NOVEL PIGGYBAC TRANSPOSASE VECTORS FOR SAFER GENE ADDITION INTO MAMMALIAN GENOMES

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

CELL AND MOLECULAR BIOLOGY

MAY 2014

By

Jesse Bruce Owens

Dissertation Committee:

Stefan Moisyadi, Chairperson
W. Steven Ward
Joe Ramos
Olivier Le Saux
Pauline Chinn

Keywords: transposon, transposase, piggyBac, DNA binding domain, zinc finger, transcription activator-like effector (TALE)

Acknowledgements

I would like to thank the members of the Moisyadi lab for their intellectual stimulation and help with the generation of ideas, for their encouragement during difficult experiments, and for their good humor throughout.

Stefan Moisyadi
Johann Urschitz
Ilko Stoytchev
Damiano Mauro
Zoia Stoytcheva
Nong Dang
Marlee Elston

Abstract

An extensive range of treatments of both inherited and acquired diseases are now possible due to our ability to permanently introduce foreign genes into chromosomes. However, the uncontrolled nature of random vector gene insertion presents a mutagenic risk. In this dissertation, a novel *piggyBac* (pB) transposon system has been developed to address potential genotoxicity issues. We hypothesized that modifications to the pB transposase could facilitate safer gene addition into mammalian genomes. In Chapter 2, we present a single-plasmid system, termed GENIE, that incorporates all of the required components for integration. These vectors are able to inactivate the pB gene after excision of the delivery transposon from the plasmid. This feature eliminates potential negative consequences that could arise from the persistence of an active pB transposase inadvertently taken up by the genome. In Chapter 3, we explore an application of our GENIE vector system by demonstrating effective knockdown of telomerase reverse transcriptase (TERT) in human immortalized cell lines. A significant hurdle for safer therapeutic gene addition is developing a method for controlling the precise location of insertion. We have developed a targetable transposase system by fusing DNA binding domains (DBDs) to pB in order to localize insertions near specific recognition sequences. In Chapter 4, we improved the GENIE vector system by fusing a GAL4 DBD to the pB transposase and demonstrated the ability of our vectors to target transgenes to predetermined sites. Gal4 recognition sites found on episomal

plasmids and on target sequences introduced into the human genome were preferentially targeted by our chimeric Gal4-pB transposase. Furthermore, a genome-wide integration analysis revealed the ability of our fusion constructs to bias integrations near endogenous Gal4 recognition sequences. In Chapter 5, we demonstrated site-specific and user-defined transposition to the CCR5 genomic safe harbor. Custom TALE DBDs were designed to bind the first intron of the human CCR5 gene. These TALE proteins were incorporated into a variety of novel targeting vectors that used both plasmid-DNA and transposase-protein relocalization to the target sequence. We used these vectors to isolate single-copy clones harboring targeted integrations.

Publications Arising from this Dissertation

Journal Publications

1. **Owens, J.B.**, Mathews, J., Davy, P., Stoytchev, I., Moisyadi, S., Allsopp, R. (2013) Effective Targeted Gene Knockdown in Mammalian Cells Using the piggyBac Transposase-based Delivery System. *Molecular therapy. Nucleic acids* 2:e137.

2. **Owens, J.B.**, Mauro, D., Stoytchev, I., Bhakta, M.S., Kim, M.S., Segal, D.J. and Moisyadi, S. (2013) Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. *Nucleic acids research***, 41**, 9197-9207.

3. **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 research*, **40**, 6978-6991.

4. Urschitz, J., Kawasumi, M., **Owens, J.B.**, Morozumi, K., Yamashiro, H., Stoytchev, I., Marh, J., Dee, J.A., Kawamoto, K., Coates, C.J., Kaminski, J.M., Pelczar, P., Yanagimachi, R., and Moisyadi, S. (2010) Helper-independent piggyBac plasmids for gene delivery approaches: strategies for avoiding potential genotoxic effects. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 8117-8122.

Journal Publications Not Directly Related to the Dissertation

1. Anderson, C.D., Urschitz, J., Khemmani, M., **Owens, J.B.**, Moisyadi, S., Shohet, R.V. and Walton, C.B. (2013) Ultrasound Directs a Transposase System for Durable Hepatic Gene Delivery in Mice. *Ultrasound Med Biol*, **39**, 2351-2361.

2. 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. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 19184-19189.

Presentations and Abstracts

1. Safer gene addition by targeting transposons to specific sites. *Oral presentation at the UH Institute for Biogenesis Research seminar series* (2013) Honolulu, HI
2. Helper Independent Self-inactivating Chimeric piggyBac Transposases for Genomic Targeting in Human Cells. *Poster presentation at The American Society of Gene and Cell Therapy 16th Annual Meeting* (2013) Salt Lake, UT
3. Helper Independent piggyBac Vectors for Gene Therapy. *Poster presentation at The American Society of Gene and Cell Therapy 13th Annual Meeting* (2010) Washington, DC
4. Novel piggyBac Vectors for Safer Gene Delivery. *Oral presentation at the UH Cell and Molecular Biology annual retreat* (2010) Honolulu, HI

Table of Contents

List of Tables

List of Figures

List of Abbreviations

µg	micrograms
µL	microliters
µm	microns (micro meters, 1 x 10^{-6} meters)
µM	micromoles per liter
AAV	adeno-associated virus
BAC	bacterial artificial chromosome
BSA	bovine serum albumin
CAG	CMV-immediate early enhancer, chicken β-actin promoter and β-globin intron
Cas	CRISPR associated
cam	chloramphenicol
CCR5	chemokine C-C motif receptor 5
CMV	cytomegalovirus
CRISPR	clustered regularly interspaced short palindromic repeats
CSH4	chicken b-globin HS4
DBD	DNA binding domain
DMSO	dimethyl sulfoxide
DNA	deoxynucleic acid
dpc	days post coitus
DpBS	Dulbecco's phosphate buffered saline
DSB	double-stranded break
EGFP	enhanced green florescent protein
ESC	embryonic stem cells
FBS	fetal bovine serum
g	grams
gDNA	genomic DNA
GFP	green florescent protein
GIN	DNA sequence including CMV, GFP, IRES, and neomycin
gRNA	guide RNA
hCG	human chorionic gonadotropin
kan	kanamycin
HEK	human embryonic kidney
HIV	human immunodeficiency virus
HPRT	hypoxanthine phosphoribosyltransferase
hr	hours
ICSI	intracytoplasmic sperm injection
i.p.	intra peritoneal
*i*pBt	insect pB transposase gene
iPCR	inverse PCR
iPSC	induced pluripotent stem cell
IRES	internal ribosome entry site

IVF	*in vitro* fertilization
LTR	long terminal repeat
MCS	multiple cloning site
mg	milligrams
min	minutes
mg	milligrams
mL	milliliters
mM	millimoles per liter
*m*pB	mouse-codon-biased pB gene
mRNA	messenger RNA
neo	neomycin
nmol	nanomoles
nm	nanometers
nM	nanomoles per liter
ori	origin of replication
pB	*piggyBac*
pBS	phosphate buffered saline
PCR	polymerase chain reaction
PVP	polyvinylpyrrolidone
Rep	reporter
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RQ	relative quantification
RT	room temperature
SB	*Sleeping Beauty*
sec	seconds
SD	standard deviation
SEM	standard error of the mean
shRNA	short hairpin RNA
SV40	simian virus 40
TALE	transcription activator-like effector
TALEN	TALE nuclease
TERC	telomerase RNA
TERT	telomerase reverse transcriptase
TRE	terminal repeat element
TRF	terminal restriction fragment
UAS	upstream activating sequence
w/v	weight by volume
ZFN	zinc finger nuclease
ZFP	zinc finger protein
ZFR	zinc-finger recombinase

Chapter 1. Introduction

1.1 Gene Addition using Viruses

A critical issue concerning gene addition using popular viral vectors is their tendency to preferentially insert cargo near actively transcribed genes (1-8) and importantly, 5 of the 20 patients enrolled in two SCID-X1 HSC gene therapy trials developed T cell acute lymphoblastic leukemia due to vector-mediated insertional mutagenesis (9-11). In this case, leukemic transformation was due to activation of endogenous genes by gammaretroviral long terminal repeats (LTRs). Self-inactivating vectors devoid of LTR promoter activity have been proposed as an alternative (12), however the semi-random integration pattern of these vectors present a mutagenic risk.

1.2 The CCR5 Genomic Safe Harbor for Gene Transfer

The ability to accurately insert therapeutic genes to genomic safe harbors in human cells has been described as the “holy grail” of transpositional research (13). Genomic safe harbors can be defined as loci well-suited for gene transfer, as integrations within these sites are not associated with adverse effects such as proto-oncogene activation or tumor suppressor inactivation. Furthermore, safe harbors should allow stable transgene expression across multiple cell types. One such putative site, chemokine C-C motif receptor 5 (CCR5), has been identified and used for integrative gene transfer. CCR5 is a member of the beta chemokine receptor family and is required for the entry of R5 tropic viral strains involved in primary infections. A homozygous Δ32 deletion in the *CCR5* gene confers resistance to HIV-

1 virus 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 (14). A clinical trial has demonstrated safety and efficacy of disrupting CCR5 via targetable nucleases, see below, as part of an anti-HIV therapeutic approach (15).

1.3 Directed Gene Transfer Approaches

Directed gene transfer approaches have been recently developed in order to resolve issues associated with random integration. Homologous recombination, the routine method for site-directed integration into embryonic stem cells for the generation of knockout animals, has been beset by low efficiencies. As an alternative, single stranded AAV vector-mediated homologous recombination has been used to correct point mutations efficiently (16-18).

1.3.1 Transcription activator-like effectors (TALEs)

TALEs are plant pathogen proteins that were recently shown to contain a DNA-binding motif (19, 20). A simple “recognition code” has been established by correlating hypervariable residues within domain repeats to the recognition of specific DNA base pairs. Binding specificity is determined by customizable arrays of amino acid repeats, with each repeat independently specifying its targeted base. TALEs can therefore be designed to bind desired sequences by arranging the appropriate repeats in tandem. Such arrays have been used to design TALE nucleases (TALENs), chimeric proteins that consist of a TALE DBD and a Fok1

nuclease domain, to induce targeted double-stranded breaks in the genome to increase the efficiency of homologous recombination (21). TALENs have been used to cause intentional mutations or insert whole genes at respective targets (22).

1.3.2 Clustered regularly interspaced short palindromic repeats (CRISPRs)

Recently, an RNA-programmable nuclease has been re-purposed for inducing directed DSBs (23, 24). The CRISPR system uses a CRISPR associated 9 (Cas9) nuclease complexed with a user-defined guide RNA (gRNA) to recognize and cut complementary sequences. Unlike TALENs, this system doesn't require custom protein design but instead simply involves the input of a 20 bp target sequence expressed as a piece of a single small gRNA.

Both TALEN and CRISPR approaches take advantage of host homology-directed repair to introduce a co-delivered donor template at the desired sequence. Importantly, efficiencies for targetable nucleases far exceed traditional homologous recombination, enabling straightforward gene transfer into such cells as human embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). However, the nuclease component can cause off-target cleavage events that result in undesired mutations and require complex genotoxicity screens (25-33). Concerns about cyto and genotoxicity remain significant obstacles to be overcome for nucleases to be safe strategies. Furthermore, gene addition using homology directed repair requires replication, thus limiting nuclease technology to dividing tissues. Although nucleases are efficient at facilitating gene transfer, development of

alternatives that do not rely on DNA double-stranded breaks and involve the associated risks of off-target mutagenesis are necessary to improve the safety of genome engineering.

1.4 Transposon-Based Vectors

Class II transposases achieve transposition by excising the transposon from the donor DNA followed by stable insertion into the recipient DNA sequence. Several transposon systems, among them *Sleeping Beauty* (SB) and pB, have been shown to efficiently deliver trangenes *in vitro* and *in vivo* (34-40). pB*,* an insect transposase isolated from the moth *Trichoplusia Ni*, is highly efficient in a broad range of organisms and is able to integrate large gene cassettes of more than 100 kb (41). pB inserts a transposon flanked by terminal repeat elements (TREs) at TTAA tetranucleotide sites and the transposase has been shown to be able to excise the transposon without leaving a DNA footprint (42). Our laboratories were the first to demonstrate that pB was the optimal transposase for transfecting mammalian cell lines (39) and generating transgenic animals (40, 43). We have since improved the plasmid constructs to be self-inactivating and to contain all transpositional machinery on a single helper-independent plasmid (GENIE) in order to reduce possible genotoxicity due to unintentional transposase integration (40). We have also shown that pB is the only transposase amenable to a Gal4-DBD fusion with little loss in activity (39).

1.5 GENIE Vector Design

We designed our pB-based GENIE constructs to address several safety issues concerning integrating vectors. One of the novel characteristics of our vector is its self-inactivating mechanism. The 3'-TRE, located within an intron of pB spatially separates the cytomegalovirus (CMV) immediate early enhancer, chicken β-actin

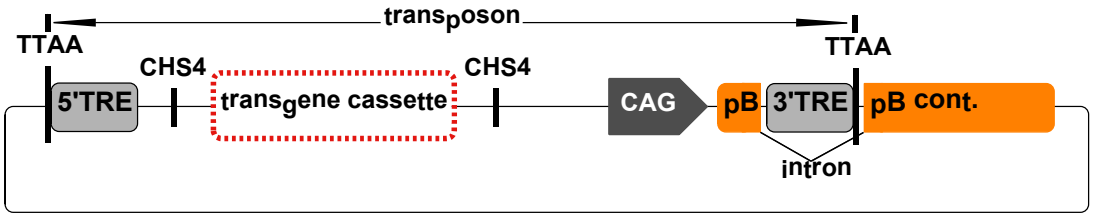


Figure 1-1 GENIE plasmid design

promoter and β-globin intron (CAG) promoter from the pB gene (Figure 1-1). Therefore, enzymatic excision of the transposon, located between the TTAA sites, from the plasmid during transposition will result in the separation of the CAG promoter from 5' end of pB. The now promoterless transposase, residing in the remaining plasmid backbone will therefore be inactive if inserted non-transpositionally into the genome, reducing the possibility of potential genotoxic effects. To prevent the CAG promoter from driving gene expression downstream of the insertion site we engineered in-frame stop codons into the chimeric intron. This is expected to stop any protein synthesis from mRNA constructs that may be erroneously synthesized (40). As an additional safety feature, two 500-bp chicken b-globin HS4 (CSH4) insulators were inserted to flank the transgene cassette to reduce transcriptional silencing and position effects imparted by chromosomal

sequences. An additional effect these insulators provide is to eliminate functional interactions of the transgene enhancer and promoter sequences with neighboring chromosomal sequences.

1.6 Stable Integration of short hairpin RNA (shRNA): An Application of the GENIE Vector

shRNA is a type of RNA that can be used to knockdown specific transcripts using the RNA interference (RNAi) mechanism. Following transcription, the shRNA forms a hairpin that mimics pri-microRNA that can be processed by Drosha. This process creates pre-shRNA which is exported from the nucleus and processed by cytoplasmic Dicer. The subsequent product is then loaded into the RNA-induced silencing complex (RISC). The antisense strand of the RNA then binds to complementary endogenous mRNA and directs RISC cleavage to this mRNA. The final result of this process is sequence specific gene silencing (44).

The telomerase enzyme is used by eukaryotic cells to add DNA sequence to the ends of chromosomes. Telomerase adds DNA repeats ("TTAGGG" in vertebrates), called telomeres, to protect coding DNA from the inherent shortening of DNA during replication. The synthesis of Okazaki fragments during replication requires an RNA template followed by a DNA strand. When replication hits the end of the chromosome, no DNA follows the RNA template causing the last fragment to not be synthesized, thus the chromosome is shortened. Also, it has been shown that free

radicals caused by oxidative stress significantly cause additional telomere shortening (45). The two molecules that make up telomerase are telomerase reverse transcriptase (TERT) and telomerase RNA (TERC). TERT is a reverse transcriptase, which uses TERC single-stranded RNA as a template for the generation of single-stranded DNA. The lengthening of telomeres by telomerase is a major factor in the potential for cells to become immortalized, as cells that continuously lose DNA after each division tend to become senescent. In fact, ~90% of known cancers have upregulated telomerase (46). Therefore we used shRNA expressed from our GENIE vector to stably knockdown TERT in immortalized cell lines. We demonstrated a successful reduction of TERT expression and associated reduced telomerase activity and length of telomeres.

1.7 DNA Binding Domain (DBD) Directed Transposition

Targetable nucleases only create the initiating double strand breaks, which are subject to multiple and uncontrolled cellular repair pathways. More efficient, safer, and predictable gene targeting can be accomplished with a targetable protein that can perform all the enzymatic steps, such as a targetable transposase. In a recent publication, we demonstrated, for the first time, DBD-directed transposition into endogenous genomic sequences (47). In this study we demonstrated that Gal4-pB fusion proteins but not native pB were capable of targeting transposition to chromosomal locations nearby upstream activating sequence (UAS) sites.

1.7.1 TALE Directed Transposition

Following this original *proof-of-principle* publication, a second targeting study explored further improvements to our plasmid architecture, including a hyperactive pB variant and customizable TALE-DBDs fused to the N-terminal of the pB transposase. These modifications enabled us to establish *the first example of targeting an integrating enzyme to a user-defined TALE-directed endogenous location* (48). Using this approach, we isolated clonal expansions of cells positive for a single insertion into the CCR5 safe-harbor.

To take advantage of the superior performance of the recently described TALE DBDs in regards to target selection and specificity we replaced the Gal4 DBD in our targetable transposon system (47) with custom TALE DBDs designed to bind the human CCR5 gene. Each TALE protein was designed to a 17-bp target site that was preceded by a required T base, and each contained at least 2 mismatches with any other site in the human genome. Regions within the genomic safe harbor locus were selected because of the high density of potential pB TTAA transposon insertion sites. We constructed two TALE DBDs for CCR5, denoted as TALC1 and TALC2. TALE design was based on framework described by Miller et al (49) and is similar to that used by Sangamo Biosciences. Our TALEs were shown to support high-affinity binding based on our gel mobility shift and cell-based reporter assays.

1.7.2 Targeted insertion site recovery

Transfections of HEK293 cells with a chimeric TALE-pB plasmid resulted in targeted insertions to the CCR5 locus as analyzed by nested PCR. Primers had been designed to be either complementary to the TRE of the transposon or anneal up- or downstream of the CCR5 gene. A total of fourteen unique insertion sites within CCR5 were recovered. Nine of the fourteen insertion sites were located within 250 bp of the TALE recognition sequence and two insertions were located ~600 bp away. In addition, three insertions were recovered far (>1200 bp) from the target sequence. Given that there are millions of TTAA sites available for random integration of the transposon into the host genome our data demonstrate that indeed TALEs display the ability to direct the pB transposase to a predetermined genomic locus.

Targeted integration of a transposon is an all-or-nothing event that is readily identified by assaying the copy number of transposon insertions. Therefore a single-insertion clone is not expected to have additional DNA modifications. In comparison, because targetable nucleases are capable of mutating the genome without introducing an identifiable insert, it remains difficult to confirm DNA integrity.
Our copy number assay (47) was used to measure the total number of insertions for correctly identified clones. This duplex quantitative PCR compares the amount of product arising from amplification from inserted sequence relative to a constant endogenous sequence and is calibrated by a sample with a known number of

insertions. Importantly, this assay requires minimal gDNA template and can provide highly accurate results for up to 90 clones in a single PCR step. Targeted clones were considered verified only if found to contain a single insertion. This ensured that clones did not contain off-target integrations.

1.8 Summary

Viruses insert transgenes semi-randomly and can potentially disrupt or deregulate genes. Using a non-viral transposon-based vector we have alleviated potential genotoxic outcomes due to unwanted transposase integration. We have used this vector in a relevant application for the successful knockdown of TERT in immortalized cells. Finally, we have developed a non-viral genomic targeting vector that can safely integrate large transgenes into a single user-defined locus.

Chapter 2. Helper-independent *piggyBac* plasmids for gene delivery approaches: strategies for avoiding potential genotoxic effects

2.1 Abstract

Efficient integration of functional genes is an essential prerequisite for successful gene delivery strategies such as cell transfection, animal transgenesis and gene therapy. Gene delivery strategies based on viral vectors are currently the most efficient. However, limited cargo capacity, host immune response, and the risk of insertional mutagenesis are limiting factors and of concern. Recently, several groups have used transposon-based approaches to deliver genes to a variety of cells. The pB transposase in particular has been shown to be well suited for cell transfection and gene therapy approaches because of its flexibility for molecular modification, large cargo capacity and high transposition activity. However, safety considerations regarding transposase gene insertions into host genomes have rarely been addressed. Here we report our results on engineering helper-independent pB plasmids. The single-plasmid gene delivery system carries both the *piggyBac* transposase (pBt) expression cassette as well as the transposon cargo flanked by terminal repeat element (TRE) sequences. Improvements to the helper-independent structure were achieved by developing new plasmids in which the pBt 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 pBt gene post-transposition are eliminated.

The results presented herein demonstrate that our helper-independent plasmids represent an important step in the development of safe and efficient gene delivery methods that should prove valuable in gene therapy and transgenic approaches.

2.2 Introduction

The ability to safely and efficiently integrate genes into a host genome is essential for successful genetic modification strategies in the context of functional genomic studies, transgenesis and gene therapy. Currently, the most commonly used vectors for permanent or transient transfer of genes in pre-clinical gene therapy trials are virus-based; however non-viral vectors are also being developed. While it is possible to achieve stable genomic integration with high efficiency using viral vectors, multiple studies have pointed out serious disadvantages. Adenoviruses, for example, have been shown to evoke host immune responses (50), while retroviruses preferentially integrate transgenes into euchromatin thereby increasing the risk of insertional mutagenesis. Viral systems are also limited in cargo size (36, 51), restricting the size and number of transgenes and their regulatory elements. Finally, biosafety considerations and production costs are additional factors to consider when using viral vectors (51, 52).

In order to avoid some of the potential drawbacks of viral systems, transposons have been tested and successfully used as an alternative in a diverse field of applications.

DNA transposons are mobile elements that generally use a “cut and paste” mechanism in which the DNA is excised by double strand cleavage from the donor molecule and consecutive integration into the acceptor molecule. Many of the initial experiments were conducted using SB (53). More recently however, pBt from the moth *Trichoplusia ni* has been widely used as a means for gene delivery in a variety of applications such as cell line transformation (39, 54-56), mutational analysis (57) and gene therapy (58). Additionally, the pB system displays several highly desirable features that are of great advantage for transgene integration; pBt is very efficient and has a higher transposition activity than any of the SB variants (39, 59, 60); the pB system allows for remobilization of transposons as they can be removed without leaving a “footprint” (42) and under certain conditions, pBt does not show any overproduction inhibition as do some members of the Tc1/mariner transposon super-family (60, 61). Finally pBt has been shown to deliver transposons up to 18kb (56), the largest cargo size described to date in any of the transposon or integrating viral systems.

Most transposition attempts for cell line transfections or gene therapy experiments with pB have used the two-plasmid donor-helper system. The donor-plasmid contains inverted terminal repeats flanking the transgene while the helper-plasmid transiently expresses the pBt enzyme, which catalyzes the insertion event from the donor plasmid to the host genome (62). Similarly, two component systems have been used for pBt-initiated transgenesis where a donor plasmid is co-injected with transposase RNA (63, 64). However, this approach is problematic because of issues

with RNA stability. The two-plasmid approach works well with cell line transfections but in our experience, this approach did not support transgenesis using Intracytoplasmic Sperm Injection (ICSI) (65).

In order to develop a pB system that is more suitable for transgenesis and gene therapy experiments, helper-independent single plasmids were constructed, containing both the donor and helper elements of the pB transposition system in the same circular construct, similar to plasmids previously described (66, 67). This cis-acting plasmid configuration ensures that both components, transposon and transposase, are delivered simultaneously to the nucleus and hence should result in an improved transposition efficiency. Here, we describe the structure and function of three such plasmids (p*i*GENIE, p*m*GENIE-2 and p*m*GENIE-3): The plasmids were designed to contain the transposase in an arrangement that prevents the activation of the enzyme in case of random, non-transpositional integration of the plasmid backbone that originates from the plasmid after transgene excision. Such a feature will be invaluable in clinical relevant settings as it increases the safety of plasmid-based research and gene therapy approaches as it prevents some of the potential genotoxic effects intrinsic in these systems.

2.3 Experimental Procedures

2.3.1 Plasmid Development.

The original pB transposase DNA was a gift from Dr. Malcolm Fraser (University of Notre Dame, IN). All the original pBt cDNA constructs were under the control of the

CMV promoter in the plasmid pcDNA3.1Δneo, derived from pcDNA3.1 (Invitrogen, Carlsbad CA) by removing most of the neomycin resistance gene. The pcDNA3.1Δneo plasmid also contained a pUC replication origin for plasmid propagation in bacteria. The CMV promoter was converted into the more active CAG (CMV-IE enhancer, chicken β-actin promoter and β-globin intron) promoter by restriction digestion with *Sna*BI and *Eco*RI (all restriction enzymes were obtained from New England Biolabs (NEB), Beverly, MA). The fragment excised by the two restriction enzymes from pcDNA3.1Δneo+pB plasmids was discarded and replaced by a CAG fragment obtained with a similar restriction digestion from plasmid pCX-EGFP (a kind gift from Dr. Junichi Miyazaki, Osaka University). After the appropriate fragments were gel purified and recovered, the CAG fragment was cloned by ligation into the pB cDNA contained in the pcDNA3.1Δneo plasmids, giving rise to pXC-insect-pB or pCX-mouse-pB (pCX-*i*pB and pCX-*m*pB) constructs.

p*i*GENIE: pCX-*i*pB was restricted with *Pvu*II, which cuts in two places between the polyA site of the pB gene and the pUC origin of replication (pUC ori), giving rise to a blunt-ended linearized product. An empty pB donor plasmid with a minimal multiple cloning site (MCS) in the transposon, flanked by the 3'- and 5'- terminal repeat elements (TREs) was restricted with *Pvu*II, which cuts outside the 3'- and 5'-TREs and removes the ampicillin resistance gene and pUC ori sequences. The appropriate fragments from both plasmids were recovered by gel purification (Zymoclean, Zymo Research, Orange, CA) and ligated to each other at 16^{o}C overnight with T4 DNA ligase (NEB, Beverly MA), with the 3'- TRE in line ahead of

the 5'-TRE. The resulting construct had the ampicillin resistance gene and the pUC ori conferred on it by the pCX-*i*pB construct. The Invitrogen RFC.1 Gateway system recombineering site was transferred by blunt-ended ligation into a *Sma*I site on the minimal MCS.

p*m*GENIE-2: The donor plasmid pXLBAC-EGFP contained a CAG-promoter-driven EGFP gene between the 3'- and 5'- TREs of the transposon. The CAG promoter terminates with an *Eco*RI site as does the following EGFP gene. There is another *Eco*RI site just 3'- of the 3'-TRE. Therefore, *Eco*RI restriction digestion and purification of the correct fragment, followed by a T4 DNA ligation, leave a CAG promoter abating the 3'-end of the 3'-TRE, with the 5'-end of the CAG promoter being flanked by the 5'-end of the 5'-TRE. On the 5'-side of the 3'-TRE in the modified pXLBAC-EGFP plasmid, now referred to as DP, there is a *Pst*I restriction site with a PsiI site further upstream. The MCS from pBluescript SK+ was restricted with *Pst*I and *Psi*I and the purified fragment was cloned into similarly restricted DP plasmid. The newly introduced MCS contained a *Kpn*I restriction site. DP plasmid was digested with *Psi*I and *Kpn*I, and the appropriate fragment with the CAG promoter was column purified after gel separation. The mouse-codon-biased pB gene (*m*pB) was digested with *Psi*I at its 3'-end, including its BGH polyA site, and digested with *Kpn*I at its 5'-end, just in front of its Kozak translation initiation site from the pcDNA3.1Δneo plasmid. After gel purification, the fragment was cloned into the corresponding DP plasmid sites, giving rise to plasmid DP-1, with the Kozak and ATG sites of the *m*pB gene fused to the 5'-end of the 3'-TRE. The Invitrogen RFC.1

recombineering site of the Gateway system was transferred from the p*i*GENIE construct described above using the flanking *Sbf*I digestion sites and cloned into a *Sbf*I restriction site situated just 3' of the CAG promoter in the transposon of DP-1. After verifying the orientation of the RCF.1 fragment, the new plasmid obtained was referred to as p*m*GENIE-2.

Two plasmid donor-helper system: The donor plasmid was built by inserting the selection cassette into the *Sma*I/*Eco*RV sites of pXLBacIIPUbnlsEGFP, derived from pBSII-ITR1. This donor plasmid is a minimal pB vector with terminal repeat elements of 308 bp and 238 bp at the 5' and 3' ends, respectively. The helper plasmid (pCX-*m*pB) was based on the *i*pBt helper plasmid described previously. The CMV promoter of *i*pBt was replaced by a CAG-promoter-driven *m*pBt gene by restriction digestion with consequent ligation.

p*m*GENIE-3: The intron to be introduced into the *m*pB gene, including the 3'-TRE in the correct orientation, was synthesized by GenScript Corp., Piscataway, NJ. The intron sequence was flanked by *m*pB sequences on either side, with correct intron splicing consensus sequences engineered into it. The synthesized sequence additionally had *Bam*HI restriction digestion sites engineered into it, matching the exact same site in the *m*pB gene and an upstream untranscribed region past the Kozak site. The fragment was introduced in the correct orientation into the DP-1 construct described above by restriction digestion with *Bam*HI and ligation. The orientation was verified by digestion and sequencing. The original 3'-TRE from the

DP-1 backbone was excised by *Eco*RI digestion from either side of it. The backbone fragment was re-ligated after gel and column purification, giving rise to plasmid DP-2. DP-2 was converted into plasmid p*m*GENIE-3 by transfer of the Invitrogen RFC.1 recombineering site of the Gateway system from the p*i*GENIE construct described above. This was achieved by *Sbf*I digestion sites flanking it and cloning into a *Sbf*I restriction site situated just 3' of the CAG promoter in the transposon of DP-2. After verifying the orientation of the RCF.1 fragment by sequencing, the new plasmid was referred to as p*m*GENIE-3.

pENTR1a with pXC-EGFP+SV40-Hygromycin+bacterial-Kanamycin resistance gene: The plasmid pMMK-1 contained the pXC-EGFP+SV40-Hygromycin+bacterial-Kanamycin resistance genes and the pUC origin of replication (pUC ori) in its transposon flanked by the 3'- and 5'-TREs of pB. In addition to expressing the EGFP and Hygromycin genes for mammalian selection, the transgene cassette also confers Kanamycin resistance on bacterial cells. The *Sna*BI restriction site in the CAG promoter of pXC-EGFP and a *Kpn*I site downstream of the mammalian Kanamycin resistance gene and pUC ori were used to digest out the selection cassette from pMMK-1. The fragment was gel purified and recovered via column purification. The plasmid pENTR1a containing the sequence of the pTandem-1 construct (Novagen, Madison, WI) was digested in the CMV IE enhancer region with *Sna*BI and at the end of the IRES sequence with *Kpn*I. The region between *Sna*BI and *Kpn*I was discarded by gel fragment purification as above and the fragment that had the pXC-EGFP+SV40-Hygromycin+mammalian-Kanamycin resistance genes

produced by *Sna*BI and *Kpn*I digestion, ligated into it with T4 DNA ligase overnight at 16^{o}C.

2.3.2 Cell Transfections.

HEK293T cells were maintained in complete DMEM supplemented with 10% heat inactivated FBS (Invitrogen). 0.5x10^{5} cells per well were seeded in 12-well plates. At 90% confluency, the cells were transfected using FuGene6 (Roche Applied Science, Dallas, TX) with 400ng of the respective circular GENIE DNA (or 100ng of helper and 200ng of donor) per well in triplicate. Twenty four hours later, 10% of the cells from each well were transferred into 10cm plates with 10ml DMEM supplemented with 10% FBS and 100 µg/ml HygromycinB (Invitrogen). The HygromycinB selection was maintained for 21 days, with the media changed every two days. Single cell colonies were established and expanded. Colonies were counted after fixing the cells with 4% Paraformaldehyde in 1xpBS for 10 minutes, and stained with 0.1% Methylene blue for 1h. We utilized two parameters, relative fold and percentage of transposition, to assess the transposition activity of the different plasmids. The relative fold was obtained by dividing the number of resistant colonies detected in cells transfected by GENIE plasmids with the colony number that resulted from random integration (i.e., controls with truncated transposases). The percentage of transposition is calculated by subtracting the number of hygromycin-resistant colonies detected in the controls from the number of resistant colonies in the presence of transposase, dividing by 0.5x10^{5} (the number of cells originally seeded before transfection), and finally multiplying by 100.

2.3.3 Immunofluorescence protocol for embryo staining.

Oocytes were washed in pBS-0.1% PVA ×3 and then fixed in pBS-0.1% PVA containing 4% paraformaldehyde at RT for 15 min. Following the removal of the zona pelucida, oocytes were washed again in pBS-3% BSA (RT, 15min×2). Cells were permeabilized and blocked with pBS-3% BSA containing 0.2% Triton X-100 at 4 °C, overnight followed by 2 wash steps in pBS-3% BSA (RT, 15min). A 1:50 dilution of the primary rabbit anti-mouse poly-clonal pB antibody (generous gift from Dr. M Fraser, Notre Dame) was used for the overnight staining at 4 °C and was followed by three pBS-3% BSA wash steps (RT, 30min). Secondary antibody staining at a dilution of 1:1000 (Santa Cruz Biotechnology, Santa Cruz CA) was performed at room temperature for 1 hour, followed by three washes with pBS-0.1% PVA. Cells were mounted in Vectashield (Vector Labs, Burlingame, CA) containing 2-5 µg/ml 4',6'–diamidino-2-phenylindole (DAPI; Invitrogen) to stain DNA.

