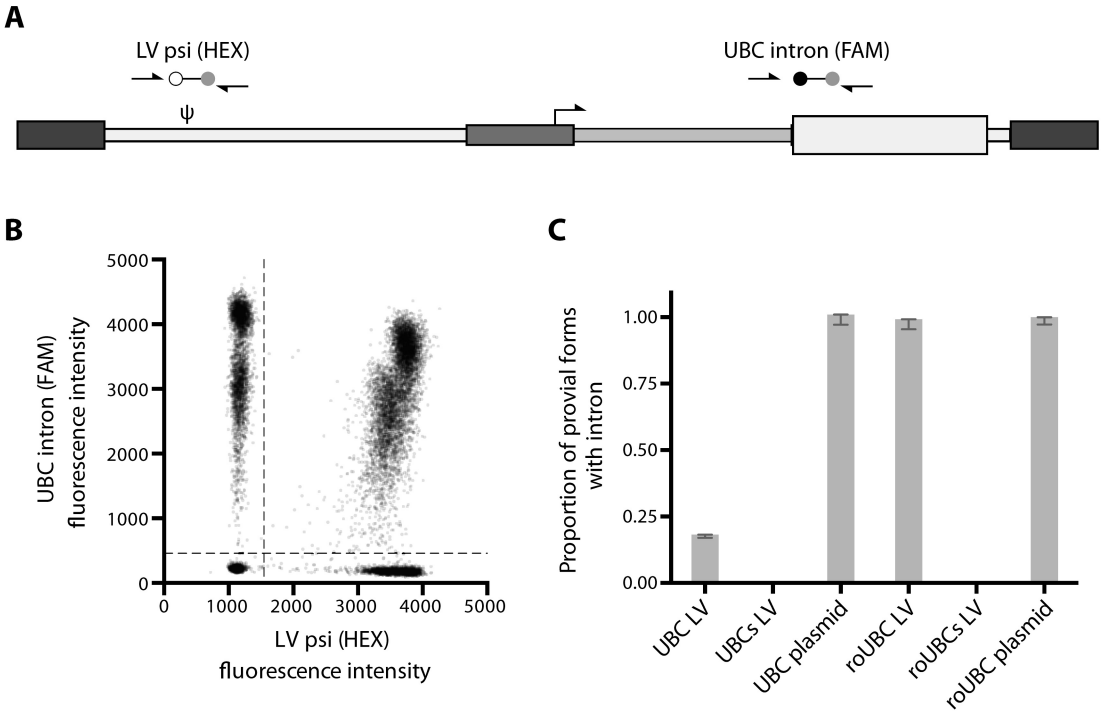


**Figure 3.** Quantitative analysis of UBC intron loss during packaging and transduction. (**A**) Duplex ddPCR strategy for quantifying UBC intron copies (FAM-UBC intron), normalized to total proviral integrations (HEX-LV psi). (**B**) Representative raw data from ddPCR, illustrating separation between positive and negative droplets. (**C**) Ratio of UBC intron copies to total proviral copies in controls and samples transduced with LV bearing UBC promoter variants. Error bars represent 95% confidence interval based on ddPCR Poisson statistics.

