Published online 5 August 2013 Nucleic Acids Research, 2013, Vol. 41, No. 19 9197–9207
doi:10.1093/nar/gkt677

Transcription activator like effector (TALE)-directed *piggyBac* transposition in human cells

Jesse B. Owens[1], Damiano Mauro[1], Ilko Stoytchev[1], Mital S. Bhakta[2], Moon-Soo Kim[2], David J. Segal[2] and Stefan Moisyadi[1,3,*]

[1]Institute for Biogenesis Research, Department of Anatomy, Biochemistry and Physiology, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, HI 96822, USA, [2]Genome Center, Department of Biochemistry and Molecular Medicine, University of California, Davis, CA 95616, USA and [3]Manoa BioSciences, Honolulu, HI 96819, USA

Received May 13, 2013; Revised July 9, 2013; Accepted July 10, 2013

ABSTRACT

Insertional therapies have shown great potential for combating genetic disease and safer methods would undoubtedly broaden the variety of possible illness that can be treated. A major challenge that remains is reducing the risk of insertional mutagenesis due to random insertion by both viral and non-viral vectors. Targetable nucleases are capable of inducing double-stranded breaks to enhance homologous recombination for the introduction of transgenes at specific sequences. However, off-target DNA cleavages at unknown sites can lead to mutations that are difficult to detect. Alternatively, the *piggyBac* transposase is able perform all of the steps required for integration; therefore, cells confirmed to contain a single copy of a targeted transposon, for which its location is known, are likely to be devoid of aberrant genomic modifications. We aimed to retarget transposon insertions by comparing a series of novel hyperactive *piggyBac* constructs tethered to a custom transcription activator like effector DNA-binding domain designed to bind the first intron of the human CCR5 gene. Multiple targeting strategies were evaluated using combinations of both plasmid-DNA and transposase-protein relocalization to the target sequence. We demonstrated user-defined directed transposition to the CCR5 genomic safe harbor and isolated single-copy clones harboring targeted integrations.

INTRODUCTION

The piggyBac (PB) transposable element can efficiently integrate transgenes into genomes of mammalian cells and organisms (1–4). This non-viral vector has several advantages over integrating viral vectors such as gamma-retroviral and lentiviral vectors, including low toxicity, larger cargo size and reduced preference for insertion into actively transcribed genes (5–8). However, insertional mutagenesis and unknown position effects that may inhibit transgene expression remain obstacles for vectors that integrate randomly (9–12). A method for user-defined directed integration would improve the safety of insertional therapies.

Engineered nucleases based on transcription activator like effector (TALE), zinc finger and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems have been used to induce DSBs at specific sites (13–17). Subsequent error-prone repair can leave desired mutations at these sites, and homology-directed repair can be exploited to introduce a co-delivered donor template. Nonetheless, cytotoxicity due to the cell's emergency response to DSBs and genotoxicity resulting from off-target cleavages and mutations remain concerns for the clinical use of nuclease-based approaches (15,18–26).

Unlike engineered nucleases, transposons perform all the enzymatic steps required for integration (27). Furthermore, we have shown that a chimeric PB transposase fused to the Gal4 DNA-binding domain (DBD) can bias integration near endogenous Gal4 recognition sequences (28). We have since modified our vector architecture to more efficiently localize transpositional activity and have incorporated a swappable custom TALE designed to bind a single genomic address.

Genomic safe harbors can be defined as loci that are well-suited for gene transfer. Integrations within these sites are not associated with adverse effects such as proto-oncogene activation or tumor suppressor inactivation. Furthermore, safe harbors may allow stable transgene expression across multiple cell types. One such putative site is chemokine C-C motif receptor 5 (CCR5) (29,30), which is required for the entry of R5 tropic HIV-1 strains involved in primary infections. A homozygous Δ32

*To whom correspondence should be addressed. Tel: +1 808 956 3118; Fax: +1 808 956 7316; Email: moisyadi@hawaii.edu

deletion in the CCR5 gene confers resistance to HIV-1 infection in humans. Disrupted CCR5 expression, naturally occurring in about 1% of the Caucasian population, does not appear to result in any significant reduction in immunity (31). Consequently, clinical trials are exploring the possibility of disrupting CCR5 via targetable nucleases as part of an anti-HIV therapeutic approach (32).

Here, we introduce novel constructs using a hyperactive PB transposase coupled with a TALE DBD to target the first intron of the human CCR5 gene and have detected stable expression of a reporter gene at this safe harbor. We identified targeted insertions in ~0.010–0.014% of total stably transfected cells. Furthermore, we demonstrate a simple PCR-based method for the identification of targeted clones containing a single transposon. This 'proof-of-concept' represents the first example of targeting an integrating enzyme to a single user-defined TALE-directed endogenous location. We anticipate that insights gained from this methodology could someday improve the safety profile for cell replacement therapies.

MATERIALS AND METHODS

Plasmid development

A detailed description of the construction of PB targeting plasmids is provided in the Supplementary Materials and Methods. Simplified illustrations of all PB targeting plasmids are depicted in Figure 1. All targeting plasmids were derived from p*mhy*GENIE-3-R6K (abbreviated hG3) that encodes a self-inactivating (33,34) hyperactive PB transposase (35) driven by the CAG [cytomegalovirus (CMV) immediate early enhancer, chicken β-actin promoter and β-globin intron] promoter. hG3-TALC1 contains a TALE DBD designed to bind a sequence in the CCR5 gene (TALC1) directly linked to the PB transposase. hGT1-TALC1 contains the Gal4 DBD linked to PB as well as second protein consisting of Gal4 linked to TALC1. hGT2-TALC1 contains Gal4 linked to TALC1 that was consecutively linked to PB. hGT3-TALC1 contains TALC1 linked to Gal4 that was consecutively linked to PB. hG3R1T1-TALC1 contains a TALE designed to bind a unique sequence in the ROSA26 gene (TALR1) linked to PB as well as TALR1 linked to TALC1. Four upstream activating sequence (UAS) arrays, each containing five Gal4 recognition sequences, were added to the plasmid backbone for hGT1-TALC1, hGT2-TALC1 and hGT3-TALC1. Four TALR1 recognition sequences were added to the backbone of hGR1T1-TALC1. All plasmids feature Gateway recombineering (Invitrogen) attR sites within the transposon for easy addition of transgene cargo. The trans-gene was made by swapping the puromycin gene in pGIPZ (Thermo Scientific) with a neomycin (GIN) gene amplified from pERV3 (Agilent Technologies). The fragment including CMV, TurboGFP, internal ribosomal entry site (IRES) and GIN was cloned into a pENTR1a shuttle plasmid (Invitrogen) and subsequently recombined into all targeting constructs.

The TALE repeat regions for TALC1, C2 and R1 were synthesized by BioBasic, Inc. The 16.5-repeat arrays were cloned by StuI/AatII digestion into pPreTALE (36), which contained truncated N- and C-termini of the naturally occurring TALE PthXo1 and flanking SfiI restriction sites. Full binding sites and protein sequences of the SfiI TALE cassettes are provided in Supplementary Figure S1.

Cell transfections

Human embryonic kidney (HEK293) cells were maintained in complete Dulbecco's modified Eagle's medium supplemented with 10% heat inactivated fetal bovine serum. Before transfection, 2×10^5 cells per well were seeded in 12-well plates. Cells at 90% confluency were transfected with 800 ng of plasmid DNA using X-tremeGENE 9 (Roche Applied Science). Cells for each transfection were maintained for 2 weeks under 200 μg/ml G418 at which point ~90% of cells were pelleted and frozen. Genomic DNA was isolated from pellets using the DNeasy kit (Qiagen) following the manufacturer's protocol. Remaining cells were grown for three additional days then frozen in liquid nitrogen. A hG3-TALC1 transfection, which was confirmed to contain positive targeted cells by PCR, was thawed and a dilution of cells was plated into a 96-well poly-D-lysine coated plate (BD Biosciences) resulting in ~56 colonies per well. After wells became >40% confluent, the cells were manually resuspended by pipetting in a total volume of 30 μl. A volume of 20 μl of the resuspension was removed for analysis using the DirectPCR Lysis Reagent (Viagen Biotech), and the remaining cells were cultured further. A well identified to contain targeted clones was expanded and single-cell sorted using serial dilution. Wells were visually monitored and 242 single-cell expansions were obtained. Clonally expanded cells were subsequently resuspended by manual pipetting and lysed for analysis.

