ARTICLE
Journal of Cellular Biochemistry 112:2361–2375 (2011)

Journal of Cellular Biochemistry

β-Globin Matrix Attachment Region Improves Stable Genomic Expression of the *Sleeping Beauty* Transposon

Lucas Sjeklocha,[1] Yixin Chen,[1] Meghan C. Daly,[1] Clifford J. Steer,[1,2] and Betsy T. Kren[1*]

[1] *Department of Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, 55455*

[2] *Department of Genetics, Cell Biology and Development, University of Minnesota Medical School, Minneapolis, Minnesota, 55455*

ABSTRACT

The liver is an attractive target for gene therapy due to its extensive capability for protein production and the numerous diseases resulting from a loss of gene function it normally provides. The *Sleeping Beauty* Transposon (*SB*-Tn)[1] system is a non-viral vector capable of delivering and mediating therapeutic transgene(s) insertion into the host genome for long-term expression. A current challenge for this system is the low efficiency of integration of the transgene. In this study we use a human hepatoma cell line (HuH-7) and primary human blood outgrowth endothelial cells (BOECs) to test vectors containing DNA elements to enhance transposition without integrating themselves. We employed the human β-globin matrix attachment region (MAR) and the Simian virus 40 (SV40) nuclear translocation signal to increase the percent of HuH-7 cells persistently expressing a GFP::Zeo reporter construct by ~50% for each element; while combining both did not show an additive effect. Interestingly, both elements together displayed an additive effect on the number of insertion sites, and in BOECs the SV40 alone appeared to have an inhibitory effect on transposition. In long-term cultures the loss of plasmid DNA, transposase expression and mapping of insertion sites demonstrated bona fide transposition without episomal expression. These results show that addition of the β-globin MAR and potentially other elements to the backbone of *SB*-Tn system can enhance transposition and expression of therapeutic transgenes. These findings may have a significant influence on the use of *SB* transgene delivery to liver for the treatment of a wide variety of disorders. J. Cell. Biochem. 112: 2361–2375, 2011. © 2011 Wiley-Liss, Inc.

KEY WORDS: BOEC; GENE THERAPY; HUH-7 CELLS; MATRIX ATTACHMENT REGION; SLEEPING BEAUTY TRANSPOSON

Viral vectors are a relatively efficient means to deliver genes to primary cells, but their use can be associated with deleterious and potentially fatal host immune responses, as well as insertional mutation and activation of certain endogenous genes [Raper et al., 2003; Ciuffi, 2008; Daniel and Smith, 2008; Vercammen et al., 2008]. Non-viral gene transfer systems offer alternatives for gene transfer in which therapeutic genes are delivered to target cells using synthetic carriers such as polymers, liposomes, or polyamines or using physical methods such as electroporation [Gresch et al., 2004]. However, such non-viral approaches in general suffer from particularly inefficient delivery to primary cells and have no specific mechanism for chromosomal integration of the transgene construct, which is essential for long-term therapeutic expression [Li and Huang, 2006; Schaffert and Wagner, 2008].

The stable transfer of foreign DNA into host genomes can be achieved in mammalian cells using a non-viral vertebrate DNA transposon (Tn) such as the *Sleeping Beauty* (*SB*) transposable

Abbreviations: 7-ADD, 7-Amino-actinomycin D; BOEC, blood outgrowth endothelial cell; CDS, coding sequence; CMV, cytomegalovirus; EMA, ethidium monoazide; FACS, fluorescent activated cell sorting; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; GFP, enhanced green fluorescent protein; HMGB1, eukaryotic high mobility group DNA-bending protein 1; HSB3, hyperactive *Sleeping Beauty* variant 3; HSB5, hyperactive *Sleeping Beauty* variant 5; IF, eukaryotic initiation factor 4A1; IR/DRs, inverted repeats/direct repeats; *lacZ*, β–galactosidase gene; MAR, matrix attachment region; PCR, polymerase chain reaction; pDNA, plasmid DNA; RT, reverse transcriptase; SV40, Simian virus 40; NLS, nuclear localization signal; NTS, nuclear targeting signal; HuH-7, human hepatoma cell line HuH-7; *SB*, *Sleeping Beauty*; Tn, transposase; UV, ultraviolet; ZeoR, zeocin resistance gene.

Conflicts of interest: The authors have none to declare.

Lucas Sjeklocha and Yixin Chen contributed equally to this work.

Additional supporting information may be found in the online version of this article.

Grant sponsor: NIH; Grant numbers: R01 DK067436, R01 HL081582, R01 HL096469.

*Correspondence to: Betsy T. Kren, Ph.D., Department of Medicine, University of Minnesota Medical School, MMC 36, 420 Delaware Street S.E., Minneapolis, Minnesota 55455. E-mail: krenx001@umn.edu

Received 3 March 2011; Accepted 19 April 2011 • DOI 10.1002/jcb.23159 • © 2011 Wiley-Liss, Inc.

Published online 25 April 2011 in Wiley Online Library (wileyonlinelibrary.com).

element [Ivics and Izsvák, 2004]. The cut-and-paste mechanism used for transposition requires both the *Sleeping Beauty* transposase and Tn elements containing inverted repeats/direct repeats (IR/DRs) flanking the transgene to be inserted into the host genome. The excision and insertion steps are mediated by the binding of the obligate *Sleeping Beauty* transposase catalytic core, which shares the conserved acidic amino acid DDE motif that is found in many evolutionarily related recombinase proteins, including the V(D)J recombinase and retroviral integrases [Ivics et al., 1997; Plasterk et al., 1999]. Several approaches have been used to improve the *SB*-Tn vector system for more efficient gene transfer and long-term expression; including modifying the transposase and the Tn IR/DRs, as well as using a single *cis* vector to carry both the expression cassette for the *SB* transposase external to the IR/DR flanked transgene for insertion [Cui et al., 2002; Geurts et al., 2003; Mikkelsen et al., 2003; Yant et al., 2004; Zayed et al., 2004; Baus et al., 2005; Jackson et al., 2006; Zhu et al., 2010]. However, to date there have been no reported studies of whether specific DNA elements in the *cis SB*-Tn vector backbone external to the IR/DRs could enhance the effective rate of transposition and/or expression of the transposed transgene. We selected two distinctly different DNA elements that individually or in combination might be anticipated to improve efficiency of the *SB*-Tn system, by aiding in nuclear translocation and retention of a plasmid DNA-based vector.

The size of most plasmid vectors remains a key barrier to efficient nuclear translocation and transgene expression [Lentacker et al., 2008]. An additional hurdle in primary cells is their limited or absent replication both in vivo and ex vivo requiring efficient integration that is not dependent on cell division for nuclear entry of the *SB*-Tn vector. To overcome the inherent inefficiency of non-viral transfection for nuclear translocation of large plasmid DNAs, we chose to include the SV40 enhancer, which contains a 72 bp DNA nuclear targeting signal (NTS) [Dean et al., 1999]. Inclusion of the 72 bp sequence or the entire SV40 enhancer has significantly enhanced transfection efficiency both in isolated primary cells and in vivo [Young et al., 2003].

A method to prolong nuclear half-life of the *cis SB*-Tn vector might also provide enhanced transposition activity. This function can be supplied by the use of a scaffold/matrix attached region (S/MAR) which is composed of DNA elements that attach the chromatin fiber to the proteinaceous network of the nucleus, termed either the nuclear matrix or nuclear scaffold [Berezney and Coffey, 1974]. S/MAR elements (MARs) are typically 70% TA rich and often contain sequences that are similar to the topoisomerase II consensus sequence [Bode et al., 1992]. The precise functions of MAR elements in vivo are largely unknown, but matrix attachment regions have been reported to increase gene expression levels as well as persistence [Harraghy et al., 2008]. The human β-globin MAR has also been implicated in the control of chromosomal activities such as DNA replication and gene expression. In fact, the human β-globin MAR is particularly effective; in one study the frequency of β-galactosidase expressing colonies was increased by up to 80% [Kim et al., 2004].

To test these two elements for improving the integration efficiency, a ubiquitous moderate level promoter, eukaryotic initiation factor 4A1 (IF) was used to drive the *Sleeping Beauty* transposase cassette of the *cis SB*-Tn vector. The basic *cis* IF vector was modified by inserting the SV40 enhancer and/or β-globin MAR 5' to the IF promoter. We assessed the effect of the DNA elements on transposition activity of the *SB*-Tns carrying either an enhanced green fluorescent protein::zeocin resistance fusion gene or a plasmid rescue construct in cultured human hepatoma cells and primary human blood outgrowth endothelial cells (BOECs). Our results indicated that inclusion of the MAR element in the vector backbone improved the transposition frequency and long-term gene expression following insertion of the Tn into the host genome.

MATERIALS AND METHODS

SLEEPING BEAUTY PLASMID CONSTRUCTS

All plasmids were constructed using standard biochemical methods with enzymes obtained from either Promega Corp., Madison, WI or New England Biolabs, Tozier, MA. Inserts and vectors were processed with Klenow polymerase to produce blunt ends if directional cloning was not possible. The basic starting vectors were the kanamycin resistant pKT2 [Cui et al., 2002]; HSB3 and HSB5 *Sleeping Beauty* transposase variants driven by mouse initiation factor promoter 4A1 which replaced the *lacZ* CDS in the parental pDRIVE01-eIF4AI(m) plasmid (InvivoGen, San Diego, CA) [Yant et al., 2004]; SV40 enhancer was from pDRIVE03-SV40enh3-Alb(b) (InvivoGen); β-globin MAR, the bacterial zeocin resistance gene and RK6 origin of replication from pCpG-MCS (InvivoGen); and the zeocin resistance::green fluorescent protein fusion gene from psiRNA4-H1GFPzeo (InvivoGen). All DNA for transfection was prepared using Qiagen EndoFree plasmid kits (Valencia, CA) kits according to the recommended protocol.

CULTURE AND TRANSFECTION OF HUH-7

The human hepatoma cell line HuH-7 [Nakabayashi et al., 1982] was grown in Dulbecco's Modified Eagle Medium (Invitrogen Corp., Carlsbad, CA) supplemented with 10% FBS (Atlanta Biologicals, Norcross, GA) and 100 uM tauroursodeoxycholic acid (EMD Chemicals, Gibbstown, NJ) at 37°C. Eighteen hours prior to transfection, the HuH-7 cells were plated at a density of 6×10^5 cells/35 mm dish in the same media and 2 μg of the largest plasmid in each series, *cis* MARSV40 GFP::zeoR (8.1 kb) or *cis* MARSV40 IFHSB5 and *cis* MARSV40 IFHSB3 (5.6 kb) were transfected using lipofectamine™ 2000 (Invitrogen Corp.) at a ratio of 1 μg DNA:2.5 μl of lipofectamine™ 2000 as previously described [Zhu et al., 2007]. For the other GFP::zeoR fusion and plasmid rescue vectors in each series and their respective controls, Amaxa™ GFP and pT2/DsRed, the same ratio of DNA to lipofectamine™ 2000 was used. But the amount of each plasmid was adjusted to be equimolar with 2 μg of the *cis* MARSV40 constructs. Transfection efficiency was determined by either counting the number of fluorescent cells in three different fields or by FACS 48 h after transfection. The cells transfected with the plasmid rescue or control constructs were split 1:5 every 5 to 7 days into 25 cm^2 flasks for a minimum of 5 passages prior to harvesting for genomic DNA isolation and processing.