2.3.4 Southern Blot Analysis.

Genomic DNA was isolated using the DNeasy kit (Qiagen, Valencia CA) following the manufacturer's protocol. Twenty micrograms of each sample were digested overnight using 10 units of *Hin*dIII (NEB) per microgram of gDNA. Twelve micrograms of gDNA was run for 75 h on a 1% agarose gel at 15 V. The gDNA was then blotted to a Hybond + nylon membrane (GE Healthcare, Piscataway NJ) for 12 h and processed for hybridization according to the method of Sambrook et al. (1989). A DIG-labeled EGFP probe was generated by PCR amplification using the

PCR DIG Probe Synthesis Kit (Roche Applied Science) and the following primers: EGFP F: (ACGTAAACGGCCACAAGTTC), EGFP R: (TGCTCAGGTAGTGGTTGTCG). PCR parameters used: initial denaturation at 94°C for 2 min, 35 cycles of 30s denaturation at 94°C, 30s annealing at 56°C, and 1min elongation at 72°C, with a final elongation for 10 min. Hybridization was performed overnight at 55°C using the DIG Easy Hyb Kit, (Roche Applied Science) and were processed according to the manufacturer's protocol. Chemiluminescent signals were visualized with an LAS-3000 imaging system (Fuji, Stamford CT).

2.3.5 Transposon Insertion Sites.

In order to characterize insertion sites, either a Vectorette kit (Sigma, St. Louis MO) or inverse PCR (iPCR) was used. Vectorette PCR was performed as outlined in the manufacturer's protocol. For iPCR, 1µg of DNA was digested with *Afl*II, *Avr*II or *Acl*I. The DNA was self-ligated with T4 DNA ligase (NEB) in dilute conditions and circularized fragments were used as templates for iPCR. For both iPCR and Vectorettes, primers directed from the terminal repeat elements into flanking genomic DNA were used: forward 5-TRE- GAG CTC CAA GCG GCG ACT GAG ATG and reverse 3-TRE- ACG CAT GAT TAT CTT TAA CGT ACG TCA CAA. PCR products were cloned into pGEM-T Easy (Promega) for sequencing according to the manufacturer's protocol, sequenced at the genomics core facility of the University of Hawaii and analyzed with NCBI's BLAST search (www.ncbi.nlm.nih.gov).

2.3.6 Detection of Non-Enzymatic Plasmid Insertion.

We performed PCR using different primers, templates and conditions. (A) Detection of EGFP in transgenic mice: PCR was performed using the Platinum PCR SuperMix (Invitrogen), with primers and cycling conditions described for the Southern blot analysis. (B) Detection of plasmid backbone insertion in transgenic mice and p*i*GENIE transformed HEK293T cells was performed by PCR using Platinum PCR SuperMix and the following parameters: initial denaturation at 94°C for 2 min followed by 35 cycles of 30s denaturation at 94°C, 30s annealing at 56°C, and 1min elongation at 72°C, with a final elongation for 10 min. Primer pairs were Beta-glob-Fwd-IDT (CGCCGGCAGGAAGGAAA) and *i*pB-R (GTAAGGGGTCCGTCAAAACA) for the detection of promoter and pB; *i*pB F (CGGAGTACCACTCGGTGAAT) and *i*pB R for the detection of insect pB; *i*pB F and 3'-TRE GENIE-A R (CCG ATA AAA CAC ATG CGT CA) for detection of insect pB and 3'-TRE. (C) Detection of transposon insertion of p*m*GENIE-2 transformed HEK293T cells: PCR was performed using the Platinum PCR SuperMix High Fidelity (Invitrogen), primers GenieA 5'-TRE_PUC ORI PCR-R (TGTGGAATTGTGAGCGGATA) and 6F: Genie-1_direct_2417-F (AGCGAGATCGTGAAG TGGAC) and the following PCR parameters: initial denaturation at 94°C for 2 min, 5 cycles of 20s denaturation at 94°C, 15s annealing at 64°C, and 20min elongation at 72°C followed by additional 35 cycles of 20s denaturation at 94°C, 15s annealing at 58°C, and 20min elongation at 72°C, with a final elongation for 20 min. Amplicons of the expected size were isolated using a Zymoclean kit, cloned into pGEM-T Easy (Promega) and sequenced. mpBt detection was performed by PCR using Platinum PCR SuperMix

and the following parameters: initial denaturation at 94°C for 2 min followed by 35 cycles of 30s denaturation at 94°C, 30s annealing at 56°C, and 1min elongation at 72°C, with a final elongation for 10 min. mpB-F (GCTCACGTTGTGGCTGTAGA) and mpB-R2 (CAGCAAGTA CGGCATCAAGA) primers were used in this reaction.

2.3.6.1 Oocyte Collection.

B6D2F1 female mice underwent ovulation induction by intra peritoneal (i.p.) injection of 5 IU pregnant mare's serum gonadotrophin (Calbiochem, San Diego, CA), followed by i.p. injection of 5 IU human chorionic gonadotropin (hCG; Calbiochem) 48 h later. Oocytes were collected from oviducts 14-16 h after the hCG injection and were then freed from cumulus cells by a 3-5 min treatment of 0.1% hyaluronidase (Sigma) dissolved in HEPES-CZB medium.

2.3.6.2 Sperm Collection. Both caudae epididymides were excised from B6D2F1 male mice. The droplet of sperm that welled up was transferred to a 1.5 ml Eppendorf tube containing HEPES-CZB medium and incubated for 15-30 min at room temperature.

2.3.6.3 Intracytoplasmic Sperm Injection.

ICSI was performed as described in detail by Kimura and Yanagimachi (65). For the experiment with p*i*GENIE and p*m*GENIE-2, a small drop of sperm suspension was mixed thoroughly with a small drop of the same volume of HEPES-CZB medium containing 24% (w/v) polyvinylpyrrolidone (PVP; MP Biomedicals, Solon, OH) and

different concentrations of circular GENIE plasmids (100, 200, 300, or 400ng/µl). As a result, the final concentration of the PVP and DNA in the medium was 50, 100, 150 or 200ng/ µl). The sperm head was separated from the midpiece and tail by applying one or more piezoelectric pulses in the medium containing DNA. After the midpiece and tail had been discarded, the head was redrawn into the pipette with DNA, and then sperm head and 0.22pg, 0.44pg, 0.66pg or 0.88pg (total amount) of plasmid were coinjected into an oocyte, respectively. For the p*m*GENIE-3 experiments, the sperm were treated with NaOH using a modified procedure based on the method described by Li *et al*. (68). Ten microliters of sperm suspension was mixed 1:10 with 10 mM NaOH solution (Sigma) in an Eppendorf tube and placed at room temperature for 1 h. The suspension was then washed once by centrifuging with HEPES-CZB medium for 3 min at 10,000 rpm, and the pellet was resuspended in HEPES-CZB medium. An aliquot of the sperm was co-incubated with different concentrations of circular p*m*GENIE-3 (0.2-20ng/µl) for 30 min at room temperature. Only sperm that had lost their tail by the NaOH-treatment were selected for ICSI. Oocytes were activated artificially by incubating them in Ca^{2+}-free CZB medium containing 10 mM $SrCl_2$ (Sigma) at 37°C and 5% CO_2 for 30 minutes. Sperm injection was then carried out with modification of the technique as described above.

2.3.7 Oocyte Culture.

Sperm-injected oocytes were transferred into CZB medium, and placed in an incubator. About 5-7 h after injection, the presence or absence of pronuclei in the oocytes was determined using an inverted microscope. After confirmation of

pronuclei presence, the oocytes were cultured in KSOM medium until they reached the 2-cell stage (24–30 h after microinjection).

2.3.7.1 Embryo Transfer.

ICSI oocytes which had reached the 2 cell stage were transferred into the oviducts of 8- to 16-week-old surrogate pseudopregnant CD-1 females at 0.5 days post coitus (dpc) that had been mated with vasectomized males of the same strain on the day before embryo transfer. Pregnant females were allowed to deliver and raise their pups.

2.4 Results

2.4.1 Helper-Independent Insect pB Vectors (p*i*GENIE) are Active in HEK293T Cells and Support Transgenesis.

As a first step, a helper-independent plasmid was constructed. This plasmid contained a CMV-early-enhancer/chicken β-actin and β-globin intron (CAG) promoter-driven insect pB transposase gene (*i*pBt), where both the CAG promoter and the pBt gene are located on the backbone, outside the terminal repeat elements (TREs) of the transposon (p*i*GENIE). The transposon for p*i*GENIE was engineered to contain the enhanced green fluorescent protein gene as a reporter, as well as a cassette with Hygromycin resistance and Kanamycin resistance genes to facilitate selection in eukaryotes and prokaryotes, respectively (Figure 2-1A).

In order to assess the transposition activity in human cells, embryonic kidney cells (HEK293T) were transfected with 400ng of p*i*GENIE plasmids. The plasmid was administered in a circular form to cells by lipofection. To contrast the efficiency of p*i*GENIE-mediated transgene integration with non-transpositional (i.e. endogenous) integration of the plasmid, cells were also transfected with 400ng of a construct lacking a functional pBt gene (p*i*GENIE/Δ*piggyBac*). Resistant clones larger than 1mm were counted after 3 weeks of selection on Hygromycin (Figure 2-1B). These experiments demonstrated that p*i*GENIE was active in mammalian cells: 5.8% of all cells used for the experiment displayed stable integration as exemplified by EGFP expression and Hygromycin resistance. Furthermore, p*i*GENIE was able to mediate transposition at levels significantly above those of random, non-transpositional integrations (Table 2-1).

Genomic DNA (gDNA) from two clones derived from two individual, stably transfected single HEK293T cells was analyzed for the number of transgene integration sites by Southern blot with a transposon-specific probe for EGFP. The gDNA was digested with restriction enzymes *Sca*I and *Bam*HI. Only *Sca*I cuts once inside the transposon to leave a 4,447bp fragment containing the EGFP gene. *Bam*HI is expected to cut human gDNA about every 7,000bp. Depending on genomic environment of the transposon insertion site, each of the individual insertions facilitated by p*i*GENIE resulted in an individual band of at least 4,447bp in Southern blots. At the used concentration, p*i*GENIE proved to be extremely competent as both clones tested carried multiple transposons (Figure 2-1C).

Only a few laboratories have employed transposases as a tool for animal transgenesis (36, 63, 64, 69, 70). These experiments have been based on supplying the transposase either as protein, mRNA or, in the helper plasmid of the donor-helper system. In order to avoid difficulties with proteins or mRNA and since, in our hands, the helper-donor approach did not yield any transgenic animals when employing ICSI, we examined if p*i*GENIE could be used as an effective tool for transgenesis. ICSI transgenesis performed with the circular form of the plasmid and fresh sperm resulted in high rates of transgenic animals. For example using 0.663pg of p*i*GENIE resulted in 18 transgenic mice, representing 69.2% of all animals born and 22.8% of oocytes injected. Surprisingly, all of the transgenic animals exhibited mosaic EGFP signals by epifluorescence (Figure 2-1D). When F_0 animals were mated to wild-type mates of the same strain, all resulting transgenic F_1 pups displayed no mosaicism. All of the F_1 animals that survived to adulthood were sacrificed at two years of age and examination of their tissues showed no signs of tumor formation. The phenomenon of mosaicism in F_0 transgenic animals is thought to be due to transgene integration after the first chromosomal replication. However, it is also possible that non-transpositional integration of the plasmid as a whole or of the backbone only (post transposition) resulted in the integration of an active *i*pBt, which ultimately led to the observed mosaicism. From cell transfection and transgenesis studies with transposons, it is evident that, in virtually every instance reported, there is some non-transpositional integration of plasmids (36, 39, 53, 60, 64, 71, 72). Such non-transpositional integrations were shown to apply to both helper and donor plasmids (36). To distinguish backbone insertions from whole

plasmid insertions, several PCR assays were performed on gDNA from 10 founder mice, and the p*i*GENIE plasmid as a control, priming for several sites within the plasmid: (1) EGFP, (2) *i*pBt, (3) a region spanning the 3'-end of the CAG promoter to *i*pBt, and finally (4) a sequence spanning *i*pBt to the 3'-TRE (Figure 2-1E). While, as expected, all of the mice were positive for EGFP, this analysis also yielded amplification products in all of the samples for the *i*pBt gene and the promoter, indicating non-transpositional insertion of the plasmid. More specifically, as no amplicons were obtained for the region spanning the *i*pBt gene and the 3'-TRE, we assumed that all of these observed non-transpositional insertions were derived from backbone segments only. pBt has been shown to excise transposons without leaving a footprint in the donor genome (42). It is therefore probable that these non-transpositionally inserted plasmid backbones originated from the re-circularized plasmid after excision of the transposon and its TREs as previously reported (39). Since insertion of an active pBt with its promoter into the genome of a host may result in hopping of the transgene—a feature advantageous for gene function studies, but undesirable for gene therapy or transgenesis experiments—a new plasmid was engineered to ameliorate this issue.

2.4.2 p*m*GENIE-2: A Mouse Codon-Biased Self- Inactivating Vector.

A new plasmid designated p*m*GENIE-2 contained the same transposon as p*i*GENIE, but differed from it in two components: (1) a mouse-codon-optimized pBt gene proven more effective in cell transfection experiments (73) replaced the *i*pBt gene; and (2) the 3'-TRE was located between the CAG promoter and pBt gene (Figure 2-

2A). Therefore, enzymatic excision of the transposon from the plasmid during transposition would result in the separation of the CAG promoter from the *m*pBt gene. The promoterless *m*pBt gene, residing in the remaining plasmid backbone should therefore be inactive if inserted non-transpositionally into the genome of the host by the DNA repair mechanism, reducing the possibility of potential genotoxic effects.

HEK293T cells were transfected with 400ng of circular p*m*GENIE-2 plasmid or the respective control (p*m*GENIE-2/Δ*piggyBac*). To compare our helper-independent system to the more commonly used two-component donor-helper system, HEK293T cells were also transfected with 100ng of helper plasmid containing the *m*pBt gene driven by the CAG promoter (pCX-*m*pB) and 200ng of donor plasmid carrying the EGFP, Hygromycin resistance and Kanamycin resistance genes. Cells were maintained under Hygromycin selection for 3 weeks to exclude the possibility of detecting persistent episomal plasmid. To validate our approach of inactivating non-transpositionally inserted pBt genes, several PCR screens were performed. Genomic DNA, extracted from mixed populations of HEK293T cells transfected with p*m*GENIE-2 were used for these reactions. For comparison, gDNA from p*i*GENIE transfected cells was used. The respective plasmids (p*i*GENIE and p*m*GENIE-2) served as positive controls for successful amplification. The obtained amplicons were subsequently analyzed by sequencing. Amplification of a region spanning the 3'-end of the CAG promoter and the *i*pBt gene resulted in products from the gDNA as well as the p*i*GENIE plasmid control (Figure 2-2B) indicating some non-

transpositional insertion of the p*i*GENIE plasmid backbone as also demonstrated in Figure 2-1E. Long-range PCR using primers that anneal within the *m*pBt gene and downstream of the 5'-TRE result in 11.2 kb products from the p*m*GENIE-2 plasmid only but not from the gDNA of p*m*GENIE-2 transfected cells (Figure 2-2B), verifying the absence of persistent episomal plasmids in the sample. It further suggests the lack of non-transpositional insertion of full-length plasmid. In contrast, the same amplification reaction gave rise to a 1.0 kb PCR product from gDNA, an amplicon absent in the plasmid control reaction (Figure 2-2B), representing amplification from the promoterless backbone that remains after transposon excision. Taken together, these findings validate our strategy of inactivating *m*pBt in p*m*GENIE-2 by separating the gene from its promoter during transposition. Thus, non-transpositional insertion of the plasmid backbone can result only in the incorporation of a promoterless transposase gene. The importance of this feature of p*m*GENIE-2 is further emphasized by the detection of PCR amplicons for pBt from non-transpositionally inserted helper plasmid in donor-helper experiments (Figure 2-7).

As a next step, p*m*GENIE-2 was evaluated in regards to its transposition activity. Colony-forming assays showed that the donor-helper system, which demonstrated efficiencies similar to those of previous experiments (39), resulted in half as many stably transfected colonies when compared to the helper-independent approach (Figure 2-2C). Furthermore, p*m*GENIE-2 was able to mediate transfection rates at levels considerably above those resulting from p*i*GENIE-mediated integration (Table 2-1).

To assess the influence of plasmid concentration on the number of simultaneous transgene integrations, HEK 293T cells were transfected with different concentrations of p*m*GENIE-2 (Figure 2-2D). Genomic DNA was extracted from clonal expansions of single transfected HEK293T cells after they had reached confluence in a T75 flask at approximately four months. Southern blot analysis of this gDNA revealed that lower plasmid amounts result in fewer transgene insertions, with the highest number of insertions at 400ng. These preliminary data suggest that p*m*GENIE-2, in the highest amount tested, is a powerful tool for *in vitro* transposition and may be better for applications, such as mutagenesis, where large numbers of inserts are desirable.

Subsequently, ICSI was performed to investigate p*m*GENIE-2's ability to support animal transgenesis, using the same experimental setup as described for p*i*GENIE. Visual assessment of the animals by epifluorescence revealed that p*m*GENIE-2 is indeed amenable to *in vivo* experiments and, contrary to p*i*GENIE, results in both mosaic and full transgenic animals (Figure 2-2E). Similarly to p*i*GENIE ICSI experiments, all transgenic F_1 animals obtained from mating F_0 to wild-type animals were true transgenics and survived to two years of age without any detectable tumorigenesis.

To assess transposon integration frequency for these transgenesis experiments, gDNA obtained from five founder mice (F_0) was analyzed by Southern blots with an EGFP-specific probe. All of the mice displayed only a few transgene insertions, with

one mouse having seven insertions while two other founders only showed the presence of one transgene (Figure 2-2F). In comparison to cell transfections, it appears that pmGENIE-2 is less efficient in inserting transposons into oocytes. However, the significant differences of the two experimental systems do not allow for such a conclusion. Transgenic founders (F_0) were mated with wild-type B6D2F1 mice to evaluate transposon transmission through the germline, again by subjecting gDNA from these F_1 animals to Southern blot analysis. Two animals (F_1-6 and F_1-8) inherited one out of two transgenes, whereas the other animals inherited all copies present in the respective founders. Moreover, a comparison of the banding patterns for F_0 and F_1 animals indicated that the transgenes were not subject to relocation by local hopping, demonstrating the validity of our approach for inactivating the transposase gene.

The Southern blot analyses for both p*m*GENIE-2 HEK293T cell transfection and ICSI-mediated mouse transgenesis indicated the presence of bands smaller than 5kb. However, *Hin*dIII restriction digestion of the gDNA used in these experiments should result in fragments of at least 5004bp or larger (Figure 2-2). We assumed that these bands originate from non-transpositional insertions of plasmid fragments that contain the EGFP transgene.

Genomic location of transposon integrations were then characterized in several individual HEK293T clones and F_1 transgenic mice by either inverse or vectorette PCR. A summary of the identified insertion sites are listed in Table 2-2. Only high-

quality sequences that showed a transposon terminal repeat element immediately adjacent to the pBt target site (TTAA) and directly followed by a genomic sequence were used. Our results were similar to other studies describing insertion patterns for pBt (59, 74). Approximately half of cell line integrations were found in intergenic regions. However, we also identified three intronic locations and one insertion into an exon of the cadherin-10 gene. In contrary, all of the transposon-gDNA junctions in the transgenic animals could be assigned to intergenic regions. While it is possible that pBt displays specific integration patterns depending on cell type, species or experimental setup the limited number of integration sites analyzed precludes such a conclusion.

Additional evidence for the functional competence of the plasmid was obtained by antibody studies with oocytes injected with p*m*GENIE-2. The time course demonstrated pBt expression as early as six hours after sperm injection with peak production at 30 hours (Figure 2-3). At this time point the transgenic embryos were at the two-cell stage and showed early EGFP expression, as previously reported (67). At 48h, the embryos were at the four-cell stage: the signal for the pBt decreased, presumably because of inactivation of the pBt gene during transposition. EGFP expression however, remained constant because of insertion of the transposon into the oocyte genome.

2.4.3 p*m*GENIE-3: A Helper-Independent piggyBac Plasmid Optimized For Gene Therapy Experiments.

Evidence of plasmid backbone insertions with p*i*GENIE (Figure 2-1E) made it necessary to modify the plasmid architecture to avoid the insertion of an active pBt gene into host genomes. Initially, we had re-engineered the plasmid so that the CAG promoter driving the pB gene resides on the transposon, while the gene itself is located on the plasmid backbone (p*m*GENIE-2). Thus the CAG promoter and pBt are separated during transposition, rendering the gene inactive. Moreover, should pBt be inserted into the genome by non-transpositional integration of the backbone, it would remain inactive unless the activity of a promoter adjacent to the site of insertion would drive its transcription in the correct orientation. However, as the CAG promoter is part of the inserted transposon, it is possible that it could potentially influence the transcription of genes downstream of the insertion site.

In the new construct p*m*GENIE-3, the 3'-TRE of the transposon resides within an intron engineered into the pBt gene near its 5'-end (Figure 2-4A). Such a design allows transcription of pBt only from pre-transposition plasmids but not from re-circularized plasmid backbones that remain after the transposition event. During transposition, pBt binds to the 3'-TRE inside the chimeric intron as well as to the 5'-TRE, forming the synaptic complex. Hence, only a truncated, inactive pBt gene would remain on the plasmid backbone after excision. This truncated pBt cannot contribute to any hopping of inserted transposons, even if integrated into the host genome by non-transpositional insertion and driven by a promoter from an adjoining

gene. Additionally, the CAG promoter left in the transposon cannot contribute to aberrant gene activation, as the small 5'-end fragment of the pBt gene and the remaining intron terminate at a UAA stop codon engineered into the construct. This design should, in effect, mitigate any influence on downstream host DNA genes.

As a first step, we evaluated the presence of an integrated pBt backbone by genomic PCR of p*m*GENIE-3 transfected HEK293T cells following methods described for p*i*GENIE transfections. Sequence analysis of the obtained amplicon indicated that the insertions derived from re-circularized pmGENIE-3 plasmids, as they contained reconstituted TTAA sites generated by transposon excision and consecutive re-circularization of the backbone plasmid. The data also confirmed that non-transpositional insertion of the p*m*GENIE-3 backbone leaves only an inactive pBt behind (Figure 2-4A and 2-5). Again, we found no evidence of full-length p*m*GENIE-3, indicating the absence of episomal pre-transpositional plasmid. More importantly, the data also indicates the lack of non-transpositional insertion of full-length plasmid.

We investigated if the new design of the plasmid would negatively influence its transposition efficiency in *in vitro* and mouse transgenesis experiments. Our results from the colony-forming assay demonstrated that p*m*GENIE-3 displayed the highest gene-delivery efficiency as compared to the less intricate p*i*GENIE and p*m*GENIE-2 plasmids (Figure 2-4B, Figure 2-6 and Table 2-1). A Southern blot analysis of HEK293T clones derived from transfected, single HEK293T cells showed a similar number of transgene insertions as observed for the other two plasmids, indicating

that the new plasmid retained transposition competence (Figure 2-4C). Insertion site data obtained for p*m*GENIE-3 provided evidence that the modified pBt gene retains its enzymatic transposition activity (Table 2-2).

ICSI transgenesis experiments were performed to assess the potential of p*m*GENIE-3 as a tool for generating transgenic animals. We introduced a modification to the pre-treatment of spermatozoa we had used for p*m*GENIE-2 transgenesis. Instead of using fresh sperm, spermatozoa were incubated with 1mM NaOH to remove their plasma membrane. This procedure has been reported to improve transgenesis efficiency by reducing chromosomal damage associated with freeze-thaw treatment of spermatozoa, ultimately resulting in higher offspring rates after embryo transfer (75). In these ICSI experiments, at the most efficient concentration of 0.7ng/µl of p*m*GENIE-3 in the circular form, 11 transgenic mice obtained, five of which were true transgenics. Similarly to p*i*GENIE and p*m*GENIE-2 ICSI experiments, all p*m*GENIE-3 transgenic F_1 animals obtained by mating the F0 animals to wild-type mates were full transgenics and are surviving to date without any detectable tumorigenesis.

2.5 Discussion

In this study, we developed novel helper-independent pB gene delivery plasmids. These plasmids are more effective for HEK293T cell transfections and ICSI transgenesis than the two-plasmid donor-helper approach. In p*i*GENIE mediated transgenesis, all of the transgenic animals were mosaic. A PCR analysis of gDNA from p*i*GENIE mice revealed random, non-transpositional insertion of the plasmid

backbone, presumably in the linearized form. pBt-mediated transposition from chromosomal DNA does not alter the donor site but instead reconstitutes the original TTAA tetranucleotide after excision of the transposon (42). In plasmids post transposition, reconstitution of the TTAA site of the remaining linear backbones should result in their re-circularization. This mechanism applies to transposon excision from plasmids in systems other than pB as well (76, 77). Re-circularized plasmids remain in the nucleus and are thought to be degraded over time. However, our data suggests that some of these plasmids are inserted into the host genome, possibly by the host DNA repair mechanism, as reported elsewhere (78). As the p*i*GENIE plasmid backbone contains an active pB gene, such insertions may result in pBt transcription in the host cell and can thus cause transgene relocation. This may contribute to the observed transgenic mosaicism and potentially to deleterious genotoxic modifications, resulting in decreased fitness of the organism.

As an alternative approach, transpositions can be catalyzed by supplying the transposase as mRNA (63). mRNA instability restricts its duration of activity and therefore the risk of introducing genotoxic effects. However, it is possible for the mRNA to undergo reverse transcription, potentially resulting in insertion of the pBt cDNA into the host genome by non-homologous recombination (79). We modified our helper-independent p*i*GENIE plasmid in two steps, thereby significantly improving its safety. The latest plasmid, p*m*GENIE-3 features an intron spanning the 3'-TRE that allows the transposase to be transcribed from unmodified plasmids only. After transposition, the remaining re-circularized plasmid backbone contains a

truncated transposase gene. Even if the re-circularized plasmid is then integrated into the host genome near an active promoter, the non-functional pBt gene cannot be expressed due to lack of an initiation codon. Additionally, after transposition, the p*m*GENIE-3 transposon retains the CAG promoter, which now can only drive the expression of a non-functional short transcript with a pre-engineered stop codon. As a consequence, the transposed promoter cannot activate neighboring host genes upon integration. These safety features integrated into p*m*GENIE-3 are of great importance for any gene therapy approach using transposase-based systems, as we and others (36, 39, 53, 60, 64) have reported a certain amount of non-transpositional plasmid integration when using pB or SB based transposition, independent of the specific system or application. The lack of full-length GENIE plasmid insertions in all of our transfection experiments, further emphasizes the importance of self-inactivating plasmids for gene delivery experiments.

To further reduce the risk of unwanted, potentially oncogenic modifications of the host genome, additional elements can be integrated into the vector design. For example, the use of a chimeric transposase consisting of the transposase itself and a DNA binding domain such as zinc finger motifs, could potentially target the transposon integration to specific DNA regions. pBt was previously shown to be amenable to such molecular modification while retaining its full transpositional activity, whereas other transposases lost their activity (39, 80). Furthermore, Chen *et al*. (81) reported using plasmids modified to carry chicken-beta-globin insulators flanking the transgene in their pBt-mediated transposition system in human

embryonic stem cells. These insulators have the potential not only to shield transcribed regions of the host genome from outside regulatory influences but also to act as barriers against position-dependent transgene silencing.

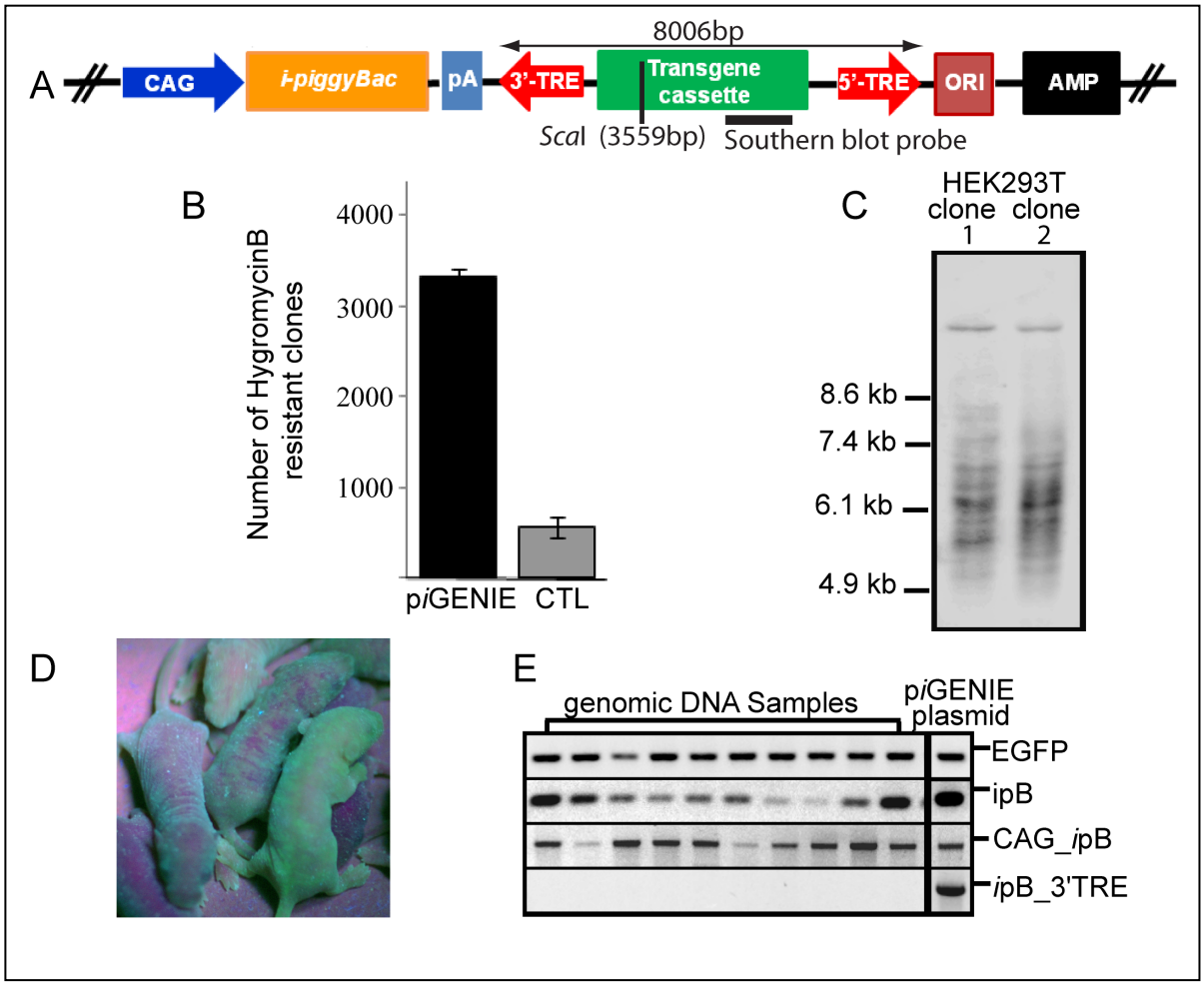
In our initial experiments with HEK293T cell transfection, a large number of transposition events per transfection were observed at the highest plasmid concentration. This result suggests that our CAG promoter-driven p*m*GENIE plasmid is a powerful tool for *in vitro* transposition. However, it may be better suited for mutagenesis experiments or other applications where large numbers of inserts are desirable. For example, a high number of transgene insertions may be of advantage during tumor reduction approaches where plasmids are used to introduce suicide genes. Here, multiple insertions of transgenes can potentially amplify the toxic effect by increasing expression of the suicide gene. Transfections with lower amounts of plasmids yielded cells with fewer transgene integrations. Hence, for transfection experiments requiring fewer inserts, using lower plasmid concentrations or modifying the plasmids to contain a less potent pBt promoter may prove beneficial.

Southern blots analyses of gDNA from transgenic animals did not indicate any concatemerized insertion of the transgenes. Such concatemerized insertions are commonly reported for transgenesis performed with pronuclear microinjections and ICSI. This phenomenon is believed to be due to ligation of the linear transgenes in a head to tail orientation and subsequent integration into the host genome by homologous recombination (36, 78, 82-84). Transgene insertions mediated by the

pBt seem to avoid such concatemers. During pB transposition a single transposon is excised from the plasmid to form a synaptic complex that appears to prevent the ligation of the transgenes into concatemers. Therefore, the cut and paste mechanism of transposases appears to ensure that only individual transgenes excised from the plasmid participate in transposition.