Copy number assay

To determine the number of transposons present in CCR5-targeted single clones, a quantitative PCR copy number assay was performed as previously described (28).

Flow cytometry

Green fluorescent protein (GFP) expression of 100 000 cells from CCR5-targeted single-cell expansions was analyzed using a FACSAria III cytometer (BD Biosciences) after 10 weeks of culture following transfection with hG3-TALC1.

Colony count assay

The 1×10^5 HEK293 cells were transfected with equal molar amounts (maximum 500 ng) of plasmid DNA for each PB targeting construct in addition to a transposase-negative control described previously (33). Cells were resuspended, diluted 1:100 and plated into 10 cm plates (1000 total cells per plate) and maintained for 3 weeks under G418 selection. The fraction of resulting GFP positive colonies >1 mm in diameter were counted using a FluorVivo 100 fluorescence imaging system (INDEC Biosystems).

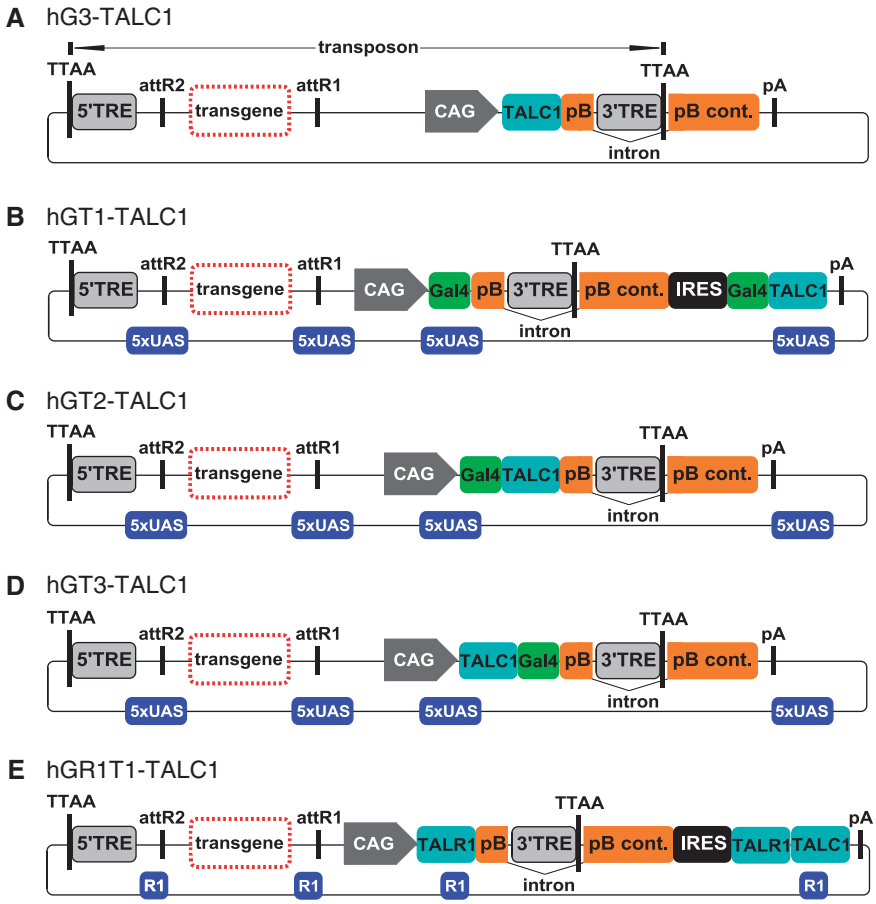

Figure 1. PB targeting plasmids. The TREs flank the PB transposon. The 3′TRE resides within an introduced intron in the PB gene leading to inactivation of the transposase on excision of the transposon. The CAG promoter drives expression of the PB targeting proteins and the IRES allows for dual expression of two proteins by the promoter. The UAS and R1 recognition sequences for the Gal4 and TALR1 DBDs were engineered into the plasmid backbones. The transgene can be Gateway recombined between the attR sites. (**A**) hG3-TALC1 (**B**) hGT1-TALC1 (**C**) hGT2-TALC2 (**D**) hGT3-TALC1 (**E**) hGR1T1-TALC1.

TALE binding assay

TALE artificial transcription factors were cloned using XhoI and AgeI, into the PGK promoter-driven mammalian expression vector pPGK-VP64, which appended an N-terminal HA epitope tag and nuclear localization sequence and a C-terminal VP64 transcriptional activation domain (36). Target sites for the TALEs were cloned between NotI and XhoI sites upstream of the SV40 promoter in pGL3-control plasmids (Promega).

In 24-well plates, HEK293T cells at 80% confluency in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, were co-transfected with 100 ng of TALE expression plasmid, 25 ng of modified pGL3-control firefly luciferase reporter plasmid containing a TALE target site and 25 ng of pRL-TK-Renilla Luciferase plasmid (as a transfection control, Promega), using Lipofectamine 2000 (Invitrogen). Cells were harvested 48 h post-transfection by removing media, washing with 500 µl of 1× phosphate buffered saline (PBS), followed by lysis in 100 µl of 1× Passive Lysis Buffer (Promega) with 1× Complete protease inhibitors (Roche). Clarified cell lysates (20 µl) were used to determine luciferase activity using DualGlo reagents (40 µl, Promega) in a Veritas microplate luminometer (Turner Biosystems). All experiments were performed in duplicate and repeated on two different days.

Targeted genomic integration site recovery

Genomic DNA or DirectPCR lysates (Viagen Biotech) from stably transfected HEK293 cells were used as template for nested PCR to identify targeted transposon insertions. Forward primers were designed to extend outward from the transposon, whereas reverse primers were designed to extend from the region adjacent to the TALC1 recognition sequence (tTTTAGCCTTACTGTT GA) found uniquely in the first intron of the human CCR5 gene. Primary PCR products obtained using the KOD Xtreme Hot Start DNA Polymerase (Novagen) were diluted 1:50 in H_20 and used as template for nested PCR. Amplification products were gel purified with the Zymoclean Gel DNA Recovery Kit (Zymo Research) and cloned into pJet1.2 (Thermo Scientific) for sequencing. Sequences were aligned to the PB transposon and human genome using BLAST to identify insertion site locations. Primer sequences are listed in Supplementary Table S1.

RESULTS

Experimental strategies for targeting PB transposition

Five unique targeting constructs were designed to localize transposition to the CCR5 locus. hG3-TALC1 (Figure 1A) includes a hyperactive PB transposase with a protein linker bound directly to a custom TALE DBD (TALC1) designed to bind a unique 17 bp sequence in the first intron of the human CCR5 gene (Figure 2A). This single plasmid includes both the transposon and transposase and is self-inactivating (33), meaning the transposase gene is rendered inactive after excision of the transposon from the plasmid. As a consequence, potentially negative effects that may develop by the persistence of an active PB gene are eliminated. hG3-TALC1 is analogous to a codon-optimized PB plasmid used to target insertions near endogenous Gal4 recognition sequences in the human genome (28).

hGT1-TALC1 (Figure 1B) was designed to include an additional 'tethering' protein on the same bicistronic coding region as the PB transposase. This plasmid includes both a Gal4 DBD linked to the CCR5-directed TALC1, and a Gal4 DBD linked to a hyperactive PB transposase. Because the backbone of the plasmid contains twenty UAS sites, the Gal4/TALC1 tethering protein is expected to bind both the TALE recognition sequence and the plasmid backbone simultaneously, thereby localizing the plasmid near CCR5 in the genome. The Gal4/PB is expected to localize the transposase to the plasmid backbone via the UAS sequences (Figure 2B). We conceived that the additional flexibility, which may be achieved for this orientation, as compared with the direct protein fusion used with hG3-TALC1, would allow for more efficient enzyme activity after CCR5 localization.