Three different culture schemes were employed for the HuH-7 cells transfected with the *SB*-Tn GFP::zeoR vectors. For selection without passaging, 24 h after transfection, 4×10^5 cells were seeded

on 25 cm^2 flasks and selected using 200 μg/ml zeocin beginning 24 h after replating. The selection media was replaced every 2 or 3 days for 4 weeks. The cells were then trypsinized, viable cell number determined by trypan blue staining and 7.5×10^5 cells were plated per 25 cm^2 flask in the presence of 200 μg/ml zeocin. Cells were harvested one day later for analysis by flow cytometry. For selection with passaging, 2×10^5 transfected cells were seeded on 25 cm^2 flasks 24 h after transfection and selected with 200 μg/ml zeocin 24 h after replating. Cells were passaged every 5 to 7 days, with reseeding of 2×10^5 cells per 25 cm^2 flask at each passage. Selection media was replaced every 2 days. For passage with no selection, an additional 2×10^5 cells/25 cm^2 flask were cultured in parallel under the same passage schedule for each of the transfected cell populations. At each passage after 24 h, an additional flask was seeded with the remaining cells up to 2×10^5 cells/25 cm^2 flask from both selected and unselected conditions.

CULTURE AND TRANSFECTION OF PRIMARY HUMAN BLOOD OUTGROWTH ENDOTHELIAL CELLS (BOECs)

Human BOECs were cultured as previously described using recovered frozen BOECs stocks previously isolated as described [Kren et al., 2007]. After one subculture, 1×10^6 BOECs were plated on collagen type I-coated 75 cm^2 flasks and the medium was changed every two or three days until confluent. Five to nine passage BOECs were used in the experiments; and 1×10^6 BOECs were seeded into each collagen I-coated 10 cm dish 16 h prior to transfection. For 12-well plates, 8×10^5 cells were seeded per well.

Human BOECs were transfected with 0.4 μg of the *SB*-Tn constructs per well of 12-well plate using Lipofectamine™ 2000 (Invitrogen) in a 1:2.5 ratio according to the manufacturer's instructions and incubated under normal conditions for 5 h. The transfecting medium was then removed and replaced with 1 ml of fresh supplemented EGM-2 medium (Lonza Inc., Riverside, CA). All cells were simultaneously transfected with 0.4 μg of the different GFP reporter constructs or *SB*-Tn GFP::zeoR fusion vector series. The average transfection frequencies for the GFP reporters were determined by counting green bright cells in a field of ~ 400 cells under fluorescent microscopy 48 h after transfection, replicated three times and confirmed by FACS analysis. One day after transfection of BOECs with the *cis SB*-Tn GFP::zeoR fusion vector series, each well from the 12-well plate was transferred to a 10 cm collagen coated plate. The following day, zeocin (Invitrogen) was added at a concentration of 10 μg/ml. The zeocin-supplemented medium was replaced every 2 or 3 days. Fifteen days after initiation of zeocin treatment, the cells were observed under fluorescent microscope to confirm GFP expression and count the number of GFP^+ cells. The plates were then stained using 1% Commassie Brilliant Blue for 20 min, and after washing, the plates were dried and photographed.

FLOW CYTOMETRY OF HUH-7 CELLS

For the transfected HuH-7 cells, 1 h prior to flow cytometry analysis (FACS) cells were washed with PBS and detached using TryplE Express (Invitrogen) centrifuged at 300 × g for 5 min at 4°C. Cells were resuspended at ~10^6 cells/ml and passed through a 40 μm nylon filter. Twenty minutes prior to analysis, cells were stained with 10 μl 7-ADD (SouthernBiotech, Birmingham, AL) per 10^6 cells, placed on ice and protected from light until FACS. Cytometric data from 3×10^4 singlet events was acquired using a FACSCalibur instrument (BD Biosciences, San Jose, CA, and the data were analyzed using FlowJo 7.5 (TreeStar Inc., Ashland, OR). Non-transfected HuH-7 cells and single color controls were used to define compensation settings and gates for single cells, dead cells, and GFP^+ cells.

PLASMID LABELING AND ANALYSES OF BOECs BY FLOW CYTOMETRY

Plasmid DNA (pDNA) of each of the *SB*-Tn GFP::ZeoR vectors was covalently labeled with ethidium monoazide (EMA) [Lampela et al., 2003]. Briefly, EMA in water [5 μg/ml] was added to the plasmid DNAs [200 ug/ml], and after 10 min incubation at room temperature the mixtures were exposed to 312 nm UV light for 2 min. Free EMA was removed by extraction 3 times with 1-butanol and labeled pDNAs were recovered by ethanol precipitation. EMA-labeled pDNA was used as a marker for intracellular delivery and maintenance of the pDNAs in the BOECs transfected using the optimized protocol. The control and transfected cells were collected 5 h, 1, 4, and 8 days after transfection. Before collection, the cells were washed with PBS, and incubated with 1 ml of 1 M NaCl for 30 s to remove any pDNA complexes bound to cell surface. Following two additional PBS washes, the cells were harvested, fixed by incubating in PBS containing 2% paraformaldehyde for 10 min, washed with PBS, and stored in the dark at 4°C prior to their analysis using a FACScan (Becton Dickinson). Expression of the GFP reporter gene was also analyzed by FACS 1 day after transfection and the GFP^+ cells collected, cultured without selection and reanalyzed 1 week later. The GFP^+ cells were again collected and cultured for an additional 7 days before FACS analysis with 10 thousand events collected for each sample.

WESTERN BLOT ANALYSES

The HuH-7 cells transfected with equimolar amounts of each of the five different *SB*-Tn GFP::zeoR constructs as outlined were harvested 2 or 8 days after transfections or after 7 weeks in culture with passage every 5 to 7 days. For western blot analysis, cells were processed using complete EDTA-free protease inhibitor cocktail (Roche Molecular Diagnostics, Indianapolis, IN) [Trembley et al., 1996]. Total protein (80 μg/lane for 2 and 8 days samples and 40 μg/lane for 7 week samples) was separated by SDS 10% PAGE, and electrophoretically transferred to nitrocellulose membranes. β-actin (lane loading control) and *SB* transposase were detected by ECL using monoclonal anti-β-actin AC15 (Sigma-Aldrich, Saint Louis, MO) or an in-house rabbit polyclonal anti-*SB* transposase [Kren et al., 2003] with HRP-conjugated secondary anti-mouse or anti-rabbit antibodies and Pierce Supersignal West Dura extended duration substrate (Rockford, IL).

The human BOECs were harvested 24 h after transfection. The cells were lysed in the presence of complete EDTA-free protease inhibitor cocktail and analyzed by immunoblotting using the same in-house polyclonal anti-*SB* transposase antibody. GFP and α-tubulin were also detected by ECL using mouse monoclonal anti-GFP sc-9996 and anti-α-tubulin sc-58667 (Santa Cruz

Biotechnology, Santa Cruz, CA), HRP-conjugated anti-mouse secondary antibodies and the same Pierce Supersignal substrate.

RESCUE VECTOR RECOVERY, TRANSFORMATION AND SEQUENCING

DNA was isolated from HuH-7 cells using the Qiagen DNeasy Blood & Tissue kit (Qiagen) according to the manufacturer's recommended protocol. All restriction endonuclease digestions, alkaline phosphatase treatment and ligations were performed using the recommended conditions. Between each step the DNA was recovered using QIAquick PCR Purification Kit (Qiagen). The isolated DNA was processed by SwaI restriction endonuclease digestion, alkaline phosphatase treatment, SspI digestion and ligation. The ligated material was eluted in 10 μl and transformed into *E. coli* GT115 (InvivoGen) via electroporation, recovered for 30 min at 37°C and then plated on [33 μg/ml] zeocin agar plates. The plates were scored for colony number after 24 h incubation at 37°C and 25 colonies were picked and plated in replicates on zeocin [33 μg/ml] or kanamycin [40 μg/ml] to determine potential false positives, e.g., retention and retransformation of the *SB*-Tn vector. Confirmation of the rescue plasmids was performed by restriction endonuclease digestion and the genomic integration sites determined by direct sequencing.

RNA ISOLATION AND RT–PCR ANALYSES

Frozen cells pellets were thawed on ice and total RNA was isolated by RNAeasy mini kit (Qiagen). The isolated RNA was treated with AMP DNase I (Invitrogen), cDNA synthesized using SuperScript III reverse transcriptase (RT) and random hexamer primers (Invitrogen) according to the manufacturer's protocols. Two microliters of the cDNA products were used as the template for PCR using the Expand Hi-Fidelity™ polymerase (Roche Molecular Diagnostics, Indianapolis, IN), dNTPs and enzyme concentrations recommended by the manufacture. PCR was performed with three sets of primers: eGFP forward: 5'-CTGGTGTTGTCCCAATTCTG-3', reverse: 5'-TACAGCTCATCCATTCCCAG-3'; transposase forward: 5'-GCGTTCTGTCTCCTAGAGATG-3', reverse: 5'-GGGTCATTGTCCATTTGGAAG-3'; GAPDH forward: 5'-CCCTTCATTGACCTCAAC-3', reverse: 5'-TTCACACCATCACAAAC-3'. Following an initial denaturation at 94°C for 5 min, the cDNAs were amplified for 18 cycles (30 s at 94°C, 20 s at 55°C, 45 s at 68°C) with a final extension of 68°C for 7 min. For HMGB1 and the GAPDH normalization control, 20ng of total RNA treated with AMP DNase I (Invitrogen), was used as template in the Titan One Tube™ RT-PCR system (Roche Molecular Diagnostics) using the GAPDH primer set listed above or the HMBG1 primer set (forward: 5'-AAGCCGAGAGGCAAAATGTC-3' and reverse: 5'-GAGGCCTCTTGGGTGCATTG-3'). Amplification was performed for 30 cycles using the temperature and cycling conditions for both the RT and cDNA amplification suggested by the manufacturer. Ten microliters of the PCR products were analyzed by electrophoresis on 1.0% agarose gels and visualized by ethidium bromide staining and UV illumination.

STATISTICAL ANALYSIS

Statistical analysis was performed using GraphPad InStat version 3.1 for OSX (GraphPad Software, San Diego, CA) for the ANOVA and Bonferroni's multiple comparison tests. Values of $P < 0.05$ were considered significant.

RESULTS

EXPRESSION LEVELS OF GFP TRANSGENE INCREASE WITH EITHER OF THE DNA ELEMENTS INCLUDED IN THE VECTOR BACKBONE

We sought to test the efficacy of two DNA elements, the SV40 enhancer and β-globin MAR elements when placed outside the IR/DR flanked Tn to modulate both the transposition frequency and efficacy of reporter gene expression. To accomplish this we designed a series of vectors that included the two different DNA elements individually or together upstream of the IF promoter driven hyperactive *Sleeping Beauty* variant 3 (HSB3) transposase (Fig. 1A) [Yant et al., 2004]. To determine the effect of selection of transposition frequency, we used a CMV HTLV driven green fluorescent protein, zeocin resistance (GFP::zeoR) fusion gene as cargo for the Tn, resulting in pDNA vectors ranging in size from 5.6 to 8.1 kb. Following transfection of the human hepatoma cell line, HuH-7 with equimolar amounts of the different GFP::zeoR *SB*-Tn vector series and the DsRed expressing control pKT2, there was no decrease in efficacy with the increased size of the different *SB*-Tn vectors. Equal molar amounts were used with ~ 45% of the cells expressing GFP or DsRed 48 h after transfection. Following selection with zeocin for 4 weeks, representative cultures showed a significant change in GFP expression levels in cells transfected with Tns carrying SV40 and MAR elements on the vector backbone (Fig. 1, B–G). The inclusion of any of the elements gave a more significant increase ($P < 0.0001$) in viable cells than the basic *cis* IF vector ($P < 0.01$) relative to the GFP::zeoR Tn only construct. Moreover, either the SV40 element alone or in combination with the MAR resulted in significantly increased numbers of viable cells over the basic *cis* IF vector (Fig. 1H). Each of the different *cis* Tns exhibited increased levels of GFP fluorescence ($P < 0.001$) relative to the Tn only control. Interestingly, in contrast to cell viability, the SV40 element alone enhanced the expression of GFP over the basic *cis* IF construct ($P < 0.05$) less than the MAR element either alone or in combination ($P < 0.01$, Fig. 1I). These data suggested that the DNA elements in the vector backbone have significant influence over transposition efficiency even though they are lost following HSB3 mediated transposition of the IR/DR flanked GFP::zeoR. As these cells were not passaged and positive selection for expression of the GFP::zeoR fusion gene was used, we transfected additional HuH-7 cells with the same *SB*-Tn GFP::zeoR fusion series and a zeocin sensitive control, SV40 IFHSB3 pKT2.