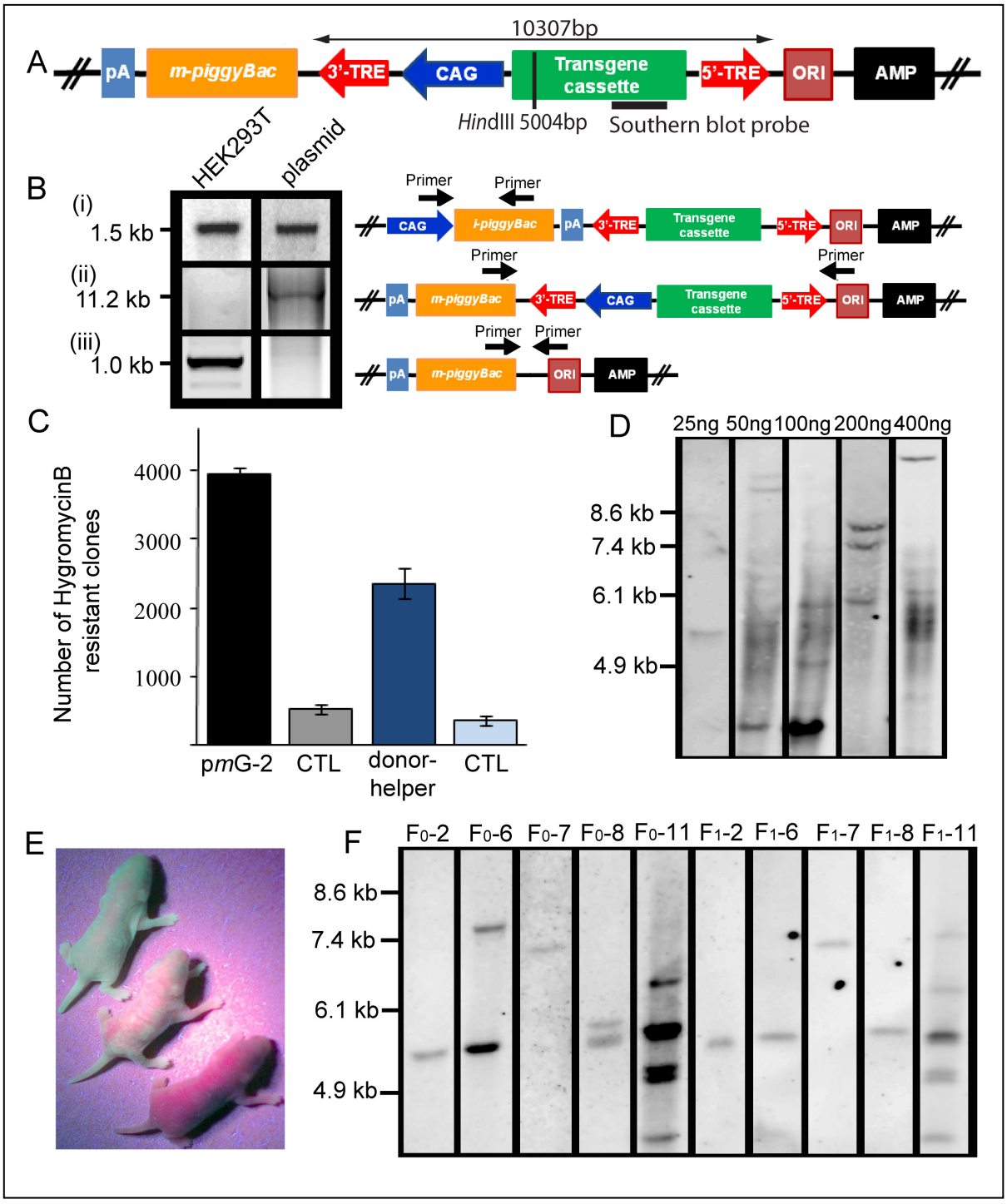
We have noted non-transpositional insertions of single plasmid constructs as well as donor and helper plasmids during our transfection and transgenesis experiments. These observations are shared by many others (36, 39, 53, 60, 64, 71, 72); however, there are reports demonstrating a complete removal of the plasmids by degradation, post transposition (85). We are presently unable to consolidate these contradictory findings; however, it is possible that the extend of this random, non-transpositional insertion of plasmids is cell and tissue specific or dependent on plasmid architecture and methods used in plasmid preparation. In *in vitro* experiments, multi-transgene insertion or local hopping of the transposon may be a lesser concern. Here, such events may potentially be genotoxic, but can result only in an increased number of dead cells, thereby reducing transfection efficiencies. In gene therapy approaches, avoiding deleterious events introduced by the plasmid or plasmid fragments is critical. Therefore, the safety features incorporated into p*m*GENIE-3 represent an important step towards using transposon-based systems for gene therapy where reporter genes will be replaced by therapeutic genes. The architecture described herein should also be applicable to transposase systems such as SB or *Tol2*, improving their safety in gene therapy experiments.

2.6 Figures



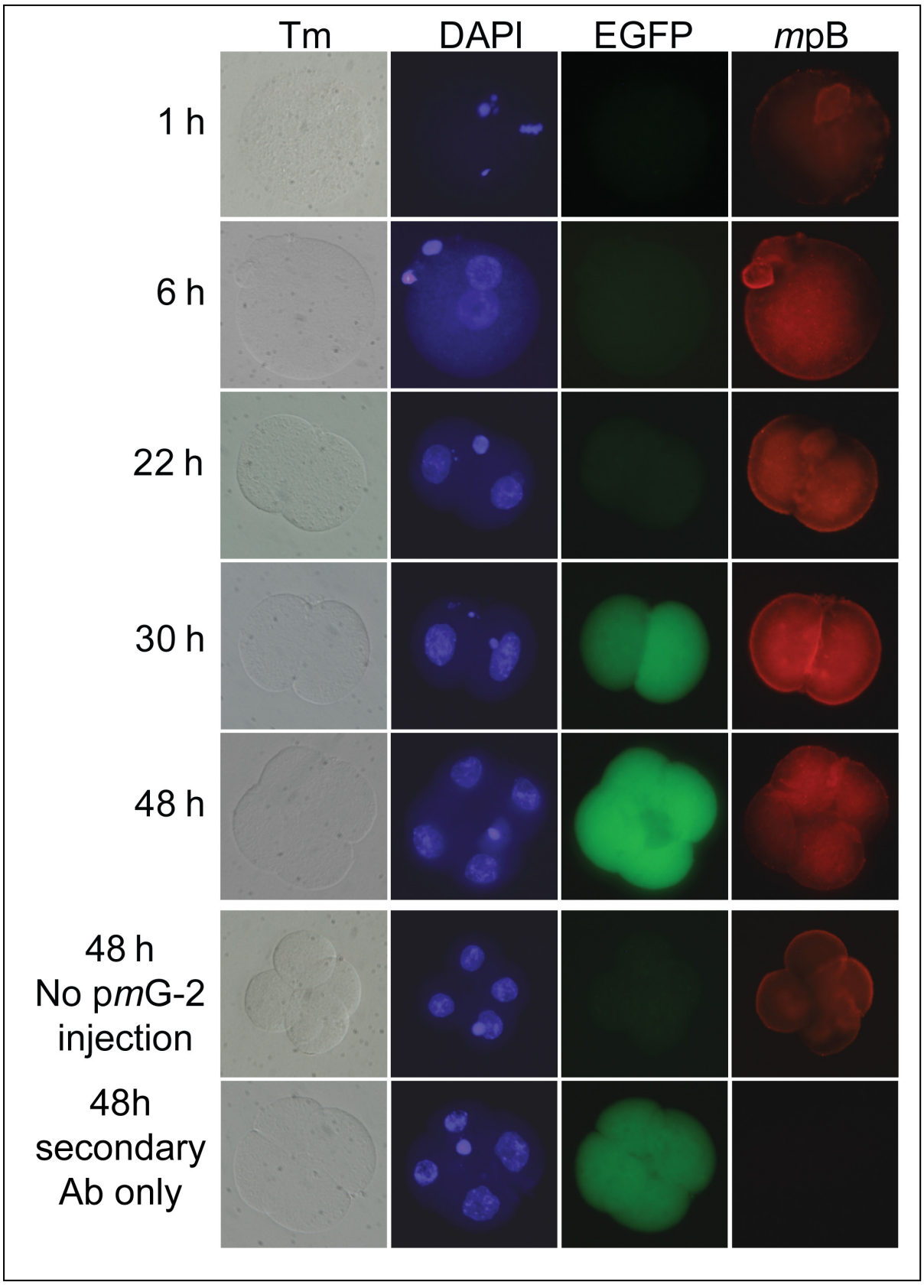
2.6.1 Figure 2-1. p*i*GENIE supports cell transfections and animal transgenesis. (A) Schematic representation of p*i*GENIE. The transposon cassette is delimited by the 3'- and 5'-TREs. Transposon size, restriction sites within the transposon and Southern blot probe location are indicated. (B) $0.5x10^5$ HEK293Tcells were transfected with 400 ng of p*i*GENIE plasmid. As control for random non-transpositional integration of the plasmid HEK293T cells were transfected with 400ng of a construct (p*i*GENIE/Δ*piggyBac*) lacking a functional pBt gene. Transposition activity was measured by counting methylene blue stained, hygromycin-resistant colonies after a three-week selection period. Data are shown

as mean values with S.D. (N=3). (C) Southern blot analysis of gDNA from a clonal expansion of transfected cells to analyze insertion events. Samples were digested with *Sca*I and *Bam*HI and probed with a DIG-labeled DNA fragment corresponding to the EGFP gene. (D) Depiction of mosaic p*i*GENIE transgenic mice. (E) Assessment of plasmid backbone insertions into the genome of ten p*i*GENIE mice by PCR analysis (p*i*GENIE plasmid served as a control). The analysis revealed that while all animals are transgenic for EGFP (top panel), they also displayed *i*pB gene insertions (second panel from the top). Moreover, amplification products for the *i*pB gene with its CAG promoter were obtained (second panel from the bottom) but not from a region spanning the *i*pB gene and the 3'-TRE (bottom panel), indicating non-transpositional insertion of the plasmid backbone only but not of the entire plasmid.

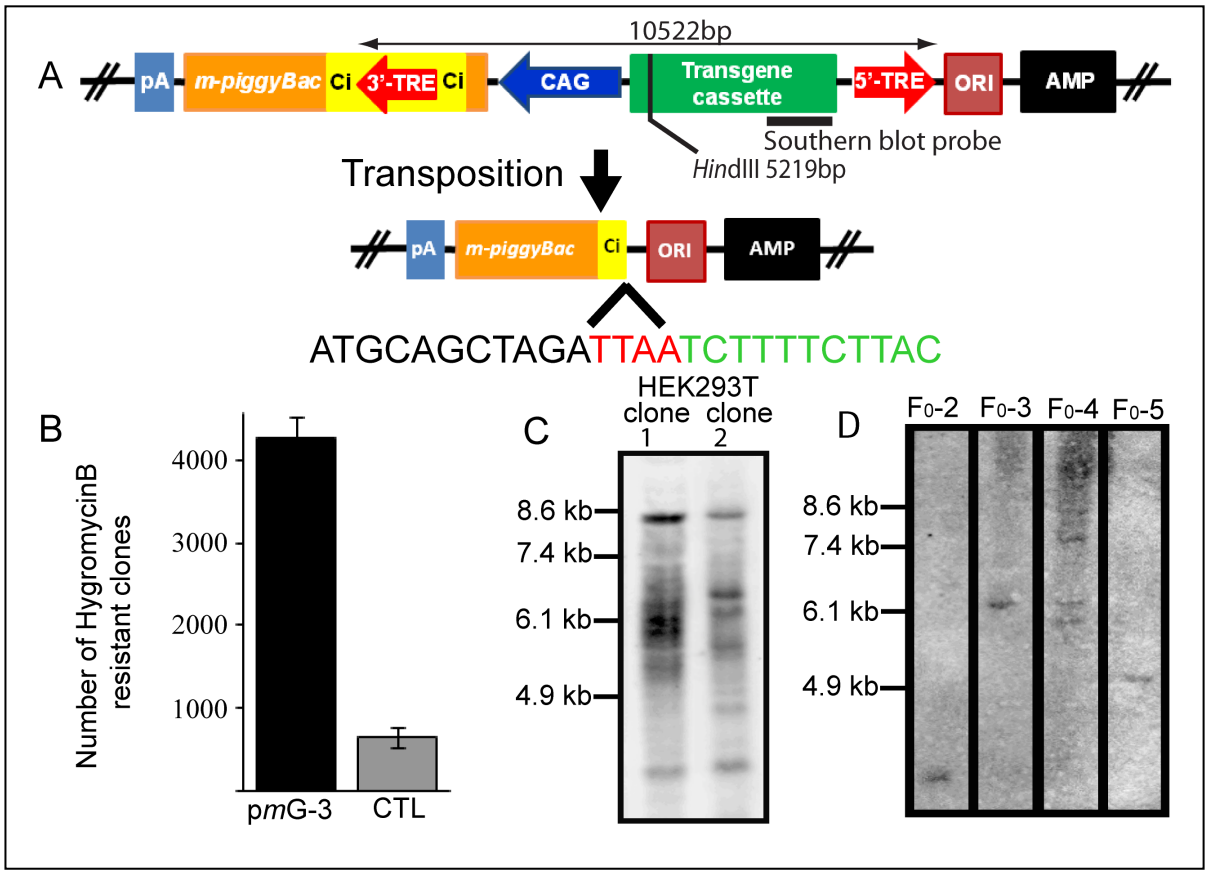


2.6.2 Figure 2-2. Analysis of p*m*GENIE-2. (A) Schematic representation of p*m*GENIE-2. The transposon cassette is delimited by the 3'- and 5'-TREs. Transposon size, restriction sites within the transposon and Southern blot probe

location are indicated. (B) The *m*pBt is separated from its CAG promoter during transposition: $0.5x10^5$ HEK293Tcells were transfected with 400 ng of p*m*GENIE-2 plasmid. As a control, HEK293T cells were transfected with 400ng p*m*GENIE-2/Δ*piggyBac* lacking a functional pBt gene. Genomic DNA, extracted from a mixed population of HEK293T cells transfected with p*i*GENIE or p*m*GENIE-2 were used for PCR analysis, where p*i*GENIE or p*m*GENIE-2 plasmid DNA served as control. (i) Amplification of a region spanning the 3'-end of the CAG promoter and the *i*pB gene resulted (schematic depiction of primers are indicated as arrows) in products from the gDNA and from the p*i*GENIE plasmid control. (ii) Long-range PCR of the complete transposon, priming within the *m*pBt gene and downstream of the 5'-TRE resulted in full length 11.2 kb products from the p*m*GENIE-2 plasmid control but not from the gDNA of p*m*GENIE-2 transfected cells. (iii) In contrast, the same amplification reaction gave rise to a smaller 1.0 kb amplicon from gDNA, but not from the plasmid. (C) Transposition activities of p*m*GENIE-2 (400ng) and the helper-donor (200ng donor and 100ng helper) plasmids (and equal amounts of their respective controls) as measured by counting hygromycin-resistant colonies. Data are shown as mean values with S.D. (N=3). (D) Representative samples of a Southern blot analysis of HEK293T cells transfected with different amounts of p*m*GENIE-2 (E) Depiction of full transgenic, mosaic and non-transgenic p*m*GENIE-2 transgenic littermates. (F) Southern blot analysis of gDNA from p*m*GENIE-2 transgenic founder and F_1 mice.


Tm
DAPI
EGFP
mpB
1 h
6 h
22 h
30 h
48 h
48 h
No pmG-2
injection
48h
secondary
Ab only

2.6.3 Figure 2-3. Time course of pB and EGFP protein expression in transgenic embryos generated by ICSI. pBt expression was visualized by immunofluorescence using a polyclonal antibody for the pBt protein and is detectable above background levels as early as 6 hours. Peak expression was observed at about 30h corresponding with early detection of EGFP transgene expression. Pictures in the Tm column represent transmission images of the developing embryos. DAPI was used to visualize nuclei.



2.6.4 Figure 2-4. Analysis p*m*GENIE-3 transposition. (A) Schematic representation of p*m*GENIE-3 and the truncation of the pBt gene during transposition. The transposon cassette for genomic integration is delimited by the 3'- and 5'-TREs. Ci represents the chimeric intron. Transposon size, restriction sites within the transposon and Southern blot probe location are indicated. (B) $0.5x10^5$ HEK293T

cells were transfected with 400ng of p*m*GENIE-3 or p*m*GENIE-3/Δ*piggyBac*. Transposition activity was measured by counting methylene blue stained, hygromycin-resistant colonies after a three-week selection period. Data are shown as mean values with S.D. (N=3). (C) Southern blot analysis of *Hin*dIII digested gDNA from clonal expansions of p*m*GENIE-3 transfected cells. Samples were probed with a DIG-labeled DNA fragment corresponding to the EGFP gene. (D) Southern blot analysis of *Hin*dIII digested gDNA from p*m*GENIE-3 transgenic mice.

Table 2-1 **Comparison of transposition efficiencies of p*i*GENIE, p*m*GENIE-2 and p*m*GENIE-3**

	p*i*GENIE	p*m*GENIE-2	p*m*GENIE-3
Relative fold *	5.8	5.8	7
Percentage of transposition †	5.5	6.7	7.6

We utilized two parameters, relative fold and percentage of transposition, to assess the transposition activity of the different plasmids.
*The number indicates the relative fold of HygromycinB resistant HEK293T done as compared to random insertions ($N - 3$).
†The number indicates the percentage of true transpositions from 0.5×10^5 cells seeded.

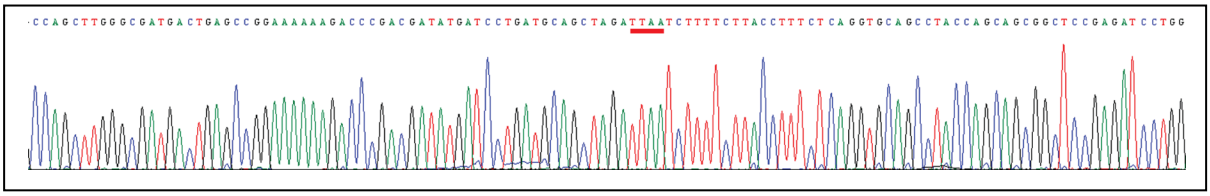
2.6.5 Table 2-1. Comparison of transposition efficiencies of p*i*GENIE, p*m*GENIE-2 and p*m*GENIE-3. We utilized two parameters, relative fold and percentage of transposition, to assess the transposition activity of the different plasmids.

Table 2-2 **Transposon integration sites identified in HEK293T cells and transgenic mice**

pmGENIE-2 Cells/Tissue	Transposition site sequence	Chr.	Location	gDNA origin
HEK 293T	TTAAGAGAAAAGACCAGAGTCTCAACATGGAATA	1	Intergenic	polyclonal
HEK 293T	TTAATACAAAATAATACTAATAAGAAAAGATGCA	15	Intergenic	polyclonal
HEK 293T	TTAAAGACAACACAAACGAAGCCGAAGACAGAGG	6	Intergenic	polyclonal
HEK 293T	TTAATGAAAGGCTAAAAGAGCATGATCTTGACCC	5	CDH10 (exon)	polyclonal
HEK 293T	TTAATCAGGATAGGGTGTACACATAACTTTTAGG	6	TFAP2A (intron)	polyclonal
HEK 293T	TTAAAAGTCAGGGGAGAGAGCAAGCAAGGCAGGG	12	Intergenic	polyclonal
HEK 293T	TTAAGACATGGTTTCTAGTCTTAAGAAATTGTTC	7	Intergenic	single clone
HEK 293T	TTAAGTGATCTGCCTGCCTCGGCCAGGGAAGAAT	8	XKR5 (intron)	single clone
HEK 293T	TTAAGCACTTCTGACACGAGGCAGGTCTTTTTTC	7	AUTS2 (intron)	single clone
F1 Mouse 7M	TTAAGGAAGGATTTATTGGCTCAGGGTTTGAGGG	11	Intergenic	single clone
F1 Mouse 6F	TTAAAACTAAAAAATGACGATGATAATAACGATA	1	Intergenic	single clone
F1 Mouse 11M	TTAAACACCAGCCATACTAAAGGACAATGCTAGG	14	Intergenic	single clone
F1 Mouse 11M	TTAAAAAGATCAGCCCTCAGTTTCAAGGTCAGCT	14	Intergenic	single clone
pmGENIE-3 Cells/Tissue	Transposition site sequence	Chr.	Location	gDNA origin
HEK 293T	TTAATGACTTTCCTCTATCCCAACACCCTGATCA	3	Intergenic	polyclonal
HEK 293T	TTAAGACTTGTTTTCTCACCTAACACATGGTCTA	1	Intergenic	polyclonal

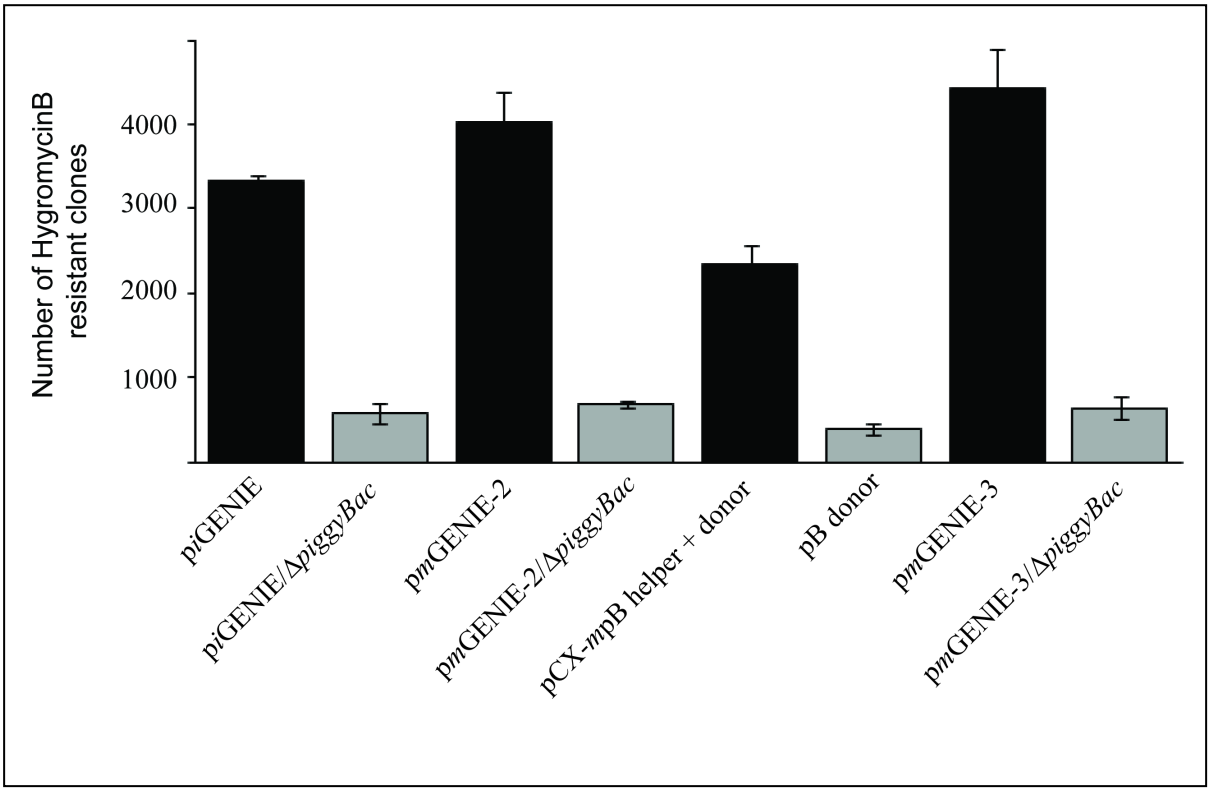
Consensus transposition sites for the *piggyBac* transposase are bolded. Inverse or vectorette PCR was performed on gDNA from p*m*GENIE-transfected cells or tail biopsies from transgenic mice as described in *SI Materials and Methods*. The pBt consensus integration site at the tetranucleotide TTAA is bolded. Chromosomal locations are noted in column 3 (Chr). Integration into *RefSeq* genes are annotated in column 4 (Location).

2.6.6 Table 2-2. Transposon integration sites identified in HEK293T cells and transgenic mice. Inverse or vectorette PCR was performed on gDNA from p*m*GENIE-transfected cells or tail biopsies from transgenic mice as described in Materials and Methods. The pBt consensus integration site at the tetranucleotide TTAA is highlighted in red. Chromosomal locations are noted in column 3 (Chr). Integration into *RefSeq* genes are annotated in column 4 (Location).

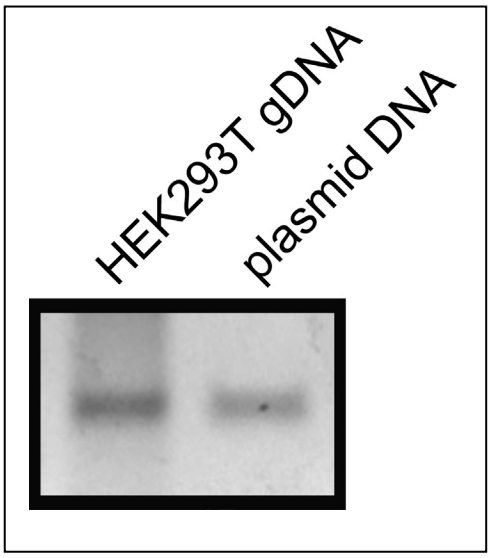


2.6.7 Figure 2-5. Evidence for pmGENIE-3 backbone insertions into the host genome. We evaluated the presence of integrated pBt backbone by genomic PCR of p*m*GENIE-3 transfected HEK293T cells. Sequence analysis of the obtained amplicon indicate that the insertions derived from re-circularized pmGENIE-3

plasmids, as they contained reconstituted TTAA sites generated (indicated by red bar) by transposon excision and consecutive re-circularization of the backbone plasmid. The data also confirmed that non-transpositional insertion of the p*m*GENIE-3 backbone (see Figure 2-4A) leaves only an inactive pBt behind.



2.6.8 Figure 2-6. Comparison of plasmid transfection efficiencies in HEK293T cells.



2.6.9 Figure 2-7. Evidence for pCX-*m*pB helper plasmid insertions into the host genome. Amplicons were obtained by amplification of the *m*pBt gene from HEK293T cells gDNA, transfected with the helper-donor system. The pCX-*m*pB helper plasmid served as a control for successful amplification.

2.7 Contributions

Research design: Stefan Moisyadi, Johann Urschitz, Jesse Owens

Transgenesis experiments: Miyuri Kawasumi, Joel Marh, Hideaki Yamashiro

Tissue culture: Ilko Stoytchev

Conducted experiments: Johann Urschitz, Miyuri Kawasumi, Jesse Owens, Hideaki Yamashiro, Ilko Stoytchev, Joel Marh

Wrote or contributed to writing: Johann Urschitz, Stefan Moisyadi, Jesse Owens

Chapter 3. Effective Targeted Gene Knockdown in Mammalian Cells Using the *piggyBac* Transposase-based Delivery System

3.1 Abstract

Non-viral gene delivery systems are rapidly becoming a desirable and applicable method to over-express genes in various types of cells. We have recently developed a pB transposase-based, helper independent and self inactivating delivery system (p*m*GENIE-3) capable of high efficiency transfection of mammalian cells including human cells. In the following study, we have assessed the potential of this delivery system to drive the expression of short hairpin RNAs (shRNA) to knock down genes in human cells. Two independent p*m*GENIE-3 vectors were developed to specifically target knock-down of an endogenous gene, telomerase reverse transcriptase (TERT), in telomerase positive human immortalized cell lines.

Compared to a transposase-deficient vector, p*m*GENIE-3 showed significantly improved short term transfection efficiency (~4-fold enhancement, 48 hour post-transfection) and long term integration efficiency (~5-fold enhancement) following antibiotic selection. We detected a significant reduction of both TERT expression and telomerase activity in both HEK293 and MCF-7 breast carcinoma cells transfected with two p*m*GENIE-3 construct targeting distinct regions of TERT. Importantly, this knock down of expression was sufficient to abrogate telomerase

function since telomeres were significantly shortened (3-4 Kb, P<0.001) in both TERT-targeted cell lines following antibiotic selection of stable integrants. Together these data show the potential of the pB non-viral delivery system to stably knock down gene expression in mammalian cells, and indicate the potential to develop novel tumor targeting therapies.

3.2 Introduction

Transposon based systems are simple and efficient transfection tools suitable for a variety of gene transfer applications (86). These plasmid-based gene delivery vehicles represent alternatives to popular integrating viral approaches. Advantages over viral vectors include tolerance by the immune system, decreased preference for integration into genes, and increased cargo capacity (1-3, 50, 87). Additionally, viral approaches are burdened by expense in production and biosafety considerations (52). Initial transposition experiments in vertebrate cells used the SB system. The SB transposon, originally reconstructed from inactivated sequences found in fish genomes, has proven to be a flexible and effective non-viral genetic tool (34, 88, 89). Recently, pB, isolated from the moth *Trichoplusia Ni*, has emerged as a highly efficient gene delivery vector for numerous *in vitro* and *in vivo* applications (36, 37, 39, 40, 67, 90). The typical 2-plasmid system includes both a "helper" plasmid encoding the transposase enzyme and a "donor" plasmid containing the intended integration sequence, such as a gene of interest, flanked by the transposon terminal repeat elements (TREs). Upon entrance to the cell, the pB transposase recognizes the TREs and excises the transposon from the donor plasmid. pB subsequently

integrates the transposon permanently into the genome at a TTAA tetranucleotide sequence. By introducing pB TREs into bacterial artificial chromosomes (BACs) it is possible to facilitate integration of large pB inserts greater than 100 kb (41). pB's unprecedented cargo capacity can allow for the integration of large genes, regulatory elements, and multiple reading frames. pB has been used to generate induced pluripotent stem cells (iPSCs) due to its ability to precisely excise its transposon from the genome without leaving a DNA "footprint" (62). This feature allows for the removal of transgenes following complete reprogramming without leaving residual sequences (38). Additionally, pB has been shown to be amenable to DNA binding domain fusions, allowing for targeted transposition to chromosomal locations (91).

We have recently designed improved self-inactivating vectors that contain all transpositional machinery on a single helper-independent plasmid, termed p*m*GENIE-3 (40). p*m*GENIE-3 encodes a specialized pB transposase containing a TRE within an engineered intron. Upon transfection, pB recognizes the TREs and subsequently excises the transposon from the plasmid, thereby truncating the transposase gene. This design renders the transposase inactive after excision. Consequentially, potentially negative genotoxic effects that may develop by the persistence of an active pB gene are eliminated.

Over the past 20 years, RNA interference, using silencing RNA molecules or short hairpin RNA molecules (shRNA) has developed into a reliable method to target and

reduce the expression of specific genes. Eukaryotic expression vectors have been shown to effectively drive expression of shRNA in mammalian cells, however the relative ability of transposase-based delivery systems, in particular the pB transposase systems, have not yet been thoroughly assessed. Telomerase is an enzymatic complex that is required for prolonged cancer cell proliferation and is present in ~90% of all human tumors (92). The activity of the telomerase complex is entirely dependent on expression of the catalytic component of telomerase, telomerase reverse transcriptase (TERT) (93, 94). Therefore, in the present study, we sought to assess the potential of a novel pB-based vector to express shRNA to specifically target TERT, in human telomerase positive cell lines.

3.3 Experimental Procedures

3.3.1 Vector Design

The p*m*GENIE-3 transposase-based delivery system has been described previously (95). p*m*GENIE-3 is compatible with Gateway recombineering cloning (Life Technologies) for simple addition of target genes into the transposon portion of the construct. The two shRNA sequences were modeled after human miR-30 and were designed to target distinct regions of the TERT mRNA, specifically, 5'-AGCAAGTTGCAAAGCAT -3' (Tert1 shRNA) and 5'-CGAGCTGCTCAGGTCTTTCTT -3' (Tert2 shRNA). The shRNAs were located 3' to a bicistronic TurboGFP and puromycin coding region driven by the cytomegalovirus (CMV) promoter. Each cassette including shRNA was excised from lentiviral pGIPZ shRNAmir plasmids (catalog #s RHS4430-101132965 and RHS4430-99161517,

Thermo Scientific) via XbaI and BglII (New England Biolabs) and cloned into an empty pENTR1A (Life Technologies) vector in preparation for Gateway recombineering into p*m*GENIE-3.

3.3.2 Cell lines and transfection

The HEK293 and MCF-7 cell lines were grown in standard culture conditions (37^{o}C, 5% CO_2) in DMEM for HEK293 and DMEM F12 for MCF-7, both in high glucose, complete media with 10% FBS. Cells were transfected using FuGENE 6, (Roche) per manufacturers' instructions. Cells were selected with Puromycin (HEK293 cells with 500 ng/ml, MCF-7 cells with 300ng/ml).

3.3.3 Measurement of TERT expression

Relative levels of TERT mRNA were assessed by real time RT-PCR using a Bio-Rad IQ cycler. Total RNA was isolated from cells using TRIzol reagent (Invitrogen). RNA cleanup and DNase treatment was done using the RNeasy Mini Kit (Qiagen). One microgram of total RNA was reverse transcribed using the iScript Advanced cDNA Synthesis Kit for RT-qPCR (Bio-Rad). The real-time PCR assays were performed in 25 µL reactions using iQ SYBR Green Supermix (Bio-Rad), 50 ng cDNA and 0.4 µM of forward and reverse primers. Analysis was done using a MyiQ Single-Color Real-Time PCR Detection System (Bio-Rad). Thermal cycling conditions were: 2 min at 50°C and 10 min at 95°C for initial denaturation, followed by 40 cycles at 95°C for 15 s and 54°C for 1 min, and a dissociation step at 95°C for 15 s, 54°C for 30 s and 95°C for 15 s. Gene relative quantification (RQ) of hTERT

was calculated as described by Livak and Schmittgen (31) and normalized with the housekeeping gene hypoxanthine phosphoribosyltransferase (HPRT). All samples were analyzed in triplicate along with no RT controls. Primers used for amplification were as follows:

hTERT F, 5'-CGTCGAGCTGCTCAGGTCTT-3' and R, 5'-AGTGCTGTCTGATTCCAATGCTT-3' (32): hHPRT F, 5'-TGACACTGGCAAAACAATGCA-3' and R, 5'-GGTCCTTTTCACCAGCAAGCT-3' (33).

3.3.4 Detection of Telomerase

Telomerase activity was measured using the TRAP assay (TRAPeze detection kit; Millipore) per manufacturers' protocol. Briefly, 1 x 10^6 cells per sample were lysed on ice in CHAPS buffer (30 min) followed by centrifugation (12,000g, 30 min). The supernatant was then used directly in TRAP reactions (PCR cycling conditions: 27 cycles of 94^{o}C/30 s, 60^{o}C/30 s, 72^{o}C/60 s).

3.3.5 Telomere length analysis

Telomere length was assessed by Southern analysis of terminal restriction fragment (TRF) length as previously described (26, 34).

3.3.10 Statistical Analysis

All measurements are shown as the mean of 3 or more independent measurements. Error bars represent standard error. P-values represent Student's t Test.