In an effort to simplify the two-protein strategy used for hGT1-TALC1, two additional constructs were designed to incorporate both the tethering protein and PB transposase on a single chain. hGT2-TALC1 (Figure 1C) features Gal4 linked to the TALC1 DBD that is subsequently linked to PB. As with hGT1-TALC1, this three-part molecule could potentially locate the plasmid to CCR5 via the two linked DBDs, Gal4 and TALC1. Additionally, because PB is also linked to Gal4, there is a potential for the transposase to be relocated to within close proximity of the plasmid backbone via the binding of Gal4 to the UAS sequences (Figure 2C). In a similar fashion, hGT3-TALC1 (Figure 1D) features the TALC1 DBD at the N-terminal of a three-part molecule linked to Gal4 that is subsequently linked to PB. The strategy for targeting for this construct is similar to that of hGT2-TALC1 except for the locations of two the DBDs are reversed (Figure 2D).

A modified version of hGT1-TALC1, called hGR1T1-TALC1 (Figure 1E), incorporates a TALE DBD in place of Gal4. Because Gal4 has a short 6 bp recognition sequence, we reasoned that our tethering constructs that use Gal4 to bind to the plasmid backbone may also be targeted to the numerous off-target Gal4 recognition sequences located in the genome. To prevent this form of unintended retargeting, we replaced the Gal4 with a TALE DBD made to bind a specific 17 bp sequence in the human ROSA26 gene (TALR1) found only once in the genome. We also replaced the UAS sites on the backbone with four TALR1 recognition sites so that the tethering molecule, consisting of TALR1 linked to TALC1, could bind both the plasmid backbone specifically and the CCR5 locus simultaneously. The purpose of TALR1 was not to target transposition to ROSA26 but to increase the specificity of binding of the TALR1/TALC1 double-DBD protein. Similar to hGT1-TALC1, the transposase was linked to TALR1 so that PB could be relocated to the plasmid backbone (Figure 2E).

A two-plasmid strategy was devised using hG3-TALC1 combined with a similar plasmid to hGT1-TALC1 described earlier in the text, called hGT1-TALC2, in which the CCR5 DBD was replaced by an alternative TALE (TALC2) designed to bind 85 bp upstream of TALC1. By using different DBDs, we reasoned that the two strategies could complement one another by allowing both PB protein (using hG3-TALC1) and plasmid DNA (using hGT1-TALC2) to locate to neighboring locations (Figure 2F). Finally, a control plasmid (hG3) was constructed containing an unfused hyperactive PB. All constructs included a bicistronic CMV promoter driven TurboGFP and GIN reporter/selection cassette within the transposon. Successful targeting for all strategies was expected to result in the excision of the transposon from the plasmid by the transposase followed by permanent introduction of the reporter/selection transgenes near the TALE recognition sequence (Figure 2G).

Activities of PB transposase and TALE DNA-binding proteins

Integration activities of each PB targeting construct were compared using a transpositional colony count assay. Non-integrated plasmid DNA is typically lost due to cell division after ~2 weeks. HEK293 cells were transfected with each plasmid and grown under G418 selection. Three weeks later, GFP positive colonies were counted (Figure 3A). Comparable activities were observed between transfections for most of the fusion constructs; however, hGT3-TALC1 had relatively diminished activity. The unfused PB expressed from hG3 was approximately twice as active as transposase linked to Gal4 or TALC1 expressed from fusion constructs. An average of 111 colonies were counted for the five fusion constructs on plates each originally seeded with 1000 cells, thus ~11% of total transfected cells received PB integrations. This integration activity is in agreement with previous hyperactive PB rates in HEK293 cells (37) and represents a 26-fold increase in activity over random integration by a transposase-negative control.

Binding activity of the three TALEs used in this study was verified using a transcription factor reporter assay (Figure 3B). We constructed expression plasmids appending a VP64 transcriptional activation domain to TALC1, TALC2 or TALR1. Reporter (Rep) plasmids were each designed to contain a single TALE-binding site located upstream of a minimal promoter driving luciferase (Rep C1, Rep C2, Rep R1). Control cotransfections of