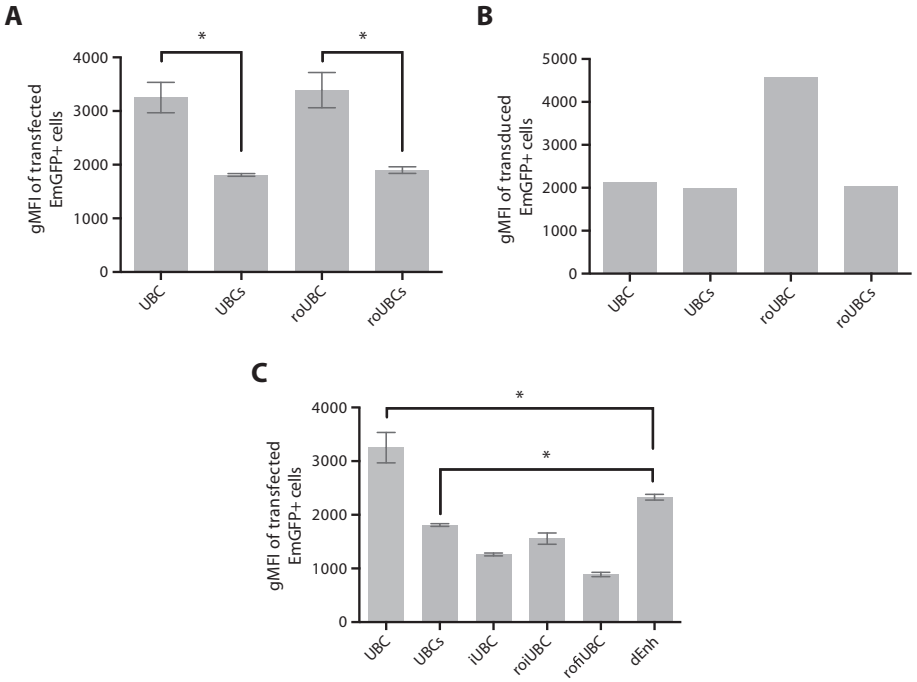


**Figure 4.** Flow cytometric expression analysis of UBC promoter variants. (**A**) Geometric mean fluorescence intensity (gMFI) of 293T cells 48 h post-transfection with expression plasmids. Error bars represent standard deviation of three biological replicates. UBC versus UBCs unpaired *t*-test $P = 0.0122$, roUBC versus roUBCs $P = 0.0134$. (**B**) gMFI of K562 cells 10 days post-transduction with CCLc lentiviral vectors bearing UBC promoter variants. Data are representative of multiple experiments. (**C**) gMFI of 293T cells 48 h post-transfection. Error bars represent SD of three biological replicates. UBC versus dEnh unpaired *t*-test $P = 0.0267$, UBCs versus dEnh $P = 0.0008$.

### Positive effect of UBC intron on expression is not through classical enhancer activity

Aside from reversal of the expression cassette, we also sought other ways to retain full expression of the UBC promoter fragment in an LV. We first investigated whether movement of the reported intronic enhancer sequence to a site immediately upstream of the promoter would lead to equivalent expression compared to the full-length UBC promoter fragment (7). Importantly, this variant lacked the intronic splice sites, which should allow its transmission in LVs. However, the resulting iUBC construct performed worse than UBCs (Figure 4C). roiUBC and rofiUBC were created and analyzed to assess whether the orientation of the enhancer sequence relative to the promoter was important, but these promoter variants expressed no better than iUBC (Figure 4C). We finally constructed dEnhUBC, in which the putative enhancer sequence was deleted, but the splicing sites were retained. This variant expressed slightly more EmGFP than UBCs, presumably due to improved nuclear export from splicing, but significantly less than UBC (Figure 4C). These results are consistent with a follow-up study on the UBC promoter fragment intron, which found that its enhancer activity was fully dependent on its position within the intron (15). This behavior, termed intron-mediated enhancement, is poorly understood.

We reasoned that if the UBC intron sequence were not a classical enhancer, then it should not increase expression from a heterologous minimal promoter. Indeed, when the intron sequence was placed in a luciferase reporter plasmid upstream of a minimal promoter in a forward or reverse orientation, no increase in luciferase expression over background was observed, in contrast to a plasmid in which a CMV enhancer sequence was placed upstream (Supplementary Figure S4). In fact, expression from these plasmids was significantly lower than from plasmids with the minimal promoter alone, consistent with the UBC intron sequence being repressive when placed outside the transcription unit. This repressive effect mirrors the reduction in expression seen when intronic sequences were placed upstream of the UBCs promoter form (Figure 4C). Interestingly, the same was true for EEF1A1 intron 1 in forward or reverse orientation (Supplementary Figure S4).

### EEF1A1 intron is maintained in proviral forms and aids in maximal expression

Because the observation of intron loss from the UBC promoter contrasts so starkly with reports on the EEF1A1 promoter fragment in LVs, we created expression vectors for transient transfection and lentiviral production with the EEF1A1 promoter fragment and an EmGFP reporter. PCR and ddPCR analysis of gDNA from transduced cells showed that nearly all vector forms retained the intron within the promoter (Figure 5B, lane 5). Extreme contrast adjustment of the gel electrophoresis image can reveal a barely detectable amount of short product at the length expected upon intron loss, but quantitative ddPCR analysis does not detect this small population of intron-lacking proviral forms (Figure 5C). Consistent with these observations and with a previous report (2), a ∼2-fold difference in expression between the intron-containing and intron-lacking promoters was observed both in transient transfection (Figure 5D) and transduction (Figure 5E) experiments, suggesting that the EEF1A1 promoter element's intron is indeed being faithfully transmitted in almost all cases.