EFFECT OF THE DNA ELEMENTS IS INDEPENDENT OF SELECTION

The three independent transfection series were scored for transfection efficiency by flow cytometry 2 days after transfection. Each transfection series was split, with one set of the cultures subjected to zeocin selection and both passaged and analyzed by FACS in parallel as outlined in "MATERIALS AND METHODS". FACS analysis indicated that no particular construct or the use of selective culture conditions resulted in a greater number of dead cells as determined by 7-ADD staining (Fig. 2A). Interestingly, in the presence or

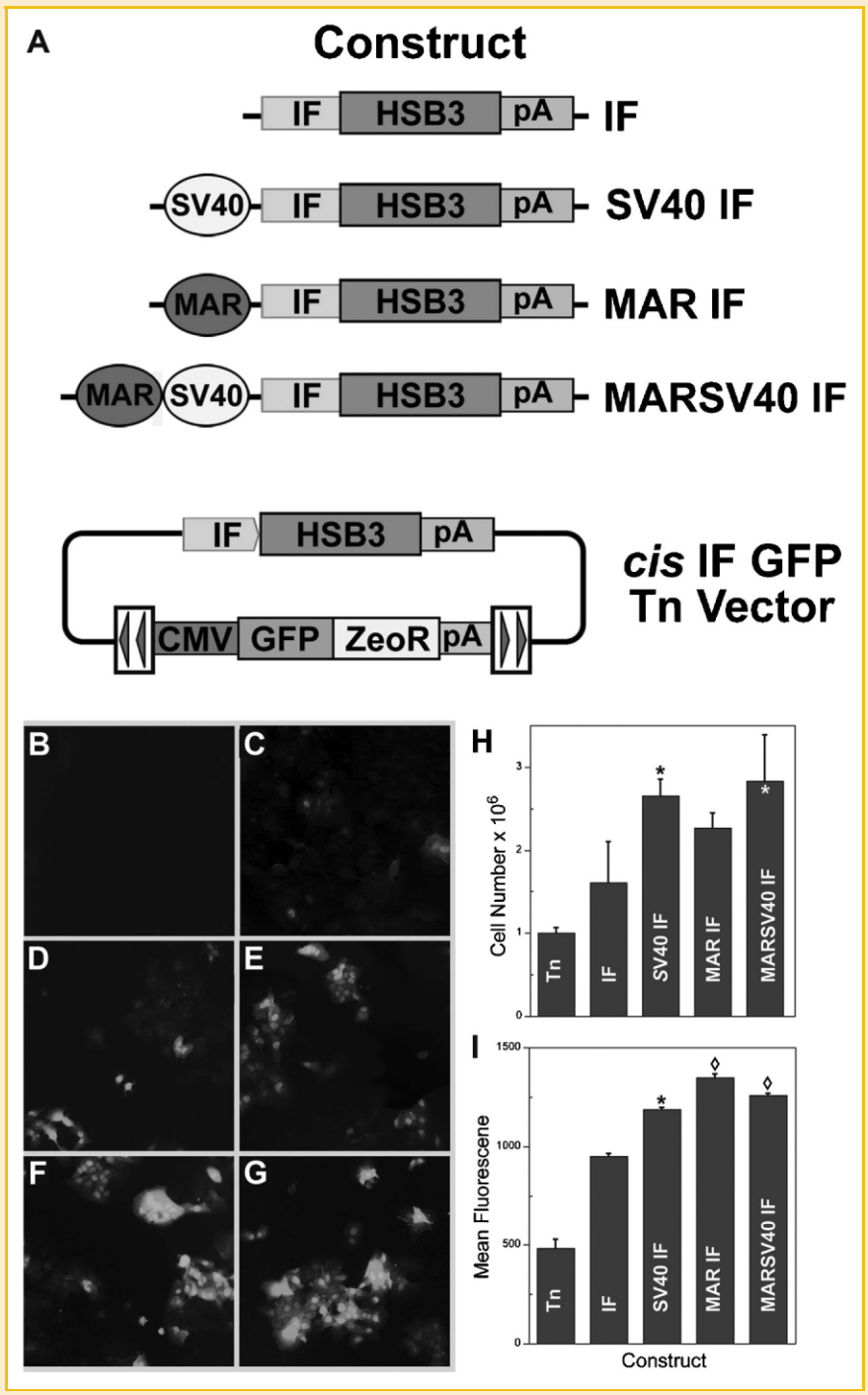

Fig. 1. A: Schematic of *Sleeping Beauty* transposon constructs for HSB3 with the different DNA elements. The basic *cis* IF GFP Tn vector is shown at bottom with the arrow indicating direction of transcription of the IF HSB3 transgene, and structure of the CMV HTLV promoter driven GFP::ZeoR fusion gene flanked by the Tn IR/DRs for transposition into host genomic DNA. Persistent expression levels of the transposed GFP are affected by vector backbone elements. Representative fluorescent micrographs of HuH-7 cells transfected with the different GFP *Sleeping Beauty* transposon constructs following selection with 200 μg/ml zeocin for 5 weeks. B: Control, non-transfected HuH-7 cells; (C) Tn GFP::ZeoR transfected HuH-7 cells; (D) *cis* IF GFP Tn; (E) *cis* SV40 IF GFP Tn; (F) *cis* MAR IF GFP Tn; and (G) *cis* MARSV40 IF GFP Tn. Magnification: B-G, 20 x. H: Cell number of zeocin resistant cells present following selection without passaging for 5 weeks. The mean ± 1 SEM from three separate transfections are. * $P < 0.05$ from Tn GFP::ZeoR transfected HuH-7 cells. I: The mean fluorescence of the transformed and zeocin selected HuH-7 cells was determined by FACS as outlined in "MATERIALS AND METHODS". The data shown is the mean ± 1 SEM. * $P < 0.05$ and ◇ $P < 0.01$ from Tn GFP::ZeoR transfected HuH-7 cells. IF, eukaryotic initiation factor 4A1 promoter; HSB3, hyperactive *Sleeping Beauty* transposase 3; pA, 3′ UTR and polyadenylation signal; SV40, simian virus 40 early enhancer; MAR, β-globin matrix attachment region; CMV, cytomegalovirus HTLV promoter; GFP, green fluorescent protein; ZeoR, zeocin resistance gene; Tn, transposon; IR/DR, *Sleeping Beauty* transposon inverted repeat/direct repeat.

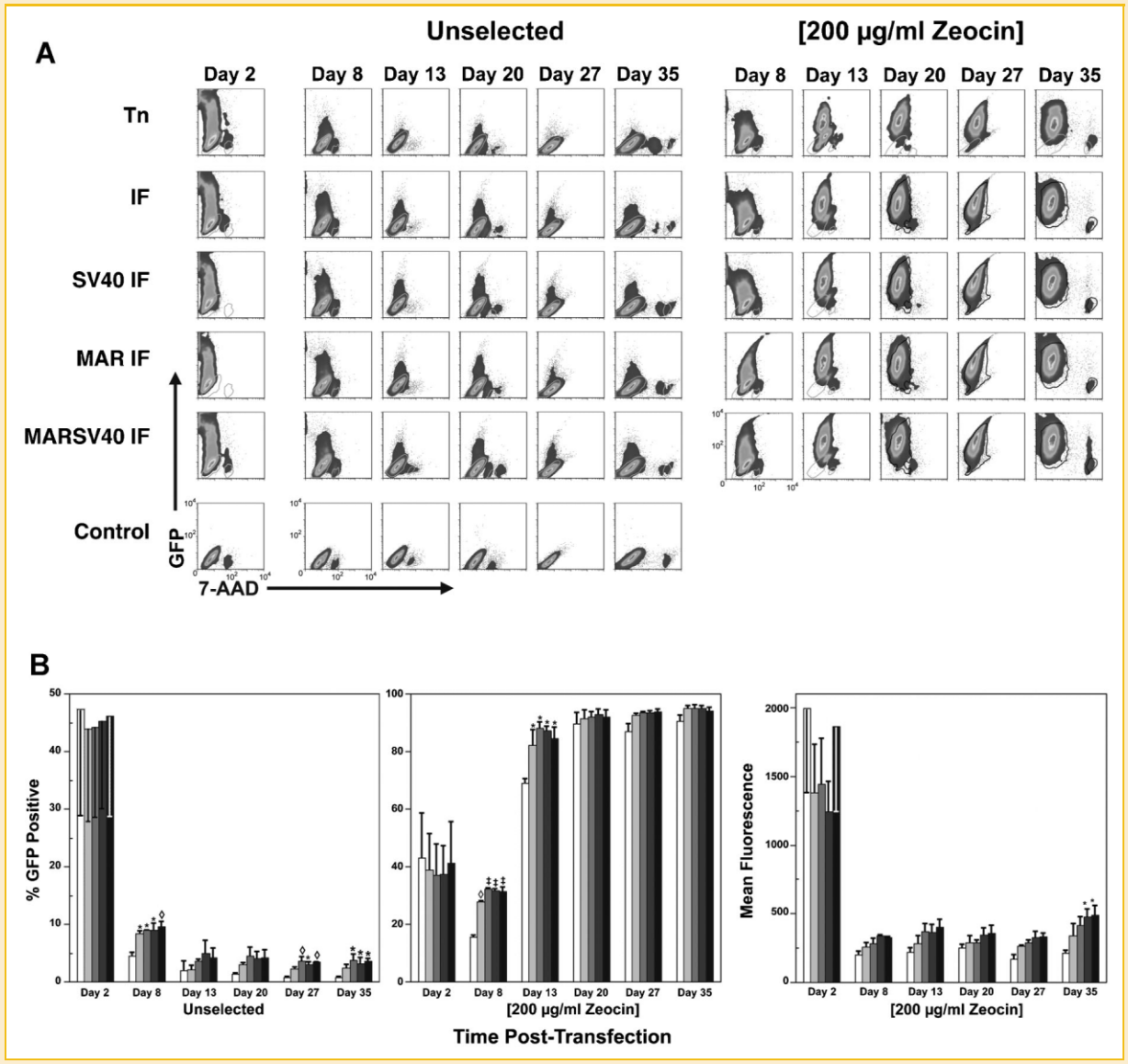

Fig. 2. A: Representative FACS analyses of HuH-7 cells transfected with GFP *Sleeping Beauty* transposon constructs in passaged cells with or without zeocin selection. The cells were transfected as outlined in "MATERIALS AND METHODS" and the transient transfection efficiency determined by FACS analysis 2 days after transfection (far left column) using 7-AAD staining to establish cell viability and the gray overlay showing the position of control HuH-7 cells. The transfected cells were split, and passaged at regular intervals (every 5 to 7 days) with or without (unselected) inclusion of 200 µg/ml zeocin in the culture media. At day 20 and beyond in the zeocin treated group, the black overlay shows the position of the Tn GFP::ZeoR transfected HuH-7 cells. Constructs are shown at left and the day after transfection is indicated above the columns. The lower left FACS plot in each group has the log scale labeled for both the X and Y axes, with the X axes representing the 7-AAD fluorescence and Y axes the GFP fluorescence. B: Inclusion of either the MAR or SV40 element in the vector backbone results in increased levels and number of HuH-7 cells expressing GFP. The percent GFP^+ cells was determined by FACS for both GFP and 7-AAD in the passaged cells cultured with (center panel) or without (unselected, left panel) 200 µg/ml zeocin. The mean ± 1 SEM from n = 3 separate transfection sets is shown. The mean ± 1 SEM fluorescence of the zeocin selected culture sets is depicted in the far left panel. The day post-transfection that the cells were analyzed is indicated below. White, GFP Tn; light gray, *cis* IF GFP Tn; gray, *cis* SV40 IF GFP Tn; dark gray, *cis* MAR IF GFP Tn vector; black, *cis* MARSV40 IF GFP Tn vector. * $P < 0.05$; ◇ $P < 0.01$; ‡ $P < 0.001$, from GFP Tn control.