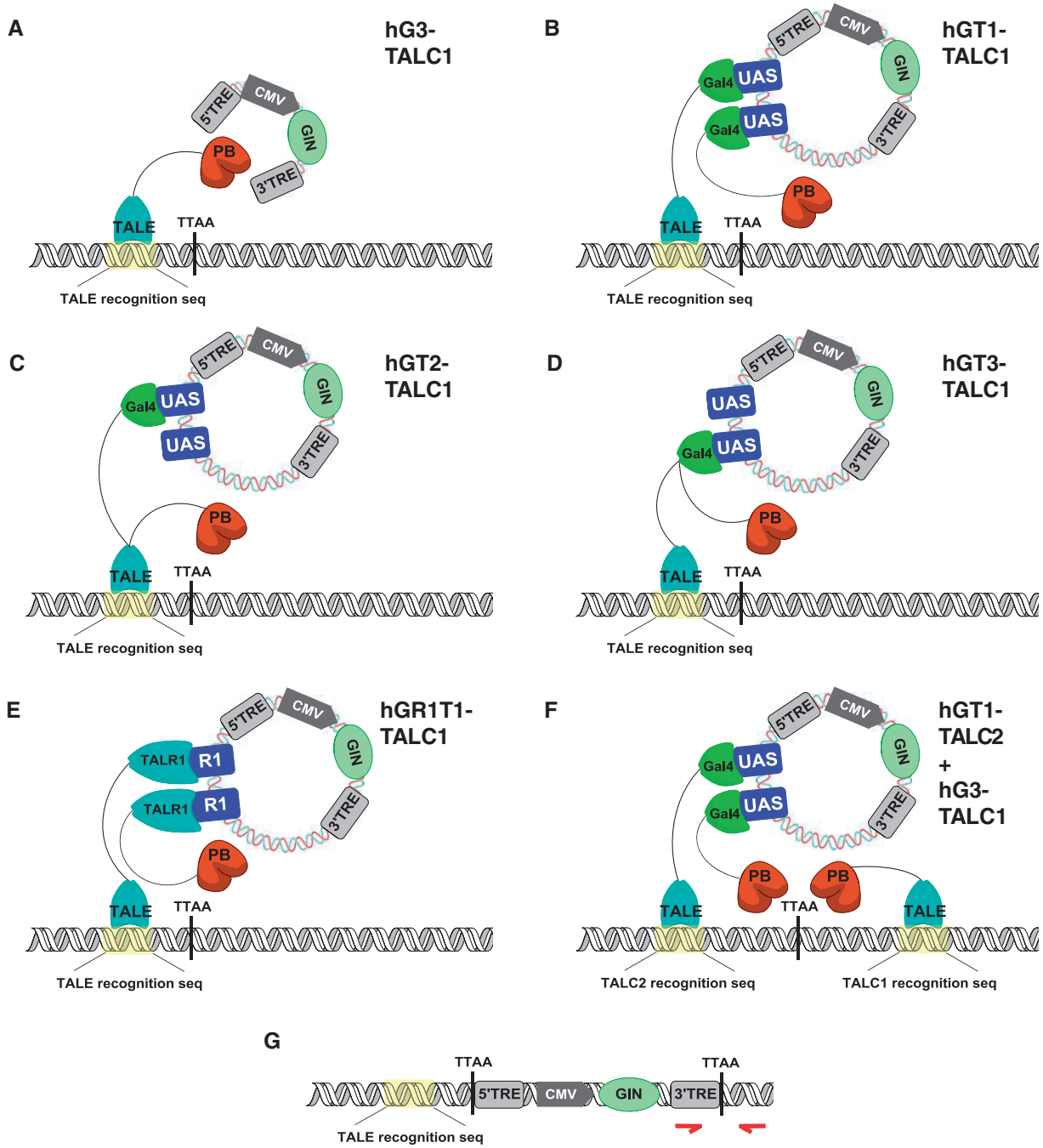

Figure 2. Schematic for various PB targeting strategies. For each strategy, all components are encoded on a single plasmid, and most components have been omitted for simplification purposes from plasmids depicted in this figure. (**A**) hG3-TALC1 encodes a hyperactive PB transposase covalently linked to a TALE designed to bind a specific sequence in the CCR5 gene (TALC1). (**B**) hGT1-TALC1 encodes a double-DBD protein including TALC1 linked to Gal4. Tethering of the plasmid to CCR5 is mediated by Gal4 binding to UAS sites found on the plasmid backbone and TALC1 binding to the genomic recognition sequence. Additionally, hGT1-TALC1 encodes a Gal4-PB fusion to draw PB to the plasmid. (**C**) hGT2-TALC1 encodes a three-part protein consisting of Gal4 linked to TALC1 followed by the PB transposase. Tethering of the plasmid to the CCR5 genomic sequence is made possible by the TALE and Gal4 segment of the protein through binding of Gal4 to UAS sites found on the plasmid backbone. In addition, PB can be relocated to CCR5 via direct linkage to TALC1. (**D**) hGT3-TALC1 is similar to hGT2-TALC1 except for the TALE and Gal4 DBDs are reversed. Similar to hGT2-TALC1, the TALE and Gal4 segment of the three-part protein mediates the relocation of the plasmid to CCR5. PB is directly linked to the duel DBDs and can therefore also be relocated to the site of interest. (**E**) hGR1T1-TALC1 encodes a double-DBD including TALC1 linked to a second TALE (TALR1) made to bind specific recognition sites introduced into the plasmid backbone. The double-DBD can therefore simultaneously bind the plasmid and CCR5. hGR1T1-TALC1 also encodes PB linked to TALR1 for the relocation of the transposase to the plasmid backbone and consequently to CCR5. (**F**) hGT1-TALC1 was modified by replacing TALC1 with a TALE made to bind upstream of TALC1 in the CCR5 gene (TALC2) to make hGT1-TALC2. By combining hG3-TALC1 and hGT1-TALC2 plasmids in a single transfection, both plasmid DNA retargeting and transposase retargeting strategies were used simultaneously to enhance transposition near CCR5. (**G**) The TALE-localized PB is expected to excise the transposon containing the reporter transgene GFP IRES neomycin (GIN) from the targeting plasmid and integrate nearby. Red arrows indicate PCR primers used to assay for targeted insertion. The depicted genomic primer *CCR5 Rev* is located 761 bp from the TALC1 recognition site.

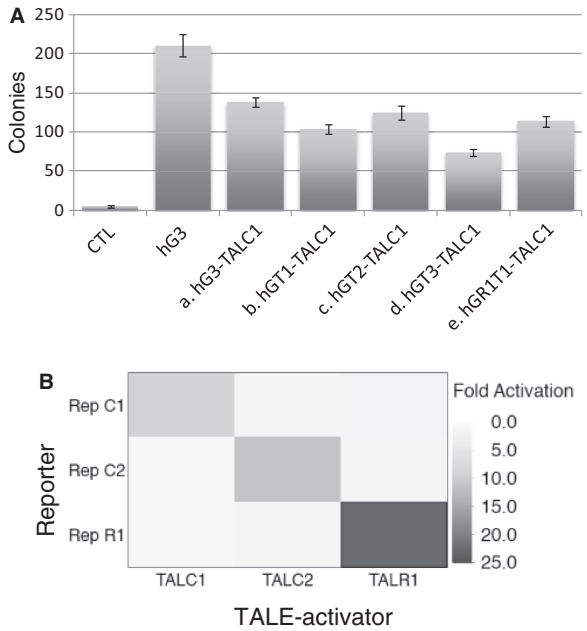

Figure 3. Verification of transposase and TALE activity. (**A**) Comparison of integration efficiencies between PB constructs transfected into HEK293 cells. One thousand cells were plated and cultured for 3 weeks before G418 resistant/GFP+ colonies were counted. Data are shown as mean values with SD ($n = 3$). (**B**) Binding activity of TALE proteins was determined using a transcription factor reporter activation assay in HEK293 cells. TALC1, TALC2 and TALR1 activators were each assayed on luciferase reporter plasmids Rep C1, Rep C2 and Rep R1, which carried a single target site for each TALE activator ($n = 4$).

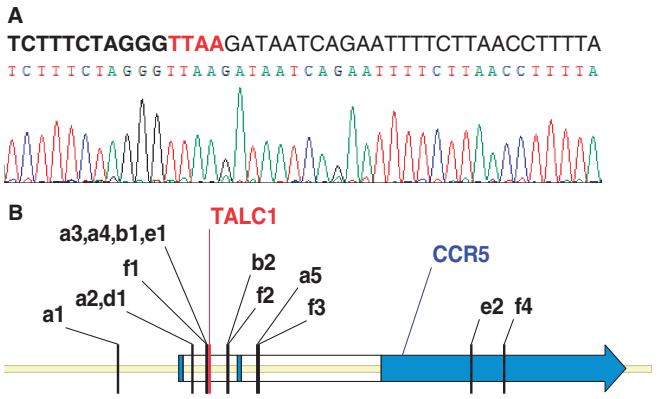

Figure 4. (**A**) Chromatogram and sequence of PCR product recovered from a representative hG3-TALC1 transfection showing the PB TRE on the left in bold, TTAA junction and flanking genomic CCR5 sequence on the right. (**B**) Locations of insertion sites recovered in the CCR5 gene. a, hG3-TALC1; b, hGT1-TALC1; c, hGT2-TALC1 (no insertions); d, hGT3-TALC1; e, hGR1T1-TALC1; f, hG3-TALC1 + hGT1-TALC2; g, hG3 (no insertions).

expression and reporter plasmids with unmatched TALE activator and target sequences resulted in background levels of luciferase activity. Cotransfection with the cognate TALE-activator and Rep C1, Rep C2, Rep R1 plasmids led to a 9, 12 and 23-fold induction of luciferase, respectively, confirming that the custom TALEs were binding and specific for their target sequence.

PB constructs mediate TALE-directed transposition to the CCR5 locus

We tested the hypothesis that our TALE-tethered PB constructs could guide transposition to regions adjacent to the TALC1 recognition site found in the CCR5 safe harbor locus. Six independent transfections for each PB targeting construct were performed and cells were subsequently selected with G418 antibiotic for 2 weeks. To estimate whether a given polyclonal population was likely to contain a high percentage of targeted clones, an initial PCR screen was performed using direct primers designed to extend from the transposon and complementary primers made to extend from the genomic CCR5 sequence. Amplification products arising from both primers included the flanking terminal repeat element (TRE) of the transposon followed by the PB canonical TTAA junction and genomic sequence of CCR5 (Figures 2G and 4A). A total of 14 unique insertion sites within CCR5 were recovered. Four of the six transfections with hG3-TALC1 gave rise to targeted insertions including one transfection resulting in two independent insertions. Two transfections each for hGT1-TALC1 and hGR1T1-TALC1 and a single transfection from hG3T3-TALC1 resulted in positive insertions. The transfections with both hG3-TALC1 and hGT1-TALC2 in combination gave rise to four insertions. No insertions were recovered from hGT2-TALC1 or hG3 control transfections (Table 1).