### Difference in intron transmission is determined by promoter exon sequences

We hypothesized that the difference in intron retention between the UBC and EEF1A1 promoters was due to sequence determinants of splicing efficiency or splicing kinetics within the introns. To test this, we swapped the introns from one promoter to the other, creating UBC (EEF1A1int) and EEF1A1(UBCint) vectors. Surprisingly, we found that the UBC (EEF1A1int) LV lost the EEF1A1 intron and expressed similar levels of EmGFP to the intronless UBCs vector, while the EEF1A1(UBCint) maintained the UBC intron and expressed significantly more EmGFP than the intronless EEF1A1s vector (Supplementary Figure S5). These results suggest that the distinct exon sequences of the two promoters are determining whether the introns are retained during lentiviral production.

### Expression cassette modification maximizes expression from UBC bidirectional vectors

We finally sought to improve expression from UBC promoter-based bidirectional vectors mediating coordinated expression of two transgenes in LVs (9,16). Because the vector design calls for a sense-strand orientation of the UBC promoter, we reasoned that the majority of proviral forms would lose the UBC intron and that reversal of the dual, divergent UBC and minimal cytomegalovirus (CMV) promoters would lead to increased expression due to intron inclusion with the UBC-promoted transgene (Figure 6A).

Genetic analysis of stably transduced 293T cells revealed that the UBC intron was lost 75% of the time from BD vectors, in which the UBC promoter is on the vector sense strand, whereas in roBD vectors, nearly all of the proviral forms contained the UBC intron (Figure 6B). This led to an increase in EGFP expression driven by the UBC promoter in stably transduced cells (Figure 6C). Surprisingly, in light of the expression data suggesting that the intron does not contain a traditional enhancer, mCherry expression driven by the minimal CMV promoter was also increased in retained UBC intron in roBD-transduced cells.

## DISCUSSION

Lentiviral gene transfer has recently advanced into clinical gene therapy trials, with multiple successes and no clinically significant adverse events, and has also shown promise in many pre-clinical studies (17–22). As therapies are developed for additional disorders, new vectors will be created bearing various genomic fragments for transgene regulation. Past promoter/transgene combinations have required the presence of introns for full activity and regulation, and it is likely that some future designs will require them as well.

Our results suggest that introns differ in terms of their likelihood of loss during vector production and transduc-