absence of zeocin selection, the percentage of GFP positive cells at passage 3 (day 13) was obviously reduced in the cultures that received the GFP::zeoR Tn relative to the *cis SB*-Tn constructs suggesting HSB3-mediated transposition of the IR/DR flanked GFP::zeoR into the host genome. We also examined the fluorescence intensity of the zeocin-selected cultures and it was readily apparent that inclusion of the MAR element increased the number of cells expressing high levels of GFP (Supplementary Figure 1). The data from all three trials are summarized in Figure 2B, showing that in the absence of selection (left panel) by the second passage significantly fewer GFP positive cells were observed in the cultures that did not receive the *cis SB*-Tn vectors. By the 5th passage, only the cultures that received the *cis* vectors with SV40 or MAR DNA elements individually or combination exhibited significantly increased

numbers of GFP^+ cells. The results for significantly increased numbers of GFP^+ cells in the cultures with the *cis SB*-Tn vectors was also observed through day 13 in the selected cultures (center panel) implicating HSB3 mediated transposition of the GFP::zeoR Tn into the host genome. Similar to the zeocin selected unpassaged cells, the mean GFP fluorescence expressed persistently was significantly increased over the basic *cis* IF GFP::zeoR Tn and was observed if the MAR element was present alone or in combination with the SV40 enhancer (Fig. 2B, right panel).

SLEEPING BEAUTY TRANSPOSASE IS NOT DETECTABLE AFTER 1 WEEK

As the cut and paste mechanism used for transposition of the IR/DR flanked Tn can generate recircularized excision products [Liu et al., 2004], the *cis* backbone could continue to express sufficient transposase to remobilize the inserted Tn. Thus, we examined the duration of expression of the transposase in the transfected cells by immunoblot. The results indicated that each of the different *cis SB*-Tn GFP::zeoR constructs expressed the transposase robustly and at equivalent levels 2 days post-transfection (Fig. 3A, left lanes 2 though 5). In contrast, by the second passage at day 8 only the cells transfected with the SV40 *cis* construct (Fig. 3A, right lane 4) had any detectable transposase expression although it was substantially reduced compared to the level at day 2 (Fig. 3A, left lane 3), and by 35 days after transfection, none of the transfected cultures expressed detectable levels of transposase (Fig. 3B). Taken together, the data suggested that modification of the vector backbone by inclusion of the two DNA elements does not significantly prolong the expression of the transposase over the basic *cis* IF vector.

INCLUSION OF BOTH ELEMENTS PRODUCES AN ADDITIVE EFFECT ON TRANSPOSITION

To establish the effect of the two vector backbone DNA elements on transposition independent of transgene expression, we generated a series of plasmid rescue vectors where the IR/DR flanked GFP::zeoR was replaced with the RK6 bacterial origin of replication and an alternate hyperactive *Sleeping Beauty*, HSB5 (Fig. 4A). In addition we also constructed a MARSV40 IFHSB3 plasmid rescue vector to compare the efficiency of the two different engineered transposase variants [Yant et al., 2004]. HuH-7 cells were transfected using equimolar amounts of the *SB*-Tn rescue vectors series, passaged 4 times and the plasmids "rescued" as outlined in "MATERIAL AND METHODS". The number of zeoR colonies from each *SB*-Tn in the series representing 2.5 μg of the isolated genomic DNA indicated the inclusion of either element individually in the vector backbone increased the mean number of colonies > 3-fold over the basic *cis* IF vector (Fig. 4B). Interestingly, inclusion of both elements was additive, significantly ($P < 0.01$) increasing the number of colonies recovered compared to the constructs carrying only MAR or SV40 enhancer alone. Moreover, we found that the HSB3 variant was ($P < 0.05$) more efficient than the HSB5 transposase in mediating genomic insertion of the IR/DR flanked Tn. These data suggested

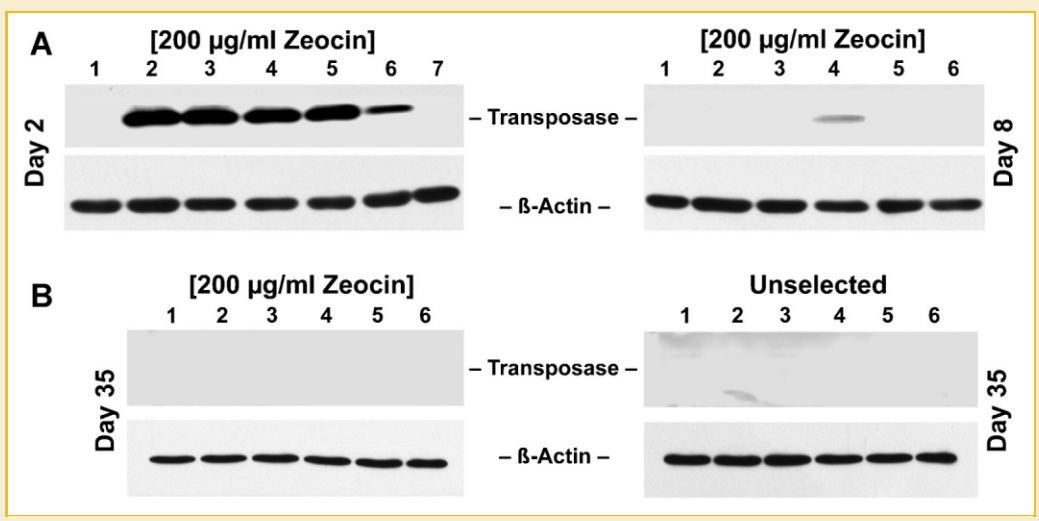

Fig. 3. Immunoblot analysis of *Sleeping Beauty* transposase expression in HuH-7 cells transfected with the GFP::ZeoR Tn vector set. A: Steady-state transposase levels and control β-actin were determined in HuH-7 cells transfected with the different vectors and cultured with zeocin for 2 or 8 days after transfection as described in "MATERIALS AND METHODS". Following harvest of the cells at 2 days or after the second passage at 8 days, total protein was separated by 10% SDS-PAGE and processed for immunoblot analysis as described in "MATERIALS AND METHODS". Proteins were detected by enhanced chemiluminescence. Day 2, lane 1, Tn; lane 2, IF; lane 3, MAR; lane 4, SV40; lane 5, MARSV40; lane 6, IF driven *Sleeping Beauty* expressing plasmid; lane 7, untransfected HuH-7 control. Day 8, lane 1, Tn; lane 2, IF; lane 3, MAR; lane 4, SV40; lane 5, MARSV40; lane 6, untransfected HuH-7 control. B: Expression of transposase is not observed in either selected or unselected cultures passaged multiple times. Cells were harvested at the time of the final FACS analysis 35 days after transfection; total protein was subjected to 10% SDS-PAGE and processed for immunoblot analysis. The steady-state levels of both *Sleeping Beauty* transposase and the lane loading control β-actin were visualized by enhanced chemiluminescence following incubation with either anti-β-actin or anti-transposase primary and HRP-conjugated secondary antibodies as outlined in "MATERIALS AND METHODS". The culture conditions used for the cells are indicated above and the identity of the proteins between immunoblots. Lane 1, Tn; lane 2, IF; lane 3, MAR; lane 4, SV40; lane 5, MARSV40; lane 6, untransfected HuH-7 control.

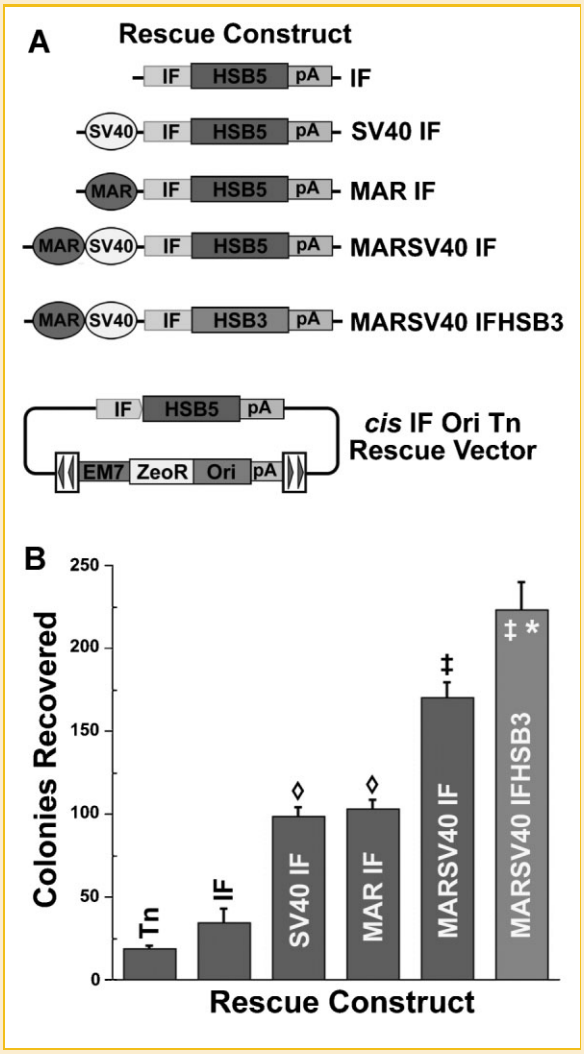

Fig. 4. Schematic of *Sleeping Beauty* transposon constructs. A: The vector backbone *Sleeping Beauty* expression cassettes for HSB5 with the different DNA elements are shown, with construct designation indicated at right. The basic *cis* IF zeoR RK6 replication origin rescue construct is shown at bottom with the arrow indicating direction of transcription of the IF HSB5 transgene, and structure of the zeoR RK6 origin rescue construct flanked by the Tn IR/DRs for transposition into the host genomic DNA. A single HSB3 construct *cis* MARSV40 IFHSB3 zeoR RK6 origin plasmid was included for comparison. B: The data shown represents the mean ± 1 SEM of the number of plasmids recovered from three independent transfection sets following transformation of *E. coli* GT115 with the processed genomic DNA isolated from the long-term cultured cells as outlined in "MATERIALS AND METHODS". ◇ $P < 0.01$; ‡ $P < 0.001$, respectively, from IF; * $P < 0.05$ from MARSV40 IF.

that each element enhances the transposition and that the dual element vector was significantly increased over the single elements alone (Fig. 2B).

We then verified the insertion sites in the recovered plasmids by direct sequencing as previously described [Zhu et al., 2007]. The results shown in Supplementary Table 1 indicated that similar to prior work, the insertion sites were distributed over multiple chromosomes, with no obvious alteration in the number of the insertions in regions associated with identified genes in the vectors containing the DNA elements. Analysis of the 200 bp sequence flanking the genomic insertion sites suggested no significant similarity other than being TA rich, consistent with the previous reports of the predominately random nature of *SB* mediated transposition [Vigdal et al., 2002; Yant et al., 2005]. Thus, inclusion of the MAR or SV40 enhancer sequence does not appear to have significantly altered the site selection for *SB*-mediated transposition.