3.4 Results

We have previously shown that the p*m*GENIE-3 transposase-based non-viral delivery system can efficiently express genes in mammalian cells both *in vitro* and *in vivo* in transgenic mice (95, 96). To assess the capacity of this system to knock-down target genes using RNA interference, we developed two vectors, p*m*GENIE-3-Trt1 and p*m*GENIE-3-Trt2, to express shRNA targeting TERT (Figure 3-1). We initially compared transfection efficiency of our actively integrating p*m*GENIE-3 to a passive, transposase-deficient pGIPZ shRNAmir vector expressing the same TERT-targeting shRNA. We found the transient transfection levels (72 hours post transfection) to be enhanced ~4-fold using the p*m*GENIE-3 vector (Figure 3-2A). Furthermore, following selection with antibiotic, we observed ~5-fold enhancement of the frequency of stably transfected cells receiving p*m*GENIE-3 based on GFP expression (Figure 3-2B).

To assess the ability of p*m*GENIE-3-Trt1 and p*m*GENIE-3-Trt2 to knock-down TERT, we stably transfected both HEK293 cells and MCF-7 cells with these constructs. We also generated control cell lines transfected with a p*m*GENIE-3 construct without shRNA (p*m*GENIE-3-Trt0). At ~2 months post-selection, we observed significant reduction in relative levels of TERT mRNA for both constructs in HEK293 and MCF-7 ($P \leq 0.001$) cells (Figure 3-3). To determine whether this level of knock-down was sufficient to have an effect on the telomerase enzymatic complex, we also measured telomerase activity in these cells. We observed significantly

(P≤0.02) reduced levels of telomerase activity with both constructs in HEK293 and MCF-7 cells, in agreement with reduced levels of TERT mRNA (Figure 3-4). These observations confirm that the non-viral p*m*GENIE-3-TERT delivery system is capable of knocking down target genes via stable expression of shRNA.

We next sought to determine if the reduced level of telomerase had a physiological effect in these cell lines. Specifically, we measured telomere length by southern analysis of terminal restriction fragments (TRFs) at ~ 2 months post-transfection (Figure 3-5). For both HEK293 and MCF-7 cells transfected with either p*m*GENIE-3-Tert1 or p*m*GENIE-3-Tert2, we observed a significant reduction in telomere length. For the HEK293 cell line, which has relatively short initial telomere length, we noted that telomeres were approaching critical length required to maintain cell viability (~3 Kb; 25,26). An analysis of proliferation rate in cells transfected with the TERT targeting vectors revealed a marked (>2-fold) reduction in growth by ~2 months post-transfection in cells (Figure 3-6).

3.5 Discussion

Transposition-based gene delivery systems are appealing tools which bypass the use of viral vectors for transgene delivery. Transposase/transposon vectors have been shown to have less of an affinity for transcriptional start sites, transcribed regions, and proto-oncogenes than lenti- and retro-viruses and their ease of production at low expense make them ideal candidates to evaluate human applications of gene therapy (74). We have modified the pB transposon system to consist of a single self-inactivating vector for highly efficient generation of stable

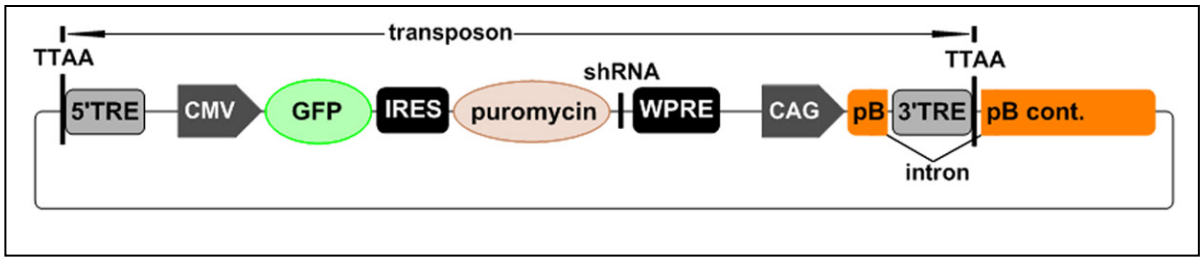
transformants in mammalian cells *in vitro* as well as in transgenic animals (95, 96). Due to our previous success at utilizing helper-independent pB constructs to stably transfect cell lines and animals with marker genes such as EGFP, we speculated that transgenes that affect the biological function of cells would have the same successful delivery and integration.

In the present study we have assessed the ability of our single plasmid pB vector, p*m*GENIE-3, to drive expression of shRNA and affect the target gene expression in human cell lines. The practical utility of silent interfering RNA to suppress cancer has been intensively studied over the past 10 years (97), however truly effective RNAi-based therapies have yet to be developed. Telomerase, specifically TERT, has been studied as a target for RNAi technology (98). Indeed, a TERT-targeting RNAi therapeutic could potentially complement other anti-sense targeting methods, such as using stable anti-sense oligonucleotides that target the telomerase RNA component (99) to dramatically reduce telomerase levels and induce cancer cell senescence or apoptosis. We show in both HEK293 and MCF-7 cells that p*m*GENIE-3 vectors designed to target the *TERT* gene transcript, effectively knock down TERT expression and reduce telomerase activity levels. Furthermore, this knock down of telomerase was sufficient to effect the physiological function of telomerase, since in both cell lines it was accompanied by significant telomere shortening ($P<0.001$). These results suggest that the pB transposase-based delivery system is an effective technique for expressing shRNA and knocking down target genes in human cell lines, including tumor cell lines.

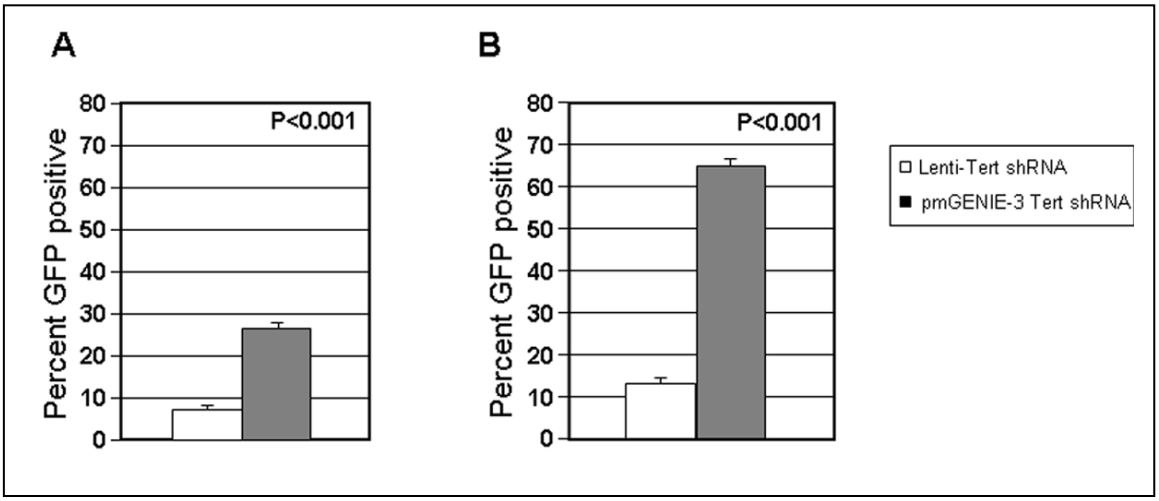
The ability of the p*m*GENIE-3 vector to efficiently knock down TERT and reduce telomerase levels, in conjunction with the high level of long term expression, indicative of genomic integration associated with this vector, suggest that pB transposase-based delivery systems may provide a new method to suppress tumor progression. We are presently undertaking experiments to assess the potential and effect of using p*m*GENIE-3 to knock down TERT and tumor cells an in vivo in murine model.

In summary, we have shown that the pB transposase-based delivery system is effective at shRNA driven targeted gene knock down in mammalian (human) cells. The ability to affect both telomerase and telomeres in a tumor cell line implies a potential utility for the pB vector in cancer therapy.

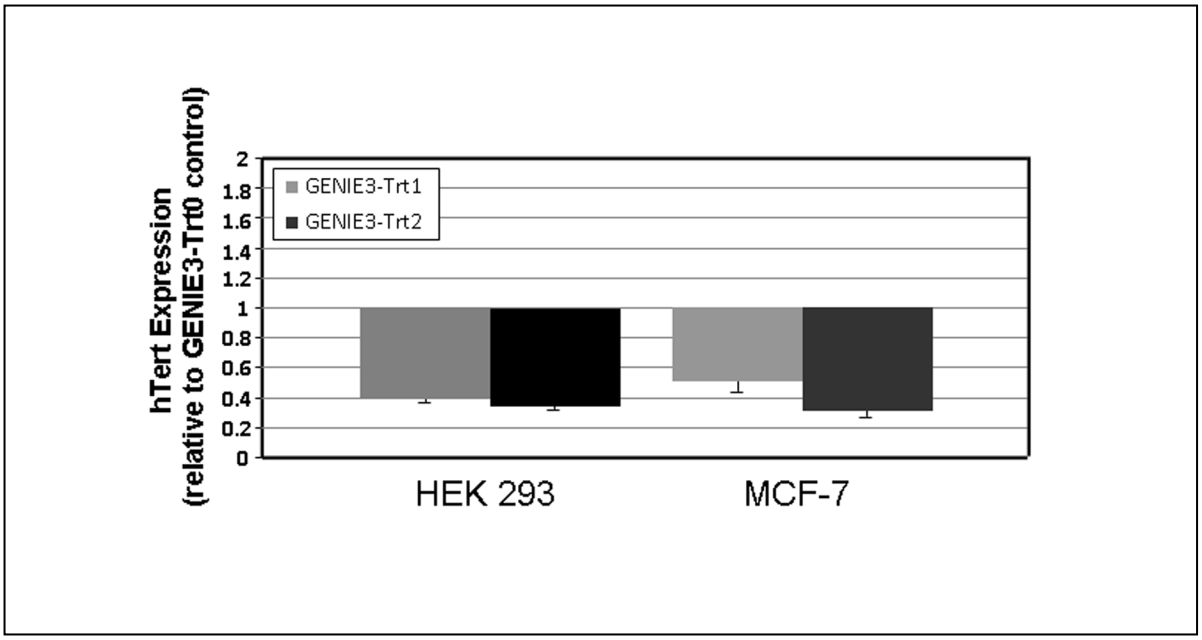
3.6 Figures



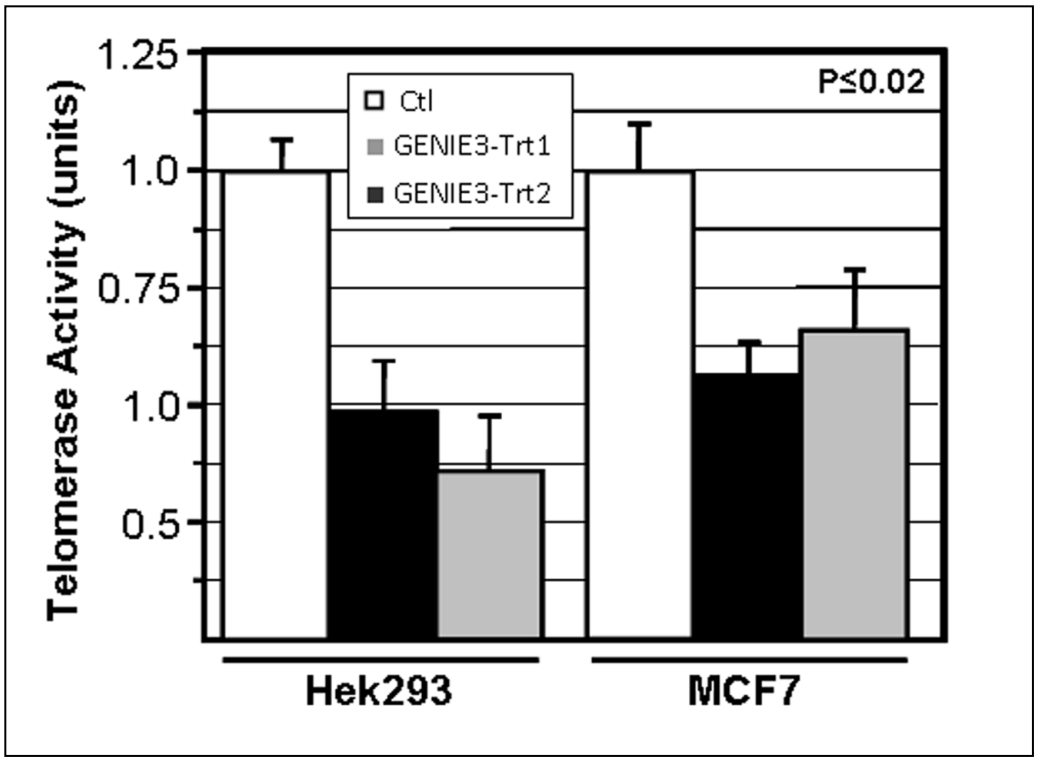
3.6.1 Figure 3-1. Map of the p*m*GENIE-3 pB transposase-based vector. Included in the vector design are ubiquitously expressed GFP to facilitate detection of transfectants, IRES-Puro to allow stable selection of integrants, and a site to allow insertion of shRNA sequences.



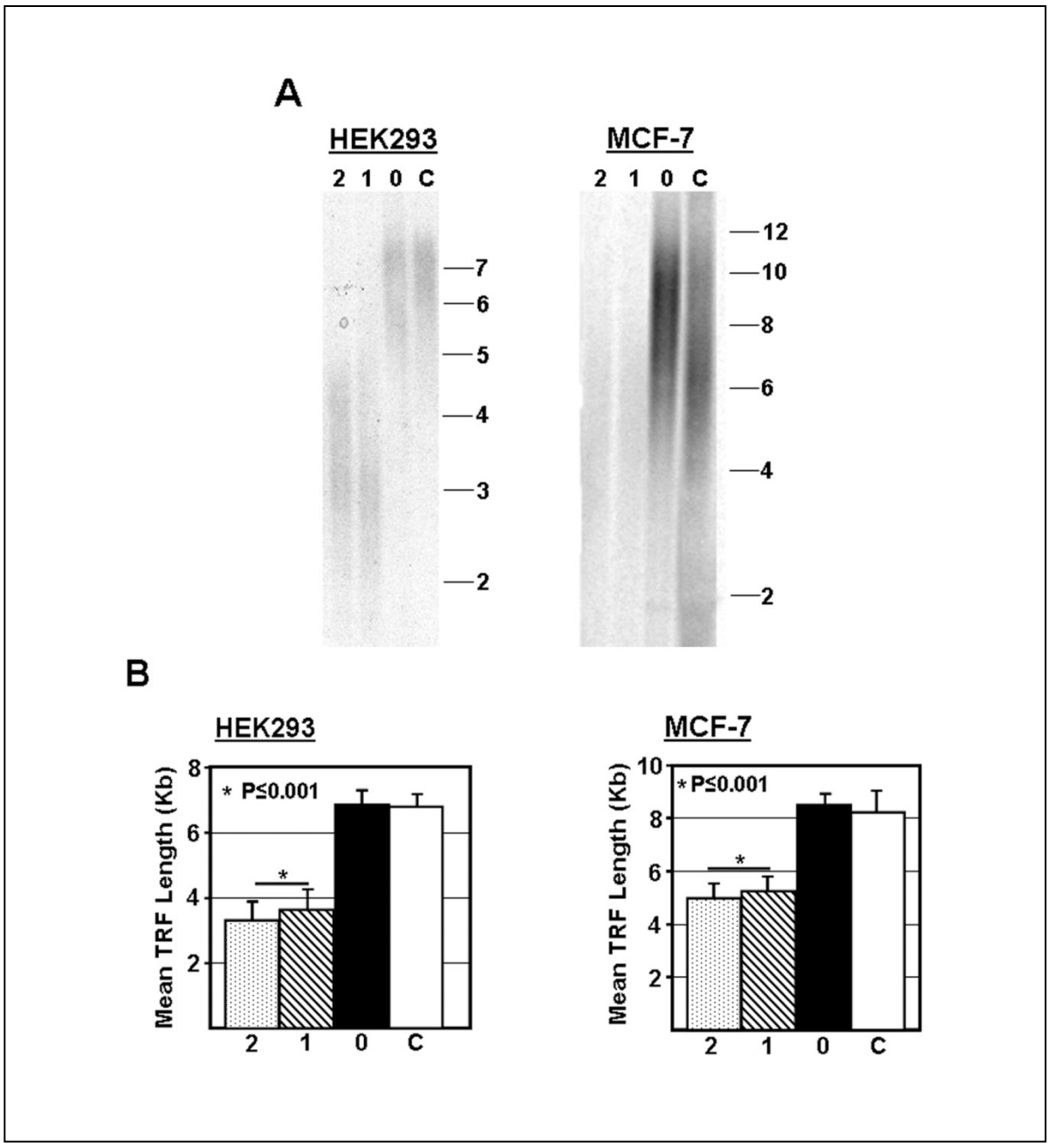
3.6.2 Figure 3-2. Comparison of the short term and long term transfection efficiency of p*m*GENIE-3 to a transposase-deficient lentiviral vector. **A.** HEK293 cells were transiently transfected (Lipofectamine) in triplicate with equal amounts of either p*m*GENIE-3 or pGIPZ shRNAmir plasmid and the fraction of GFP positive cells (the ratio of GFP+ to GFP- cells) was assessed by FACS analysis at 72 hours post-transfection. **B.** HEK293 cells were transfected as in **A**, followed by 4 week selection with Puromycin. Values represent the number of total GFP+ cells relative to a control reference sample transfected with an empty p*m*GENIE-3 vector grown without selection. Standard error bars and P-value (student's t Test) are shown.



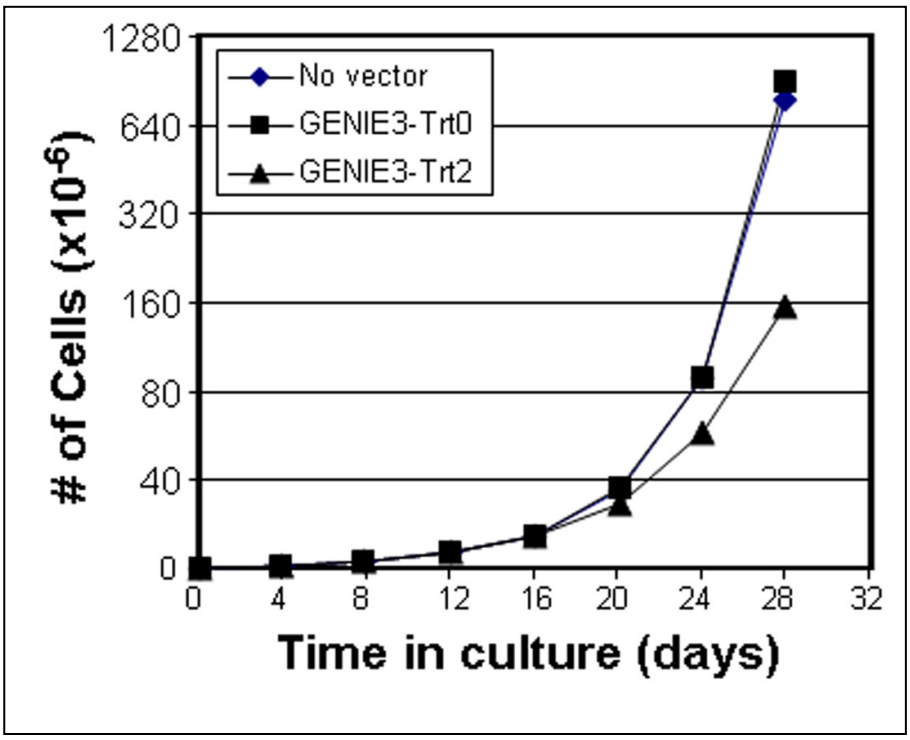
3.6.3 Figure 3-3. Analysis of the potential of p*m*GENIE-3 to perform RNAi mediated knock down of gene expression in human cells. Human cell lines (HEK293 and MCF-7) were transfected with p*m*GENIE-3 vectors expressing shRNA targeting hTERT. Following Puromycin selection, the relative expression of hTERT was assessed using real time RT-PCR. Expression values, normalized to Gapdh, are shown relative to corresponding control cells (transfected with Trt0). All analyses were performed in triplicate; standard error bars are shown.



3.6.4 Figure 3-4. Analysis of p*m*GENIE-3 mediated hTERT knockdown of telomerase activity in human cell lines. Telomerase activity was measured using the TRAP assay in the same cells as assessed in Figure 3. All analyses were performed in triplicate; standard error bars are shown.



3.6.5 Figure 3-5. Analysis of the physiological effect of p*m*GENIE-3 mediated gene knock down in human cells. Following Puromycin selection, telomere length analysis was performed on the same cells that were transfected with the hTERT targeting p*m*GENIE-3 in Figure 3. **A.** Sample blots of southern analysis of TRF length for HEK293 and MCF-7. C-no treatment control; 0- p*m*GENIE-3-Trt0; 1- p*m*GENIE-3-Trt1; 2- p*m*GENIE-3-Trt2. **B.** Quantitative analysis of mean TRF length for all samples. All analyses were performed in triplicate; standard error bars are shown. P value represents Student's t Test.



3.6.6 Figure 3-6. Effect of TERT-targeting shRNA expression on growth rate of HEK293 cells. Cells were transfected with p*m*GENIE-3 followed by continuous growth in the presence of Puromycin (2µg/mL) starting at 72 hours post-transfection. At 30 days post-transfection (~25 population doublings), Equal numbers (250,000) of control cells (untransfected/Puro free or p*m*GENIE-3-Trt0 transfected) or TERT-targeted cells (p*m*GENIE-3-Trt2 transfected) were seeded in 100mm dishes and allowed to grow for 4 weeks. Cells were passaged a total of 3 times (1:20 split). Cells were counted every fourth day.

3.7 Contributions

Research design: Rich Allsopp, Jesse Owens

Conducted experiments: Jesse Owens, Juanita Mathews, Philip Davy, Ilko Stoytchev

Performed data analysis: Rich Allsopp, Jesse Owens

Wrote or contributed to writing: Jesse Owens, Rich Allsopp, and Stefan Moisyadi

Chapter 4. Chimeric *piggyBac* transposases for genomic targeting in human cells

4.1 Abstract

Integrating vectors such as viruses and transposons insert transgenes semi-randomly and can potentially disrupt or deregulate genes. For these techniques to be of therapeutic value a method for controlling the precise location of insertion is required. The pB transposase is an efficient gene transfer vector active in a variety of cell types and proven to be amenable to modification. Here we present the design and validation of chimeric pB proteins fused to the Gal4 DNA binding domain (DBD) with the ability to target transgenes to predetermined sites. Upstream activating sequence (UAS) Gal4 recognition sites harbored on recipient plasmids were preferentially targeted by the chimeric Gal4-pB transposase in human cells. To analyze the ability of these pB fusion proteins to target chromosomal locations, UAS sites were randomly integrated throughout the genome using the SB transposon. Both N- and C-terminal Gal4-pB fusion proteins but not native pB were capable of targeting transposition nearby these introduced sites. A genome-wide integration analysis revealed the ability of our fusion constructs to bias 24% of integrations near endogenous Gal4 recognition sequences. This work provides a powerful approach to enhance the properties of the pB system for applications such as genetic engineering and gene therapy.

4.2 Introduction

The ability of integrating vectors to permanently introduce foreign genes into chromosomes has resulted in major advances in the fields of genetic engineering, functional genomics, and gene therapy. For these techniques to be of value in the clinical setting it is imperative that insertions occur at known safe loci in order to avoid deregulation of the cell due to deleterious integrations and to control expression of transgenes. Commonly used viral vectors have been shown to preferentially insert their cargo near transcriptional start sites (1-3) and there has been increasing concern for the implications of insertional mutagenesis (7, 8, 11). Thus, the safety of insertional therapies would be improved by the ability to target vector integration to a specific genomic safe harbor.

Cys_2His_2 zinc finger proteins (ZFPs) can bind to specific sequences by inserting an alpha-helix into the major groove of the DNA double helix. These DBDs are specific for 6 bp to 18 bp DNA sites and can now be easily designed in a few weeks and be made to target almost any location in the genome (100-104).

By fusing ZFPs to activator or repressor domains novel zinc finger effectors have been used to upregulate or downregulate transcription (105-108). Recently, zinc finger nucleases (ZFNs), chimeric proteins that consist of a ZFP and a Fok1 nuclease domain, have proven to be effective in a variety of applications such as gene disruption, transgene integration and the generation of knockout mice (109-113). By inducing targeted double-stranded breaks (DSBs) and using the host cell's

repair machinery ZFNs have been used to cause intentional mutations or insert whole genes at respective targets (114-117). However, the nuclease component of ZFNs can cause off-target cleavage events that result in undesired mutations. Concerns about cyto and genotoxicity remain significant obstacles to be overcome for this to be a safe strategy (27-29, 109, 118).

An alternative approach has been to directly fuse DNA integrating enzymes to ZFPs in an attempt to localize activity of the vector to a specific genomic location (119). For example, ZFP-HIV-1 integrase fusions packaged in virions showed promise with preferential targeting in both plasmids and genomic DNA (gDNA), albeit at low levels (120-124). Programmable recombinases using zinc fingers bound to a catalytic domain have been shown to precisely integrate transgenes at predetermined sites (125-128). However, the catalytic domains for these proteins are sequence specific, thus targeting is limited to sites containing the required sequences. Steps have been taken to alter the sequence specificity of these catalytic domains in order to allow integration into novel sites (129, 130).

Kaminski et al. suggested using transposases fused to a DBD as a method for directing transgene insertion and it was shown previously that modifications to the yeast retrotransposase Ty5 could influence its target selection profile (51, 131-133). Others have shown in *E. coli* that the ISY100 transposase bound to a ZFP from the mouse transcription factor Zif268 could target transgenes near the expected binding site on recipient plasmids 48% of the time (134). In addition, the prokaryotic mobile

element IS30 fused to the Gli1 transcription factor is able to target extrachromosomal plasmids in zebrafish embryos (135).

The SB transposase has shown activity in mammalian cells and has been used for diverse non-viral applications (34, 88, 89). However, the direct fusion of DBDs to SB has lead to complete or significant reductions in transposase activity (39, 136, 137). Despite this, Yant et. al demonstrated DBD-SB fusion protein mediated transgene targeting at efficiencies of 18-33% to specific sites within plasmids in human cells. However, this group was unable to detect targeted integration when coupling SB with a DBD specific to an endogenous chromosomal location.

pB, an insect transposase isolated from the moth *Trichoplusia Ni*, is highly efficient in a broad range of organisms including yeast and mouse as well as human cell lines and is able to integrate relatively large cassettes of more than 100kb. pB inserts transgenes at TTAA tetranucleotide sites and the transposase has been shown to be able to excise the transposon without leaving a DNA footprint (36-38, 40-42, 56, 73, 138-141). Previously, we have shown that pB is amenable to a Gal4 DBD fusion with little loss in activity (39). Furthermore, we demonstrated targeting to a plasmid recipient harboring the UAS Gal4 recognition site in *Aedes aegypti* embryos (142, 143).

A wide range of applications would directly benefit from safe targeted transgenic insertion. We have tested both N- and C- terminal fusions of the classic Gal4 DBD to

the pB transposase and assessed for targeted integration in human cells near the UAS DNA binding sequence in a genomic setting. The aim of our experiments was to demonstrate the ability of a chimeric pB transposase fused to a DNA binding domain to target the genome. This proof of principle is important for transposition research and will serve as a basis for future improvements that one day may lead to safer transpositional gene therapy treatments in a clinical setting.

4.3 Experimental Procedures

4.3.1 Plasmid Development

The mammalian codon biased pB transposase was a gift from Dr. Allan Bradley. The backbone for all pB plasmids was the self-inactivating helper-independent p*m*GENIE-plasmid system described earlier (40). The Gal4 DBD, (amino acids: KLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTK RSPLTRAHLTEVESRLERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNV NKDAVTDRLASVETDMPLTLRQHRISATSSSEESSNKGQRQLTVS) was introduced via homologous recombination along with the linker: KLGGGAPAVGGGPK (142). The transgene for all p*m*GENIE constructs was a PCR amplified fragment from pERV3 (Agilent Technologies) including the bacterial and eukaryotic promoter driven neomycin gene. This fragment was TA-cloned into the pENTR1a vector (Invitrogen) that had been digested with HincII and EcoRI and then t-tailed using a terminal transferase (New England Biolabs). The ligation product, including the neomycing gene, was Gateway (Invitrogen) recombineered between the attR sites in the p*m*GENIE plasmid transposon.

The CMV-SB11 helper plasmid was a gift from Dr. Perry Hackett. To create the SB UAS donor plasmid pT2/HB cam UAS hygro, we used an intermediate plasmid (pELO4) that contains the chloramphenicol gene. The UAS site was isolated from the pGDV1-UAS plasmid by restriction digestion with HindIII and BamHI (142), and cloned into the pELO4 vector adjacent to its chloramphenicol gene. Two identical 650 bp TTAA rich regions, custom synthesized from GenScript and excised from the pUC57 shuttle vector using BamHI or HindIII, were ligated to both sides of the UAS target sequence to make pELO4 cam TTAA-region UAS TTAA-region. The hygromycin gene from pEGSH (Agilent Technologies) was PCR amplified and ligated into the pT2/HB SB donor plasmid with BglII and HindIII and a unique NheI site was introduced with the PCR primers used. The pT2/HB hygro construct was linearized with NheI and a pELO4 PCR product, including cam, UAS and TTAA-rich regions, was added via In-Fusion (Clontech). This plasmid was then used to generate hygromycin resistant cell lines containing the SB transposon and UAS. The UAS-negative control SB donor plasmid was made by PCR amplifying all but the UAS site from the pT2/HB cam UAS hygro construct. This inverse PCR product was re-ligated to form the control plasmid which retained the TTAA rich regions but had the UAS removed.

The pT2/HB cam UAS hygro construct was reduced in size by self ligation using XmnI and SmaI, and used as the recipient in the plasmid into plasmid experiment. In order to eliminate false positives during the plasmid into plasmid assay, the bacterial

suicide ccdB gene was added to the backbone of the p*m*GENIE plasmids by In-Fusion. This strategy prevented recovery of double resistant colonies that contained delivery plasmid backbone resulting from non-transpositional insertion. All restriction enzymes were purchased from New England Biolabs.

4.3.2 Cell Transfections

Human embryonic kidney (HEK293) cells were maintained in complete DMEM supplemented with 10% heat inactivated FBS (Invitrogen) and prior to transfection 0.5×10^5 cells per well were seeded in 12-well plates. To make the SB UAS and SB control cell lines lacking a UAS, 100ng of CMV-SB11 helper and 200ng of the appropriate pT2/HB SB donor plasmid were transfected using FuGene6 (Roche Applied Science) into cells at 90% confluency. Stable cell lines were obtained after 30 days of culture in 100 µg⁄mL HygromycinB (Invitrogen) and were cultured for an additional 3 weeks. We subjected these stable SB transposon containing cell lines to a second round of transfection with 400ng of helper-independent p*m*GENIE construct containing the neomycin selection cassette within the transposon. p*m*GENIE constructs with N- or C-terminal Gal4 fusion or native pB were transfected in triplicates into two independent SB UAS cell lines. Stable G418 resistant lines were obtained after 30 days and equal numbers of cells for each experiment were pelleted and frozen. Genomic DNA was isolated using the DNeasy kit (Qiagen) following the manufacturer's protocol. A standard colony count assay was performed using all three constructs and a transposase-negative control as described previously (40).

4.3.3 Plasmid into Plasmid Assay

HEK293 cells were transfected with 750ng each of UAS recipient plasmid pT2/HB cam UAS hygro and helper-independent p*m*GENIE delivery plasmids containing either of Gal4-pB fusions or native pB. Cells were grown without selection for 3 days then pelleted and episomal plasmids were isolated using the Zyppy plasmid miniprep kit (Zymo Research). Lucigen 10G Elite *E. coli* cells were electroporated with the isolated plasmid DNA and plated on double antibiotic cam^R/kan^R plates for selection. Colonies were screened by duplex PCR to simultaneously confirm the integration of the delivery neomycin transposon as well as the excision of the transposon from the backbone of the delivery p*m*GENIE plasmid. Plasmid DNA from positive double antibiotic resistant colonies was purified by miniprep and sequenced using the primer pB *5TRE*: ACG GAT TCG CGC TAT TTA GA which extends from the pB transposon into the adjacent sequence. Obtained sequences were aligned to the recipient plasmid pT2/HB cam UAS hygro to determine the insertion site and distance from the UAS site. The efficiency by which pB delivery transposons were integrated into recipient SB plasmids by the respective pB transposase was calculated by dividing the total number of correctly aligned sequences recovered by the percent of colonies screened. The percent colonies screened was calculated by dividing the total number of colonies PCR screened by the total colony count. (n=3, mean ± SD) Statistics include two-sided, two-sample Student's t-test assuming equal variance $P = 0.01$.

4.3.4 Targeted Genomic Integration Site Recovery

Genomic DNA was extracted from pooled clones of stably double transfected HEK293 cells and nested PCR was performed using forward primers designed to extend from either terminal repeat element (TRE) of the delivery pB transposon, while the reverse primers were designed to the SB UAS transposon target. Because of the repetitive nature of the UAS and surrounding sequence we did not obtain PCR products extending through the UAS and therefore designed two sets of reverse primers to identify insertions upstream and downstream of the UAS. See Supplementary Table 4-4 for primer sequences. Primary PCR products obtained using KOD Xtreme Hot Start DNA Polymerase (Novagen) were diluted 1:100 in H_2O and used as template for nested PCR using Easy-A High-Fidelity Polymerase (Agilent Technologies). Amplification products were either first gel purified with Zymoclean Gel DNA Recovery Kit (Zymo Research) or directly TA cloned into the pGEM-T Easy Vector (Promega). Unique clones were verified by colony PCR, plasmid DNA was purified by miniprep and then sequenced with Sp6 or T7 primers. Sequences were aligned to the SB UAS transposon and distance to the UAS and exact insertion site locations were recorded.