Two of the observed insertion sites were recovered from multiple transfections. One site, located 24 bp upstream of the TALC1 recognition sequence, was recovered from two independent hG3-TALC1 transfections as well as from hGT1-TALC1 and hGR1T1-TALC1 transfections. Additionally, one site, located 221 bp upstream of the TALC1 sequence, was targeted by both hG3-TALC1 and hGT3-TALC1. In all, 9 of the 14 insertion sites were located within 250 bp of the TALC1 sequence, and two insertions were located 639 and 659 bp away. In addition, three insertions at distances of 1231, 3495 and 3991 bp were recovered far from the target sequence. (Figure 4B and Table 1). This represents the first evidence that an integrating enzyme can be made to target a transgene to a genomic location using a user-defined TALE.

Isolation of CCR5 targeted clones

Successful cell replacement therapy using this approach will require that rare targeted clones be identified from the original polyclonal transfection for subsequent use. hG3-TALC1 gave rise to the highest number of insertions as analyzed by our initial screen (Table1); therefore, we chose a single polyclonal population that produced one of these insertions to attempt to identify and clonally expand safely modified cells.

Cells originating from a hG3-TALC1 transfection were plated into a single 96-well plate. One week later, each well was found to contain an average of 56 colonies (Figure 5A). Each well was resuspended, and a fraction

Table 1. Transposon insertions recovered in CCR5

Construct	Insertion site #	Transfection	Distance to TALC1 (bp)	Transposon orientation	Flanking CCR5 sequence
hG3-TALC1	a1	#2	1231	For	TTAATCAATGCCTT
	a2	#3	221	For	TTAAAACTCTTTAG
	a3	#1	24	For	TTAAGATAATCAGA
	a4	#4	24	For	TTAAGATAATCAGA
	a5	#2	639	For	TTAAAGGGAGCAA
hGT1-TALC1	b1	#5	24	Rev	TTAAGATAATCAGA
	b2	#1	236	Rev	TTAAGCTCAACTTA
hGT2-TALC1	c0				
hGT3-TALC1	d1	#3	221	Rev	TTAAAACTCTTTAG
hGR1T1-TALC1	e1	#5	24	For	TTAAGATAATCAGA
	e2	#1	3495	For	TTAAAAGGAAGTTA
hG3-TALC1 + hGT1-TALC2	f1	#4	37	For	TTAATAGCAACTCT
	f2	#6	247	Rev	TTAAAAGGAAGAAC
	f3	#5	659	For	TTAATAACTAACAA
	f4	#3	3991	For	TTAAATGAGAAGGA
hG3	g0				

of the cells were removed and lysed for direct PCR analysis. Using an identical PCR as the initial screen, a positive well was identified to have the same insertion as previously obtained that was located 24 bp upstream from the TALC1 recognition sequence. In a final step, this positive well was single-cell sorted and direct lysis templates from 242 single-cell expansions were screened by PCR. A total of five wells (2%) were verified to have targeted insertions. This frequency of 1/48 positive wells parallels the expected frequency of 1/56 positive colonies from which the single-cell expansions arose. Lysates from 20 clonal expansions isolated from a control hG3 transfection did not give rise to PCR products (Figure 5B). Two positive clones (293-1 and 293-2) were expanded for further analysis. A quantitative PCR copy number assay revealed that both clones contained a single transposon insertion (Figure 5C). Position effects caused by neighboring CCR5 genomic sequences could lead to silencing of the transgene. As analyzed by flow cytometry, robust GFP expression from targeted clones was detected beyond 10 weeks of culture (Figure 5D). Populations expanded from clones 293-1 and 293-2 were found to be 99.9 and 98.0% GFP positive, respectively.

Targeting efficiencies of hG3-TALC1 and hGT1-TALC1

An initial PCR screen was used to estimate the relative targeting efficiencies of the five TALC1-directed PB constructs (Table 1). The two most promising constructs, hG3-TALC1 and hGT1-TALC1, resulted in more than one insertion within 250 bp of the TALC1 genomic recognition site. As described earlier in the text, hG3-TALC1 transfection #1 was plated into wells on a 96-well plate and a single well containing a targeted colony was identified. Each well contained an average of 56 colonies; therefore, we identified about one in 5376 correctly modified cells. This represents 0.019% of total stably transfected cells. To gain a better understanding of the number of targeted cells present in our polyclonal populations, cells originating from a single hGT1-TALC1 transfection and two hG3-TALC1 transfections were seeded into additional wells. The hGT1-TALC1 transfection #1 was plated into 960 wells, and an average of eleven colonies per well was counted. hG3-TALC1 transfections #1 and #2 were each plated into 480 wells, and averages of 16 colonies per well were counted. A single positive well was identified by PCR for all three transfections. By including the data from the first plating of hG3-TALC1 transfection #1, we determined that the percentage of targeted cells found for hG3-TALC1 transfections #1 and #2 was 0.015 and 0.013, respectively, or 0.014 combined. The hGT1-TALC1 transfection resulted in 0.010% of targeted cells (Table 2).

DISCUSSION

The ability of viruses to efficiently introduce therapeutic transgenes permanently into cellular chromosomes has led to reliable treatments for a diverse set of genetic diseases (38). A system that could safely direct insertions to genomic safe harbors would overcome the strong preference for the disruption of active genes that burdens viral-based approaches (5,7,8), thereby transforming the gene therapy field.

In an effort to reduce the risks of random viral insertion, Papapetrou *et al.* (39) has defined criteria for *de novo* safe harbor sequences in the genome based on their position relative to contiguous coding genes, microRNAs and ultraconserved regions. This strategy involves clonally expanding cells containing random integrations followed by identifying all genomic insertion sites. Only clones containing a single insertion that is located within these 'safe' regions are selected. This strategy does not require screening for insertions at specific sites, which may reduce the necessary number of clones. Drawbacks include the requirement for the identification of random insertion sites in the genome for each clone as well as ambiguity about the selected safe harbor. These *de novo* safe harbor sequences may perform unidentified cellular functions and local chromosomal