OPTIMIZATION OF PRIMARY CELL TRANSFECTION CONDITIONS

In contrast to established cell lines, primary cells such as BOECs are difficult to transfect efficiently [Kren et al., 2007]. Multiple transfection conditions to optimize the uptake of larger plasmids were therefore tested. We varied the amount of transfection mixture, tested a series of different concentrations of cells, finding the optimum efficiency using 0.2 μg plasmid, 1 μl lipofectamine in 25 μl optimum medium with 4×10^4 cells per 2 cm^2 well. Using these conditions, we tested a series of different sized GFP plasmids ranging from 4.9 to 9.3 kb. Using fluorescent microscopy, transfection efficiency determined from randomly selected fields was as high as 40% in primary human BOECs transfected using a 4.9 kb GFP expressing Tn without compromising phenotype or proliferative capacity of the BOECs one day after transfection (Fig. 5A). Although decreased efficiency of transformation was observed as size of the plasmid increased, at least 20% of these primary cells were transfected with a 9.2 kb GFP expressing vector (Fig. 5B). However, FACS analysis indicated an approximately 20% increase in efficiency over that determined by counting the GFP^+ cells (Fig. 5C).

MAR INCLUSION INCREASES LONG-TERM EXPRESSION OF THE GFP REPORTER

Using the improved transfection method and the GFP::zeoR *cis* vector series shown in Figure 1 we examined if the SV40 enhancer's nuclear DNA targeting sequence or the ß-globin MAR affected the persistent expression of transgene in BOECs in vitro. The *SB*-Tn GFP::zeoR IR/DR flanked Tn vector and the basic *cis* IF HSB3 were transfected into BOECs and GFP^+ cells were counted 1 day, 1 week and 4 weeks after transfection. Without transposase, the ratio of transgene expressing cells declined in the first week from 70 to 4%, and continued to decrease to less than 0.03% one month later. With the *cis* IFHSB3 vector expressing transposase, the number of BOECs expressing GFP dropped dramatically in the first week from 50 to 1.5%, but then was maintained at levels around 0.1% through one month (data not shown).

We then tested the *SB*-Tn GFP::zeoR IR/DR flanked Tn vector and the *cis* IF HSB3 vector series containing the two elements individually or in combination in a colony-forming assay as described in "MATERIALS AND METHODS". After selection for 2 weeks, the colonies with more than 50 cells were counted under fluorescent microscopy, and stained to visualize the colonies (Fig. 6A). The *cis* IF vector significantly ($P < 0.05$) increased the number of zeoR GFP^+ colonies over the GFP::zeoR Tn from 2.5 ± 0.7 to 12 ± 1.4 (Fig. 6B). Inclusion of the MAR element further increased

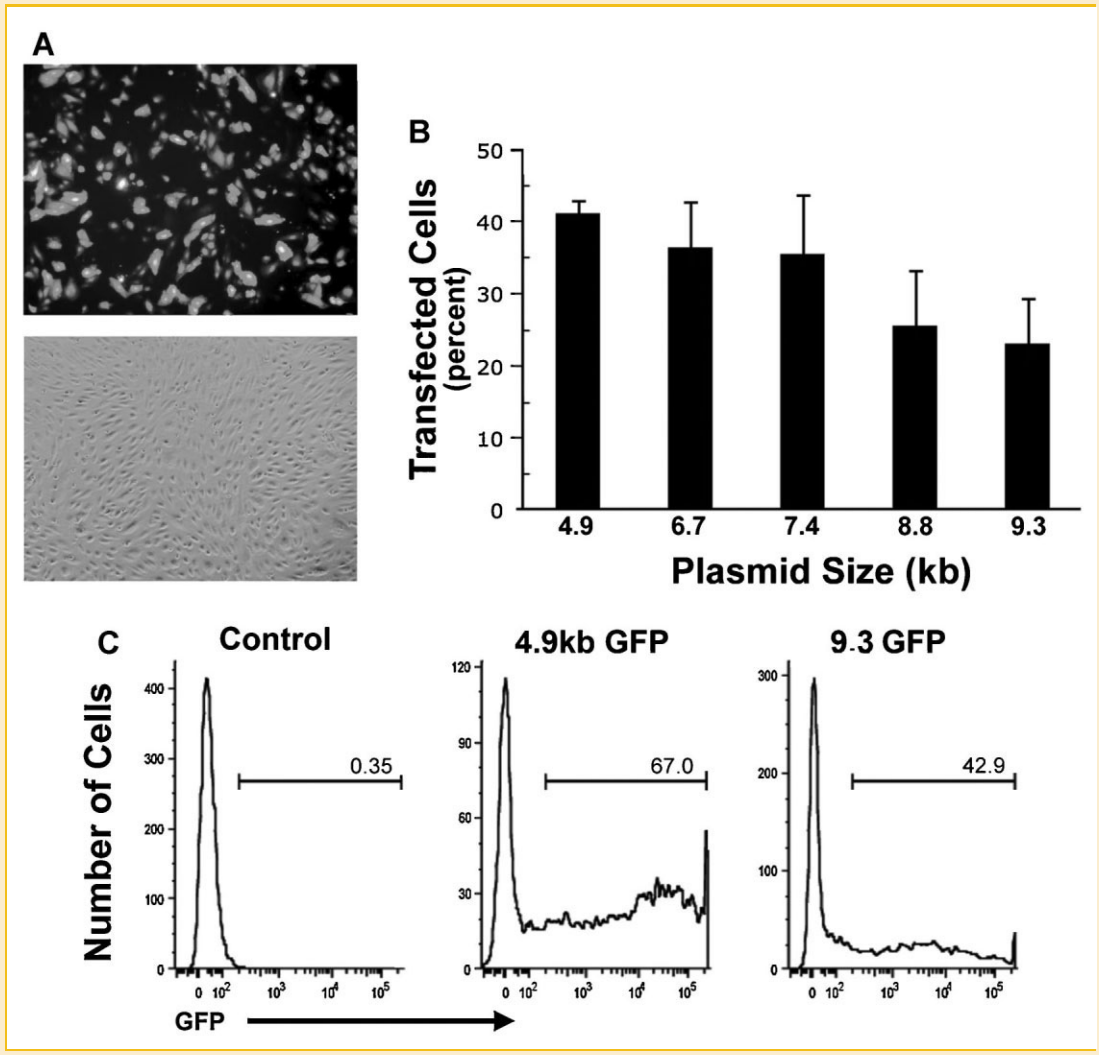

Fig. 5. Optimization of transfection of primary BOEC cells. A: Representative micrographs showing the expression of GFP by fluorescent microscopy (upper panel) and retained phenotype of the BOECs by phase microscopy (lower panel) 24 h after transfection as outlined in "MATERIALS AND METHODS". Magnification was 20X for both panels. B: Data is the mean ± 1 SEM of three independent transfections after scoring a minimum of 400 cells 24 h post transfection. The size of the GFP expressing plasmid used is indicated below the graph. C: FACS confirmation of the transfection efficiency using 4.9 and 9.3 kb GFP expressing constructs. FACS was used to confirm the percent BOECs determined by fluorescent microscopy using the smallest (4.9 kb) and largest (9.3 kb) GFP constructs 24 h following transfection. The number of GFP positive cells is indicated on the Y axes; and GFP fluorescence with the % of transfected cells indicated above the bar on the X axes. Left panel, control untransfected BOECs; middle panel, BOECs transfected with a 4.9 kb GFP expressing construct; right panel, 9.3 kb GFP expressing construct used for transfection.

the number of GFP+ colonies, to 23 ± 6.1 over the basic *cis* IF vector alone ($P < 0.05$). Interestingly, in contrast to the HuH-7 cells, in BOECs the *cis* SV40IF *SB*-Tn give the lowest clone numbers, suggesting a potential for an inhibitory effect on the HSB3-mediated genomic insertion. However, the *cis SB*-Tn with both elements gave the highest level of GFP^+ colonies, 3-fold greater ($P < 0.01$) than observed with only the IF promoter even though the vector was 1 kb larger (Fig. 6). The additive effect of combining both elements indicated that the decrease in colony number observed with the SV40IF *cis* vector was not due to a direct inhibitory affect of the SV40 enhancer element. Taken together, the results suggested that inclusion of the MAR element in the vector backbone increased the long-term expression of *SB*-Tn.

PERSISTENT EXPRESSION OF GFP IN BOECs DOES NOT RESULT FROM PROLONGED EPISOMAL FUNCTION OF pDNA VECTORS FROM MAR ELEMENTS

Given the clear difference between the results obtained with vectors carrying the single MAR or SV40 elements in the primary human BOECs, we strove to find a mechanistic explanation for the superiority of the MAR element in generating GFP^+zeoR colonies. In addition to the other functions of MARs, literature reports have demonstrated that MAR(s) can function as a *cis*-acting element for episomal maintenance of pDNA in cultured eukaryotic cells [Wong et al., 2009]. To determine if this was the cause of the apparent superiority of the MAR element containing vectors in promoting colony formation, we covalently labeled all the different *cis SB*-Tn

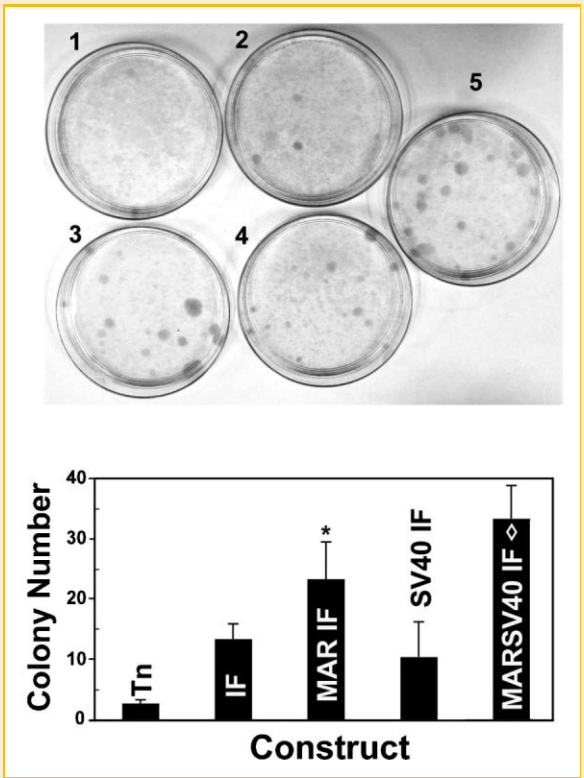

Fig. 6. The MAR element significantly enhances the number of GFP$^+$, zeocin resistant colonies recovered. The BOECs were transfected as outlined in "MATERIALS AND METHODS" with the same GFP Tn control and *cis* HSB3 GFP Tn vector series used in the HuH-7 cells. Forty eight hours after transfection the BOECs were subjected to zeocin selection for 2 weeks, followed by fixation and Giemsa staining to establish the number of colonies exhibiting persistent expression of the GFP::zeoR fusion gene. A: A single transfection set of BOECs is shown following the 2-week selection with zeocin and Giemsa staining to detect viable colonies. 1, Tn transfected; 2, IF transfected; 3, MAR IF transfected; 4, SV40 IF transfected; 5, MARSV40 IF transfected BOECs. B: Data shown is the mean ± 1 SEM from three independent transfections followed by 2-week zeocin selection periods. * $P < 0.05$; ‡ $P < 0.01$; respectively from the *cis* IF GFP Tn vector.

plasmids and the GFP::zeoR Tn only vector with ethidium monoazide (EMA) [Lampela et al., 2003], and subjected the transfected BOECs to FACS as described in "MATERIALS AND METHODS". The results showed that the numbers of EMA-positive cells decreased rapidly over time, with a greater than 50% decrease in positive cells between 5 h, 1 and 4 days (Fig. 7A). There were almost no EMA-positive cells 8 days after transfection, and of particular note no significant increase was observed with the *cis SB*-Tns having the MAR element at any time point sampled. The results suggested that long-term expression of GFP and increased numbers of colonies formed in the transfected BOECs is not mediated by the MAR element on the *SB*-Tn vector backbone maintaining the relevant pDNAs in an episomal state.