4.3.5 Copy Number Assay

Monoclonal expansions of double-resistant hygromycin/G418 HEK293 cells which had been first transfected with SB UAS (pT2/HB cam UAS hygro) then with N- or C-terminal Gal4-pB, or pB control p*m*GENIE plasmids (n=5 each) were seeded into 96 well plates at low density and visually verified. gDNA from expanded clones was

isolated for Southern blot and qPCR copy number assay. In order to standardize the qPCR copy number assay a Southern blot was performed to identify a known number of insertions for representative pB and SB experiments. Genomic DNA was isolated using the DNeasy kit (Qiagen) following the manufacturer's protocol. 20µg of each sample was digested overnight using 10 units of HindIII per µg of gDNA. 12µg of digested gDNA was run for 75 h on a 1% agarose gel at 15V. The gDNA was then blotted to a Hybond + nylon membrane (GE Healthcare) for 12 h and processed for hybridization according to the method of Sambrook et al. (1989). A DIG-labeled probe was generated by PCR amplification using the PCR DIG Probe Synthesis Kit (Roche Applied Science) and the following primers: pB *Southern Forward* ACGTAAACGGCCACAAGTTC, pB *Southern Reverse* TGCTCAGGT AGTGGTTGTCG. *SB Southern Forward* AACTCGTTTTTCAACTACTCC-ACA, *SB Southern Reverse* ACTGTCGGGCGTACACAAAT. PCR parameters used: initial denaturation at 94°C for 2 min, 35 cycles of 30 s denaturation at 94°C, 30 s annealing at 56°C, and 1 min elongation at 72°C, with a final elongation for 10 min. Hybridization was performed overnight at 55°C using the DIG Easy Hyb Kit, (Roche Applied Science) and were processed according to the manufacturer's protocol. Chemiluminescent signals were visualized with an LAS-3000 imaging system (Fujifilm).

qPCR copy number assays were performed by duplex Taqman real-time PCR, where one assay interrogates the transgene copy number, while the other assay (to RNaseP) serves as a reference. Primers and probes were custom designed (to the

5'TRE for pB and the 3'TRE for SB) or pre-made (RNAseP) and were supplied by Applied Biosystems. The primer and probe sequences are as follows: pB *5TRE Copy Forward* GTGACACTTACCGCATTGACAAG, pB *5TRE Copy Reverse* GCTGTGCATTTAGGACATCTCAGT, pB *Reporter* ACGCCTCACGGGAGCTC, *SB 3TRE Copy Forward* CTCGTTTTTCAACTACTCCACAAATTTCT, *SB 3TRE Copy Reverse* ACAATTGTTGGAAAAATGACTTGTGTCA, *SB Reporter* TTTGGCAAGTCAGTTAGGACATCTA. The assays were performed according to the TaqMan copy number assay protocol (Applied Biosystems) using the Step-One-Plus real-time PCR machine in a 20µl reaction volume containing 50ng DNA. A minimum of 4 replicates per sample was assayed. One sample with known transgene copy number (as determined by Southern blot analysis) was included. The copy number assays were normalized to RNAseP, known to occur in two copies in the genome (Applied Biosystems). The results were analyzed using the software CopyCaller v1.0 (Applied Biosystems).

4.3.6 Western Blotting

Cells were cultured as described. After 48 h of incubation, cells were lysed with lysis buffer supplemented with Set III protease inhibitors (Calbiochem). 60µg of total protein was resolved on a precast SDS–polyacrylamide gel (Bio-Rad). Expression of pB proteins was determined by Western blotting using a mouse monoclonal anti-*piggyBac* non-purified antibody. Binding of primary antibody was detected using an anti-mouse horseradish peroxidase-conjugated secondary antibody (Santa Cruz Biotechnology). Bands were visualized with an LAS-3000 imaging system (Fujifilm).

4.3.7 Non-Restrictive Linear Amplification-Mediated (nrLAM) PCR and 454 Sequencing

For the off-target insertion analysis we adapted non-restrictive LAM PCR (144). Briefly, 1µg of gDNA from double-resistant hygromycin/G418 HEK293 cells for both Gal4 fusion samples as well as pB control (n=4 each) was used as template for linear PCR using single primers for linear amplification extending from the pB-TREs into flanking genomic sequence. See Supplementary Table 4-5 for primer sequences. Single stranded linkers were ligated to these linear PCR products and nested PCR was performed to amplify the flanking genomic sequence. GS-FLX sequencing primers were added by PCR and samples were sequenced using a 454 GS-FLX Titanium sequencer (Life Sciences) in accordance with the manufacturer's protocol by the University of Hawaii Advanced Studies in Genomics, Proteomics and Bioinformatics (ASGpB) unit. Resulting sequences were trimmed and demultiplexed using CLC Genomics Workbench version 4.7 (CLC Bio). The reads were mapped against the human genome reference, version GRCh37.63, using the short read alignment component (bwa-short) of the Burrows-Wheeler Aligner (145), selecting for reads that align over at least 80% of the sequence with a minimum of 90% similarity. Distances of insertion sites to endogenous Gal4 recognition sites were obtained using custom scripts, which are available upon request. Gal4 recognition sequences were defined as CGGNNNNNNNNNNCCG and a total of 56,898 sites were identified in the human genome. The position weight matrix for the Gal4 DBD is depicted in Supplementary Table 4-6. Distances to transcriptional start sites as well

as other annotations were obtained using the Homer bioinformatics tool (146) available online at: http://biowhat.ucsd.edu/homer/ngs/annotation.html. Of the 66,414 integration sites recovered using nrLAM PCR, 7004 sites aligned to unique genomic locations.

4.4 Results

4.4.1 Chimeric Gal4 pB directs transposition into plasmid targets

In a previous publication we described highly efficient pB plasmids which maintained 92% activity of integration after addition of the Gal4 DBD (39). Here we have tested the hypothesis that by tethering the pB transposase to Gal4 we are able to target integration of transposons near UAS recognition sites in mammalian cells. In the first set of experiments we used a plasmid into plasmid approach. The recipient plasmid contained a chloramphenicol gene (cam^R) and a UAS site that consisted of 5 recognition sequences described earlier (142). The delivery plasmid contained the pB transposase, with or without a Gal4 DBD, and a transposon delivery cassette harboring the neomycin (kan^R) gene for bacterial selection. Integration of the delivery cassette into the recipient plasmid conferred double resistance to cam^R/kan^R when transformed into *E. coli* (Figure 4-1A).

Recipient and delivery plasmids were transfected into HEK293 cells and plasmid DNA was isolated 3 days later. Total isolates were electroporated into *E. coli* and plated on cam^R/kan^R for selection. Colonies were screened by colony PCR to

confirm enzymatic excision of the transposon from the delivery plasmid. Plasmid DNA from positive clones was purified and sequenced.

To identify transposon insertion sites within the recipient plasmid we sequenced out from the delivery cassette into the adjoining plasmid DNA. We flanked both sides of the UAS target with a 650 bp region in which 65 TTAA sites were spaced 10 bp apart. This design would allow us to analyze distance requirements for integration from the UAS. For example, preferential integration at a certain distance from the UAS might be evidence for spatial protein tension during integration or be the result of Gal4 linker length. Only sequences where the pB TRE was immediately followed by a TTAA and consecutively flanked by recipient sequence were considered as verified transpositional insertions. A total of 182 verified sequences were recovered for the Gal4 fusions and native pB experiments. The efficiency of total integrations of the chimeric Gal4 protein was over three times that of native pB control (Figure 4-1B). This was expected because the Gal4 DBD is thought to bring the pB protein and recipient plasmid together via UAS target binding (Figure 4-1A). Furthermore an increase in total integrations is suggestive of targeting to sites nearby the UAS on the recipient plasmid.

By analyzing the distance of integration sites on the recipient plasmid relative to the UAS site, we found that 87% of N-terminal Gal4-pB (nGal4-pB) insertions were located within a region 800 bp up- or downstream of the UAS while only 59% of the pB control insertions fell within this region. Similarly, the C-terminal fused pB (pB-

cGal4) also had a significant ability (77%) to integrate near the UAS site (Table 4-1). More specifically, 47% of the nGal4-pB and 32% of the pB-cGal4 directed integrations were within 250 bp of the UAS compared to 21% for the native pB control. In addition, 39% of these nGal4-pB mediated insertions were detected within 250 bp upstream of the UAS site (Figure 4-1C). In contrast, the pB-cGal4 sample displayed a more evenly spaced distribution of insertion sites. Native pB frequently integrated farther upstream or downstream from the UAS compared to both fusion proteins (Figure 4-1C). In summary this data shows that in our plasmid model the addition of a Gal4 DBD to either N- or C-terminal end of the pB protein confers a propensity for transposition near the UAS site.

4.4.2 Genomic targeting of the chimeric Gal4 pB

In order to determine whether targeting could be achieved in a genomic setting we used a recipient transposon containing the UAS and TTAA-rich regions flanked by the TREs of SB (pT2/HB cam UAS hygro). This allowed for random integration of SB transposon targets into the genome of HEK293 cells which in turn were targeted by the chimeric Gal4-pB. The host repair machinery can be used to uptake fragments of exogenous DNA sequences into cells and integrants can be isolated following selection. However, this process is relatively inefficient and only the sequence for the selection marker is required for cell survival. In order to ensure that an intact sequence was efficiently integrated at a large number of genomic loci we used the SB system for transposition of the UAS target. By using this approach we avoided the possibility of pB recognizing its own TREs and consequentially excising the

target transposon. We used the two plasmid SB approach in which the recipient transposon contained the mammalian selection gene hygromycin and the UAS site flanked by TTAA-rich regions on one plasmid, as well as the SB11 transposase encoded on a second plasmid (Figure 4-2A).

HEK293 cells were transfected with both SB plasmids and stable lines were obtained after 4 weeks of culture under hygromycin selection. Two stable polyclonal expansions of HEK293 cell lines harboring the SB UAS transposon were transfected with delivery plasmid expressing N- or C-terminal Gal4-fused pB transposases or the native pB transposase. As with the plasmid into plasmid experiment, these delivery plasmids integrated a pB transposon containing the antibiotic resistance gene neomycin and conferred G418 resistance. Genomic DNA was isolated from 6 replicates each for the 3 experiments. To demonstrate the requirement of the UAS for Gal4 directed targeting, control transfections with a SB recipient transposon plasmid that lacked the UAS target sequence were performed and hygromycin resistant HEK293 cells were subjected to a second round of transfections with the pB delivery plasmids described above.

In order to detect targeted genomic insertion we used nested PCR with the forward set of primers complementary to the pB transposon and the reverse primers extending from the SB transposon target (Figure 4-2A,B). As expected we did not obtain any PCR products from the UAS-negative control samples indicating that neither Gal4 fusion nor native pB transposase targeted the SB recipient transposon

alone (Figure 4-2D). Furthermore, we did not detect any targeted PCR products for the 6 UAS-positive cell populations that had been transfected with native pB control transposase. In contrast, for all UAS-positive populations tested with a Gal4-pB fusion, nested PCR products were obtained and sequenced (Figure 4-2B). Within the 8000 bp region that we analyzed all 49 unique integrations recovered localized within 1300 bp of the UAS site, with the vast majority (96%) found less than 800 bp up- or down-stream from the UAS (Figure 4-2C and Table 4-2). 95% of sequenced PCR product bands were verified SB transposon insertions; the remaining 5% resulted from non-specific primer binding. Both N- and C-terminal fusions displayed a similar targeting efficiency and the integration profile within the TTAA-rich region flanking the UAS site appeared somewhat random without a predictable integration distance from the UAS. We also identified a number of hot-spots where the same insertion site was found across multiple samples; seven sites that were targeted three or more times and for one site, 193 bp upstream of the UAS site, we detected a total of four integrations by both nGal4-pB and pB-cGal4. Furthermore, eight loci shared integrations from both fusion constructs (Figure 4-2C).

4.4.3 Transposon copy number and off-target analysis

In order to assess the number of possible transposon targets per cell, monoclonal expansions were established from HEK293 cells transfected with both SB and pB transposons described above. Individual clones from nGal4-pB, pB-cGal4, and native pB samples were subjected to Southern blot (n=1) and qPCR (n=5 each) analysis. All samples tested contained either 1 or 2 UAS-SB transposons (Figure 4-

4A). We additionally analyzed the number of pB mediated integrations per clonal sample by qPCR. Here, the majority contained 3 to 5 delivery transposons with an average number of 7 integrations per cell (Figure 4-4B).

The protein levels of pB, nGal4-pB and pB-cGal4 were determined by Western blot (Figure 4-3A). Both fusion proteins expressed at similar levels and ran at a higher molecular weight than native pB. The ability to form $G418^R$ colonies in a standard colony count assay (40) was used to estimate transposase efficiency. The three constructs produced similar numbers of colonies with a slight reduction in efficiency for nGal4-pB compared to native pB control (Figure 4-3B) re-confirming that the Gal4 fusion did not inactivate the pB transposase.

Non-restrictive LAM PCR can be used to amplify genomic sequences flanking known insertion cassettes such as transposons to identify off-target insertion sites. Single stranded adaptors were ligated onto linear PCR products made from primers designed to extend away from the TREs of the pB delivery transposon. Nested PCR was used to amplify the TRE-gDNA junction and products were subjected to 454 pyrosequencing. Verified insertion sites included TRE sequence followed by the TTAA tetranucleotide followed by genomic sequence (Figure 4-4C). Four independent transfections for each of the 3 experiments were pooled and analyzed using published human genome annotations. A total of 66,414 integration sites were recovered from the nrLAM sequencing, many of which were repeated reads indicating good sequencing coverage. 7004 of the integration sites aligned to unique

genomic locations. An off-target analysis was performed on the integration sites identified from the 12 polyclonal samples confirming previous integration profiles for pB reported by us and others (Figure 4-4D) (36, 56, 60, 147-150). pB showed a slight preference for integration near the 5' end (within 10KB of transcriptional start sites) and 3' end (within 10KB of polyA sites) of genes as well as a preference for introns, possibly due to the large size and number of TTAA sites found within introns. The frequencies of insertions recovered within known genes can be found in Table 4-3. When compared to viruses such as HIV, pB displayed a more random insertion site distribution and targeted genes much less frequently (148-150). However, pB mediated insertions into genes were significantly more common when compared to a random insertion pattern.

4.4.4 Gal4 pB biases integrations near endogenous Gal4 recognition sites

A comparison of off-target integrations for the three constructs revealed that the nGal4-pB sample profile was shifted; we observed an increased number of integrations into exons, polyA sites and transcription start sites, presumably because of altered preference of integration due to Gal4 binding (Figure 4-4D). To estimate the efficiency of the Gal4-pB fusion protein's ability to insert near endogenous Gal4 recognition sites, we annotated the 56,898 UAS-like sites found in the human genome and examined insertions that were located in the vicinity of these sites. Gal4 binds tightly to a specific 6 bp binding site defined as CGGNNNNNNNNNN NCCG. Sequence variability or alterations in the number of variable (N) bp greatly reduces binding affinity (151). We counted the number of insertions that occurred in

20 bp increments from 0 to 10,000 bp from each Gal4 recognition site. 32% of nGal4-pB transpositions landed within 1.8kb of endogenous Gal4 sites compared to 8% for native pB and 23% were within 0.8kb compared to 5% for native pB (Figure 4-4D). The cumulative percentage of integrations recovered (Figure 4-4E) dramatically increases up until 1800 bp for nGal4-pB but not for pB-cGal4 or native pB. A histogram displaying the percentage of total integrations that occurred within 400 bp intervals shows increased insertions recovered in regions up to 1800 bp from Gal4 sites for nGal4-pB but not for pB-cGal4 or native pB (Figure 4-4F). While pB-cGal4 targeted the exogenous UAS-SB transposon almost as efficiently as nGal4-pB, we were surprised to find no preferential targeting of the endogenous consensus sites by pB-cGal4, which had an overall integration profile that resembled that of native pB.

4.5 Discussion

Traditionally, integrating vectors such as viruses have been used to insert transgenes semi-randomly and have led to deleterious effects due to integration at unwanted sites (152, 153). To address this problem we have designed novel proteins encompassing the classical Gal4 DBD fused to a pB transposase in an attempt to bias genomic insertion to specific sites within the genome. We have demonstrated targeting to UAS recipient plasmids in human cells using a tethered Gal4-pB and shown that integration preferentially occurs near the UAS recognition sequence (Table 4-1). Gal4 is a tight binding Zn_2/Cys_6 zinc finger with a 6 bp binding site that occurs not only in our inserted UAS sites, but also at many endogenous

human loci (151). Despite the numerous target sites for Gal4, we were successful in showing that genomic targeting can be achieved near introduced target UAS sites. It is important to assay transpositional events that occur on genomic DNA because histone-associated DNA may influence transposition as compared to naked DNA. We hence stably introduced our recipient transposon to ensure that the subtleties of the genomic environment could be accounted for during Gal4 directed transposition.

In our experiments merely integrating SB transposons containing TTAA sites without a UAS was not sufficient for Gal4-pB targeting. Both the UAS recognition sequence as well as the Gal4 DBD fused to the pB transposase was required for enhanced genomic targeting (Figure 4-2B, 4-2D and Table 4-2). It should be noted that during the preparation of this publication a similar methodology was published. In that publication the pB transposase was fused to CHK2-ZFP to direct integration (154). However, of the targeted single clones that were isolated in that study, greater than 20% of negative control clones showed evidence of native pB targeting into SB transposon targets containing the CHK2 recognition sequence. We did not observe such insertions in our negative control experiments, which was expected because it seems unlikely that an unmodified pB would preferentially integrate into one of the less than 200 TTAA sites found on a SB transposon target given the greater than 10 million possible TTAA sites available in the human genome. Currently, the discrepancy between these studies is unclear.

Given that there were only 1-2 inserted exogenous UAS targets per cell (Figure 4-4A) but millions of available TTAA sequences throughout each cell's genome, it is remarkable that we have been able to detect a bias for targeted transgene insertion. In this study, both N- and C-terminal Gal4 fusions but not native pB preferentially integrated within 800 bp of the UAS site. This information could prove important in the future design of alternative DBD-pB fusion proteins. Despite the even distribution of TTAA sites in the regions flanking the UAS site, one every 10 bp, certain hotspots were targeted frequently. A possible explanation for this observation is that the physical structure of the DNA places restraints on some of the potential integrations while other TTAA sites are more readily available for transposition. If indeed this is the case, this phenomenon may explain why the same sites were targeted in repeated transfections with the same plasmid and why the N-terminal fused pB transposase shared hotspots with C-terminal modified pB (Figure 4-2C).

An extensive off-target analysis using nrLAM PCR and 454 pyrosequencing revealed that fusing the Gal4 DBD to the C-terminal of pB, in contrast to an N-terminal fusion, did not significantly modify pB's off-target integration profile (Figure 4-4D). It may be that pB retains its intrinsic ability to bind DNA and is therefore able to bind to one of the many available TTAA sites within the genome thereby mediating off-target integration despite its fusion to the Gal4 DBD. While the same mechanism should apply for the N-terminal fusion of Gal4 to pB, we noticed a different integration pattern: in addition to targeting the UAS-SB transposon, nGal4-pB directed integration within 0.8kb of endogenous genomic Gal4 recognition sites

23% of the time compared to 5% for native pB controls, and within 1.8kb of endogenous sites 32% of the time compared to 8% for controls. Because 8% of integrations occur near Gal4 sites at random, and the percentage of total recovered sites for nGal4-pB is 32%, the difference of 24% represents the percentage of targeted integrations due to the presence of the Gal4 DBD. Given that there was an average of 7 pB insertions per cell (Figure 4-4B) we estimate that, on average, 1.7 insertions per cell (24% of 7 pB insertions) were targeted to within 1.8kb of a Gal4 site by nGal4-pB and that 5.3 insertions landed at random TTAA sites. We detected a variable number of pB insertions per cell (Figure 4-4B) and recognize that it is possible to have cells with or without any targeted insertions and with or without any random off-target insertions. It is entirely possible that the targeting efficiency would be reduced should the number of available recognition sites be decreased. Because the Gal4-fused pB transposase remains functionally active, there may be a higher probability of encountering and inserting at a random off-target TTAA site before a TTAA site near a single unique recognition sequence. Further experiments are needed to determine the level of influence of the number of possible target sequences in relation to the targeting efficiency of a DBD-pB fusion.

It is not evident to us why nGal4-pB but not pB-cGal4 targeted endogenous sites so much more effectively. The protein levels for the two transposases are comparable (Figure 4-3A) and the efficiency of integration for nGal4-pB was not higher than pB-cGal4 (Figure 4-3B). It is possible that this phenomenon may be due to spatial/steric interactions between Gal4 and pB. For example, in the N-terminal Gal4-pB

configuration the linker and pB transposase extend from the C-terminal side of the Gal4 DBD. This configuration is similar to the orientation of the natural activation domain in the full length wild-type Gal4 protein. However, this model of favorable fusion architecture does not explain why both fusions to the pB transposase apparently target the exogenous UAS on the SB transposon with similar efficiency. Neither does it explain why both constructs are able to integrate near the UAS on recipient plasmids. One explanation could be that pB-cGal4 may only efficiently bind naked DNA such as episomal plasmids. Exogenous UAS sites were chromatinized however, and pB-cGal4 retained the ability to target these sites, indicating that pB-cGal4 can bind DNA associated with histones. An alternative explanation for the difference in exogenous targeting between the N- and C-terminal fusions is that the introduced UAS sequence, like classic UAS sequence arrays, is made up of repeated Gal4 recognition sequences, and that the C-terminal fusion requires a number of sites in tandem for efficient binding. It is possible that, in addition to binding to UAS arrays, the N-terminal fusion also effectively binds endogenous monomeric Gal4 recognition sequences.

Although both chimeric proteins were effective in biasing integration toward exogenous or endogenous Gal4 binding sites, the number of off-target integrations identified in this study remains high and points to the need for a system where binding of a specific DBD is a prerequisite for pB mediated transposition. It may be possible to achieve this by mutational molecular evolution in which the activity of the pB protein is made to be dependent on its DBD. Redesign of the dimerization

interface of ZFNs has reduced off-target toxicity by reducing binding of the protein dimers in solution (155). Perhaps mutations in pB's dimerization domain could inhibit activity in solution but retain the ability of the protein to dimerize and integrate should the dimers unite at a DBD target sequence. It is clear that modifications such as this will be necessary for the chimeric pB strategy to mature into a viable method for genetic engineering and therapy.

Φ-C31 integrase mediates efficient DNA delivery to recipient plasmids using site-specific recombination between its *attP* and *attB* recognition sites and to pseudo *attP* sites of which there is an estimated 370 in the mammalian genome. A study looking at 196 independent genomic integration events revealed that 80% of insertions occurred near Φ -C31 *attP* sequence motifs, with 7.5% of integrations at a unique site on chromosome 19 (156). Unfortunately a high proportion of cells expressing Φ -C31 integrase were found to have numerous chromosomal abnormalities including various translocations (156, 157).

The adeno-associated virus (AAV) is able to insert its genome into a specific site on chromosome 19 (AAVS1) through the activity of its Rep78 protein. Hybrid adenovirus/AAV vectors have been developed to target transgenes to this site (158). Coinfection of human hematopoietic cells with two helper-dependent adenovirus vectors containing the *rep78* gene on one vector and a GFP reporter on the second resulted in 30% of integration into the AAVS1 region (159).

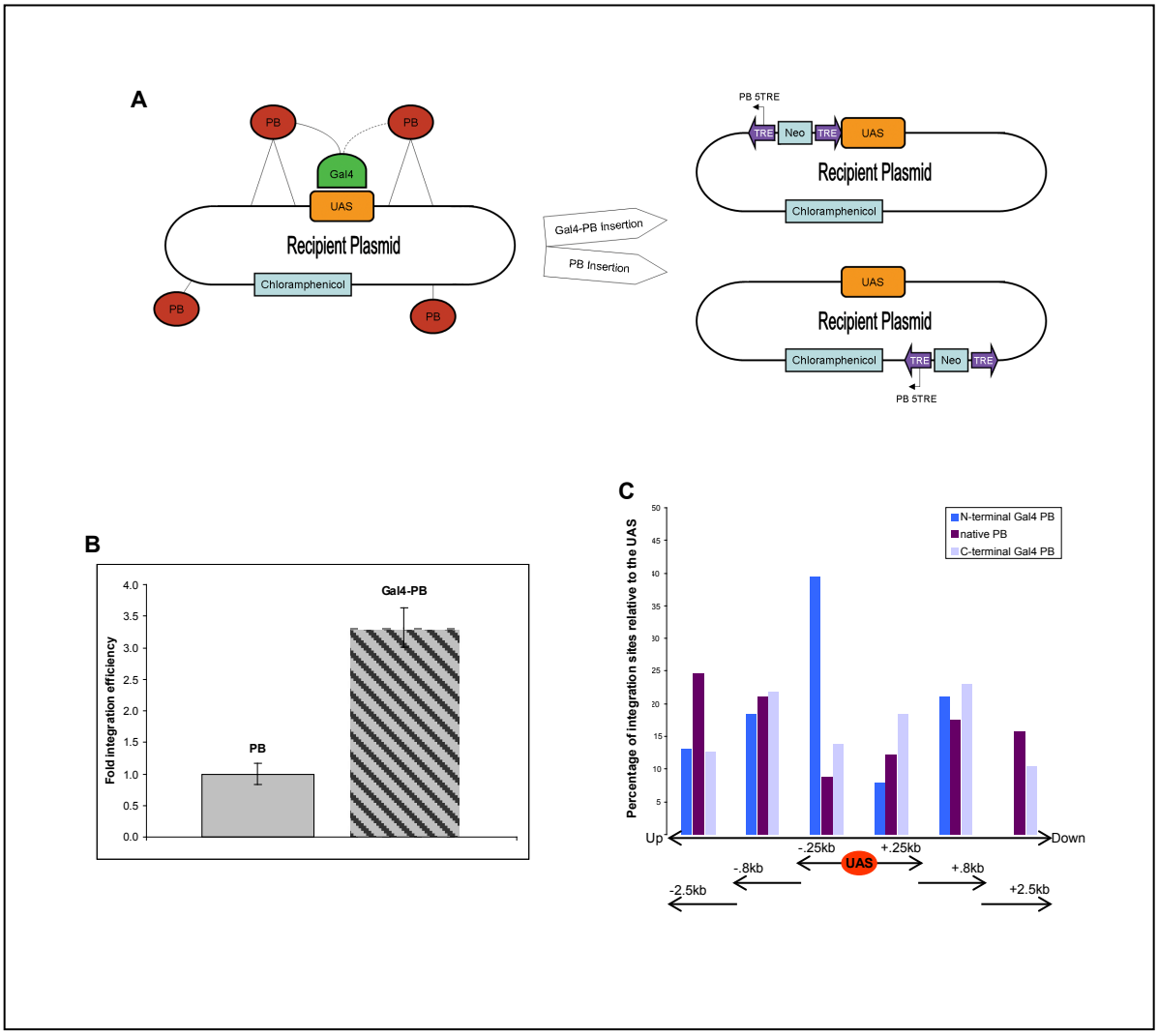
Gersbach et al. recently described zinc-finger recombinase (ZFR) fusion proteins with high specificity in targeting and with few off-target consequences (160). This group was able to efficiently integrate transgenes into a specific target sequence harbored on pB transposons randomly integrated within the genome. The utility of this approach however is hindered by the necessity to introduce target sites into the genome of interest due to the fact that the catalytic domain of the integrase retains strict sequence specificity for its native recognition sequence.

The plasmid based pB system has many advantages over viral techniques including tolerance by the immune system, large cargo capacity, as well as inexpensive and simple preparation. pB was shown to tolerate an assortment of fusion domains (39, 73, 142) while retaining high integration efficiencies. Unlike recombinase based approaches, which are restrained by a specific DNA binding sequence, it should be possible to replace Gal4 with any custom DBD. In order to reduce the number of off-target integrations, a DBD with a unique 18 bp+ recognition sequence could be used to tailor targeting of the chimeric transposase to virtually any predefined genomic site.

The need for highly specific endogenous integration is paramount and our future focus is to explore the use of highly specific DBDs to target a natural safe harbor in the human genome. One of the advantages of this strategy would be the ability to isolate single clones with safe single targeted insertions for use in a wide variety of cell replacement therapies. By expanding clones and verifying both the presence of

a targeted insertion and the absence of multiple insertions it would be possible to control gene silencing resulting from position effects and provide a means for avoiding detrimental mutations. The current study has laid the groundwork for using DBD-pB transposase fusion proteins for directed genomic integration. We anticipate that further improvements to this versatile framework will ultimately permit researchers to safely target a genetic cassette to any location within the genome.

4.6 Figures



4.6.1 Figure 4-1. (**A**) Schematic for the plasmid into plasmid experiment. Both the delivery plasmid containing the delivery transposon and transposase coding sequence as well as the recipient plasmid containing the chloramphenicol gene (camR) and the UAS were transfected into HEK293 cells. The tethering of Gal4 to the pB transposase (red circles) is thought to restrict integration to TTAA sites found near the UAS recognition sequence. Native pB proteins are free to integrate throughout the recipient plasmid. Delivery transposons contain the 5'TRE and 3'TRE for pB (purple arrows) and the neomycin gene (Neo) and confer kanR to the recipient plasmid. Recovered camR/kanR plasmids were sequenced with pB 5TRE (black arrows) in order to identify insertion sites. (**B**) Plasmid into plasmid integration efficiency of pB vs. Gal4-pB. Increased total integrations into the

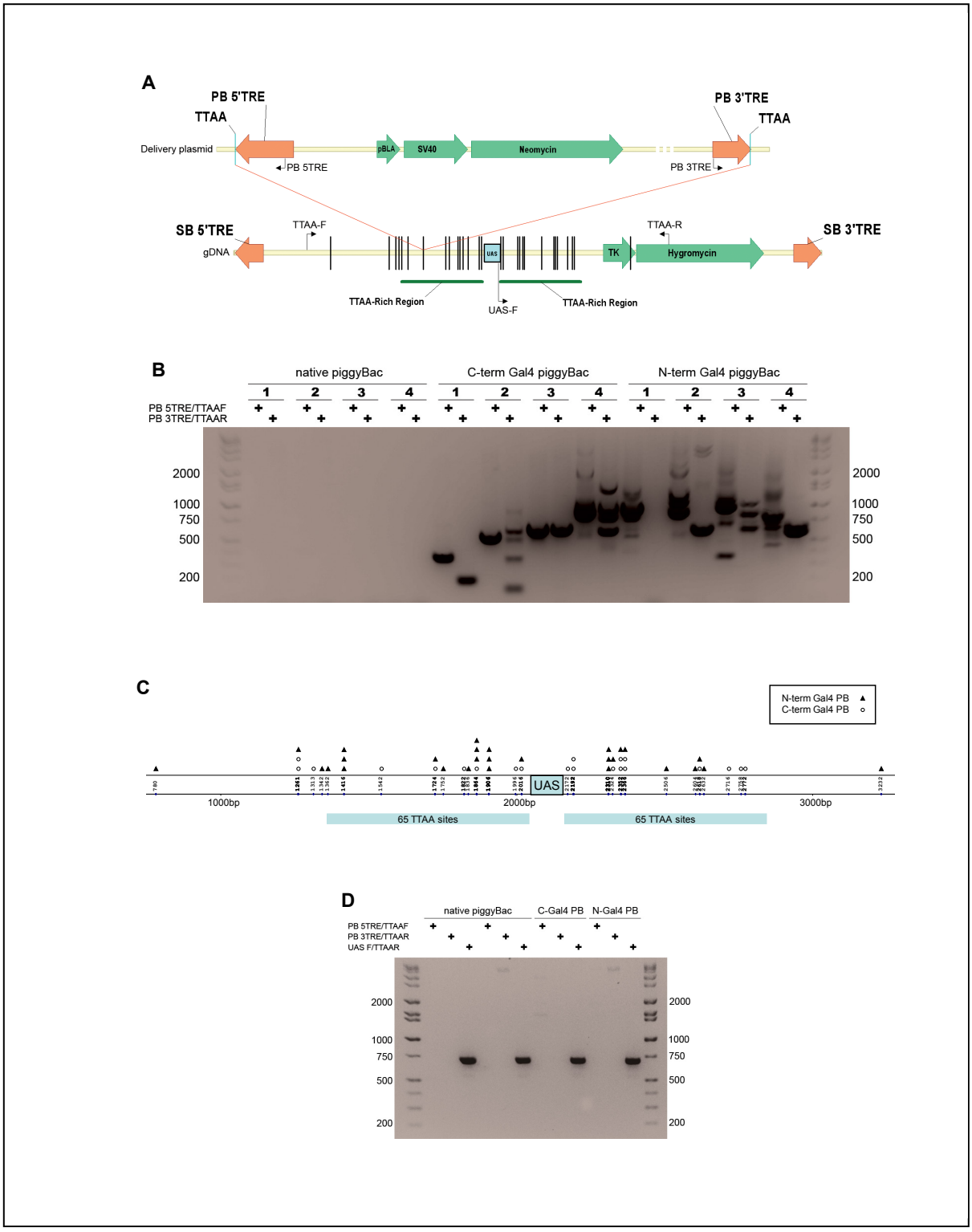
recipient plasmid were observed by fusing the Gal4 DBD to pB. (**C**) Percentages of integration sites recovered at increasing distances from the UAS. N- and C-terminal Gal4 pB integrate closer to the UAS on the recipient plasmid compared to native pB.