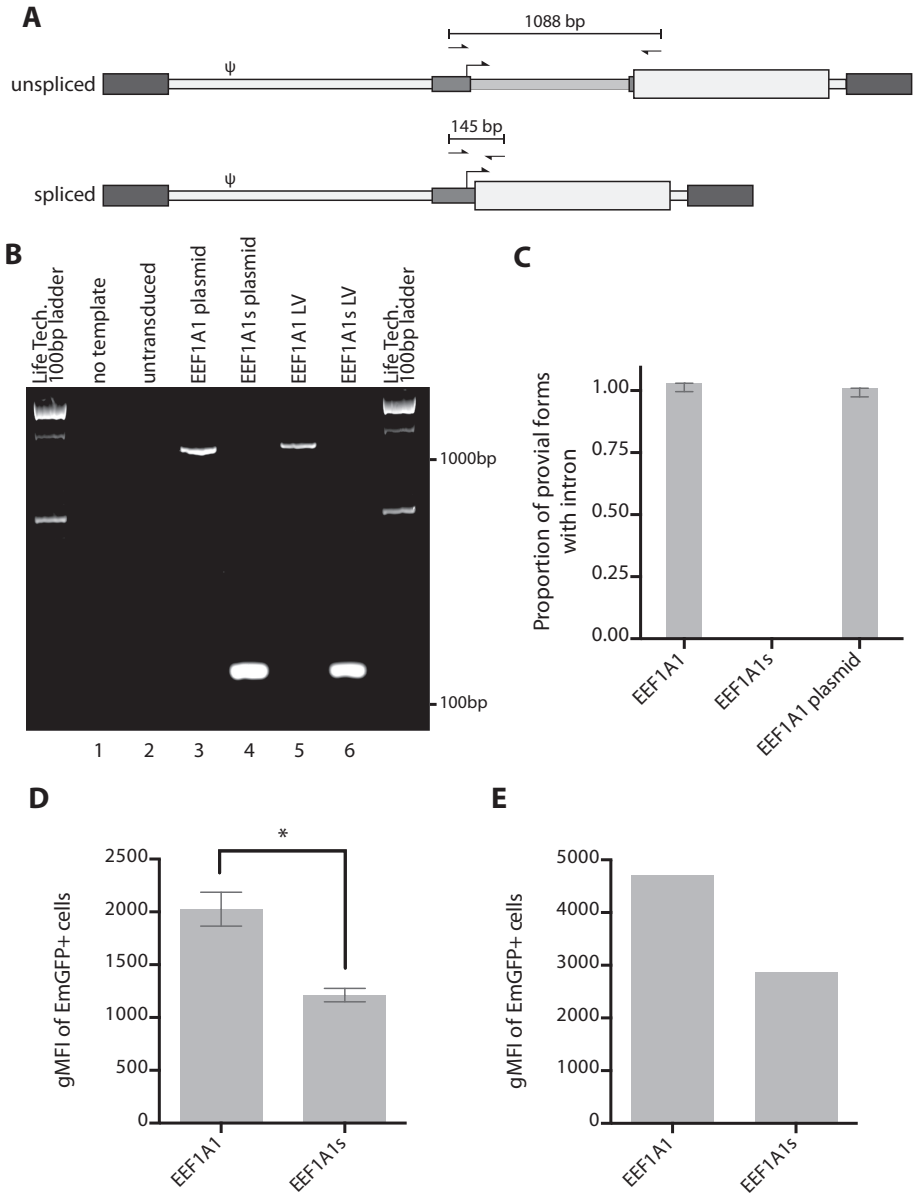


**Figure 5.** EEF1A1 analysis. (**A**) Diagrams of lentiviral vectors bearing EEF1A1 promoter variants. (**B**) Gel electrophoresis of PCR product amplifying across EEF1A1 intron in stably transduced K562 cells, greater than 2 weeks post-transduction. (**C**) ddPCR quantification of the ratio of intron copies to proviral copies in samples analyzed in (**B**). Error bars represent 95% confidence interval. (**D**) gMFI of transiently transfected 293T cells 48 h post-transfection with expression plasmids, measured by flow cytometry. Error bars represent SD of three biological replicates. Unpaired *t*-test $P = 0.0064$. (**E**) gMFI of stably transduced K562 cells 10 days post-transduction, measured by flow cytometry.

tion. While the human UBC promoter fragment was missing its intron in most proviral forms, the human EEF1A1 promoter fragment was not similarly affected. Inclusion of the UBC intron requires that the transgene cassette be reversed to avoid the processing of splicing machinery, but the EEF1A1 intron is maintained in almost every proviral form even though it is theoretically exposed to the spliceosome. A previous study indicates that the hybrid CAG promoter is also maintained throughout vector production and transduction (4). An intronless version of the EEF1A1 promoter has moved into clinical trials for both adenosine-deaminase-deficient severe combined immunodeficiency (ADA-SCID) and X-linked SCID (SCID-X1), and preclinical studies suggested that it will drive sufficient transgene expression for therapeutic effect. Our data indicate that a full EEF1A1 promoter containing intron 1 leads to roughly 2-fold higher transgene expression, an increase that could be necessary or beneficial for future vector designs. Overall, these results illustrate the importance of full genetic characterization of retroviral vectors, as known or unknown introns can lead to transduced cells bearing highly variant vector forms. In the area of gene therapy, where product characterization is important from a regulatory standpoint, this variation is unlikely to find acceptance.