In order to rule out an effect of drug selection on the results obtained with the single SV40 enhancer element, we used FACS analysis to study the ratio of GFP$^+$ cells over time following transfection of the BOECs with the different *cis* GFP::zeoR *SB*-Tns and control pDNAs. On the first day after transfection, 40–50% of the transfected human BOECs expressed GFP (Fig. 7B, upper row). The GFP$^+$ cells from each *SB*-Tn transfected culture were collected and cultured for an additional seven days and re-analyzed by FACS. Less than 1% of the collected BOECs expressed GFP one week later; however, the cells transfected with *cis* vectors with the MAR element had twice the number of GFP$^+$ cells than those without. The GFP$^+$ cells collected on day 8 were placed in culture for 1 week and re-analyzed indicating that about half of the BOECs-transfected with the *cis* GFP::zeoR *SB*-Tns without the MAR were GFP$^+$. In contrast, more than 90% of the BOECs transfected with the *SB*-Tn constructs carrying the ß-globin MAR were GFP$^+$, suggesting that the MAR element mediates persistent expression of the GFP::zeoR transgene cargo.

EFFICIENCY OF TRANSPOSITION IS ASSOCIATED WITH TRANSPOSASE EXPRESSION

Having excluded the possibility that the MAR enhanced persistent GFP expression by maintaining episomal pDNA, we investigated whether the difference between the single SV40 element *cis* vector and the dual MARSV40 *cis* SB-Tn was potentially due to transposase expression. Twenty-four hours after transfection, the steady-state mRNA level for transposase was increased for *cis SB*-Tns with the either the SV40 or MAR element compared to the IF promoter alone (Fig. 8A, lane 2 versus lanes 3 to 5). The level of GFP transcript produced by the *SB*-Tns was similar indicating the differences were not likely due to variance in transfection efficacy [Izsvák et al., 2000; Mikkelsen et al., 2003]. While the rate of transposition has been associated with transposase expression, in our study, the *SB*-Tn with the SV40 element had the fewest GFP$^+$ colonies, suggesting post-transcriptional regulation of the transposase. Thus we investigated the steady-state levels of HSB3 transposase by immunoblot analyses in the BOECs transfected with the four different *cis* vectors. The results of the western blot analysis (Fig. 8B) showed that the level of the transposase was decreased in BOECs transfected with the *cis SB*-Tn carrying the SV40 element alone and increased with the *cis SB*-Tn MAR element with or without the SV40 element. This was consistent with the decreased number of GFP$^+$ colonies observed with the *cis* SV40 IF GFP::zeoR *SB*-Tn. Taken together these findings suggested that in BOECs, integration efficiency was associated with levels of transposase expression.

Previous studies have shown that protein expression can be regulated through the ubiquitin pathway [Zhu et al., 2007]. To determine whether the transposase expressed by the *cis* SV40 IF GFP::zeoR *SB*-Tn vector was regulated by the same proteolytic pathway in the BOECs, we carried out studies using MG132, a well-known proteosome inhibitor. MG132 was added to the medium 24 h after transfection, the cells cultured an additional 24 h before collection and analyzed by western blot. Bands of transposase did not change in these three lanes compared with the α-tubulin control. The transposase protein levels did not increase following treatment with 0.5 or 5 μM MG132 (Fig. 8C, lanes 2 and 3). This indicated that the decline in transposase expression with the SV40 enhancer was unrelated to proteosomal degradation. Taken together, the results of RT-PCR and proteosomal inhibitor treatment suggested that translational mechanisms were responsible for the difference in transposase levels observed.

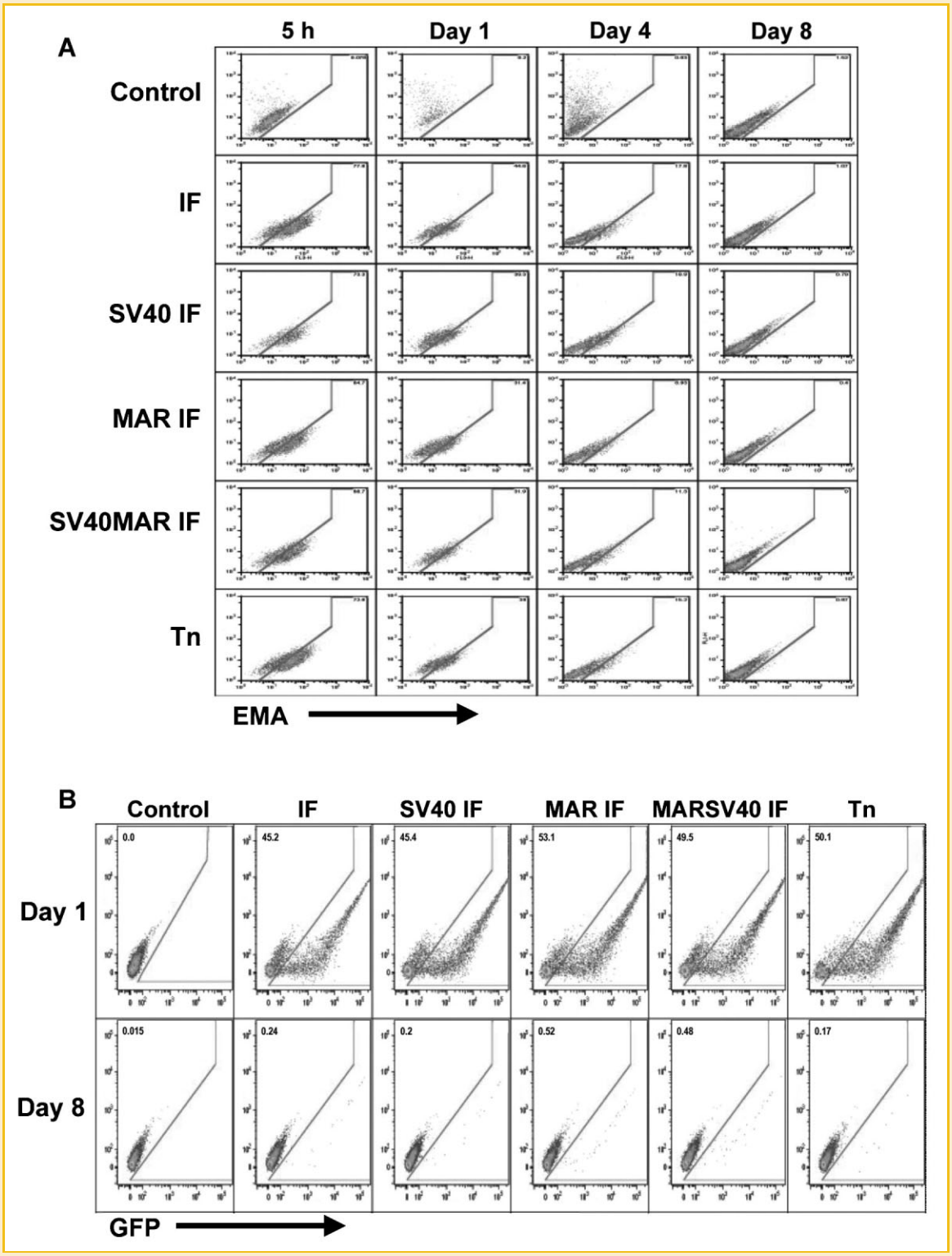

Fig. 7. Persistent expression of GFP is not from pDNA vectors maintained in hBOECs by the β-globin MAR. The control Tn GFP::ZeoR vector as well as the different *cis* HSB3 GFP Tn vectors were covalently labeled using EMA as outlined in "MATERIALS AND METHODS". Following transfection of BOECs with EMA labeled constructs, the cultures were subjected to FACS analysis (A) 5 h, 1, 4, and 8 days after transfection for detection of the EMA labeled pDNA. The X axes represent the EMA fluorescence with the pink line indicating the gating used with the EMA positive cells to the right of the line. Time after transfection is indicated above the columns and the construct used for transfection at left with control representing untransfected BOECs. B: BOECs transfected with unlabeled control Tn and the *cis* HSB3 vector series as outlined in "MATERIALS AND METHODS" were also subjected to FACS 1 day after transfection and the GFP^+ cells collected, cultured for an additional 7 days in the absence of zeocin selection and reanalyzed by FACS on day 8. The percentage of GFP positive cells is indicated in the upper right hand corner of each plot; the X axes represent the GFP fluorescence and the pink line shows the gating used with the GFP^+ cells to the right of the line. The day the BOECs were analyzed is indicated at left and the construct used for transfection indicated above each column.

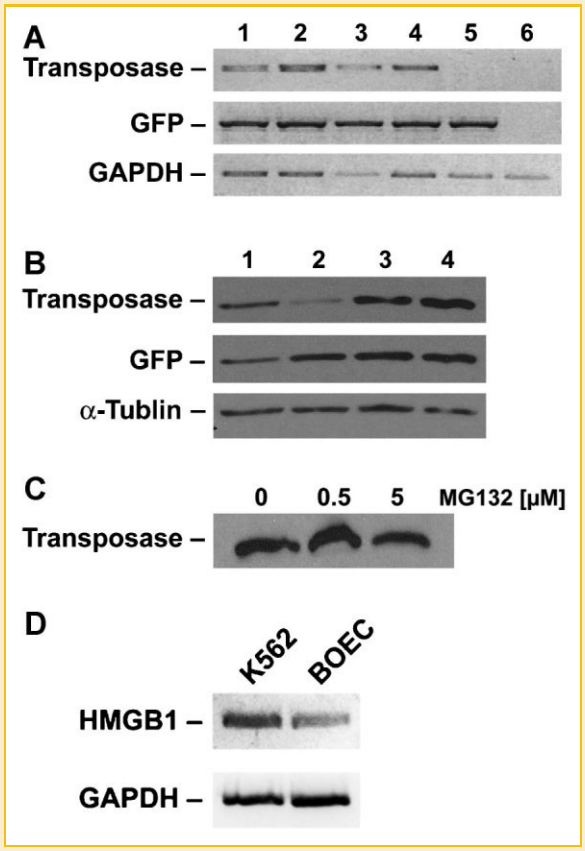

Fig. 8. Transcript and protein levels for *Sleeping Beauty* transposase are decreased in BOECs with a single SV40 enhancer element. A: RNA was isolated from a portion of the harvested transfected cells and subjected to RT-PCR analysis to determine steady-state transcript levels for the transposase, GFP transgene, and GAPDH as control. Following PCR amplification using the primer pairs and conditions outlined in "MATERIALS AND METHODS", the PCR products were subjected to agarose gel electrophoresis and visualized using ethidium bromide staining at 302 nm. The gene identity of the PCR amplicons is shown at left. Lane 1, IF; lane 2, SV40; lane 3, MAR; lane 4, MARSV40; lane 5, Tn; lane 6, untransfected BOEC control. B: BOECs transfected with the GFP::zeoR Tn construct set were harvested 24 h after transfection and processed for immunoblot analysis as outlined in "MATERIALS AND METHODS". The isolated total protein was subjected to 10% SDS-PAGE followed by transfer and incubation with primary anti-*Sleeping Beauty* transposase or anti-GFP or anti-α-tubulin antibodies. After incubation with secondary-HRP conjugated antibodies, the proteins were detected by enhanced chemiluminescence with the identity of the protein indicated at right. Lane 1, IF; lane 2, SV40; lane 3, MAR; lane 4, MARSV40. C: BOECs were transfected with the SV40 construct and after 24 h the proteosome inhibitor MG132 was added and the cells cultured for an additional 16 h before harvesting and processing for immunoblot analysis as outlined in "MATERIALS AND METHODS". Following incubation with the primary anti-transposase and secondary-HRP conjugated secondary antibody, the transposase was detected by enhanced chemiluminescence. The concentration of MG132 that was included in the culture media is indicated. D: Steady-state transcripts of the host-encoded cofactor HMGB1 were isolated from transfected BOECs and the K562 erythroid cell line and subjected to RT-PCR for HMGB1 mRNA levels using the conditions outlined in "MATERIALS AND METHODS". The HMGB1 PCR products and those from GAPDH amplified in parallel were separated by agarose gel electrophoresis as described. The gene identity of the PCR amplicon is indicated at the left and the cells used for the template RNA are shown above.