Table 1 Distances of recovered plasmid into plasmid insertions sites from UAS targets

Plasmid to Plasmid	N-terminal Gal4 PB	Native PB	C-terminal Gal4 PB
insertions < 800bp from UAS	87%[a]	59%	77%[b]
insertions < 250bp from UAS	47%[c]	21%	32%[d]

Both N- and C-terminal Gal4 PB fusion constructs significantly biased integration near the UAS compared to native PB by the Fisher's exact test. P-values: a=0.002, b=0.01, c=0.005, d=0.05

4.6.2 Table 4-1.



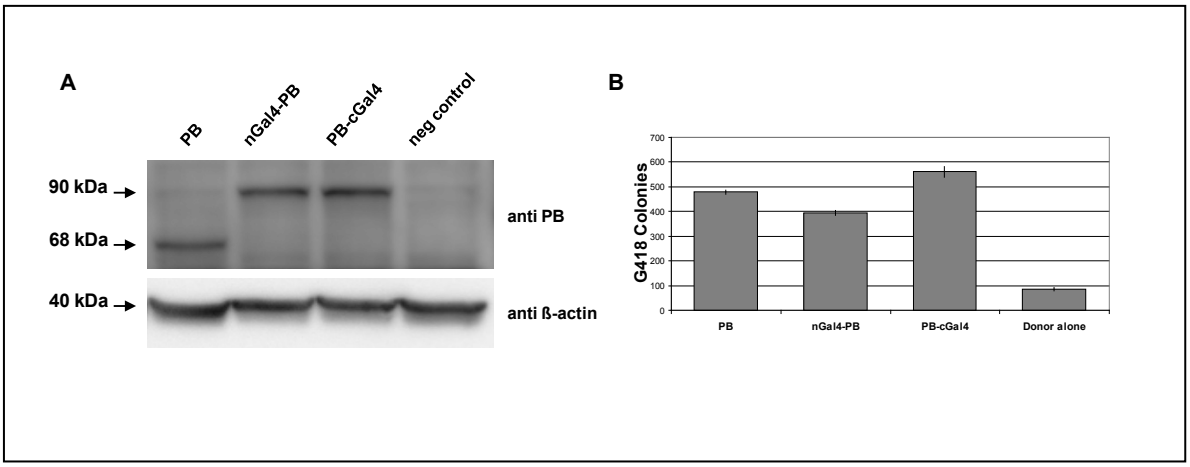
4.6.3 Figure 4-2. (**A**) Schematic for the genomic DNA targeting experiment. Helper-independent pmGENIE plasmids containing both the delivery transposon as well as the pB coding sequence were used in both the plasmid into plasmid and genomic targeting experiments. The neomycin gene is driven by both bacterial (pBLA) and eukaryotic (SV40)

promoters. Sleeping Beauty and the SB recipient transposons were encoded on different plasmids. The recipient transposon harboring the UAS target was first integrated into gDNA by SB and stable integrants were selected for with hygromycin. A second transfection was performed with the pmGENIE delivery plasmid containing both the pB transposase and the pB transposon with the neoR gene. Insertions by the chimeric Gal4-pB transposase that had been directed to the vicinity of the UAS target (red lines) were detected by nested PCR (primers represented by black arrows). Black vertical lines represent actual distances of insertions recovered from the UAS on the SB transposon. UAS-F and TTAA-R served as positive control primers to verify the presence of the SB transposon target. (**B**) Evidence for genomic targeting and the requirement of Gal4. Genomic DNA from hygroR/G418R cell populations transfected first with SB11 and UAS-SB transposon then next with delivery plasmids containing pB, pB-cGal4 or nGal4-pB was isolated and analyzed by PCR. Shown is a representative gel displaying 4 of 6 independent samples each of nested products recovered for both Gal4 fusions but not native pB. (**C**) Schematic map of the UAS-SB target transposon showing integrations of pB donor elements using the Gal4-pB chimeric transposase. Open circles and closed triangles represent insertions by C terminal and N terminal Gal4-pB respectively. The vertical numbers represent the nucleotide location of targeted TTAA sites on the UAS-SB transposon. The UAS was flanked on both sides by 65 TTAA sites spaced 10bp apart. (**D**) Evidence for the requirement of the UAS. Stable cells transfected with the UAS-negative SB recipient transposon were re-transfected with pB delivery plasmids. Shown is a control gel displaying PCR products for positive control UAS F/TTAAR but not products from targeting primers for 2 native pB samples and both Gal4 fusion samples (HI-LO DNA Marker, Bionexus).

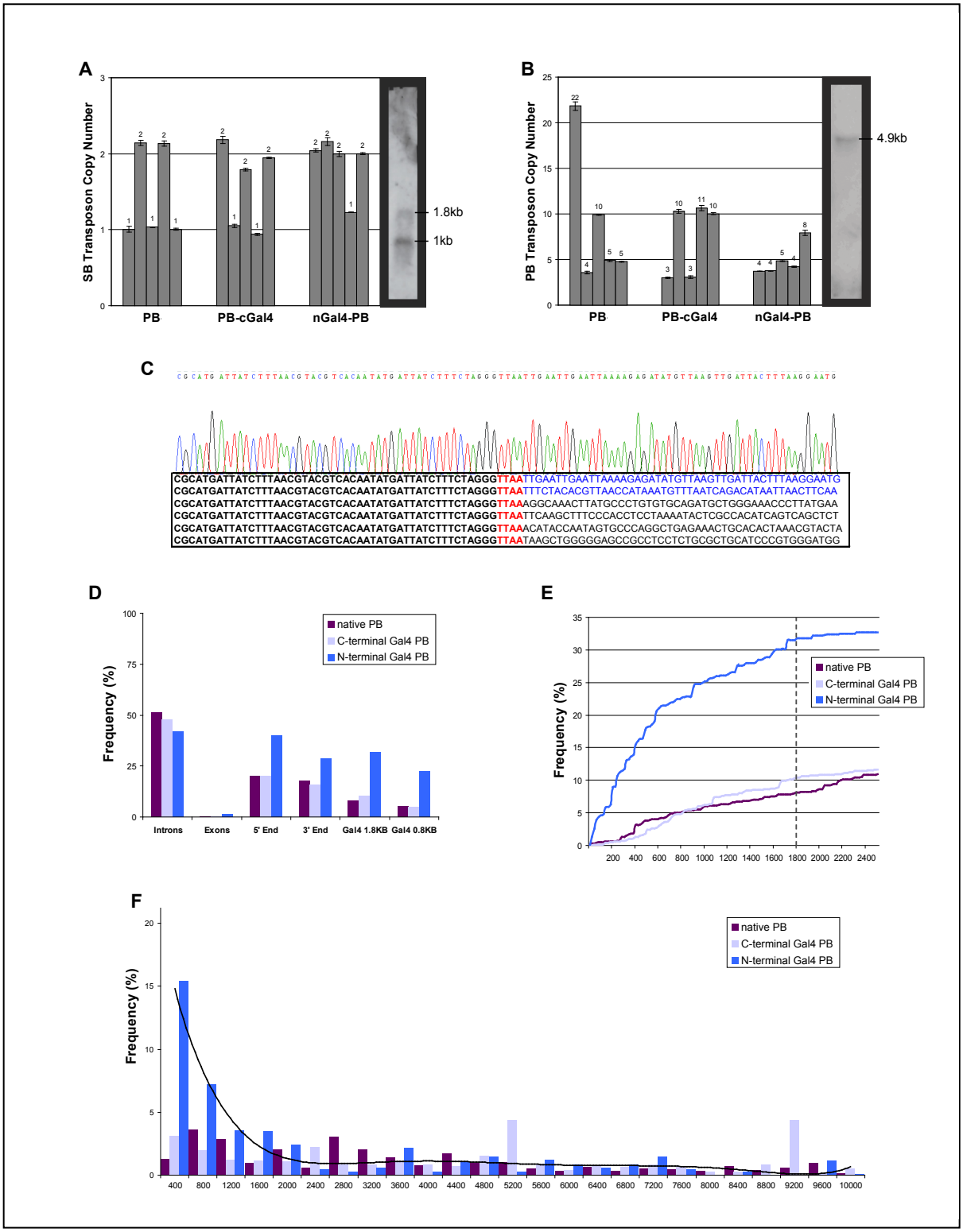
Table 2 Total genomic insertions into the SB recipient transposon recovered by nested PCR

HEK 293 Cell Line	N-terminal Gal4 PB	Native PB	C-terminal Gal4 PB
UAS +	28	0	21
UAS -	0	0	0

4.6.4 Table 4-2.



4.6.5 Figure 4-3. (**A**) Western blot comparison of protein levels of native pB, N-terminal Gal4 pB, and C-terminal Gal4 pB. The larger size for nGal4-pB and pB-cGal4 samples is due to the Gal4 DBD fusion to the transposase. The negative control used was untransfected HEK293 cells. (**B**) Comparison of plasmid integration efficiencies in HEK293T cells. Number of G418 resistant colonies from transfections with transposase-negative control, native pB, C-terminal Gal4 pB, and N-terminal Gal4 pB (n=3, mean ± SD).


A
SB Transposon Copy Number
1.8kb
1kb
PB
PB-cGal4
nGal4-PB
B
PB Transposon Copy Number
4.9kb
PB
PB-cGal4
nGal4-PB
C
CGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATTGAATTGAATTAAAAGAGATATGTTAAGTTGATTACTTTAAGGAATG
CGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATTTCTACACGTTAACCATAAATGTTTAATCAGACATAATTAACTTCAA
CGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAAAGGCAAACTTATGCCCTGTGTGCAGATGCTGGGAAACCCTTATGAA
CGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATTCAAGCTTTCCCACCTCCTAAAATACTCGCCACATCAGTCAGCTCT
CGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAAACATACCAATAGTGCCCAGGCTGAGAAACTGCACACTAAACGTACTA
CGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATAAGCTGGGGGAGCCGCCTCCTCTGCGCTGCATCCCGTGGGATGG
D
native PB
C-terminal Gal4 PB
N-terminal Gal4 PB
Frequency (%)
Introns
Exons
5' End
3' End
Gal4 1.8KB
Gal4 0.8KB
E
native PB
C-terminal Gal4 PB
N-terminal Gal4 PB
Frequency (%)
F
native PB
C-terminal Gal4 PB
N-terminal Gal4 PB
Frequency (%)

4.6.6 Figure 4-4. (**A**, **B**) Copy number assays for number of SB and pB transposon integrations. gDNA from 5 single clones each for N and C terminal Gal4 fusion and native pB experiments was analyzed by duplex Taqman real-time PCR. The numbers above the bars represent the estimated copy number for each sample. The Southern blot shown on the right of each graph was applied as a standard of known number of transposon integrations and was used to calibrate the qPCR data. (**C**) Sequences recovered from a representative sample showing the pB TRE on the left in bold, TTAA, and flanking sequence on the right. The top 2 lines with flanking sequence in blue show nested PCR products that align to the genomic UAS-SB recipient transposon. The bottom 4 lines with flanking sequence in black show recovered nrLAM and 454 sequences representing off-target events with alignments to various locations in the human genome. (**D**) The frequency of insertion sites recovered from nrLAM PCR that land within introns and exons, within a 10KB window surrounding transcriptional start sites (5' end) or polyA termination sites (3' end), and +/- 1.8kb and +/- 0.8kb of endogenous Gal4 recognition sites. (**E**) The cumulative percentage of total integrations from 0 to 2,400 bp from endogenous recognition sequences. The frequency of insertions for native pB and pB-cGal4 increased linearly. nGal4-pB insertion frequency increased logarithmically until 1800 bp and then increased linearly. (**F**) Histogram displaying the percentage of total integrations that occurred within 400 bp intervals from 0 to 10,000 bp from endogenous Gal4 recognition sequences. The black line represents the best fit curve for nGal4-pB.

Table 3 Frequencies of integration into intragenic regions and transcriptional start sites of RefSeq genes

	% of Integrations				
Genomic Location	Random[a]	PB	PB-cGal4	nGal4-PB	HIV[a,b,c]
In RefSeq genes	33.2	**60.8**	**57.8**	**61.5**	83.4
+/- 5KB from start site	5.4	**19.9**	**19.9**	**39.9**	11.4

Frequency of off-target integrations into genes and regions near transcriptional start sites, recovered from nrLAM PCR, compared to random and viral integration. Results from this study are boldfaced. [a]Values from the work of Yant et al. (72), [b]Adjusted values from the work of Narezkina et al. (71) and reported in Yant et al. (72), [c]Adjusted values from the work of Schroder et al. (73) and reported in Yant et al. (72), HIV human immunodeficiency virus.

4.6.6 Table 4-3.

Nested SB and PB Plasmid Primers

UAS-R Primary	CCTGCAGGTCGGAGTACTGT
UAS-F Primary	GCGCTCGCTAGAGTCTCC
TTAA-F Primary	TCCGGCCTTTATTCACATTC
TTAA-R Primary	GGCTAAATACGGAAGGATCTGA
UAS-R	GAGACTCTAGCGAGCGCAAC
UAS-F	ACCTGCAGGCATGGAAGC
TTAA-F	GCGTGTTACGGTGAAAACCT
TTAA-R	CAAAGGGAAAACTGTCCAATG
PB 5TRE Primary	CGACTACGCACTAGCCAACA
PB 5TRE	ACGGATTCGCGCTATTTAGA
PB 3TRE Primary	GGTGCACGAGGTAAGAGAGG
PB 3TRE	CCGATAAAACACATGCGTCA

4.6.7 Table 4-4.

nrLAM PCR Primers

5'TRE Linear1	CGACTACGCACTAGCCAACA
5'TRE Linear2	CACGCGGTCGTTATAGTTCA
3'TRE Linear1	GGTGCACGAGGTAAGAGAGG
3'TRE Linear2	TGTCTGGGAGTCCCTCTCAC
5'TRE Primary	CTTACCGCATTGACAAGCAC
3'TRE Primary	GGGAGTCCCTCTCACCAAA
Linker Primary	GACCCGGGAGATCTGAATTC
5'TRE For Nested	ACGGATTCGCGCTATTTAGA
3'TRE Rev Nested	CCGATAAAACACATGCGTCA
Linker Nested	AGTGGCACAGCAGTTAGG
Single stranded adaptor	5'-P-CCTAACTGCTGTGCCACTGAATTCAGATCTCCCGGGTddC -3'

4.6.8 Table 4-5.

Position weight matrix of Gal 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	0	0	0	¼	.	.	.	.	.	.	.	.	.	¼	0	0	0
C	1	0	0	¼	.	.	.	.	.	.	.	.	.	¼	1	1	0
G	0	1	1	¼	.	.	.	.	.	.	.	.	.	¼	0	0	1
T	0	0	0	¼	.	.	.	.	.	.	.	.	.	¼	0	0	0

4.6.9 Table 4-6.

4.7 Contributions

Research design: Jesse Owens, Stefan Moisyadi

Conducted experiments: Jesse Owens, Johann Urschitz, Ilko Stoytchev, and Nong C. Dang

Performed data analysis: Jesse Owens, Stefan Moisyadi

Wrote or contributed to writing: Jesse Owens, Johann Urschitz, Stefan Moisyadi

Chapter 5. Transcription activator like effector (TALE)-directed *piggyBac* transposition in human cells

5.1 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 (DSBs) 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 pB 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 pB constructs tethered to a custom transcription activator like effector (TALE) DNA-binding domain (DBD) 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.

5.2 Introduction

The pB transposable element can efficiently integrate transgenes into genomes of mammalian cells and organisms (36, 37, 39, 161). This nonviral 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 (1-3, 87). However, insertional mutagenesis and unknown position effects that may inhibit transgene expression remain obstacles for vectors that integrate randomly (7, 8, 162, 163). A method for user-defined directed integration would improve the safety of insertional therapies.

Engineered nucleases based on TALE, zinc finger, and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems have been used to induce DSBs at specific sites (21, 23, 24, 109, 114). 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 (25-29, 114, 118, 164-166).

Unlike engineered nucleases, transposons perform all the enzymatic steps required for integration (42). Furthermore, we have shown that a chimeric pB transposase fused to the Gal4 DBD can bias integration near endogenous Gal4 recognition

sequences (47). 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) (167, 168), which is required for the entry of R5 tropic HIV-1 strains involved in primary infections. A homozygous Δ32 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 (14). Consequently, clinical trials are exploring the possibility of disrupting CCR5 via targetable nucleases as part of an anti-HIV therapeutic approach (169).

Here we introduce novel constructs utilizing 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 approximately 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.

5.3 Experimental Procedures

5.3.1 Plasmid development

Simplified illustrations of all pB targeting plasmids are depicted in Figure 5-1. All targeting plasmids were derived from p*mhy*GENIE-3-R6K (abbreviated hG3) that encodes a self-inactivating (40, 43) hyperactive pB transposase (170) 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 transgene was made by swapping the puromycin gene in pGIPZ (Thermo Scientific) with a neomycin gene amplified from pERV3 (Agilent Technologies). The fragment including CMV, TurboGFP, IRES, and neomycin (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 (171), which contained truncated N- and C-termini of the naturally occurring TALE PthXo1 and flanking SfiI restriction sites. Full binding sites of the SfiI TALE cassettes are provided in Figure 5-6.

5.3.2 Cell transfections

Human embryonic kidney (HEK293) cells were maintained in complete DMEM supplemented with 10% heat inactivated fetal bovine serum. Prior to transfection, $2x10^5$ 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 two weeks under 200 µg/ml G418 at which point about 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, that 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 approximately 56 colonies per well. After wells became greater than 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.

5.3.3 Copy number assay

In order to determine the number of transposons present in CCR5-targeted single clones, a quantitative PCR copy number assay was performed as previously described (47).

5.3.4 Flow cytometry

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

5.3.5 Colony count assay

$1x10^5$ 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 (40). Cells were resuspended, diluted 1:100, and plated into 10 cm plates (1000 total cells per plate) and maintained for three weeks under G418 selection. The fraction of resulting GFP positive colonies greater than 1 mm in diameter were counted using a FluorVivo 100 fluorescence imaging system (INDEC Biosystems).

5.3.6 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 (171). 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 DMEM 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 1x DpBS, followed by lysis in 100 µL of 1x Passive Lysis Buffer (Promega) with 1x 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.

5.3.7 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 (tTTTAGCCTTACTGTTGA) 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 in order to identify insertion site locations. Primer sequences are listed in Table 5-3.

5.4 Results

5.4.1 Experimental strategies for targeting *piggyBac* transposition

Five unique targeting constructs were designed to localize transposition to the CCR5 locus. hG3-TALC1 (Figure 5-1A) employs 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 5-2A). This single plasmid includes both the transposon and transposase and is self-inactivating (40), 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 (47).

hGT1-TALC1 (Figure 5-1B) 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 5-2B). We conceived that the additional flexibility which may be achieved for this orientation, as

compared to 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 5-1C) 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 5-2C). In a similar fashion, hGT3-TALC1 (Figure 5-1D) 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 5-2D).

A modified version of hGT1-TALC1, called hGR1T1-TALC1 (Figure 5-1E) 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. In order 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 5-2E).

A two-plasmid strategy was devised using hG3-TALC1 combined with a similar plasmid to hGT1-TALC1 described above, 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 5-2F). Finally, a control plasmid (hG3) was constructed containing an unfused hyperactive pB. All constructs included a bicistronic CMV promoter driven TurboGFP and neomycin 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 5-2G).

5.4.2 Activities of *piggyBac* 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 approximately two weeks. HEK293 cells were transfected with each plasmid and grown under G418 selection. Three weeks later, GFP positive colonies were counted (Figure 5-3A). 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 approximately 11% of total transfected cells received pB integrations. This integration activity is in agreement with previous hyperactive pB rates in HEK293 cells (172) 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 5-3B). 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 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.

5.4.3 *piggyBac* 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 two weeks. In order to estimate if 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 (Figure 5-2G and 5-4A). A total of fourteen 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 5-1).

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. Nine of the fourteen insertion sites were located within 250 bp of the TALC1 sequence and two insertions were located 639 bp and 659 bp away. In addition, three insertions at distances of 1231 bp, 3495 bp, and 3991 bp were recovered far from the target sequence. (Figure 5-4B and Table 5-1). 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.

5.4.4 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 (Table 5-1), 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 5-5A). Each well was resuspended and a fraction 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 twenty clonal expansions isolated from a control hG3 transfection did not give rise to PCR products (Figure 5-5B). 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 5-5C). 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 ten weeks of culture (Figure 5-5D). Populations expanded from clones 293-1 and 293-2 were found to be 99.9% and 98.0% GFP positive, respectively.

5.4.5 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 5-1). 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 above, 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 1 in 5,376 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 sixteen 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 5-2).

5.5 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 (173). 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 (1-3) thereby transforming the gene therapy field.

In an effort to reduce the risks of random viral insertion, Papapetrou et al. 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 (174). 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 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, a Human Immunodeficiency Virus (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 to wild-type HIV integrase (121). 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 (175). 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, employing 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 (47, 133-137, 142, 154, 176). Recently, endogenous transpositional targeting has been achieved (47, 176). 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, SB, and *Tol2* transposons near both minimal Rep binding sequences (15726 sites per human genome) and consensus Rep binding sequences (2134 sites per human genome) (176). 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 (47). 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 (177). 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 bp - 3991 bp near the TALE recognition sequence. Two "hot-spots", 24 bp 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 5-1). It is likely that these polyclonal populations contained additional targeted insertions, however the percentage of targeted cells was very 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 (178). 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 (170, 172) 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.

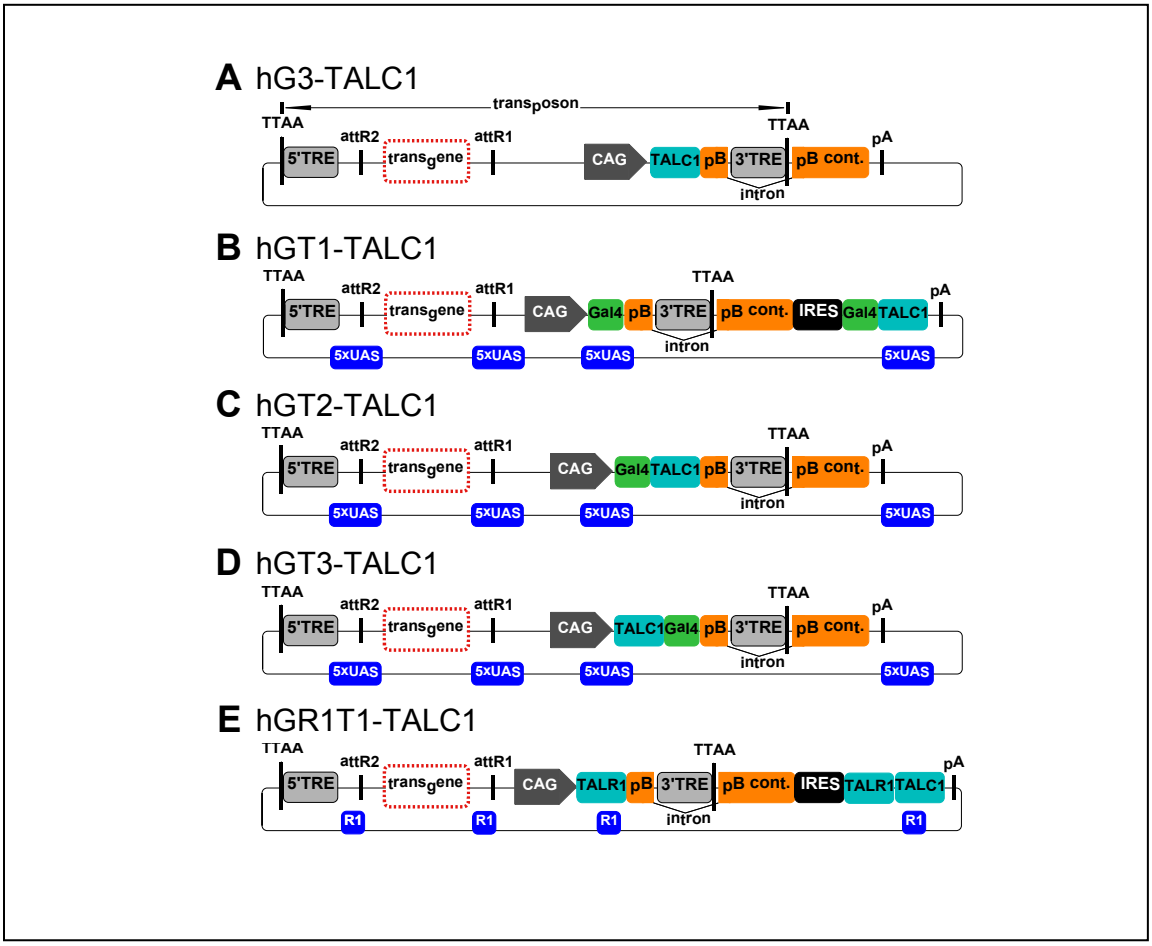
The pB transposase is autonomously functional in our system and therefore is able to integrate into many genomic locations. A major improvement to the system would be to make the localization or binding of pB to the genomic target a *required* event for transposition. This might be achieved by mutating 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. Upon 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. described excision competent/integration defective transposases with mutations in pB's DBD (179). 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 (33, 180-184) and easy-to-implement modifications to both zebrafish and rat genomes have become a possibility (185, 186). 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 (170). 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 (27, 29, 33, 181).

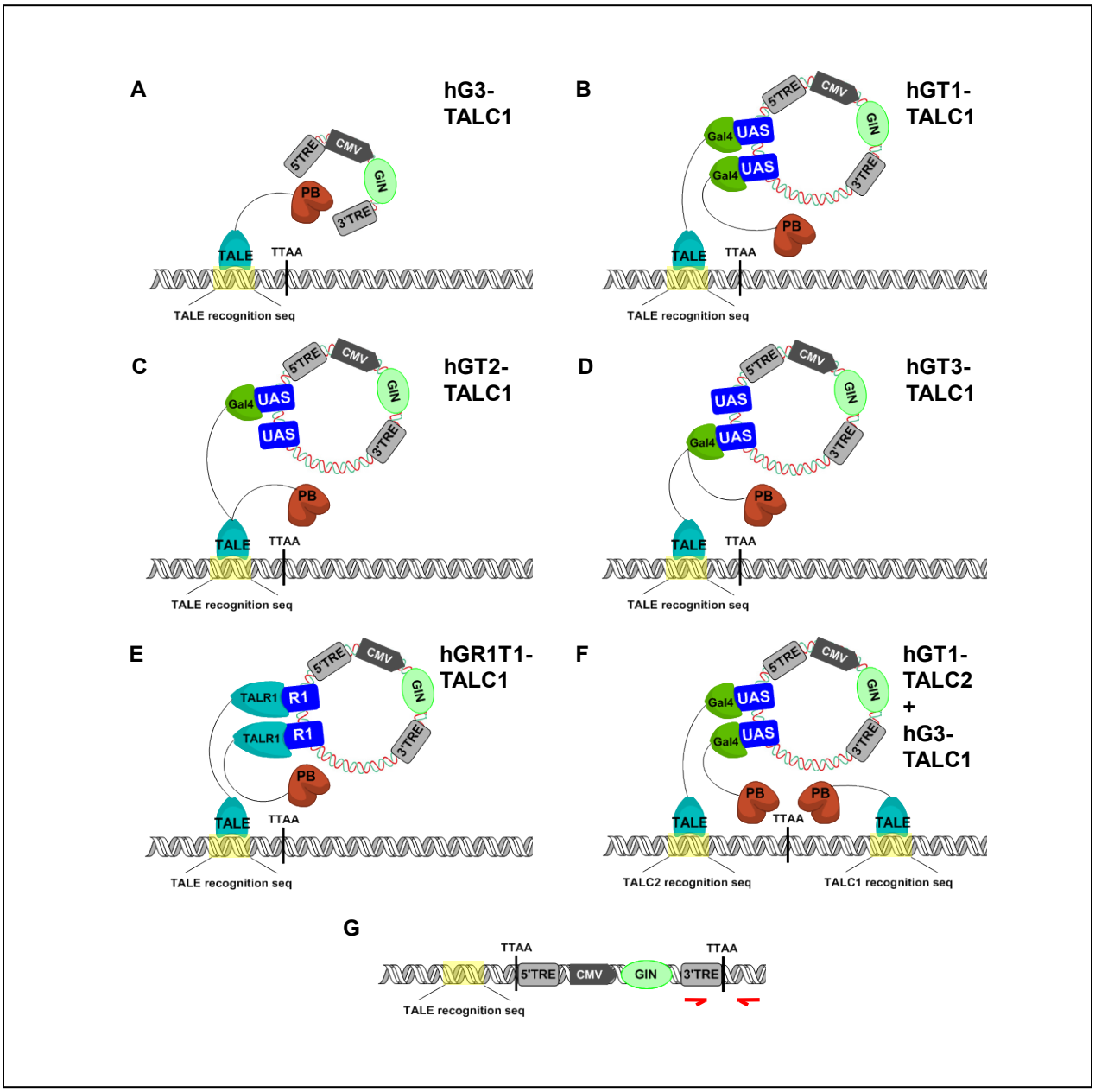
The pB system can permanently introduce large cassettes (>100 kb) encoding numerous components such as multiple transgenes, insulators, and inducible or endogenous promoters (187). The current study has laid the groundwork for enhancing this system by allowing researchers to potentially target integrations 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 in order to protect the cargo

from unknown chromosomal position effects and to circumvent accidental mutation of target cells.

5.6 Figures

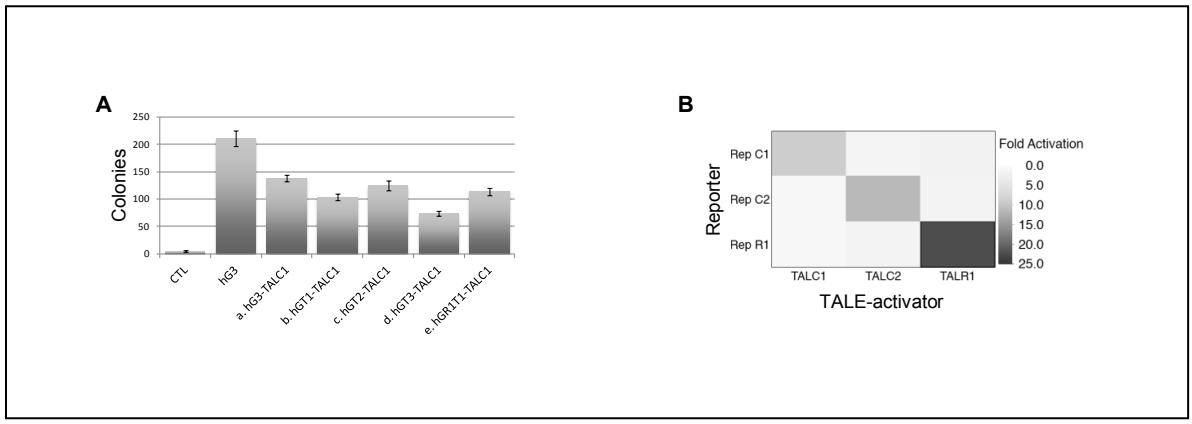


5.6.1 Figure 5-1. pB targeting plasmids. The terminal repeat elements (TREs) flank the pB transposon. The 3'TRE resides within an introduced intron in the pB gene leading to inactivation of the transposase upon excision of the transposon. The CAG promoter drives expression of the pB targeting proteins and the internal ribosomal entry site (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

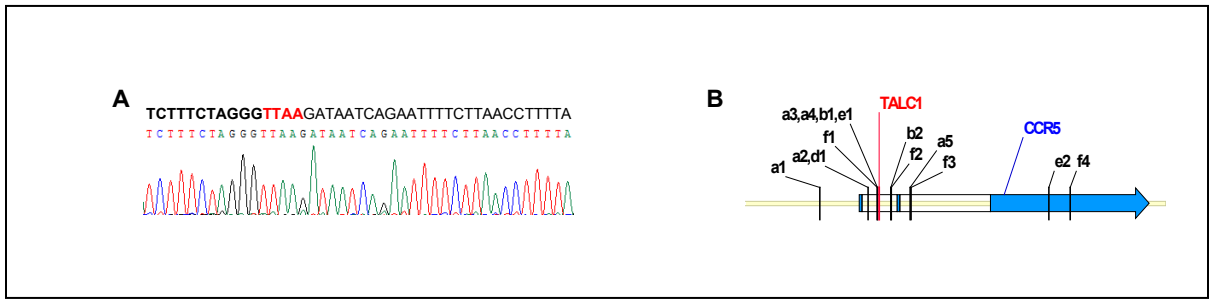


5.6.2 Figure 5-2. Schematic for various pB targeting strategies. For each strategy, all components are encoded on a single plasmid and most components have been omitted for simplification purposes from plasmids depicted in this figure. (**A**) hG3-TALC1 encodes a hyperactive pB transposase covalently 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.



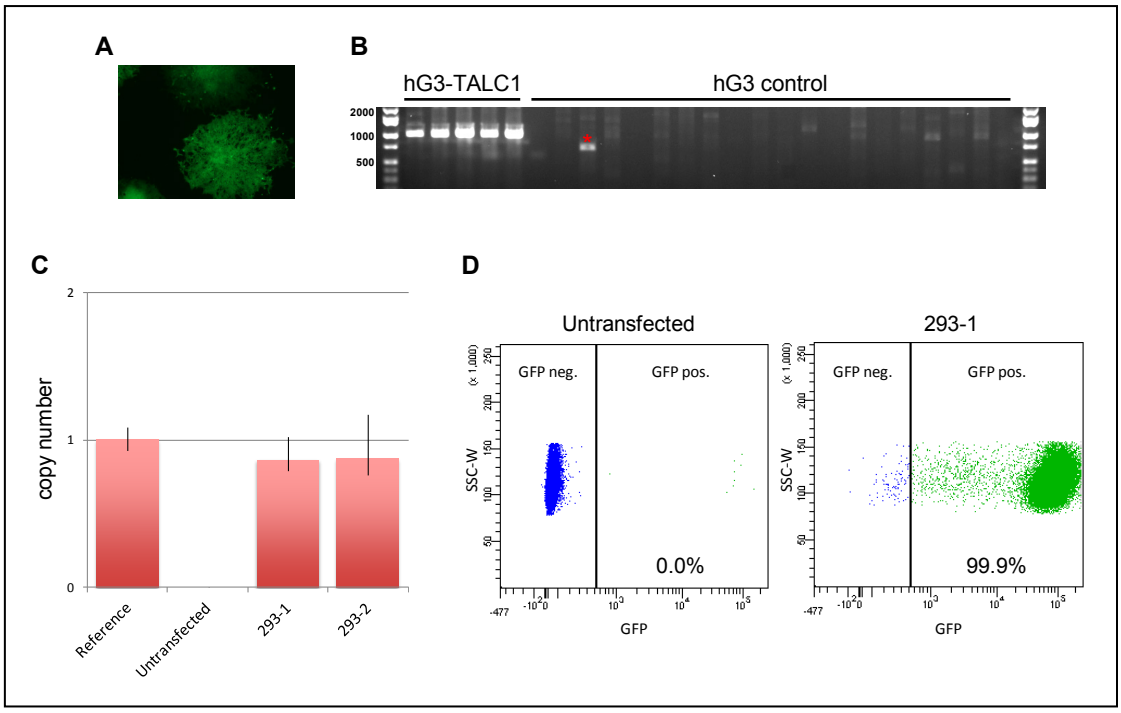
5.6.3 Figure 5-3. Verification of transposase and TALE activity. (**A**) Comparison of integration efficiencies between pB constructs transfected into HEK293 cells. One thousand cells were plated and cultured for three weeks before G418 resistant/GFP+ colonies were counted. Data are shown as mean values 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).