It has been reported that antisense RNA targeted to splice donor or acceptor sites can prevent splicing of pri-

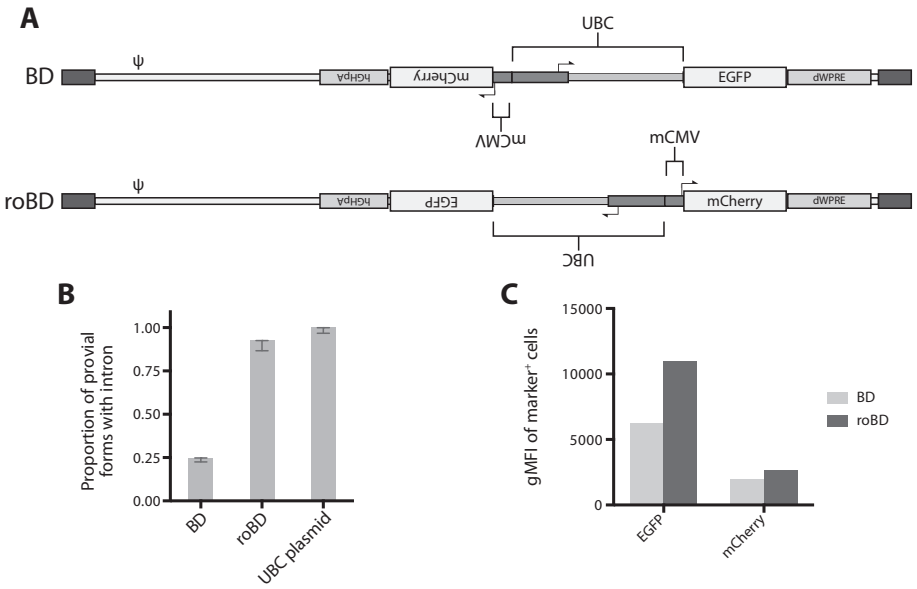


**Figure 6.** Bidirectional vector analysis. (**A**) Vector schematics. (**B**) ddPCR analysis of intron loss in BD and roBD vectors. (**C**) gMFI of stably transduced 293T cells 2 weeks post-transduction, measured by flow cytometry. Error bars represent 95% confidence interval.

mary transcripts (23). We therefore attempted to inhibit splicing of UBC vector genomes using U6-driven plasmids expressing 50 nt antisense sequences to either the splice donor or splice acceptor site during lentiviral production. Unfortunately, these constructs did not lead to higher expression from UBC vectors upon transduction when used alone or in combination (data not shown). It is possible that this strategy could lead to retention of other introns in LVs, but it was ineffective for the UBC intron in our experiments.

We found that the UBC promoter intron does indeed increase expression, as previously reported, but that the enhancer-like activity within the intron sequence is dependent on its location inside the intron. This could be paralleled by future vector designs incorporating transgenes with endogenous introns for full activity, in which regulatory activity of intronic sequences might similarly not be mobile. Further research is also warranted to investigate why the UBC intronic sequences have a positive effect on expression when present within the transcription unit, but a negative effect when placed upstream of the promoter. This would likely have important implications for both endogenous gene regulation and transgene regulation for gene therapy and genetic engineering. Importantly, our data suggest that such introns are relatively safe payloads for integrating vectors, as they probably will not transactivate nearby promoters in the manner that has caused adverse events and subclinical clonal expansion in clinical gene therapy trials (18,24–27).

The UBC intron and EEF1A1 intron 1 do not differ noticeably at the sequence level in terms of their adherence to canonical splice donor, acceptor and branch point sites, and our data from vectors in which the introns are swapped indicate that the sequence determinants of intron loss are not within the introns themselves but within the exonic sequences of the UBC and EEF1A1 promoters. This would be unfortunate if true generally, as potential modifications to vectors to alter splicing would be limited dramatically in the majority of exons that are coding sequences. Biologically speaking, it is unsurprising, as exons in the human genome are known to contain exonic splicing enhancers as well as exonic splicing suppressors/silencers. These sequences control the efficiency of splicing of human introns, most of which are thought to be suboptimally defined (28).