TRANSPOSITION EFFICIENCY CORRELATES WITH HOST COFACTOR HMGB1 LEVELS

The results indicated that even with the optimized transfection protocol and improved transposition efficiency with the *SB*-Tns carrying the MAR elements of transposase the long-term transgene expression achieved was significantly lower in the BOECs than observed with other cell lines and primary cells. HMGB1 (eukaryotic high mobility group) is a host-encoded cofactor of required for efficient SB-mediated transposition of the IR/DR flanked transgene in mouse cells [Zayed et al., 2003]. To determine basal expression of HMGB1, mRNA was isolated from both BOECs and the K562 human erythroid cell line, with high levels of SB-mediated transposition. The isolated RNA was subjected to RT-PCR as outlined in "MATERIALS AND METHODS" using GAPDH as the normalization control. The results indicated that the level of HMGB1 transcript was substantially reduced in the human BOECs relative to the K562 cell line, suggesting that low HMGB1 levels may be related to decreased SB-mediated transposition observed in the BOECs (Fig. 8D).

DISCUSSION

High rates of transposition and persistent transgene expression must be achieved in considering the use of *Sleeping Beauty* transposon for successful gene therapy. However, inclusion of DNA elements such as strong enhancers within the transgene cargo may achieve high levels of transgene expression but can result in reduced cell survival and/or activation of endogenous host genes [Walisko et al., 2008; Zhu et al., 2010]. In contrast, DNA elements on the vector backbone are lost when the Tn is inserted into the host genome, thus eliminating the potential for modification of endogenous gene expression and reduced transposition efficiency with increased Tn size [Izsvák et al., 2000; Geurts et al., 2003]. While the larger *cis SB*-Tn vectors may reduce transfection efficiency, the first order delivery kinetics, increased frequency of transposition [Kren et al., 2003; Mikkelsen et al., 2003] and increased fidelity of transposition [Huang et al., 2010] suggest significant advantages of the *cis* over the two-plasmid *trans* system for primary cells.

To compensate for the low transfection activity observed with non-viral delivery agents in quiescent and/or primary cells, multiple approaches to overcoming the nuclear barrier have been developed [Lentacker et al., 2008]. The simplest have exploited DNA sequences incorporated into the vectors such as transcription factor binding sites or the SV40 NTS to provide efficient nuclear translocation of the transfected pDNAs [Dean et al., 1999; Miller and Dean, 2008]. The incorporation of a single SV40 enhancer element into the *cis SB*-Tn vector backbone significantly enhanced the efficiency of transposition in the HuH-7 cells with or without drug selection. This effect was not due to an increased transfection efficiency of the HuH-7 cell line. Interestingly, in contrast to results observed in the HuH-7 cell line, inclusion of the SV40 element provided no advantage relative to the basic vector alone in BOECs. No increase in the rate of transfection, number of GFP^{+}zeoR colonies formed or number of cells persistently expressing GFP by flow cytometry was observed in the BOECs.

Previous studies on *SB* transposition activity indicated that it is affected not only by the sequence of the IR/DRs and activity of transposase but also by their molecular ratio [Mikkelsen et al., 2003]. Within a certain range of concentrations, the increasing amount of transposase normally present in cells results in a higher level of transposition. However, excessive transposase leads to the inhibition of transposition activity, a phenomenon termed ''overproduction inhibition'' [Wilson et al., 2005]. We examined the levels of transposase expressed by the different constructs to determine differential modulation by the SV40 enhancer. If fact, there was no change in the initial steady-state expression levels of the *SB* transposase in HuH-7 cells. However, the duration of transposase expression was prolonged suggesting a longer episomal persistence of the *SB*-Tns carrying a vector with the NTS. In contrast, inclusion of the SV40 enhancer did enhance the initial steady-state transcript levels in BOECs, although surprisingly, the increase was associated with reduced protein levels. This provided one potential reason for the very different results observed using the SV40 element alone in the two cell lines.

MARs comprise one of the few classes of eukaryotic noncoding DNA that are involved in the attachment of chromatin to nuclear matrix, chromatin remodeling, and transcription regulation [Gluch et al., 2008]. They have been shown to enhance transgene expression, inhibit DNA methylation at the insertion site, and produce persistent episomal transgene expression. Inclusion of a MAR element binding protein within the transgene cargo biased the Tn insertion site to the vicinity of MAR sequences [Ivics et al., 2007]. Additionally, MAR binding sites are TA rich and associate with DNA curvature [Fiorini et al., 2006], both important DNA sequence characteristics for *SB*-mediated insertion sites [Vigdal et al., 2002; Liu et al., 2005; Geurts et al., 2006], suggesting that the known tethering of MARs to the nuclear matrix via MAR binding protein(s) might enhance the frequency of *SB*-mediated transposition. In contrast to results with the SV40 element, inclusion of a single β-globin MAR element in the *cis* IF GFP::zeoR Tn significantly increased the rate of transposition over the basic *cis* IF *SB*-Tn in both types of cells, with or without zeocin selection. Notably, when the elements were combined an additive effect was observed despite the larger size.

The additive effect of combining both elements was much more pronounced in the primary BOECs than with the plasmid rescue constructs in the HuH-7 cells. In BOECs, the MAR element restored transposase protein levels potentially offsetting the negative impact of the SV40 element on transposition activity. The further increase observed in BOECs was most likely due to the NTS carried by the SV40 enhancer. Use of the rescue plasmids also indicated that the HSB3 hyperactive transposase was significantly more active than the HSB5 transposase in mediating insertion events, demonstrating the utility of this approach for evaluating the activity of the different transposases independently of transgene expression. By inclusion of arbitrary non-coding DNA fragments, the size of the plasmid rescue vector cargo can be increased in size to model any transgene providing a mechanism for rapidly screening the efficacy of the different transposase variants and vector designs.

The beneficial effect of the MAR element on the persistent GFP::zeoR transgene expression in both cell lines is of particular note. Although MARs are known to enhance transgene expression when directly flanking a construct in *cis*, they are also known to provide episomal pDNA maintenance in eukaryotic cells [Wong et al., 2009]. The later possibility was investigated using EMA labeled *cis* MAR *SB*-Tn vectors and long-term culture and indicated that there was no increased episomal function of the MAR containing pDNAs. The significant increase with the single MAR element over the basic *cis* IF GFP::zeoR *SB*-Tn observed in both cell types may be due to tethering of the *cis SB*-Tn vector to the nuclear matrix via interaction of the MAR and MARBP(s), increasing the likelihood of transposase catalyzed integration in the nucleus [Yant et al., 2002]. A potential explanation for the increased gene expression from MAR-containing vectors is that the MAR directed the insertion of the IR/DR flanked transgenes into regions of the host genome less susceptible to epigenetic silencing or open euchromatin regions [Muller et al., 2001; Molto et al., 2009]. However, no specific bias toward a particular chromosome or location was noted in the limited sample set. The plasmid rescue vectors attempt to assay all insertion sites, and not only those ''effective'' transpositions that lead to persistent expression of the transgene. It is possible that the ''effective'' insertion sites would be biased as previously observed [Ivics et al., 2007].

Although the MAR containing vectors showed significantly enhanced persistent gene expression of the transposed transgene, decreased expression in the nonselected cultures over time was noted by FACS analysis. This was most likely due to epigenetic silencing via methylation of the Tn cargo transgene which may in fact, be partially triggered by the very high level of CpGs present in the GFP coding sequences [Zhu et al., 2007] or an effect of the viral enhancer/promoter [Gill et al., 2009]. Although addition of insulators flanking the transgene constructs has been the traditional approach, recent work has suggested that the use of all mammalian DNA sequences for construction of transgenes dramatically reduces the level of post-integration epigenetic modification [Chen et al., 2008]. Thus, transgenes for therapeutic use if comprised entirely of mammalian DNA sequences should be less susceptible to epigenetic silencing than the viral promoter driven GFP::zeoR fusion reporter used in these studies.

Although inclusion of the SV40 enhancer and β-globin MAR in the vector backbone enhanced the rate of transposition and long-term expression of the GFP::zeoR transgene, less than 1% of the primary BOEC cells persistently expressed GFP indicative of successful transposition. In contrast, K562 cells that have been one of the most successful cell lines for *SB*-mediated transposition achieving rates as high as 60% of cells expressing the DsRed reporter gene following Tn integration. Interestingly, we characterized the expression of transcript encoding the host co-factor HMGB1, which has been shown to be rate limiting for *SB*-mediated transposition [Zayed et al., 2003; de Silva et al., 2010]. The levels of HMGB1 mRNA were substantially lower in the BOECs than K562 cells, suggesting that HMGB1 might be a rate limiting factor for transposition.

During the time that these studies were in progress other advances in achieving increased transposition activity using the *SB*-Tn system in primary cells and in vivo have been reported. In particular, a new hyperactive *SB*100X that is clearly superior to any previous

transposase in achieving high levels of Tn insertion ex vivo has been developed [Mates et al., 2009; Xue et al., 2009]. Coupling these improvements with the additional enhancement provided by the inclusion of the SV40MAR elements on the *cis* vector backbone may result in the ability to achieve the high levels of persistent transgene expression demonstrated in other cell types. Finally, for hepatocyte directed gene therapy in vivo, the novel sub 50 nm nanocapsule delivery system using asialoorasomuccoid (ASOR) as the targeting ligand is able to specifically deliver plasmids as large as 12 kb to hepatocytes providing effective transfection of the majority of hepatocytes [Kren et al., 2009]. The rapid progress of the *SB*-Tn from the bench to clinic in a little more than a decade suggests that this non-viral integrating vector system paired with the ASOR nanocapsule delivery in vivo may finally be able to achieve the goal of effective clinical gene therapy for hepatocyte based disorders [Williams, 2008].

ACKNOWLEDGMENTS

The authors are grateful to Phillip Y-P Wong and Laura Angelus for technical help. They would also like to acknowledge the assistance of the Flow Cytometry Core Facility of the Masonic Cancer Center, a comprehensive cancer center designated by the National Cancer Institute and supported by NIH grant P30 CA77598. This work was supported in part by NIH grants [R01 DK067436] to BTK and [R01 HL081582 and R01 HL096469] to CJS.

REFERENCES

Baus J, Liu L, Heggestad AD, Sanz S, Fletcher BS. 2005. Hyperactive transposase mutants of the *Sleeping Beauty* transposon. Mol Ther 12: 1148–1156.