5.6.4 Figure 5-4. (**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, and g. hG3 (no insertions).

construct	is #	trans-fection	distance to TALC1	transposon orientation	flanking CCR5 sequence
hG3-TALC1	a1	#2	1231 bp	For.	**TTAA**TCAATGCCTT
	a2	#3	221 bp	For.	**TTAA**AACTCTTTAG
	a3	#1	24 bp	For.	**TTAA**GATAATCAGA
	a4	#4	24 bp	For.	**TTAA**GATAATCAGA
	a5	#2	639 bp	For.	**TTAA**AGGGAGCAA
hGT1-TALC1	b1	#5	24 bp	Rev.	**TTAA**GATAATCAGA
	b2	#1	236 bp	Rev.	**TTAA**GCTCAACTTA
hGT2-TALC1	c0	--	--	--	--
hGT3-TALC1	d1	#3	221 bp	Rev.	**TTAA**AACTCTTTAG
hGR1T1-TALC1	e1	#5	24 bp	For.	**TTAA**GATAATCAGA
	e2	#1	3495 bp	For.	**TTAA**AAGGAAGTTA
hG3-TALC1 + hGT1-TALC2	f1	#4	37 bp	For.	**TTAA**TAGCAACTCT
	f2	#6	247 bp	Rev.	**TTAA**AAGGAAGAAC
	f3	#5	659 bp	For.	**TTAA**TAACTAACAA
	f4	#3	3991 bp	For.	**TTAA**ATGAGAAGGA
hG3	g0	--	--	--	--

5.6.5 Table 5-1.

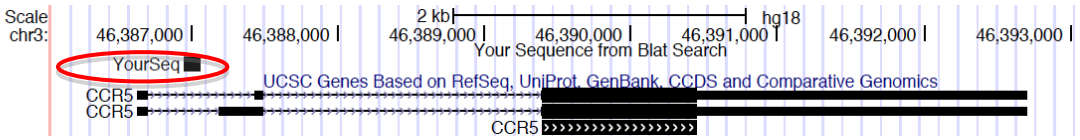


5.6.6 Figure 5-5: (**A**) Cells from a hG3-TALC1 transfection were plated into a 96-well plate and one week later, individual non-overlapping colonies were established for counting. 40x magnification. (**B**) Nested PCR for the identification of transposition near CCR5. Expected products arose from 5 positive clones identified from hG3-TALC1 transfection, but not clones from hG3 control transfection. ***** 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 ten weeks of culture. Flow cytometry analysis displaying GFP positive events for both untransfected HEK293 cells and an expansion of clone 293-1.

	hG3-TALC1 #1	hG3-TALC1 #2	hG3-TALC1 combined	hGT1-TALC1 #1
total cells screened	13,686	7,622	21,308	10,476
positive wells	2	1	3	1
% targeted cells	0.015	0.013	0.014	0.010

5.6.7 Table 5-2.

TALC1 AND TALC2 SITES



Intron 1 of human CCR5 chr3:46,411,950-46,412,060 (hg19)

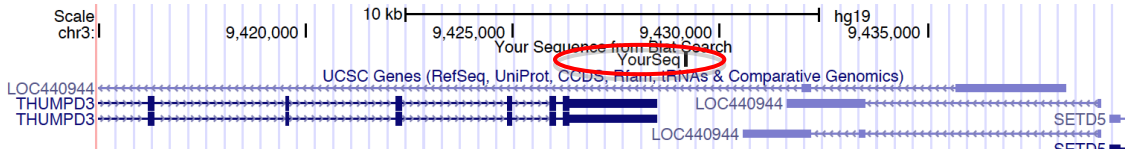
TALEC2 (16.5)

TALEC1 (16.5)

GAAGTCCAGGATCCCCCTCTACATTTAAAGTT

GGTTTAAGTTGGCTTTAATTAATAGCAACTC**TTAA**GATAATCAGAATTTTCTTAACCTTTTAGCCTTACTGTTGAAAAG

TALR1 SITE

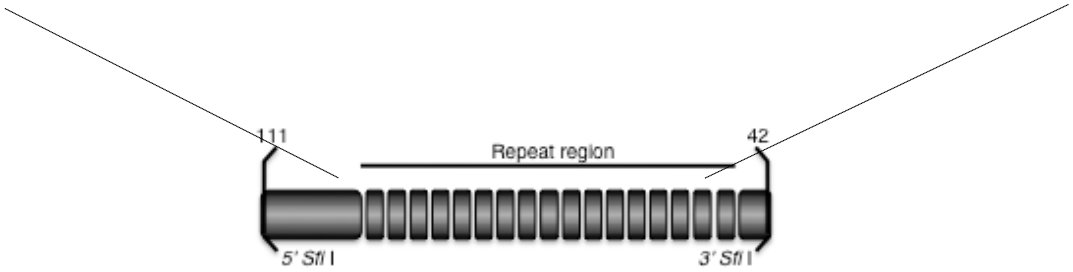


Human ROSA26 locus chr3:9,429,120-9,429,190 (hg19)

TALER1 (16.5)

TGAAG**TTAA**GGAAATCTGCTTCCTGTCCTAGCATGCCCATTATCCCAGCCATACAGAT**TTAA**TACCAGGAG

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	16.5
TALC1	T	T	T	A	G	C	C	T	T	A	C	T	G	T	T	G	A
	NG	NG	NG	NI	NN	HD	HD	NG	NG	NI	HD	NG	NN	NG	NG	NN	NI
TALC2	C	C	A	G	G	A	T	C	C	C	C	C	T	C	T	A	C
	HD	HD	NI	NN	NN	NI	NG	HD	HD	HD	HD	HD	NG	HD	NG	NI	HD
TALR1	C	C	T	A	G	C	A	T	G	C	C	C	A	T	T	A	T
	HD	HD	NG	NI	NN	HD	NI	NG	NN	HD	HD	HD	NI	NG	NG	NI	NG



5.6.8 Figure 5-6. TALC1, C2, and R2 binding sites.

Primer sequences

CCR5 For Primary	GGGCACGTAATTTTGCTGTT
CCR5 For	ACCGCCAAGAGAGCTTGATA
CCR5 Rev Primary	TCCCTCTTGTCTGGAGGAAA
CCR5 Rev	TCAGAAGGCATCTCACTGGA
PB 5TRE Primary	CGACTACGCACTAGCCAACA
PB 5TRE	ACGGATTCGCGCTATTTAGA
PB 3TRE Primary	GGTGCACGAGGTAAGAGAGG
PB 3TRE	CCGATAAAACACATGCGTCA

5.6.9 Table 5-3.

5.7 Contributions

Research design: Jesse Owens, Stefan Moisyadi

Conducted experiments: Jesse Owens, Damiano Mauro, Ilko Stoytchev, and Mital Bhakta

Performed data analysis: Jesse Owens, Stefan Moisyadi

Wrote or contributed to writing: Jesse Owens, Stefan Moisyadi

Chapter 6. Discussion

Non-viral gene therapy approaches hold enormous promise for individuals with genetic disorders that are difficult to treat due to a lack of pharmaceuticals. Realizing the full potential of integrating vectors will require reliable methods for performing gene targeting. A major challenge is reducing the risk of insertional mutagenesis due to random insertion. Importantly, the propensity of viral vectors to preferentially integrate into active genes is a major concern.

Targetable nucleases based on clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas9) and transcription activator like effector (TALE) nuclease (TALEN) systems are capable of inducing double stranded breaks (DSBs). These DSBs can 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.

We have recently improved the non-viral pB system and have demonstrated applications for these vectors. Unlike nucleases, pB can autonomously perform all excision and integration steps without activating the cellular DNA damage response pathway. Furthermore, nuclease-mediated homology directed repair occurs at the S and G2 cell cycle stages of replication. This requirement for cell division excludes possible therapies for the majority of tissues for nuclease-based approaches. Our unique self-inactivating pB vectors (GENIE) are intended to reduce cell genotoxicity.

Following pB expression and excision of the transposon, the transposase becomes inactivated and therefore cannot persist should it become accidentally integrated into the genome. We have demonstrated multiple applications for this single-plasmid vector (40, 43, 47, 48, 188, 189), including the stable knockdown of telomerase in cancer cells in order to inhibit growth described in Chapter 3 (189). GENIE vectors also effectively integrate into liver cells *in-vivo* when loaded onto microbubbles that are subsequently destroyed using ultrasound at the intended target tissue (188). The highest reported rates of transgenic animal production have been achieved using this vector (43) and we are currently using it to introduce transgenes for the recovery of damaged ischemic heart tissue (unpublished). Although GENIE plasmids can potentially integrate at any TTAA site in the genome, we have shown that these vectors display less of a preference for integration into genes compared to viruses (47).

In an effort to avoid vector-based mutagenesis, we have developed an alternative strategy to nuclease-based approaches using chimeric pB-DBD fusion proteins and demonstrated the ability to deliver genes to predefined genomic locations. Chimeric pB-DBD proteins have been shown to be able to direct integrations to plasmid recipients harboring the upstream activating sequence (UAS) of the Gal4 recognition site in *Aedes aegypti embryos* (80). Of significance, our lab was the first to report no loss in activity for such chimeric fusion proteins (39). In contrast, fusion of DBDs to other transposases such as *Sleeping Beauty* (*SB*) were shown to result in a significant or even complete reduction of transpositional activity (39, 136, 137).

To further explore the potential of using a chimeric pB for gene therapy we analyzed the ability of Gal4-pB fusion proteins to target integration near UAS recognition sites in mammalian cells (91). We used a plasmid into plasmid approach in which the recipient plasmid contained a chloramphenicol gene and a UAS target site. The delivery plasmid contained the pB transposase, with or without a Gal4-DBD and a transposon delivery cassette harboring the neomycin (kanR) gene for bacterial selection. Integration of the delivery cassette into the recipient plasmid conferred double resistance to camR/kanR when transformed into *E. coli*. Transfection experiments in HEK293 cells suggested that, in our plasmid model, the addition of a Gal4-DBD to either N- or C-terminal end of the pB protein confers a propensity for transposition near the UAS site. In order to determine whether targeting could be achieved in a genomic setting, we used a recipient transposon containing the UAS flanked by the TREs of SB. This allowed for random integration of SB transposon targets into the genome of HEK293 cells, which in turn were targeted by the chimeric Gal4-pB. In order to detect targeted genomic insertion we used nested PCR with the forward set of primers complementary to the pB transposon and the reverse primers extending from the SB transposon target. As expected, we did not detect any targeted PCR products for the UAS-positive cell populations that had been transfected with native pB control transposase, clearly demonstrating that both the UAS recognition sequence, as well as the Gal4-DBD fused to the pB transposase were required for enhanced genomic targeting. Within the 8000 bp region that we analyzed, all of the 49 unique integrations recovered localized within 1300 bp of the

UAS site with the vast majority (96%) found <800 bp up- or down-stream from the UAS site. Gal4 is a Zn2/Cys6 zinc finger with a 6 bp binding site that occurs not only in our inserted UAS sites, but also at many endogenous human loci (151). To estimate the efficiency of the Gal4-pB fusion protein's ability to insert near endogenous Gal4 recognition sites, we annotated the 56898 UAS-like sites found in the human genome and examined insertions located in the vicinity of these sites. A high-throughput insertion site analysis of more than 7000 unique genomic integration sites revealed that the cumulative percentage of integrations recovered dramatically increases up until 1800bp for nGal4–pB but not for pB–cGal4 or native pB. In summary, we were able to demonstrate the ability of our fusion constructs to bias 24% of integrations near endogenous Gal4 recognition sequences and that the N-terminal configuration of the chimeric pB-DBD is superior to the C-terminal conformation (91). This represent the first demonstration of directed pB insertion to both introduced and endogenous recognition sequences in the genome.

This *proof of principle* encouraged us to develop a TALE-based targetable transposon system that has recently been used by us to direct integration into a single genomic address in order to provide a basis for safe clinical therapy. We further improved the pB system by designing a series of novel hyperactive pB constructs tethered to custom TALEs designed to bind a *unique site* in the human CCR5 genomic safe harbor. Multiple targeting strategies were evaluated using combinations of both plasmid-DNA and transposase-protein re-localization to the target sequence. Importantly, we successfully isolated cells containing single

targeted insertions (48). Given that the genome contains millions of potential off-target integration sites available for random integration of the transposon, these data demonstrate that indeed, TALEs display the ability to direct the pB transposase to a predetermined genomic locus. To our knowledge, this is the first demonstration of a transposase facilitating TALE-mediated targeted gene transfer into a single user-defined endogenous genomic location.

Currently the most useful application for gene targeting with pB is *ex-vivo* gene addition. This type of therapy could theoretically be carried out by the following steps: 1) Primary cells with the capacity to clonally expand, such as induced pluripotent stems cells (iPSCs), are isolated from a patient with a genetic deficiency. 2) These cells are transfected with targeting GENIE plasmids containing a transposon cargo encoding an expression cassette for the gene of interest. These vectors could be targeted to preferentially insert their cargo into a genomic safe harbor in order to provide uninhibited expression and eliminate risks of insertional mutagenesis at undesired sites. 3) Clones are single-cell expanded. 4) Clones are screened for inserts at the desired target sequence. Clones are also screened for the number of insertions. 5) Clones containing only a single insertion for which it is targeted are used for downstream therapies such as differentiating into a useful cell-type from corrected iPSCs. One such example includes a study using parkinsonian rats that exhibited robust long term engraftment of midbrain dopamine neurons efficiently derived from human iPSCS. Following transplant, these rats showed complete rescue of amphetamine induced rotational asymmetry (190).

A major improvement to our targetable pB system would be a method for identifying correctly targeted from off-target cells. We are currently working on a "event detection" approach that will use nano-flare probes that can be used to FACS live cells expressing specific mRNA sequences (191). The safe harbor ROSA26 transcriptional region will be targeted such that the insertion cassette will be appended to an endogenous non-coding mRNA. A nano-flare probe will span the junction flanking the insertion so that cells transcribing the uniquely modified mRNA will fluoresce. No activator/repressor domains or reporter/selection elements are required; therefore this event detection strategy could demonstrate feasibility for clinical applications.

Possibly the most important result from these studies is the creation of the first known base-line efficiency for targetable transposition to a user-defined sequence. Because only a fraction of cells receive targeted inserts, this technology is not yet ready for use *in-vivo*, as the majority of cells would receive off-target integrations. However the major step that is required for improvements using directed evolution, is the base-line activity for which evolution can begin to change the properties of our transposase. This activity is absolutely required to kick-start improvements to our system. Someday it may be possible for us to generate targetable transposases that require binding to a specific sequence in order to mediate insertion. Should such a vector be achievable, we envision our targetable pB to have broad potential for ameliorating risks involved with current gene therapy approaches. The current

bottleneck is not the efficiency of gene addition but rather safety issues associated with these approaches. Therefore this work directly relates to our ultimate goal of enabling gene therapy to be a universal practice for combating not just terminal illnesses but a broad-spectrum of disease.

References

1. Daniel R, Smith JA. Integration site selection by retroviral vectors: molecular mechanism and clinical consequences. Hum Gene Ther. 2008;19(6):557-68. Epub 2008/06/07. doi: 10.1089/hum.2007.148. PubMed PMID: 18533894; PubMed Central PMCID: PMC2940482.

2. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS biology. 2004;2(8):E234. Epub 2004/08/18. doi: 10.1371/journal.pbio.0020234. PubMed PMID: 15314653; PubMed Central PMCID: PMC509299.

3. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science (New York, NY. 2003;300(5626):1749-51. Epub 2003/06/14. doi: 10.1126/science.1083413. PubMed PMID: 12805549.

4. Kang Y, Moressi CJ, Scheetz TE, Xie L, Tran DT, Casavant TL, et al. Integration site choice of a feline immunodeficiency virus vector. Journal of virology. 2006;80(17):8820-3. Epub 2006/08/17. doi: 10.1128/JVI.00719-06. PubMed PMID: 16912328; PubMed Central PMCID: PMC1563849.

5. Li H, Malani N, Hamilton SR, Schlachterman A, Bussadori G, Edmonson SE, et al. Assessing the potential for AAV vector genotoxicity in a murine model. Blood. 2011;117(12):3311-9. Epub 2010/11/26. doi: 10.1182/blood-2010-08-302729. PubMed PMID: 21106988; PubMed Central PMCID: PMC3069673.

6. Nakai H, Montini E, Fuess S, Storm TA, Grompe M, Kay MA. AAV serotype 2 vectors preferentially integrate into active genes in mice. Nat Genet. 2003;34(3):297-302. Epub 2003/06/05. doi: 10.1038/ng1179. PubMed PMID: 12778174.

7. Baum C. Insertional mutagenesis in gene therapy and stem cell biology. Current opinion in hematology. 2007;14(4):337-42. Epub 2007/05/31. doi: 10.1097/MOH.0b013e3281900f01. PubMed PMID: 17534158.

8. Fehse B, Roeder I. Insertional mutagenesis and clonal dominance: biological and statistical considerations. Gene Ther. 2008;15(2):143-53. Epub 2007/11/02. doi: 10.1038/sj.gt.3303052. PubMed PMID: 17972922.

9. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. The Journal of clinical investigation. 2008;118(9):3132-42. Epub 2008/08/09. doi: 10.1172/JCI35700. PubMed PMID: 18688285; PubMed Central PMCID: PMC2496963.

10. Howe SJ, Mansour MR, Schwarzwaelder K, Bartholomae C, Hubank M, Kempski H, et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. The Journal of clinical investigation. 2008;118(9):3143-50. Epub 2008/08/09. doi: 10.1172/JCI35798. PubMed PMID: 18688286; PubMed Central PMCID: PMC2496964.

11. Deichmann A, Hacein-Bey-Abina S, Schmidt M, Garrigue A, Brugman MH, Hu J, et al. Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. The Journal of clinical investigation. 2007;117(8):2225-32. Epub 2007/08/03. doi: 10.1172/JCI31659. PubMed PMID: 17671652; PubMed Central PMCID: PMC1934585.

12. Yu SF, von Ruden T, Kantoff PW, Garber C, Seiberg M, Ruther U, et al. Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. Proceedings of the National Academy of Sciences of the United States of America. 1986;83(10):3194-8. Epub 1986/05/01. PubMed PMID: 3458176; PubMed Central PMCID: PMC323479.

13. Hackett PB, Largaespada DA, Cooper LJ. A transposon and transposase system for human application. Mol Ther. 2010;18(4):674-83. Epub 2010/01/28. doi: 10.1038/mt.2010.2. PubMed PMID: 20104209; PubMed Central PMCID: PMC2862530.

14. Lobritz MA, Ratcliff AN, Arts EJ. HIV-1 Entry, Inhibitors, and Resistance. Viruses. 2010;2(5):1069-105. Epub 2010/05/01. doi: 10.3390/v2051069. PubMed PMID: 21994672; PubMed Central PMCID: PMC3187606.

15. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med. 2014;370(10):901-10. Epub 2014/03/07. doi: 10.1056/NEJMoa1300662. PubMed PMID: 24597865.

16. Paulk NK, Wursthorn K, Wang Z, Finegold MJ, Kay MA, Grompe M. Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo. Hepatology. 2010;51(4):1200-8. Epub 2010/02/18. doi: 10.1002/hep.23481. PubMed PMID: 20162619; PubMed Central PMCID: PMC3136243.

17. Miller DG, Wang PR, Petek LM, Hirata RK, Sands MS, Russell DW. Gene targeting in vivo by adeno-associated virus vectors. Nature biotechnology. 2006;24(8):1022-6. Epub 2006/08/01. doi: 10.1038/nbt1231. PubMed PMID: 16878127.

18. Vasileva A, Jessberger R. Precise hit: adeno-associated virus in gene targeting. Nature reviews Microbiology. 2005;3(11):837-47. Epub 2005/11/02. doi: 10.1038/nrmicro1266. PubMed PMID: 16261169.

19. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science (New York, NY. 2009;326(5959):1509-12. Epub 2009/11/26. doi: 10.1126/science.1178811. PubMed PMID: 19933107.

20. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science (New York, NY. 2009;326(5959):1501. Epub 2009/11/26. doi: 10.1126/science.1178817. PubMed PMID: 19933106.

21. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic acids research. 2011;39(1):359-72. Epub 2010/08/12. doi: 10.1093/nar/gkq704. PubMed PMID: 20699274; PubMed Central PMCID: PMC3017587.

22. Gaj T, Gersbach CA, Barbas CF, 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology. 2013;31(7):397-405. Epub 2013/05/15. doi: 10.1016/j.tibtech.2013.04.004. PubMed PMID: 23664777; PubMed Central PMCID: PMC3694601.

23. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science (New York, NY. 2013;339(6121):823-6. Epub 2013/01/05. doi: 10.1126/science.1232033. PubMed PMID: 23287722.

24. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science (New York, NY. 2013;339(6121):819-23. Epub 2013/01/05. doi: 10.1126/science.1231143. PubMed PMID: 23287718.

25. Cornu TI, Cathomen T. Quantification of zinc finger nuclease-associated toxicity. Methods Mol Biol. 2010;649:237-45. Epub 2010/08/04. doi: 10.1007/978-1-60761-753-2_14. PubMed PMID: 20680838.

26. Olsen PA, Gelazauskaite M, Randol M, Krauss S. Analysis of illegitimate genomic integration mediated by zinc-finger nucleases: implications for specificity of targeted gene correction. BMC molecular biology. 2010;11:35. Epub 2010/05/13. doi: 10.1186/1471-2199-11-35. PubMed PMID: 20459736; PubMed Central PMCID: PMC2875229.

27. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nature biotechnology. 2011;29(9):816-23. Epub 2011/08/09. doi: 10.1038/nbt.1948. PubMed PMID: 21822255.

28. Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA. Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. Nucleic acids research. 2011;39(1):381-92. Epub 2010/09/17. doi: 10.1093/nar/gkq787. PubMed PMID: 20843781; PubMed Central PMCID: PMC3017618.

29. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods. 2011;8(9):765-70. Epub 2011/08/09. doi: 10.1038/nmeth.1670. PubMed PMID: 21822273; PubMed Central PMCID: PMC3164905.

30. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nature biotechnology. 2013;31(9):822-6. Epub 2013/06/25. doi: 10.1038/nbt.2623. PubMed PMID: 23792628; PubMed Central PMCID: PMC3773023.

31. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic acids research. 2011;39(21):9283-93. Epub 2011/08/05. doi: 10.1093/nar/gkr597. PubMed PMID: 21813459; PubMed Central PMCID: PMC3241638.

32. Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S, et al. Knockout rats generated by embryo microinjection of TALENs. Nature biotechnology. 2011;29(8):695-6. Epub 2011/08/09. doi: 10.1038/nbt.1940. PubMed PMID: 21822240.

33. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nature biotechnology. 2009;27(9):851-7. Epub 2009/08/15. doi: 10.1038/nbt.1562. PubMed PMID: 19680244.

34. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell. 1997;91(4):501-10. Epub 1997/12/09. PubMed PMID: 9390559.

35. Montini E, Held PK, Noll M, Morcinek N, Al-Dhalimy M, Finegold M, et al. In vivo correction of murine tyrosinemia type I by DNA-mediated transposition. Mol Ther. 2002;6(6):759-69. Epub 2002/12/25. PubMed PMID: 12498772.

36. Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell. 2005;122(3):473-83. PubMed PMID: 16096065.

37. Feschotte C. The piggyBac transposon holds promise for human gene therapy. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(41):14981-2. Epub 2006/10/04. doi: 10.1073/pnas.0607282103. PubMed PMID: 17015820; PubMed Central PMCID: PMC1622765.

38. Yusa K, Rad R, Takeda J, Bradley A. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat Methods. 2009;6(5):363-9. Epub 2009/04/02. doi: 10.1038/nmeth.1323. PubMed PMID: 19337237; PubMed Central PMCID: PMC2677165.

39. Wu SC, Meir YJ, Coates CJ, Handler AM, Pelczar P, Moisyadi S, et al. piggyBac is a flexible and highly active transposon as compared to sleeping beauty, Tol2, and Mos1 in mammalian cells. Proc Natl Acad Sci U S A. 2006;103(41):15008-13. Epub 2006/09/29. doi: 0606979103 [pii]

10.1073/pnas.0606979103. PubMed PMID: 17005721; PubMed Central PMCID: PMC1622771.

40. Urschitz J, Kawasumi M, Owens J, Morozumi K, Yamashiro H, Stoytchev I, et al. Helper-independent piggyBac plasmids for gene delivery approaches: strategies for avoiding potential genotoxic effects. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(18):8117-22. Epub 2010/04/21. doi: 10.1073/pnas.1003674107. PubMed PMID: 20404201; PubMed Central PMCID: PMC2889585.

41. Li MA, Turner DJ, Ning Z, Yusa K, Liang Q, Eckert S, et al. Mobilization of giant piggyBac transposons in the mouse genome. Nucleic acids research. 2011. Epub 2011/09/29. doi: 10.1093/nar/gkr764. PubMed PMID: 21948799.

42. Mitra R, Fain-Thornton J, Craig NL. piggyBac can bypass DNA synthesis during cut and paste transposition. EMBO J. 2008;27(7):1097-109. Epub 2008/03/21. doi: 10.1038/emboj.2008.41. PubMed PMID: 18354502; PubMed Central PMCID: PMC2323262.

43. Marh J, Stoytcheva Z, Urschitz J, Sugawara A, Yamashiro H, Owens JB, et al. Hyperactive self-inactivating piggyBac for transposase-enhanced pronuclear microinjection transgenesis. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(47):19184-9. Epub 2012/10/25. doi: 10.1073/pnas.1216473109. PubMed PMID: 23093669; PubMed Central PMCID: PMC3511126.

44. Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, et al. Structural basis for double-stranded RNA processing by Dicer. Science (New York, NY. 2006;311(5758):195-8. Epub 2006/01/18. doi: 10.1126/science.1121638. PubMed PMID: 16410517.

45. Richter T, von Zglinicki T. A continuous correlation between oxidative stress and telomere shortening in fibroblasts. Experimental gerontology. 2007;42(11):1039-42. Epub 2007/09/18. doi: 10.1016/j.exger.2007.08.005. PubMed PMID: 17869047.

46. Rodier F, Kim SH, Nijjar T, Yaswen P, Campisi J. Cancer and aging: the importance of telomeres in genome maintenance. Int J Biochem Cell Biol. 2005;37(5):977-90. Epub 2005/03/04. doi: 10.1016/j.biocel.2004.10.012. PubMed PMID: 15743672.

47. Owens JB, Urschitz J, Stoytchev I, Dang NC, Stoytcheva Z, Belcaid M, et al. Chimeric piggyBac transposases for genomic targeting in human cells. Nucleic acids research. 2012;40(14):6978-91. Epub 2012/04/12. doi: 10.1093/nar/gks309. PubMed PMID: 22492708; PubMed Central PMCID: PMC3413120.

48. Owens JB, Mauro D, Stoytchev I, Bhakta MS, Kim MS, Segal DJ, et al. Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. Nucleic acids research. 2013;41(19):9197-207. Epub 2013/08/08. doi: 10.1093/nar/gkt677. PubMed PMID: 23921635; PubMed Central PMCID: PMC3799441.

49. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nature biotechnology. 2011;29(2):143-8. Epub 2010/12/24. doi: 10.1038/nbt.1755. PubMed PMID: 21179091.

50. Muruve DA, Barnes MJ, Stillman IE, Libermann TA. Adenoviral gene therapy leads to rapid induction of multiple chemokines and acute neutrophil-dependent hepatic injury in vivo. Hum Gene Ther. 1999;10(6):965-76. Epub 1999/05/01. doi: 10.1089/10430349950018364. PubMed PMID: 10223730.

51. Kaminski J, Summers JB. Delivering zinc fingers. Nature biotechnology. 2003;21(5):492-3. Epub 2003/05/02. doi: 10.1038/nbt0503-492b. PubMed PMID: 12721569.

52. Kay MA, Glorioso JC, Naldini L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nat Med. 2001;7(1):33-40. PubMed PMID: 11135613.

53. Ivics Z, Li MA, Mates L, Boeke JD, Nagy A, Bradley A, et al. Transposon-mediated genome manipulation in vertebrates. Nat Methods. 2009;6(6):415-22. Epub 2009/05/30. doi: nmeth.1332 [pii]

10.1038/nmeth.1332. PubMed PMID: 19478801.

54. Saridey SK, Liu L, Doherty JE, Kaja A, Galvan DL, Fletcher BS, et al. PiggyBac Transposon-based Inducible Gene Expression In Vivo After Somatic Cell Gene Transfer. Mol Ther. 2009. Epub 2009/10/08. doi: mt2009234 [pii]

10.1038/mt.2009.234. PubMed PMID: 19809403.

55. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature. 2009;458(7239):766-70. Epub 2009/03/03. doi: 10.1038/nature07863. PubMed PMID: 19252478.

56. Lacoste A, Berenshteyn F, Brivanlou AH. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. Cell Stem Cell. 2009;5(3):332-42. Epub 2009/09/08. doi: S1934-5909(09)00347-6 [pii]

10.1016/j.stem.2009.07.011. PubMed PMID: 19733544.

57. Wang W, Bradley A, Huang Y. A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. Genome Res. 2009;19(4):667-73. Epub 2009/02/24. doi: gr.085621.108 [pii]

10.1101/gr.085621.108. PubMed PMID: 19233961; PubMed Central PMCID: PMC2665785.

58. Kang Y, Zhang XY, Jiang W, Wu CQ, Chen CM, Gu JR, et al. The piggyBac transposon is an integrating non-viral gene transfer vector that enhances the efficiency of GDEPT. Cell Biol Int. 2009;33(4):509-15. Epub 2009/04/09. PubMed PMID: 19353779.

59. Liang Q, Kong J, Stalker J, Bradley A. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. Genesis. 2009;47(6):404-8. Epub 2009/04/25. doi: 10.1002/dvg.20508. PubMed PMID: 19391106.

60. Wilson MH, Coates CJ, George AL, Jr. PiggyBac Transposon-mediated Gene Transfer in Human Cells. Mol Ther. 2007;15(1):139-45. PubMed PMID: 17164785.

61. Lampe DJ, Grant TE, Robertson HM. Factors affecting transposition of the Himar1 mariner transposon in vitro. Genetics. 1998;149(1):179-87. Epub 1998/05/28. PubMed PMID: 9584095; PubMed Central PMCID: PMC1460121.

62. Fraser MJ, Ciszczon T, Elick T, Bauser C. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. Insect Mol Biol. 1996;5(2):141-51. Epub 1996/05/01. PubMed PMID: 8673264.

63. Mates L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat Genet. 2009;41(6):753-61. Epub 2009/05/05. doi: ng.343 [pii]

10.1038/ng.343. PubMed PMID: 19412179.

64. Dupuy AJ, Clark K, Carlson CM, Fritz S, Davidson AE, Markley KM, et al. Mammalian germ-line transgenesis by transposition. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(7):4495-9. PubMed PMID: 11904379.

65. Kimura Y, Yanagimachi R. Intracytoplasmic sperm injection in the mouse. Biol Reprod. 1995;52(4):709-20. Epub 1995/04/01. PubMed PMID: 7779992.

66. Haueter S, Kawasumi M, Asner I, Brykczynska U, Cinelli P, Moisyadi S, et al. Genetic vasectomy-overexpression of Prm1-EGFP fusion protein in elongating spermatids causes dominant male sterility in mice. Genesis.48(3):151-60. Epub 2010/01/23. doi: 10.1002/dvg.20598. PubMed PMID: 20095053.

67. Shinohara ET, Kaminski JM, Segal DJ, Pelczar P, Kolhe R, Ryan T, et al. Active integration: new strategies for transgenesis. Transgenic Res. 2007;16(3):333-9. PubMed PMID: 17340207.

68. Li C, Mizutani E, Ono T, Wakayama T. An efficient method for generating transgenic mice using NaOH-treated spermatozoa. Biol Reprod. 2010;82(2):331-40. Epub 2009/10/09. doi: 10.1095/biolreprod.109.078501. PubMed PMID: 19812303.

69. Suganuma R, Pelczar P, Spetz JF, Hohn B, Yanagimachi R, Moisyadi S. Tn5 transposase-mediated mouse transgenesis. Biol Reprod. 2005;73(6):1157-63. PubMed PMID: 16079303.

70. Coussens M, Yamazaki Y, Moisyadi S, Suganuma R, Yanagimachi R, Allsopp R. Regulation and effects of modulation of telomerase reverse transcriptase expression in primordial germ cells during development. Biol Reprod. 2006;75(5):785-91. PubMed PMID: 16899651.