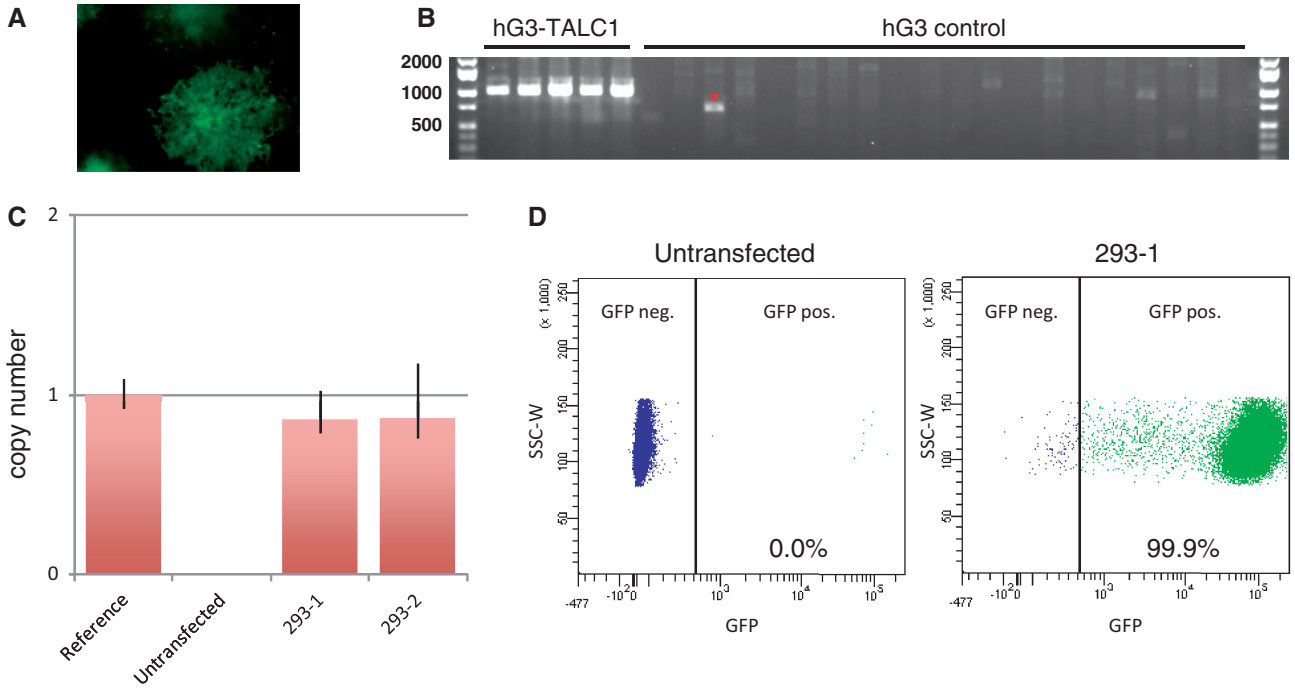

Figure 5. (**A**) Cells from a hG3-TALC1 transfection were plated into a 96-well plate, and 1 week later, individual non-overlapping colonies were established for counting; 40× magnification. (**B**) Nested PCR for the identification of transposition near CCR5. Expected products arose from five positive clones identified from hG3-TALC1 transfection, but not clones from hG3 control transfection. The asterisk denotes a sequenced non-specific PCR product. (**C**) Transposon copy number for clones 293-1 and 293-2. Quantitative PCR predictions were calibrated using a reference HEK293 cell line known to contain a single copy transposon. (**D**) Cells positive for GFP reporter gene targeting to CCR5 displayed sustained expression past 10 weeks of culture. Flow cytometry analysis displaying GFP positive events for both untransfected HEK293 cells and an expansion of clone 293–1.

Table 2. Targeted cells recovered from hG3-TALC1 and hGT1-TALC1

Construct	hG3-TALC1 #1	hG3-TALC1 #2	hG3-TALC1 combined	hGT1-TALC1 #1
Total cells screened	13 686	7622	21 308	10 476
Positive wells	2	1	3	1
% Targeted cells	0.015	0.013	0.014	0.010

position effects at these sites are unknown. These issues may necessitate individual characterization of each clone and are likely to be resolved by targeting a specific well-characterized safe harbor.

In an attempt to redirect viral insertions to a known sequence, an HIV integrase fused to a zinc finger, designed to bind the *erbB-2* gene, has been shown to increase targeted integration into the genome by 10-fold compared with wild-type HIV integrase (40). A non-viral alternative approach has been to use zinc-finger recombinases (ZFRs) consisting of a custom designed zinc-finger DBD and recombinase catalytic domain. Insertion site preference can be altered by zinc-finger binding but is restricted by sequence requirements dictated by the native recombinase. Using directed evolution, unique catalytic domains have been produced that are able to tolerate additional core sequences, theoretically allowing ZFRs to target up to 3.77×10^7 unique genomic sites (41). ZFRs display high targeting efficiencies (8.3–14.2%) of stably transfected cells but are limited by target site inflexibility and low total integration efficiencies (0.14–0.31%).

Targetable transposition, using chimeric proteins consisting of a DBD fused to a transposase, can be used to preferentially insert transgenes near a specific sequence. A variety of DBDs have been used to bias transposon integration on recipient plasmids in various cell types (28,42–49). Recently, endogenous transpositional targeting has been achieved (28,42). The Rep DBD, known to target the wild-type adeno-associated virus (AAV) to a region on human chromosome 19 called AAVS1, was used to bias integrations of PB, *Sleeping Beauty* and *Tol2* transposons near both minimal Rep binding sequences (15 726 sites per human genome) and consensus Rep binding sequences (2134 sites per human genome) (42). Previously, we demonstrated that a Gal4-PB transposase fusion was able to bias 24% of integrations near endogenous Gal4 recognition sequences; however, these targets were found in numerous genomic locations and, like the Rep DBD, its recognition sequence was pre-defined (28). Moreover, in our preliminary study, single targeted clones were not isolated. Here, we have evaluated the ability of a variety of vector architectures to localize transposition near a

user-defined TALE recognition sequence found in the CCR5 gene. TALEs are simple to generate and can be designed to specifically bind almost any sequence (50). By fusing a TALE to the PB transposase using a direct protein linker (hG3-TALC1 and hGT3-TALC1) or by tethering the TALE to the plasmid backbone (hGT1-TALC1, hGT3-TALC1 and hGR1T1-TALC1), or by combining both strategies (hG3-TALC1 + hGT1-TALC2), we achieved user-defined directed integration into the genome. We targeted an endogenous genomic safe harbor and recovered multiple insertion sites within this region ranging from 24 to 3991 bp near the TALE recognition sequence. Two 'hot-spots', 24 and 221 bp away, were targeted multiple times and most insertions (9/14) clustered within 250 bp of the TALE sequence. Rare targeted clones positive for a single CCR5-targeted insertion were isolated and stable GFP reporter expression was confirmed for these cells.

The aim of these experiments was to demonstrate the ability of our novel PB constructs to target a user-defined genomic address. Although we did successfully obtain targeted integrants, these primary experiments necessitate a number of improvements to the system. After transfection with our TALE-directed PB construct, we performed a simple pre-plating step into a single 96-well plate followed by PCR analysis. This allowed for the isolation of a small pool (56 colonies) of potentially targeted cells before single-cell sorting. Although we were successful in identifying positive single-cell expansions in 1/48 (2%) of wells, it would be desirable to omit the pre-screening step used in these experiments. Although we were able to identify targeted insertion sites for all constructs except hGT2-TALC1, many transfection replicates did not give rise to detectable insertions (Table 1). It is likely that these polyclonal populations contained additional targeted insertions; however, the percentage of targeted cells was low and therefore was not detectable by our PCR screen. Moreover, the targeting efficiencies of total stably transfected cells observed for our constructs were 0.010–0.014%, which are significantly lower than nuclease-based approaches used to target CCR5 (51). These efficiencies might be improved by performing additional experiments aimed at optimizing transposase expression level by assaying a range of transfection concentrations. The hyperactive pB transposase is exceptionally efficient at integration (35,37) and does not rely on rate-limiting host-factors, as do alternative retargeting strategies. This system, currently in early stages, is ideally suited for improvements to efficiency.

The PB transposase is autonomously functional in our system and therefore is able to integrate into many genomic locations. A major improvement to the system would be to make the localization or binding of PB to the genomic target a 'required' event for transposition. This might be achieved by mutating the native PB DBD domain such that the transposase would be inhibited from binding off-target sequences and consequently rely on a user-supplied DBD, such as a custom TALE, for transposition. Furthermore, modifications to the dimerization domain could prevent PB from dimerizing in solution. On colocalization of both dimers at the genomic target sequence via attached TALEs, catalytic activity could theoretically be restored. Modifications such as these would be anticipated to not only eliminate off-target integrations for targeted clones but also increase the total number of targeted cells due to the limited number of transposons being prevented from getting 'soaked up' by the rest of the genome. Recently, Li *et al.* (52) described excision competent/integration defective transposases with mutations in PB's catalytic core. Interestingly, the integration activity of these mutants can be rescued by fusing a custom zinc-finger DBD to the transposase. However, integrations were not associated with the recognition sites of the custom zinc-fingers, as genomic targeting using these fusion proteins was unsuccessful. Nevertheless, these PB mutants could potentially serve as a framework for future studies into site-required transposition.