Berezney R, Coffey DS. 1974. Identification of a nuclear protein matrix. Biochem Biophys Res Commun 60:1410–1417.

Bode J, Kohwi Y, Dickinson L, Joh T, Klehr D, Mielke C, Kohwi-Shigematsu T. 1992. Biological significance of unwinding capability of nuclear matrix-associating DNAs. Science 255:195–197.

Chen ZY, Riu E, He CY, Xu H, Kay MA. 2008. Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. Mol Ther 16:548–556.

Ciuffi A. 2008. Mechanisms governing lentivirus integration site selection. Curr Gene Ther 8:419–429.

Cui Z, Geurts AM, Liu G, Kaufman CD, Hackett PB. 2002. Structure-function analysis of the inverted terminal repeats of the *Sleeping Beauty* transposon. J Mol Biol 318:1221–1235.

Daniel R, Smith JA. 2008. Integration site selection by retroviral vectors: Molecular mechanism and clinical consequences. Hum Gene Ther 19:557–568.

de Silva S, Lotta LT Jr, Burris CA, Bowers WJ. 2010. Virion-associated co-factor HMGB1 facilitates transposition from the HSV/*Sleeping Beauty* amplicon vector platform. Hum Gene Ther 21:1615–1622.

Dean DA, Dean BS, Muller S, Smith LC. 1999. Sequence requirements for plasmid nuclear import. Exp Cell Res 253:713–722.

Fiorini A, Gouveia Fde S, Fernandez MA. 2006. Scaffold/Matrix Attachment Regions and intrinsic DNA curvature. Biochemistry (Mosc) 71:481–488.

Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, Bell JB, Largaespada DA, Hackett PB. 2003. Gene transfer into genomes of human cells by the *Sleeping Beauty* transposon system. Mol Ther 8:108–117.

Geurts AM, Hackett CS, Bell JB, Bergemann TL, Collier LS, Carlson CM, Largaespada DA, Hackett PB. 2006. Structure-based prediction of insertion-site preferences of transposons into chromosomes. Nucleic Acids Res 34: 2803–2811.

Gill DR, Pringle IA, Hyde SC. 2009. Progress and prospects: The design and production of plasmid vectors. Gene Ther 16:165–171.

Gluch A, Vidakovic M, Bode J. 2008. Scaffold/matrix attachment regions (S/MARs): Relevance for disease and therapy. Handb Exp Pharmacol 186:67–103.

Gresch O, Engel FB, Nesic D, Tran TT, England HM, Hickman ES, Körner I, Gan L, Chen S, Castro-Obregon S, Hammermann R, Wolf J, Müller-Hartmann H, Nix M, Siebenkotten G, Kraus G, Lun K. 2004. New non-viral method for gene transfer into primary cells. Methods 33:151–163.

Harraghy N, Gaussin A, Mermod N. 2008. Sustained transgene expression using MAR elements. Curr Gene Ther 8:353–366.

Huang X, Haley K, Wong M, Guo H, Lu C, Wilber A, Zhou X. 2010. Unexpectedly high copy number of random integration but low frequency of persistent expression of the *Sleeping Beauty* transposase after *trans* delivery in primary human T cells. Hum Gene Ther 21:1577–1590.

Ivics Z, Izsvák Z. 2004. Transposable elements for transgenesis and insertional mutagenesis in vertebrates: A contemporary review of experimental strategies. Methods Mol Biol 260:255–276.

Ivics Z, Hackett PB, Plasterk RH, Izsvák Z. 1997. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91:501–510.

Ivics Z, Katzer A, Stuwe EE, Fiedler D, Knespel S, Izsvák Z. 2007. Targeted *Sleeping Beauty* transposition in human cells. Mol Ther 15:1137–1144.

Izsvák Z, Ivics Z, Plasterk RH. 2000. *Sleeping Beauty*, a wide host-range transposon vector for genetic transformation in vertebrates. J Mol Biol 302:93–102.

Jackson DA, Juranek S, Lipps HJ. 2006. Designing nonviral vectors for efficient gene transfer and long-term gene expression. Mol Ther 14:613–626.

Kim JM, Kim JS, Park DH, Kang HS, Yoon J, Baek K, Yoon Y. 2004. Improved recombinant gene expression in CHO cells using matrix attachment regions. J Biotechnol 107:95–105.

Kren BT, Ghosh SS, Linehan CL, Roy-Chowdhury N, Hackett PB, Roy-Chowdhury J, Steer CJ. 2003. Hepatocyte-targeted delivery of Sleeping Beauty mediates efficient, inheritable gene transfer in vivo. Gene Ther Mol Biol 7:231–240.

Kren BT, Yin W, Key NS, Hebbel RP, Steer CJ. 2007. Blood outgrowth endothelial cells as a vehicle for transgene expression of hepatocyte-secreted proteins via *Sleeping Beauty*. Endothelium 14:97–104.

Kren BT, Unger GM, Sjeklocha L, Trossen AA, Korman V, Diethelm-Okita BM, Reding MT, Steer CJ. 2009. Nanocapsule-delivered *Sleeping Beauty* mediates therapeutic Factor VIII expression in liver sinusoidal endothelial cells of hemophilia A mice. J Clin Invest 119:2086–2099.

Lampela P, Elomaa M, Ruponen M, Urtti A, Mannisto PT, Raasmaja A. 2003. Different synergistic roles of small polyethylenimine and Dosper in gene delivery. J Control Release 88:173–183.

Lentacker I, Vandenbroucke RE, Lucas B, Demeester J, De Smedt SC, Sanders NN. 2008. New strategies for nucleic acid delivery to conquer cellular and nuclear membranes. J Control Release 132:279–288.

Li SD, Huang L. 2006. Gene therapy progress and prospects: Non-viral gene therapy by systemic delivery. Gene Ther 13:1313–1319.

Liu G, Aronovich EL, Cui Z, Whitley CB, Hackett PB. 2004. Excision of *Sleeping Beauty* transposons: Parameters and applications to gene therapy. J Gene Med 6:574–583.

Liu G, Geurts AM, Yae K, Srinivasan AR, Fahrenkrug SC, Largaespada DA, Takeda J, Horie K, Olson WK, Hackett PB. 2005. Target-site preferences of *Sleeping Beauty* transposons. J Mol Biol 346:161–173.

Mates L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela DP, Schmitt A, Becker K, Matrai J, Ma L, Samara-Kuko E, Gysemans C, Pryputniewicz D, Miskey C, Fletcher B, Vandendriessche T, Ivics Z, Izsvák Z. 2009. Molecular evolution of a novel hyperactive *Sleeping Beauty* transposase enables robust stable gene transfer in vertebrates. Nat Genet 41:753–761.

Mikkelsen JG, Yant SR, Meuse L, Huang Z, Xu H, Kay MA. 2003. Helper-Independent *Sleeping Beauty* transposon-transposase vectors for efficient nonviral gene delivery and persistent gene expression in vivo. Mol Ther 8:654–665.

Miller AM, Dean DA. 2008. Cell-specific nuclear import of plasmid DNA in smooth muscle requires tissue-specific transcription factors and DNA sequences. Gene Ther 15:1107–1115.

Molto E, Fernandez A, Montoliu L. 2009. Boundaries in vertebrate genomes: Different solutions to adequately insulate gene expression domains. Brief Funct Genomic Proteomic 8:283–296.

Muller K, Heller H, Doerfler W. 2001. Foreign DNA integration. Genome-wide perturbations of methylation and transcription in the recipient genomes. J Biol Chem 276:14271–14278.

Nakabayashi H, Taketa K, Miyano K, Yamane T, Sato J. 1982. Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. Cancer Res 42:3858–3863.

Plasterk RH, Izsvák Z, Ivics Z. 1999. Resident aliens: The Tc1/mariner superfamily of transposable elements. Trends Genet 15:326–332.

Raper SE, Chirmule N, Lee FS, Wivel NA, Bagg A, Gao GP, Wilson JM, Batshaw ML. 2003. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. Mol Genet Metab 80:148–158.

Schaffert D, Wagner E. 2008. Gene therapy progress and prospects: Synthetic polymer-based systems. Gene Ther 15:1131–1138.

Trembley JH, Ebbert JO, Kren BT, Steer CJ. 1996. Differential regulation of cyclin B1 RNA and protein expression during hepatocyte growth in vivo. Cell Growth Differ 7:903–916.

Vercammen E, Staal J, Beyaert R. 2008. Sensing of viral infection and activation of innate immunity by toll-like receptor 3. Clin Microbiol Rev 21:13–25.

Vigdal TJ, Kaufman CD, Izsvák Z, Voytas DF, Ivics Z. 2002. Common physical properties of DNA affecting target site selection of *Sleeping Beauty* and other Tc1/mariner transposable elements. J Mol Biol 323:441–452.

Walisko O, Schorn A, Rolfs F, Devaraj A, Miskey C, Izsvák Z, Ivics Z. 2008. Transcriptional activities of the *Sleeping Beauty* transposon and shielding its genetic cargo with insulators. Mol Ther 16:359–369.

Williams DA. 2008. *Sleeping Beauty* vector system moves toward human trials in the United States. Mol Ther 16:1515–1516.

Wilson MH, Kaminski JM, George AL Jr. 2005. Functional zinc finger/*Sleeping Beauty* transposase chimeras exhibit attenuated overproduction inhibition. FEBS Lett 579:6205–6209.

Wong SP, Argyros O, Coutelle C, Harbottle RP. 2009. Strategies for the episomal modification of cells. Curr Opin Mol Ther 11:433–441.

Xue X, Huang X, Nodland SE, Mates L, Ma L, Izsvák Z, Ivics Z, LeBien TW, McIvor RS, Wagner JE, Zhou X. 2009. Stable gene transfer and expression in cord blood-derived $CD34^+$ hematopoietic stem and progenitor cells by a hyperactive *Sleeping Beauty* transposon system. Blood 114:1319–1330.

Yant SR, Ehrhardt A, Mikkelsen JG, Meuse L, Pham T, Kay MA. 2002. Transposition from a gutless adeno-transposon vector stabilizes transgene expression in vivo. Nat Biotechnol 20:999–1005.

Yant SR, Park J, Huang Y, Mikkelsen JG, Kay MA. 2004. Mutational analysis of the N-terminal DNA-binding domain of *Sleeping Beauty* transposase: Critical residues for DNA binding and hyperactivity in mammalian cells. Mol Cell Biol 24:9239–9247.

Yant SR, Wu X, Huang Y, Garrison B, Burgess SM, Kay MA. 2005. High-resolution genome-wide mapping of transposon integration in mammals. Mol Cell Biol 25:2085–2094.

Young JL, Benoit JN, Dean DA. 2003. Effect of a DNA nuclear targeting sequence on gene transfer and expression of plasmids in the intact vasculature. Gene Ther 10:1465–1470.

Zayed H, Izsvák Z, Khare D, Heinemann U, Ivics Z. 2003. The DNA-bending protein HMGB1 is a cellular cofactor of *Sleeping Beauty* transposition. Nucleic Acids Res 31:2313–2322.

Zayed H, Izsvák Z, Walisko O, Ivics Z. 2004. Development of hyperactive *Sleeping Beauty* transposon vectors by mutational analysis. Mol Ther 9:292–304.

Zhu J, Kren BT, Park CW, Bilgim R, Wong PY, Steer CJ. 2007. Erythroid-specific expression of β-globin by the *Sleeping Beauty* transposon for sickle cell disease. Biochemistry 46:6844–6858.

Zhu J, Park CW, Sjeklocha L, Kren BT, Steer CJ. 2010. High-level genomic integration, epigenetic changes, and expression of *Sleeping Beauty* transgene. Biochemistry 49:1507–1521.