We hypothesize that the difference in frequency of loss between these introns is linked to the speed at which they are spliced, which can be largely determined by exonic sequences. Future experiments could assess the splicing kinetics of these two genetic elements, the speed of which would be predicted to correlate inversely with intron transmission. While new work has examined the kinetics of transcript splicing and release from chromatin, the sequence determinants of the range of rates observed for different transcripts are not yet understood (29). A better understanding of the determinants of splicing kinetics could direct the modification of the UBC promoter fragment to decrease splicing speed sufficiently to get intron-containing genomic RNA into vector particles, while maintaining efficient splicing during transgene expression.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGMENTS

The authors acknowledge Oz Davis and Navdeep Saini for technical assistance, Roger Hollis, Douglas Black, Chintan Parekh and Sanjeet Patel for helpful discussions and the reviewers of this manuscript for their thoughtful comments. The flow cytometry core at the Broad Stem Cell Research Center was invaluable for these experiments.

## FUNDING

The National Heart, Lung and Blood Institute at the National Institutes of Health [2P01 HL073104]; the Philip J. Whitcome Predoctoral Fellowship from the UCLA Molecular Biology Institute [to A.R.C.]; the Ruth L. Kirschstein National Research Service Award [GM007185 to A.R.C.]; the Broad Stem Cell Research Center (BSCRC)—California Institute of Regenerative Medicine (CIRM) [TG2-01169 to E.H.G.]. Funding for open access charge: The National Heart, Lung and Blood Institute at the National Institutes of Health [2P01 HL073104].

*Conflict of interest statement.* None declared.