Targetable nucleases have been used to insert transgenes into endogenous genes and safe harbor loci in embryonic and induced pluripotent stem cells (53–58), and easy-to-implement modifications to both zebrafish and rat genomes have become a possibility (59,60). One of the benefits of using transposase-based genomic targeting over nuclease-based techniques is that integration via the class II transposon cut-and-paste mechanism is readily identified by assaying the copy number of transposon insertions. Therefore, a single-insertion clone is not expected to have additional DNA modifications (35). In comparison, because targetable nucleases are capable of mutating the genome without introducing an identifiable insert, it remains difficult to confirm the DNA integrity of modified cells. Genomic screens used to attempt to identify off-target nuclease mutations are complex and limited in coverage (23,24,54,57).

The PB system can permanently introduce large cassettes (>100 kb) encoding numerous components such as multiple transgenes, insulators and inducible or endogenous promoters (61). The current study has laid the groundwork for enhancing this system by allowing researchers to potentially target integrations to nearly any genomic region. This system is especially applicable for cell-replacement therapies where safe single-targeted insertions could be verified *ex vivo*, and cells could subsequently be amplified and re-infused into patients. We envision targeted transposition could be used to intentionally disrupt endogenous coding regions or to direct insertions to user-defined genomic safe harbors to protect the cargo from unknown chromosomal position effects and to circumvent accidental mutation of target cells.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr Johann Urschitz for expert advice in the design and trouble-shooting of the experiments. They thank Dr Zoia Stoytcheva for her invaluable advice in subcloning and molecular biology techniques and Dr Philip Davy for assisting with the flow cytometry

analysis. They also thank Dr Joseph Kaminski for critical reading of the manuscript and the Sanger Institute, Cambridge, England for the hyperactive *piggyBac* plasmid.

FUNDING

Funding for open access charge: National Institutes of Health Grants [5P20RR024206, R01GM083158-01A1].

Conflict of interest statement. S.M. is the founder of Manoa BioSciences.

REFERENCES

1. Di Matteo,M., Matrai,J., Belay,E., Firdissa,T., Vandendriessche,T. and Chuah,M.K. (2012) PiggyBac toolbox. *Methods Mol. Biol.*, **859**, 241–254.
2. Ding,S., Wu,X., Li,G., Han,M., Zhuang,Y. and Xu,T. (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*, **122**, 473–483.
3. Feschotte,C. (2006) The piggyBac transposon holds promise for human gene therapy. *Proc. Natl Acad. Sci. USA*, **103**, 14981–14982.
4. Wu,S.C., Meir,Y.J., Coates,C.J., Handler,A.M., Pelczar,P., Moisyadi,S. and Kaminski,J.M. (2006) PiggyBac is a flexible and highly active transposon as compared to sleeping beauty, Tol2, and Mos1 in mammalian cells. *Proc. Natl Acad. Sci. USA*, **103**, 15008–15013.
5. Daniel,R. and Smith,J.A. (2008) Integration site selection by retroviral vectors: molecular mechanism and clinical consequences. *Hum. Gene Ther.*, **19**, 557–568.
6. Di Matteo,M., Belay,E., Chuah,M.K. and Vandendriessche,T. (2012) Recent developments in transposon-mediated gene therapy. *Expert Opin. Biol. Ther.*, **12**, 841–858.
7. Mitchell,R.S., Beitzel,B.F., Schroder,A.R., Shinn,P., Chen,H., Berry,C.C., Ecker,J.R. and Bushman,F.D. (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.*, **2**, E234.
8. Wu,X., Li,Y., Crise,B. and Burgess,S.M. (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science*, **300**, 1749–1751.
9. Baum,C. (2007) Insertional mutagenesis in gene therapy and stem cell biology. *Curr. Opin. Hematol.*, **14**, 337–342.
10. Fan,H. and Johnson,C. (2011) Insertional oncogenesis by non-acute retroviruses: implications for gene therapy. *Viruses*, **3**, 398–422.
11. Fehse,B. and Roeder,I. (2008) Insertional mutagenesis and clonal dominance: biological and statistical considerations. *Gene Ther.*, **15**, 143–153.
12. Romano,G., Marino,I.R., Pentimalli,F., Adamo,V. and Giordano,A. (2009) Insertional mutagenesis and development of malignancies induced by integrating gene delivery systems: implications for the design of safer gene-based interventions in patients. *Drug News Perspect.*, **22**, 185–196.
13. Li,T., Huang,S., Jiang,W.Z., Wright,D., Spalding,M.H., Weeks,D.P. and Yang,B. (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.*, **39**, 359–372.
14. Perez,E.E., Wang,J., Miller,J.C., Jouvenot,Y., Kim,K.A., Liu,O., Wang,N., Lee,G., Bartsevich,V.V., Lee,Y.L. *et al.* (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.*, **26**, 808–816.
15. Porteus,M.H. and Baltimore,D. (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science*, **300**, 763.
16. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
17. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
18. Olsen,P.A., Gelazauskaite,M., Randol,M. and Krauss,S. (2010) Analysis of illegitimate genomic integration mediated by zinc-finger nucleases: implications for specificity of targeted gene correction. *BMC Mol. Biol.*, **11**, 35.
19. Pruett-Miller,S.M., Connelly,J.P., Maeder,M.L., Joung,J.K. and Porteus,M.H. (2008) Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. *Mol. Ther.*, **16**, 707–717.
20. Cornu,T.I., Thibodeau-Beganny,S., Guhl,E., Alwin,S., Eichtinger,M., Joung,J.K. and Cathomen,T. (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol. Ther.*, **16**, 352–358.
21. Bohne,J. and Cathomen,T. (2008) Genotoxicity in gene therapy: an account of vector integration and designer nucleases. *Curr. Opin. Mol. Ther.*, **10**, 214–223.
22. Cornu,T.I. and Cathomen,T. (2010) Quantification of zinc finger nuclease-associated toxicity. *Methods Mol. Biol.*, **649**, 237–245.
23. Pattanayak,V., Ramirez,C.L., Joung,J.K. and Liu,D.R. (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat. Methods*, **8**, 765–770.
24. Gabriel,R., Lombardo,A., Arens,A., Miller,J.C., Genovese,P., Kaeppel,C., Nowrouzi,A., Bartholomae,C.C., Wang,J., Friedman,G. *et al.* (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.*, **29**, 816–823.
25. Gupta,A., Meng,X., Zhu,L.J., Lawson,N.D. and Wolfe,S.A. (2011) Zinc finger protein-dependent and -independent contributions to the *in vivo* off-target activity of zinc finger nucleases. *Nucleic Acids Res.*, **39**, 381–392.
26. Radecke,S., Radecke,F., Cathomen,T. and Schwarz,K. (2010) Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. *Mol. Ther.*, **18**, 743–753.
27. Mitra,R., Fain-Thornton,J. and Craig,N.L. (2008) piggyBac can bypass DNA synthesis during cut and paste transposition. *EMBO J.*, **27**, 1097–1109.
28. Owens,J.B., Urschitz,J., Stoytchev,I., Dang,N.C., Stoytcheva,Z., Belcaid,M., Maragathavally,K.J., Coates,C.J., Segal,D.J. and Moisyadi,S. (2012) Chimeric piggyBac transposases for genomic targeting in human cells. *Nucleic Acids Res.*, **40**, 6978–6991.
29. Lombardo,A., Cesana,D., Genovese,P., Di Stefano,B., Provasi,E., Colombo,D.F., Neri,M., Magnani,Z., Cantore,A., Lo Riso,P. *et al.* (2011) Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods*, **8**, 861–869.
30. van Rensburg,R., Beyer,I., Yao,X.Y., Wang,H., Denisenko,O., Li,Z.Y., Russell,D.W., Miller,D.G., Gregory,P., Holmes,M. *et al.* (2013) Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. *Gene Ther.*, **20**, 201–214.
31. Lobritz,M.A., Ratcliff,A.N. and Arts,E.J. (2010) HIV-1 entry, inhibitors, and resistance. *Viruses*, **2**, 1069–1105.
32. Maier,D.A., Brennan,A.L., Jiang,S., Binder-Scholl,G.K., Lee,G., Plesa,G., Zheng,Z., Cotte,J., Carpenito,C., Wood,T. *et al.* (2013) Efficient clinical scale gene modification via zinc finger nuclease-targeted disruption of the HIV co-receptor CCR5. *Hum. Gene Ther.*, **24**, 245–258.
33. Urschitz,J., Kawasumi,M., Owens,J., Morozumi,K., Yamashiro,H., Stoytchev,I., Marh,J., Dee,J.A., Kawamoto,K., Coates,C.J. *et al.* (2010) Helper-independent piggyBac plasmids for gene delivery approaches: strategies for avoiding potential genotoxic effects. *Proc. Natl Acad. Sci. USA*, **107**, 8117–8122.
34. Marh,J., Stoytcheva,Z., Urschitz,J., Sugawara,A., Yamashiro,H., Owens,J.B., Stoytchev,I., Pelczar,P., Yanagimachi,R. and Moisyadi,S. (2012) Hyperactive self-inactivating piggyBac for transposase-enhanced pronuclear microinjection transgenesis. *Proc. Natl Acad. Sci. USA*, **109**, 19184–19189.
35. Yusa,K., Zhou,L., Li,M.A., Bradley,A. and Craig,N.L. (2011) A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl Acad. Sci. USA*, **108**, 1531–1536.
36. Meckler,J.F., Bhakta,M.S., Kim,M.S., Ovadia,R., Habrian,C.H., Zykovich,A., Yu,A., Lockwood,S.H., Morbitzer,R., Elsaesser,J. *et al.* (2013) Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res.*, **41**, 4118–4128.