71. Kang Y, Zhang X, Jiang W, Wu C, Chen C, Zheng Y, et al. Tumor-directed gene therapy in mice using a composite nonviral gene delivery system consisting of the piggyBac transposon and polyethylenimine. BMC Cancer. 2009;9:126. Epub 2009/04/29. doi: 1471-2407-9-126 [pii]

10.1186/1471-2407-9-126. PubMed PMID: 19397814.

72. Cooper LJ. Off-the-shelf T-cell therapy. Blood.116(23):4741-3. Epub 2010/12/04. doi: 116/23/4741 [pii]

10.1182/blood-2010-10-308379. PubMed PMID: 21127184.

73. Cadinanos J, Bradley A. Generation of an inducible and optimized piggyBac transposon system. Nucleic acids research. 2007;35(12):e87. Epub 2007/06/20. doi: 10.1093/nar/gkm446. PubMed PMID: 17576687; PubMed Central PMCID: PMC1919496.

74. Galvan DL, Nakazawa Y, Kaja A, Kettlun C, Cooper LJ, Rooney CM, et al. Genome-wide mapping of PiggyBac transposon integrations in primary human T cells. J Immunother. 2009;32(8):837-44. Epub 2009/09/16. doi: 10.1097/CJI.0b013e3181b2914c. PubMed PMID: 19752750; PubMed Central PMCID: PMC2796288.

75. Li C, Mizutani E, Ono T, Wakayama T. An Efficient Method for Generating Transgenic Mice Using NaOH-Treated Spermatozoa. Biol Reprod. 2009. Epub 2009/10/09. doi: biolreprod.109.078501 [pii]

10.1095/biolreprod.109.078501. PubMed PMID: 19812303.

76. Capecchi MR. The new mouse genetics: altering the genome by gene targeting. Trends Genet. 1989;5(3):70-6. Epub 1989/03/01. PubMed PMID: 2660363.

77. Niu Y, Liang S. Progress in gene transfer by germ cells in mammals. J Genet Genomics. 2008;35(12):701-14. Epub 2008/12/24. doi: S1673-8527(08)60225-8 [pii]

10.1016/S1673-8527(08)60225-8. PubMed PMID: 19103425.

78. Smith K. Theoretical mechanisms in targeted and random integration of transgene DNA. Reprod Nutr Dev. 2001;41(6):465-85. Epub 2002/07/20. PubMed PMID: 12126294.

79. Wilber A, Frandsen JL, Geurts JL, Largaespada DA, Hackett PB, McIvor RS. RNA as a source of transposase for Sleeping Beauty-mediated gene insertion and expression in somatic cells and tissues. Mol Ther. 2006;13(3):625-30. Epub 2005/12/22. doi: S1525-0016(05)01668-0 [pii]

10.1016/j.ymthe.2005.10.014. PubMed PMID: 16368272.

80. Maragathavally KJ, Kaminski JM, Coates CJ. Chimeric Mos1 and piggyBac transposases result in site-directed integration. Faseb J. 2006. PubMed PMID: 16877528.

81. Chen YT, Furushima K, Hou PS, Ku AT, Deng JM, Jang CW, et al. PiggyBac transposon-mediated, reversible gene transfer in human embryonic stem cells. Stem Cells Dev. 2009. Epub 2009/09/11. doi: 10.1089/scd.2009.0118. PubMed PMID: 19740021.

82. Bishop JO. Chromosomal insertion of foreign DNA. Reprod Nutr Dev. 1996;36(6):607-18. Epub 1996/01/01. doi: S0926528797818850 [pii]. PubMed PMID: 9021872.

83. Bishop JO, Smith P. Mechanism of chromosomal integration of microinjected DNA. Mol Biol Med. 1989;6(4):283-98. Epub 1989/08/01. PubMed PMID: 2695741.

84. Brinster RL, Chen HY, Trumbauer ME, Yagle MK, Palmiter RD. Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(13):4438-42. PubMed PMID: 3892534.

85. Hackett PB, Largaespada DA, Cooper LJ. A transposon and transposase system for human application. Mol Ther.18(4):674-83. Epub 2010/01/28. doi: mt20102 [pii]

10.1038/mt.2010.2. PubMed PMID: 20104209.

86. Katter K, Geurts AM, Hoffmann O, Mates L, Landa V, Hiripi L, et al. Transposon-mediated transgenesis, transgenic rescue, and tissue-specific gene expression in rodents and rabbits. Faseb J. 2013;27(3):930-41. Epub 2012/12/01. doi: 10.1096/fj.12-205526. PubMed PMID: 23195032; PubMed Central PMCID: PMC3574282.

87. Di Matteo M, Belay E, Chuah MK, Vandendriessche T. Recent developments in transposon-mediated gene therapy. Expert opinion on biological therapy. 2012;12(7):841-58. Epub 2012/06/12. doi: 10.1517/14712598.2012.684875. PubMed PMID: 22679910.

88. Ivics Z, Izsvak Z. Non-viral Gene Delivery with the Sleeping Beauty Transposon System. Hum Gene Ther. 2011. Epub 2011/08/27. doi: 10.1089/hum.2011.143. PubMed PMID: 21867398.

89. Izsvak Z, Ivics Z. Sleeping beauty transposition: biology and applications for molecular therapy. Mol Ther. 2004;9(2):147-56. Epub 2004/02/05. doi: 10.1016/j.ymthe.2003.11.009. PubMed PMID: 14759798.

90. Saridey SK, Liu L, Doherty JE, Kaja A, Galvan DL, Fletcher BS, et al. PiggyBac transposon-based inducible gene expression in vivo after somatic cell gene transfer. Mol Ther. 2009;17(12):2115-20. Epub 2009/10/08. doi: 10.1038/mt.2009.234. PubMed PMID: 19809403; PubMed Central PMCID: PMC2814386.

91. Owens JB, Urschitz J, Stoytchev I, Dang NC, Stoytcheva Z, Belcaid M, et al. Chimeric piggyBac transposases for genomic targeting in human cells. Nucleic acids research. 2012. Epub 2012/04/12. doi: 10.1093/nar/gks309. PubMed PMID: 22492708.

92. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, et al. Specific association of human telomerase activity with immortal cells and cancer. Science (New York, NY. 1994;266(5193):2011-5. Epub 1994/12/23. PubMed PMID: 7605428.

93. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, et al. Extension of life-span by introduction of telomerase into normal human cells. Science (New York, NY. 1998;279(5349):349-52. Epub 1998/02/07. PubMed PMID: 9454332.

94. Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, Lingner J, et al. Telomerase catalytic subunit homologs from fission yeast and human. Science (New York, NY. 1997;277(5328):955-9. Epub 1997/08/15. PubMed PMID: 9252327.

95. Urschitz J, Kawasumi M, Owens J, Morozumi K, Yamashiro H, Stoytchev I, et al. Helper-independent piggyBac plasmids for gene delivery approaches: strategies for avoiding potential genotoxic effects. Proceedings of the National Academy of Sciences of the United States of America.107(18):8117-22. Epub 2010/04/21. doi: 1003674107 [pii]

10.1073/pnas.1003674107. PubMed PMID: 20404201.

96. Kirkby NS, Leadbeater PD, Chan MV, Nylander S, Mitchell JA, Warner TD. Anti-platelet Effects of Aspirin Vary with Level of P2Y(12) Receptor Blockade Supplied by Either Ticagrelor or Prasugrel. Journal of thrombosis and haemostasis : JTH. 2011. Epub 2011/08/05. doi: 10.1111/j.1538-7836.2011.04453.x. PubMed PMID: 21812912.

97. Boudreau RL, Davidson BL. Generation of hairpin-based RNAi vectors for biological and therapeutic application. Methods in enzymology. 2012;507:275-96. Epub 2012/03/01. doi: 10.1016/B978-0-12-386509-0.00014-4. PubMed PMID: 22365779.

98. Zaffaroni N, Pennati M, Folini M. Validation of telomerase and survivin as anticancer therapeutic targets using ribozymes and small-interfering RNAs. Methods Mol Biol.

2007;361:239-63. Epub 2006/12/19. doi: 10.1385/1-59745-208-4:239. PubMed PMID: 17172716.

99. Dikmen ZG, Gellert GC, Jackson S, Gryaznov S, Tressler R, Dogan P, et al. In vivo inhibition of lung cancer by GRN163L: a novel human telomerase inhibitor. Cancer research. 2005;65(17):7866-73. Epub 2005/09/06. doi: 10.1158/0008-5472.CAN-05-1215. PubMed PMID: 16140956.

100. Gonzalez B, Schwimmer LJ, Fuller RP, Ye Y, Asawapornmongkol L, Barbas CF, 3rd. Modular system for the construction of zinc-finger libraries and proteins. Nature protocols. 2010;5(4):791-810. Epub 2010/04/03. doi: 10.1038/nprot.2010.34. PubMed PMID: 20360772; PubMed Central PMCID: PMC2855653.

101. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Molecular cell. 2008;31(2):294-301. Epub 2008/07/29. doi: 10.1016/j.molcel.2008.06.016. PubMed PMID: 18657511; PubMed Central PMCID: PMC2535758.

102. Isalan M, Klug A, Choo Y. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. Nature biotechnology. 2001;19(7):656-60. Epub 2001/07/04. doi: 10.1038/90264. PubMed PMID: 11433278; PubMed Central PMCID: PMC2677679.

103. Segal DJ, Dreier B, Beerli RR, Barbas CF, 3rd. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(6):2758-63. Epub 1999/03/17. PubMed PMID: 10077584; PubMed Central PMCID: PMC15842.

104. Mandell JG, Barbas CF, 3rd. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. Nucleic acids research. 2006;34(Web Server issue):W516-23. Epub 2006/07/18. doi: 10.1093/nar/gkl209. PubMed PMID: 16845061; PubMed Central PMCID: PMC1538883.

105. Beerli RR, Segal DJ, Dreier B, Barbas CF, 3rd. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(25):14628-33. Epub 1998/12/09. PubMed PMID: 9843940; PubMed Central PMCID: PMC24500.

106. Beerli RR, Barbas CF, 3rd. Engineering polydactyl zinc-finger transcription factors. Nature biotechnology. 2002;20(2):135-41. Epub 2002/02/01. doi: 10.1038/nbt0202-135. PubMed PMID: 11821858.

107. Blancafort P, Chen EI, Gonzalez B, Bergquist S, Zijlstra A, Guthy D, et al. Genetic reprogramming of tumor cells by zinc finger transcription factors. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(33):11716-21. Epub 2005/08/06. doi: 10.1073/pnas.0501162102. PubMed PMID: 16081541; PubMed Central PMCID: PMC1187960.

108. Beltran AS, Russo A, Lara H, Fan C, Lizardi PM, Blancafort P. Suppression of breast tumor growth and metastasis by an engineered transcription factor. PLoS ONE. 2011;6(9):e24595. Epub 2011/09/21. doi: 10.1371/journal.pone.0024595. PubMed PMID: 21931769; PubMed Central PMCID: PMC3172243.

109. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nature biotechnology. 2008;26(7):808-16. Epub 2008/07/01. doi: 10.1038/nbt1410. PubMed PMID: 18587387.

110. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. Science (New York, NY. 2009;325(5939):433. Epub 2009/07/25. doi: 10.1126/science.1172447. PubMed PMID: 19628861; PubMed Central PMCID: PMC2831805.

111. Moehle EA, Rock JM, Lee YL, Jouvenot Y, DeKelver RC, Gregory PD, et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(9):3055-60. Epub 2007/03/16. doi: 10.1073/pnas.0611478104. PubMed PMID: 17360608; PubMed Central PMCID: PMC1802009.

112. Zhu C, Smith T, McNulty J, Rayla AL, Lakshmanan A, Siekmann AF, et al. Evaluation and application of modularly assembled zinc-finger nucleases in zebrafish. Development. 2011;138(20):4555-64. Epub 2011/09/23. doi: 10.1242/dev.066779. PubMed PMID: 21937602; PubMed Central PMCID: PMC3177320.

113. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nature biotechnology. 2008;26(6):695-701. Epub 2008/05/27. doi: 10.1038/nbt1398. PubMed PMID: 18500337; PubMed Central PMCID: PMC2502069.

114. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. Science (New York, NY. 2003;300(5620):763. Epub 2003/05/06. doi: 10.1126/science.1078395. PubMed PMID: 12730593.

115. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature. 2005;435(7042):646-51. Epub 2005/04/05. doi: 10.1038/nature03556. PubMed PMID: 15806097.

116. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(3):1156-60. Epub 1996/02/06. PubMed PMID: 8577732; PubMed Central PMCID: PMC40048.

117. Wolfe SA, Grant RA, Pabo CO. Structure of a designed dimeric zinc finger protein bound to DNA. Biochemistry. 2003;42(46):13401-9. Epub 2003/11/19. doi: 10.1021/bi034830b. PubMed PMID: 14621985.

118. Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, et al. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger

nucleases. Mol Ther. 2008;16(2):352-8. Epub 2007/11/21. doi: 10.1038/sj.mt.6300357. PubMed PMID: 18026168.

119. Coates CJ, Kaminski JM, Summers JB, Segal DJ, Miller AD, Kolb AF. Site-directed genome modification: derivatives of DNA-modifying enzymes as targeting tools. Trends in biotechnology. 2005;23(8):407-19. Epub 2005/07/05. doi: 10.1016/j.tibtech.2005.06.009. PubMed PMID: 15993503.

120. Su K, Wang D, Ye J, Kim YC, Chow SA. Site-specific integration of retroviral DNA in human cells using fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc-finger protein E2C. Methods. 2009;47(4):269-76. Epub 2009/02/03. doi: 10.1016/j.ymeth.2009.01.001. PubMed PMID: 19186211; PubMed Central PMCID: PMC2695809.

121. Tan W, Dong Z, Wilkinson TA, Barbas CF, 3rd, Chow SA. 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. Journal of virology. 2006;80(4):1939-48. Epub 2006/01/28. doi: 10.1128/JVI.80.4.1939-1948.2006. PubMed PMID: 16439549; PubMed Central PMCID: PMC1367172.

122. Tan W, Zhu K, Segal DJ, Barbas CF, 3rd, Chow SA. Fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc finger protein E2C direct integration of viral DNA into specific sites. Journal of virology. 2004;78(3):1301-13. Epub 2004/01/15. PubMed PMID: 14722285; PubMed Central PMCID: PMC321411.

123. Bushman FD. Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(20):9233-7. Epub 1994/09/27. PubMed PMID: 7937746; PubMed Central PMCID: PMC44786.

124. Katz RA, Merkel G, Skalka AM. Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: in vitro activities and incorporation of a fusion protein into viral particles. Virology. 1996;217(1):178-90. Epub 1996/03/01. doi: 10.1006/viro.1996.0105. PubMed PMID: 8599202.

125. Akopian A, He J, Boocock MR, Stark WM. Chimeric recombinases with designed DNA sequence recognition. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(15):8688-91. Epub 2003/07/03. doi: 10.1073/pnas.1533177100. PubMed PMID: 12837939; PubMed Central PMCID: PMC166373.

126. Bolusani S, Ma CH, Paek A, Konieczka JH, Jayaram M, Voziyanov Y. Evolution of variants of yeast site-specific recombinase Flp that utilize native genomic sequences as recombination target sites. Nucleic acids research. 2006;34(18):5259-69. Epub 2006/09/28. doi: 10.1093/nar/gkl548. PubMed PMID: 17003057; PubMed Central PMCID: PMC1635253.

127. Gordley RM, Gersbach CA, Barbas CF, 3rd. Synthesis of programmable integrases. Proceedings of the National Academy of Sciences of the United States of America.

2009;106(13):5053-8. Epub 2009/03/14. doi: 10.1073/pnas.0812502106. PubMed PMID: 19282480; PubMed Central PMCID: PMC2654808.

128. Gordley RM, Smith JD, Graslund T, Barbas CF, 3rd. Evolution of programmable zinc finger-recombinases with activity in human cells. Journal of molecular biology. 2007;367(3):802-13. Epub 2007/02/10. doi: 10.1016/j.jmb.2007.01.017. PubMed PMID: 17289078.

129. Gaj T, Mercer AC, Gersbach CA, Gordley RM, Barbas CF, 3rd. Structure-guided reprogramming of serine recombinase DNA sequence specificity. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(2):498-503. Epub 2010/12/29. doi: 10.1073/pnas.1014214108. PubMed PMID: 21187418; PubMed Central PMCID: PMC3021078.

130. Gersbach CA, Gaj T, Gordley RM, Barbas CF, 3rd. Directed evolution of recombinase specificity by split gene reassembly. Nucleic acids research. 2010;38(12):4198-206. Epub 2010/03/03. doi: 10.1093/nar/gkq125. PubMed PMID: 20194120; PubMed Central PMCID: PMC2896519.

131. Zhu Y, Dai J, Fuerst PG, Voytas DF. Controlling integration specificity of a yeast retrotransposon. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(10):5891-5. Epub 2003/05/06. doi: 10.1073/pnas.1036705100. PubMed PMID: 12730380; PubMed Central PMCID: PMC156297.

132. Kaminski JM, Huber MR, Summers JB, Ward MB. Design of a nonviral vector for site-selective, efficient integration into the human genome. Faseb J. 2002;16(10):1242-7. Epub 2002/08/03. doi: 10.1096/fj.02-0127hyp. PubMed PMID: 12153992.

133. Brady TL, Schmidt CL, Voytas DF. Targeting integration of the Saccharomyces Ty5 retrotransposon. Methods Mol Biol. 2008;435:153-63. Epub 2008/03/29. doi: 10.1007/978-1-59745-232-8_11. PubMed PMID: 18370074.

134. Feng X, Bednarz AL, Colloms SD. Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. Nucleic acids research. 2010;38(4):1204-16. Epub 2009/12/08. doi: 10.1093/nar/gkp1068. PubMed PMID: 19965773; PubMed Central PMCID: PMC2831304.

135. Szabo M, Muller F, Kiss J, Balduf C, Strahle U, Olasz F. Transposition and targeting of the prokaryotic mobile element IS30 in zebrafish. FEBS Lett. 2003;550(1-3):46-50. Epub 2003/08/26. PubMed PMID: 12935884.

136. Yant SR, Huang Y, Akache B, Kay MA. Site-directed transposon integration in human cells. Nucleic acids research. 2007;35(7):e50. Epub 2007/03/09. doi: 10.1093/nar/gkm089. PubMed PMID: 17344320; PubMed Central PMCID: PMC1874657.

137. Ivics Z, Katzer A, Stuwe EE, Fiedler D, Knespel S, Izsvak Z. Targeted Sleeping Beauty transposition in human cells. Mol Ther. 2007;15(6):1137-44. Epub 2007/04/12. doi: 10.1038/sj.mt.6300169. PubMed PMID: 17426709.

138. Claeys Bouuaert C, Chalmers RM. Gene therapy vectors: the prospects and potentials of the cut-and-paste transposons. Genetica. 2010;138(5):473-84. Epub 2009/08/04. doi: 10.1007/s10709-009-9391-x. PubMed PMID: 19649713.

139. Chen YT, Furushima K, Hou PS, Ku AT, Deng JM, Jang CW, et al. PiggyBac transposon-mediated, reversible gene transfer in human embryonic stem cells. Stem Cells Dev. 2010;19(6):763-71. Epub 2009/09/11. doi: 10.1089/scd.2009.0118. PubMed PMID: 19740021; PubMed Central PMCID: PMC3135255.

140. Kim A, Pyykko I. Size matters: versatile use of PiggyBac transposons as a genetic manipulation tool. Molecular and cellular biochemistry. 2011;354(1-2):301-9. Epub 2011/04/26. doi: 10.1007/s11010-011-0832-3. PubMed PMID: 21516337.

141. Jang CW, Behringer RR. Transposon-mediated transgenesis in rats. CSH protocols. 2007;2007:pdb prot4866. Epub 2007/01/01. doi: 10.1101/pdb.prot4866. PubMed PMID: 21356954.

142. Maragathavally KJ, Kaminski JM, Coates CJ. Chimeric Mos1 and piggyBac transposases result in site-directed integration. Faseb J. 2006;20(11):1880-2. Epub 2006/08/01. doi: 10.1096/fj.05-5485fje. PubMed PMID: 16877528.

143. Kolb AF, Coates CJ, Kaminski JM, Summers JB, Miller AD, Segal DJ. Site-directed genome modification: nucleic acid and protein modules for targeted integration and gene correction. Trends in biotechnology. 2005;23(8):399-406. Epub 2005/06/29. doi: 10.1016/j.tibtech.2005.06.005. PubMed PMID: 15982766.

144. Paruzynski A, Arens A, Gabriel R, Bartholomae CC, Scholz S, Wang W, et al. Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. Nature protocols. 2010;5(8):1379-95. Epub 2010/07/31. doi: 10.1038/nprot.2010.87. PubMed PMID: 20671722.

145. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60. Epub 2009/05/20. doi: 10.1093/bioinformatics/btp324. PubMed PMID: 19451168; PubMed Central PMCID: PMC2705234.

146. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell. 2010;38(4):576-89. Epub 2010/06/02. doi: 10.1016/j.molcel.2010.05.004. PubMed PMID: 20513432; PubMed Central PMCID: PMC2898526.

147. Huang X, Guo H, Tammana S, Jung YC, Mellgren E, Bassi P, et al. Gene transfer efficiency and genome-wide integration profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. Mol Ther. 2010;18(10):1803-13. Epub 2010/07/08. doi: 10.1038/mt.2010.141. PubMed PMID: 20606646; PubMed Central PMCID: PMC2951558.

148. Narezkina A, Taganov KD, Litwin S, Stoyanova R, Hayashi J, Seeger C, et al. Genome-wide analyses of avian sarcoma virus integration sites. Journal of virology.

2004;78(21):11656-63. Epub 2004/10/14. doi: 10.1128/JVI.78.21.11656-11663.2004. PubMed PMID: 15479807; PubMed Central PMCID: PMC523270.

149. Yant SR, Wu X, Huang Y, Garrison B, Burgess SM, Kay MA. High-resolution genome-wide mapping of transposon integration in mammals. Molecular and cellular biology. 2005;25(6):2085-94. Epub 2005/03/04. doi: 10.1128/MCB.25.6.2085-2094.2005. PubMed PMID: 15743807; PubMed Central PMCID: PMC1061620.

150. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002;110(4):521-9. Epub 2002/08/31. PubMed PMID: 12202041.

151. Liang SD, Marmorstein R, Harrison SC, Ptashne M. DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. Molecular and cellular biology. 1996;16(7):3773-80. Epub 1996/07/01. PubMed PMID: 8668194; PubMed Central PMCID: PMC231373.

152. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science (New York, NY. 2003;302(5644):415-9. Epub 2003/10/18. doi: 10.1126/science.1088547. PubMed PMID: 14564000.

153. Li Z, Dullmann J, Schiedlmeier B, Schmidt M, von Kalle C, Meyer J, et al. Murine leukemia induced by retroviral gene marking. Science (New York, NY. 2002;296(5567):497. Epub 2002/04/20. doi: 10.1126/science.1068893. PubMed PMID: 11964471.

154. Kettlun C, Galvan DL, George AL, Jr., Kaja A, Wilson MH. Manipulating piggyBac Transposon Chromosomal Integration Site Selection in Human Cells. Mol Ther. 2011;19(9):1636-44. Epub 2011/07/07. doi: 10.1038/mt.2011.129. PubMed PMID: 21730970.

155. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nature biotechnology. 2007;25(7):786-93. Epub 2007/07/03. doi: 10.1038/nbt1317. PubMed PMID: 17603476.

156. Chalberg TW, Portlock JL, Olivares EC, Thyagarajan B, Kirby PJ, Hillman RT, et al. Integration specificity of phage phiC31 integrase in the human genome. Journal of molecular biology. 2006;357(1):28-48. Epub 2006/01/18. doi: 10.1016/j.jmb.2005.11.098. PubMed PMID: 16414067.

157. Liu J, Jeppesen I, Nielsen K, Jensen TG. Phi c31 integrase induces chromosomal aberrations in primary human fibroblasts. Gene Ther. 2006;13(15):1188-90. Epub 2006/05/05. doi: 10.1038/sj.gt.3302789. PubMed PMID: 16672982.

158. Recchia A, Parks RJ, Lamartina S, Toniatti C, Pieroni L, Palombo F, et al. Site-specific integration mediated by a hybrid adenovirus/adeno-associated virus vector. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(6):2615-20. Epub 1999/03/17. PubMed PMID: 10077559; PubMed Central PMCID: PMC15817.

159. Wang H, Lieber A. A helper-dependent capsid-modified adenovirus vector expressing adeno-associated virus rep78 mediates site-specific integration of a 27-kilobase transgene cassette. Journal of virology. 2006;80(23):11699-709. Epub 2006/09/22. doi: 10.1128/JVI.00779-06. PubMed PMID: 16987973; PubMed Central PMCID: PMC1642588.

160. Gersbach CA, Gaj T, Gordley RM, Mercer AC, Barbas CF, 3rd. Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. Nucleic acids research. 2011. Epub 2011/06/10. doi: 10.1093/nar/gkr421. PubMed PMID: 21653554.

161. Di Matteo M, Matrai J, Belay E, Firdissa T, Vandendriessche T, Chuah MK. PiggyBac toolbox. Methods Mol Biol. 2012;859:241-54. Epub 2012/03/01. doi: 10.1007/978-1-61779-603-6_14. PubMed PMID: 22367876.

162. Fan H, Johnson C. Insertional oncogenesis by non-acute retroviruses: implications for gene therapy. Viruses. 2011;3(4):398-422. Epub 2011/10/14. doi: 10.3390/v3040398. PubMed PMID: 21994739; PubMed Central PMCID: PMC3186009.

163. Romano G, Marino IR, Pentimalli F, Adamo V, Giordano A. 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 & perspectives. 2009;22(4):185-96. Epub 2009/06/19. doi: 10.1358/dnp.2009.22.4.1367704. PubMed PMID: 19536363.

164. Pruett-Miller SM, Connelly JP, Maeder ML, Joung JK, Porteus MH. Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. Mol Ther. 2008;16(4):707-17. Epub 2008/03/13. doi: 10.1038/mt.2008.20. PubMed PMID: 18334988.

165. Bohne J, Cathomen T. Genotoxicity in gene therapy: an account of vector integration and designer nucleases. Current opinion in molecular therapeutics. 2008;10(3):214-23. Epub 2008/06/07. PubMed PMID: 18535928.

166. Radecke S, Radecke F, Cathomen T, Schwarz K. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. Mol Ther. 2010;18(4):743-53. Epub 2010/01/14. doi: 10.1038/mt.2009.304. PubMed PMID: 20068556; PubMed Central PMCID: PMC2862519.

167. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat Methods. 2011;8(10):861-9. Epub 2011/08/23. doi: 10.1038/nmeth.1674. PubMed PMID: 21857672.

168. van Rensburg R, Beyer I, Yao XY, Wang H, Denisenko O, Li ZY, et al. Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. Gene Ther. 2013;20(2):201-14. Epub 2012/03/23. doi: 10.1038/gt.2012.25. PubMed PMID: 22436965; PubMed Central PMCID: PMC3661409.

169. Maier DA, Brennan AL, Jiang S, Binder-Scholl GK, Lee G, Plesa G, et al. Efficient Clinical Scale Gene Modification via Zinc Finger Nuclease-Targeted Disruption of the HIV Co-receptor CCR5. Hum Gene Ther. 2013;24(3):245-58. Epub 2013/01/31. doi: 10.1089/hum.2012.172. PubMed PMID: 23360514.

170. Yusa K, Zhou L, Li MA, Bradley A, Craig NL. A hyperactive piggyBac transposase for mammalian applications. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(4):1531-6. Epub 2011/01/06. doi: 10.1073/pnas.1008322108. PubMed PMID: 21205896; PubMed Central PMCID: PMC3029773.

171. Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A, et al. Quantitative analysis of TALE-DNA interactions suggests polarity effects. Nucleic acids research. 2013;41(7):4118-28. Epub 2013/02/15. doi: 10.1093/nar/gkt085. PubMed PMID: 23408851; PubMed Central PMCID: PMC3627578.

172. Doherty JE, Huye LE, Yusa K, Zhou L, Craig NL, Wilson MH. Hyperactive piggyBac gene transfer in human cells and in vivo. Hum Gene Ther. 2012;23(3):311-20. Epub 2011/10/14. doi: 10.1089/hum.2011.138. PubMed PMID: 21992617; PubMed Central PMCID: PMC3300075.

173. Herzog RW, Cao O, Srivastava A. Two decades of clinical gene therapy--success is finally mounting. Discovery medicine. 2010;9(45):105-11. Epub 2010/03/03. PubMed PMID: 20193635; PubMed Central PMCID: PMC3586794.

174. Papapetrou EP, Lee G, Malani N, Setty M, Riviere I, Tirunagari LM, et al. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nature biotechnology. 2011;29(1):73-8. Epub 2010/12/15. doi: 10.1038/nbt.1717. PubMed PMID: 21151124; PubMed Central PMCID: PMC3356916.

175. Gaj T, Mercer AC, Sirk SJ, Smith HL, Barbas CF, 3rd. A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. Nucleic acids research. 2013;41(6):3937-46. Epub 2013/02/09. doi: 10.1093/nar/gkt071. PubMed PMID: 23393187; PubMed Central PMCID: PMC3616721.

176. Ammar I, Gogol-Doring A, Miskey C, Chen W, Cathomen T, Izsvak Z, et al. Retargeting transposon insertions by the adeno-associated virus Rep protein. Nucleic acids research. 2012;40(14):6693-712. Epub 2012/04/24. doi: 10.1093/nar/gks317. PubMed PMID: 22523082; PubMed Central PMCID: PMC3413126.

177. Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, et al. A library of TAL effector nucleases spanning the human genome. Nature biotechnology. 2013;31(3):251-8. Epub 2013/02/19. doi: 10.1038/nbt.2517. PubMed PMID: 23417094.

178. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nature biotechnology. 2007;25(11):1298-306. Epub 2007/10/30. doi: 10.1038/nbt1353. PubMed PMID: 17965707.

179. Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, et al. piggyBac transposase tools for genome engineering. Proceedings of the National Academy of Sciences of the United States of America. 2013. Epub 2013/06/01. doi: 10.1073/pnas.1305987110. PubMed PMID: 23723351.

180. Benabdallah BF, Allard E, Yao S, Friedman G, Gregory PD, Eliopoulos N, et al. Targeted gene addition to human mesenchymal stromal cells as a cell-based plasma-

soluble protein delivery platform. Cytotherapy. 2010;12(3):394-9. Epub 2010/03/25. doi: 10.3109/14653240903583803. PubMed PMID: 20331411.

181. Zou J, Maeder ML, Mali P, Pruett-Miller SM, Thibodeau-Beganny S, Chou BK, et al. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell. 2009;5(1):97-110. Epub 2009/06/23. doi: 10.1016/j.stem.2009.05.023. PubMed PMID: 19540188; PubMed Central PMCID: PMC2720132.

182. Zou J, Sweeney CL, Chou BK, Choi U, Pan J, Wang H, et al. Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. Blood. 2011;117(21):5561-72. Epub 2011/03/18. doi: 10.1182/blood-2010-12-328161. PubMed PMID: 21411759; PubMed Central PMCID: PMC3110021.

183. DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH, et al. 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. 2010;20(8):1133-42. Epub 2010/05/29. doi: 10.1101/gr.106773.110. PubMed PMID: 20508142; PubMed Central PMCID: PMC2909576.

184. Perez-Pinera P, Ousterout DG, Brown MT, Gersbach CA. Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. Nucleic acids research. 2012;40(8):3741-52. Epub 2011/12/16. doi: 10.1093/nar/gkr1214. PubMed PMID: 22169954; PubMed Central PMCID: PMC3333879.

185. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG, 2nd, et al. In vivo genome editing using a high-efficiency TALEN system. Nature. 2012;491(7422):114-8. Epub 2012/09/25. doi: 10.1038/nature11537. PubMed PMID: 23000899; PubMed Central PMCID: PMC3491146.

186. Tong C, Huang G, Ashton C, Wu H, Yan H, Ying QL. Rapid and cost-effective gene targeting in rat embryonic stem cells by TALENs. J Genet Genomics. 2012;39(6):275-80. Epub 2012/07/04. doi: 10.1016/j.jgg.2012.04.004. PubMed PMID: 22749015.

187. Li MA, Turner DJ, Ning Z, Yusa K, Liang Q, Eckert S, et al. Mobilization of giant piggyBac transposons in the mouse genome. Nucleic acids research. 2011;39(22):e148. Epub 2011/09/29. doi: 10.1093/nar/gkr764. PubMed PMID: 21948799; PubMed Central PMCID: PMC3239208.

188. Anderson CD, Urschitz J, Khemmani M, Owens JB, Moisyadi S, Shohet RV, et al. Ultrasound directs a transposase system for durable hepatic gene delivery in mice. Ultrasound Med Biol. 2013;39(12):2351-61. Epub 2013/09/17. doi: 10.1016/j.ultrasmedbio.2013.07.002. PubMed PMID: 24035623; PubMed Central PMCID: PMC3838570.

189. Owens JB, Mathews J, Davy P, Stoytchev I, Moisyadi S, Allsopp R. Effective Targeted Gene Knockdown in Mammalian Cells Using the piggyBac Transposase-based Delivery System. Molecular therapy Nucleic acids. 2013;2:e137. Epub 2013/12/12. doi: 10.1038/mtna.2013.61. PubMed PMID: 24326734.

190. Kriks S, Shim JW, Piao J, Ganat YM, Wakeman DR, Xie Z, et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature. 2011;480(7378):547-51. Epub 2011/11/08. doi: 10.1038/nature10648. PubMed PMID: 22056989; PubMed Central PMCID: PMC3245796.

191. Seferos DS, Giljohann DA, Hill HD, Prigodich AE, Mirkin CA. Nano-flares: probes for transfection and mRNA detection in living cells. J Am Chem Soc. 2007;129(50):15477-9. Epub 2007/11/24. doi: 10.1021/ja0776529. PubMed PMID: 18034495; PubMed Central PMCID: PMC3200543.