## REFERENCES

1. Zufferey,R., Dull,T., Mandel,R.J., Bukovsky,A., Quiroz,D., Naldini,L. and Trono,D. (1998) Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.*, **72**, 9873–9880.
2. Schambach,A., Bohne,J., Chandra,S., Will,E., Margison,G.P., Williams,D.A. and Baum,C. (2006) Equal potency of gammaretroviral and lentiviral SIN vectors for expression of O6-methylguanine-DNA methyltransferase in hematopoietic cells. *Mol. Ther.*, **13**, 391–400.
3. Dull,T., Zufferey,R., Kelly,M., Mandel,R.J., Nguyen,M., Trono,D. and Naldini,L. (1998) A third-generation lentivirus vector with a conditional packaging system. *J. Virol.*, **72**, 8463–8471.
4. Ramezani,A., Hawley,T.S. and Hawley,R.G. (2000) Lentiviral vectors for enhanced gene expression in human hematopoietic cells. *Mol. Ther.*, **2**, 458–469.
5. Zaiss,A.-K., Son,S. and Chang,L.-J. (2002) RNA 3′ readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. *J. Virol.*, **76**, 7209–7219.
6. Logan,A.C., Lutzko,C. and Kohn,D.B. (2002) Advances in lentiviral vector design for gene-modification of hematopoietic stem cells. *Curr. Opin. Biotechnol.*, **13**, 429–436.
7. Bianchi,M., Crinelli,R., Giacomini,E., Carloni,E. and Magnani,M. (2009) A potent enhancer element in the 5′-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. *Gene*, **448**, 88–101.
8. Lois,C., Hong,E.J., Pease,S., Brown,E.J. and Baltimore,D. (2002) Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*, **295**, 868–872.
9. Kamata,M., Liang,M., Liu,S., Nagaoka,Y. and Chen,I.S.Y. (2010) Live cell monitoring of hiPSC generation and differentiation using differential expression of endogenous microRNAs. *PLoS ONE*, **5**, e11834.
10. Joglekar,A.V., Hollis,R.P., Kuftinec,G., Senadheera,S., Chan,R. and Kohn,D.B. (2013) Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol. Ther.*, **21**, 1705–1717.
11. Zufferey,R., Nagy,D., Mandel,R.J., Naldini,L. and Trono,D. (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.*, **15**, 871–875.
12. Hanawa,H., Kelly,P.F., Nathwani,A.C., Persons,D.A., Vandergriff,J.A., Hargrove,P., Vanin,E.F. and Nienhuis,A.W. (2002) Comparison of various envelope proteins for their ability to pseudotype lentiviral vectors and transduce primitive hematopoietic cells from human blood. *Mol. Ther.*, **5**, 242–251.
13. Maetzig,T., Galla,M., Brugman,M.H., Loew,R., Baum,C. and Schambach,A. (2010) Mechanisms controlling titer and expression of bidirectional lentiviral and gammaretroviral vectors. *Gene Ther.*, **17**, 400–411.
14. Tsien,R.Y. (1998) The green fluorescent protein. *Annu. Rev. Biochem.*, **67**, 509–544.
15. Bianchi,M., Crinelli,R., Giacomini,E., Carloni,E., Radici,L. and Magnani,M. (2013) Yin Yang 1 intronic binding sequences and splicing elicit intron-mediated enhancement of ubiquitin C gene expression. *PLoS ONE*, **8**, e65932.
16. Amendola,M., Venneri,M.A., Biffi,A., Vigna,E. and Naldini,L. (2005) Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters. *Nat. Biotechnol.*, **23**, 108–116.
17. Cartier,N., Hacein-Bey-Abina,S., Bartholomae,C.C., Veres,G., Schmidt,M., Kutschera,I., Vidaud,M., Abel,U., Dal Cortivo,L., Caccavelli,L. *et al.* (2009) Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science*, **326**, 818–823.
18. Cavazzana-Calvo,M., Payen,E., Negre,O., Wang,G., Hehir,K., Fusil,F., Down,J., Denaro,M., Brady,T., Westerman,K. *et al.* (2010) Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. *Nature*, **467**, 318–322.
19. Aiuti,A., Biasco,L., Scaramuzza,S., Ferrua,F., Cicalese,M.P., Baricordi,C., Dionisio,F., Calabria,A., Giannelli,S., Castiello,M.C. *et al.* (2013) Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science*, **341**, 1233151.
20. Biffi,A., Montini,E., Lorioli,L., Cesani,M., Fumagalli,F., Plati,T., Baldoli,C., Martino,S., Calabria,A., Canale,S. *et al.* (2013) Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science*, **341**, 1233158.
21. Romero,Z., Urbinati,F., Geiger,S., Cooper,A.R., Wherley,J., Kaufman,M.L., Hollis,R.P., de Assin,R.R., Senadheera,S., Sahagian,A. *et al.* (2013) β-globin gene transfer to human bone marrow for sickle cell disease. *J. Clin. Invest.*, **123**, 3317–3330.
22. Carbonaro,D.A., Zhang,L., Jin,X., Montiel-Equihua,C., Geiger,S., Carmo,M., Cooper,A., Fairbanks,L., Kaufman,M.L., Sebire,N.J. *et al.* (2014) Preclinical demonstration of lentiviral vector-mediated correction of immunological and metabolic abnormalities in models of adenosine deaminase deficiency. *Mol. Ther.*, **22**, 607–622.
23. Morcos,P.A. (2007) Achieving targeted and quantifiable alteration of mRNA splicing with Morpholino oligos. *Biochem. Biophys. Res. Commun.*, **358**, 521–527.
24. Hacein-Bey Abina,S., von Kalle,C., Schmidt,M., McCormack,M.P., Wulffraat,N., Leboulch,P., Lim,A., Osborne,C.S., Pawliuk,R., Morillon,E. *et al.* (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, **302**, 415–419.
25. Hacein-Bey-Abina,S., Garrigue,A., Wang,G.P., Soulier,J., Lim,A., Morillon,E., Clappier,E., Caccavelli,L., Delabesse,E., Beldjord,K. *et al.* (2008) Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.*, **118**, 3132–3142.
26. Howe,S.J., Mansour,M.R., Schwarzwaelder,K., Bartholomae,C., Hubank,M., Kempski,H., Brugman,M.H., Pike-Overzet,K., Chatters,S.J., de Ridder,D. *et al.* (2008) Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.*, **118**, 3143–3150.
27. Stein,S., Ott,M.G., Schultze-Strasser,S., Jauch,A., Burwinkel,B., Kinner,A., Schmidt,M., Krämer,A., Schwäble,J., Glimm,H. *et al.* (2010) Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat. Med.*, **16**, 198–204.
28. Zheng,Z.-M. (2004) Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J. Biomed. Sci.*, **11**, 278–294.
29. Pandya-Jones,A., Bhatt,D.M., Lin,C.-H., Tong,A.-J., Smale,S.T. and Black,D.L. (2013) Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. *RNA*, **19**, 811–827.