37. Doherty,J.E., Huye,L.E., Yusa,K., Zhou,L., Craig,N.L. and Wilson,M.H. (2012) Hyperactive piggyBac gene transfer in human cells and *in vivo*. *Hum. Gene Ther.*, **23**, 311–320.
38. Herzog,R.W., Cao,O. and Srivastava,A. (2010) Two decades of clinical gene therapy–success is finally mounting. *Discov. Med.*, **9**, 105–111.
39. Papapetrou,E.P., Lee,G., Malani,N., Setty,M., Riviere,I., Tirunagari,L.M., Kadota,K., Roth,S.L., Giardina,P., Viale,A. *et al.* (2011) Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. *Nat. Biotechnol.*, **29**, 73–78.
40. Tan,W., Dong,Z., Wilkinson,T.A., Barbas,C.F. 3rd and Chow,S.A. (2006) Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *J. Virol.*, **80**, 1939–1948.
41. Gaj,T., Mercer,A.C., Sirk,S.J., Smith,H.L. and Barbas,C.F. 3rd (2013) A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. *Nucleic Acids Res.*, **41**, 3937–3946.
42. Ammar,I., Gogol-Doring,A., Miskey,C., Chen,W., Cathomen,T., Izsvak,Z. and Ivics,Z. (2012) Retargeting transposon insertions by the adeno-associated virus Rep protein. *Nucleic Acids Res.*, **40**, 6693–6712.
43. Feng,X., Bednarz,A.L. and Colloms,S.D. (2010) Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucleic Acids Res.*, **38**, 1204–1216.
44. Brady,T.L., Schmidt,C.L. and Voytas,D.F. (2008) Targeting integration of the *Saccharomyces* Ty5 retrotransposon. *Methods Mol. Biol.*, **435**, 153–163.
45. Yant,S.R., Huang,Y., Akache,B. and Kay,M.A. (2007) Site-directed transposon integration in human cells. *Nucleic Acids Res.*, **35**, e50.
46. Ivics,Z., Katzer,A., Stuwe,E.E., Fiedler,D., Knespel,S. and Izsvak,Z. (2007) Targeted sleeping beauty transposition in human cells. *Mol. Ther.*, **15**, 1137–1144.
47. Maragathavally,K.J., Kaminski,J.M. and Coates,C.J. (2006) Chimeric Mos1 and piggyBac transposases result in site-directed integration. *FASEB J.*, **20**, 1880–1882.
48. Szabo,M., Muller,F., Kiss,J., Balduf,C., Strahle,U. and Olasz,F. (2003) Transposition and targeting of the prokaryotic mobile element IS30 in zebrafish. *FEBS Lett.*, **550**, 46–50.
49. Kettlun,C., Galvan,D.L., George,A.L. Jr, Kaja,A. and Wilson,M.H. (2011) Manipulating piggyBac transposon chromosomal integration site selection in human cells. *Mol. Ther.*, **19**, 1636–1644.
50. Kim,Y., Kweon,J., Kim,A., Chon,J.K., Yoo,J.Y., Kim,H.J., Kim,S., Lee,C., Jeong,E., Chung,E. *et al.* (2013) A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.*, **31**, 251–258.
51. Lombardo,A., Genovese,P., Beausejour,C.M., Colleoni,S., Lee,Y.L., Kim,K.A., Ando,D., Urnov,F.D., Galli,C., Gregory,P.D. *et al.* (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.*, **25**, 1298–1306.
52. Li,X., Burnight,E.R., Cooney,A.L., Malani,N., Brady,T., Sander,J.D., Staber,J., Wheelan,S.J., Joung,J.K., McCray,P.B. Jr *et al.* (2013) piggyBac transposase tools for genome engineering. *Proc. Natl Acad. Sci. USA.*, **18**, E2279–E2287.
53. Benabdallah,B.F., Allard,E., Yao,S., Friedman,G., Gregory,P.D., Eliopoulos,N., Fradette,J., Spees,J.L., Haddad,E., Holmes,M.C. *et al.* (2010) Targeted gene addition to human mesenchymal stromal cells as a cell-based plasma-soluble protein delivery platform. *Cytotherapy*, **12**, 394–399.
54. Zou,J., Maeder,M.L., Mali,P., Pruett-Miller,S.M., Thibodeau-Beganny,S., Chou,B.K., Chen,G., Ye,Z., Park,I.H., Daley,G.Q. *et al.* (2009) Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell*, **5**, 97–110.
55. Zou,J., Sweeney,C.L., Chou,B.K., Choi,U., Pan,J., Wang,H., Dowey,S.N., Cheng,L. and Malech,H.L. (2011) Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. *Blood*, **117**, 5561–5572.
56. DeKelver,R.C., Choi,V.M., Moehle,E.A., Paschon,D.E., Hockemeyer,D., Meijsing,S.H., Sancak,Y., Cui,X., Steine,E.J., Miller,J.C. *et al.* (2010) Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.*, **20**, 1133–1142.
57. Hockemeyer,D., Soldner,F., Beard,C., Gao,Q., Mitalipova,M., DeKelver,R.C., Katibah,G.E., Amora,R., Boydston,E.A., Zeitler,B. *et al.* (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.*, **27**, 851–857.
58. Perez-Pinera,P., Ousterout,D.G., Brown,M.T. and Gersbach,C.A. (2012) Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Res.*, **40**, 3741–3752.
59. Bedell,V.M., Wang,Y., Campbell,J.M., Poshusta,T.L., Starker,C.G., Krug,R.G. 2nd, Tan,W., Penheiter,S.G., Ma,A.C., Leung,A.Y. *et al.* (2012) *In vivo* genome editing using a high-efficiency TALEN system. *Nature*, **491**, 114–118.
60. Tong,C., Huang,G., Ashton,C., Wu,H., Yan,H. and Ying,Q.L. (2012) Rapid and cost-effective gene targeting in rat embryonic stem cells by TALENs. *J. Genet Genomics*, **39**, 275–280.
61. Li,M.A., Turner,D.J., Ning,Z., Yusa,K., Liang,Q., Eckert,S., Rad,L., Fitzgerald,T.W., Craig,N.L. and Bradley,A. (2011) Mobilization of giant piggyBac transposons in the mouse genome. *Nucleic Acids Res.*, **39**, e148.