2017

# An optimized, broadly applicable piggyBac transposon induction system

Zongtai Qi
*Washington University School of Medicine in St. Louis*

Michael Nathaniel Wilkinson
*Washington University School of Medicine in St. Louis*

Xuhua Chen
*Washington University School of Medicine in St. Louis*

Sumithra Sankararaman
*Washington University School of Medicine in St. Louis*

David Mayhew
*Washington University School of Medicine in St. Louis*

***See next page for additional authors***

*Published online 13 January 2017*

*Nucleic Acids Research, 2017, Vol. 45, No. 7* ***e55***
*doi: 10.1093/nar/gkw1290*

# An optimized, broadly applicable *piggyBac* transposon induction system

**Zongtai Qi, Michael Nathaniel Wilkinson, Xuhua Chen, Sumithra Sankararaman, David Mayhew and Robi David Mitra***

Department of Genetics and Center for Genome Sciences and Systems Biology, Washington University, School of Medicine, St. Louis, MO 63108, USA

Received April 11, 2016; Revised November 01, 2016; Editorial Decision December 10, 2016; Accepted January 1, 2017

## ABSTRACT

**The *piggyBac* (PB) transposon has been used in a number of biological applications. The insertion of PB transposons into the genome can disrupt genes or regulatory regions, impacting cellular function, so for many experiments it is important that PB transposition is tightly controlled. Here, we systematically characterize three methods for the post-translational control of the PB transposon in four cell lines. We investigated fusions of the PB transposase with ERT2 and two degradation domains (FKBP-DD, DHFR-DD), in multiple orientations, and determined (i) the fold-induction achieved, (ii) the absolute transposition efficiency of the activated construct and (iii) the effects of two inducer molecules on cellular transcription and function. We found that the FKBP-DD confers the PB transposase with a higher transposition activity and better dynamic range than can be achieved with the other systems. In addition, we found that the FKBP-DD regulates transposon activity in a reversible and dose-dependent manner. Finally, we showed that Shld1, the chemical inducer of FKBP-DD, does not interfere with stem cell differentiation, whereas tamoxifen has significant effects. We believe the FKBP-based PB transposon induction will be useful for transposon-mediated genome engineering, insertional mutagenesis and the genome-wide mapping of transcription factor binding.**

## INTRODUCTION

Transposons are genetic elements that are able to mobilize throughout a host genome autonomously. Due to their mobility, DNA transposons and retrotransposons have been widely used as tools for generating mutation libraries and for delivering non-viral gene constructs into the genomes of a variety of organisms (1). The use of DNA transposons in mammalian genomes, however, was initially impeded by a lack of active transposable elements. The synthetic resurrection of *Sleeping Beauty*, a Tc1/mariner-like transposable element of the salmanoid fish genome, initiated the development of transposon technologies for use in mammalian cells (2). Since then, several other transposon systems have been developed (3–5), including the *piggyBac* (PB) system, derived from the cabbage looper moth *Trichoplusia ni* (6). This system is widely used because the PB transposon consistently exhibits high transposition efficiencies in different cell lines and organisms (7), can transpose cargos of up to 100 kb without a significant reduction in efficiency (8), mobilizes without leaving footprint mutations at the excision site, and is amenable to molecular engineering (9). In addition, the PB system is highly efficient for germ line insertional mutagenesis in mice and lacks overproduction inhibition (10). These unique characteristics make the *piggyBac* transposon an invaluable tool for a wide range of applications, including stable gene delivery (11), transgene excision (12), insertional mutagenesis (13) and the mapping of transcription factor (TF) binding in eukaryotic genomes (14–16). For many of these applications, the PB transposon is engineered to act as a gene trap that will disrupt genes when inserted into the genome (13), so it is useful to be able to tightly control PB transposition so that insertion only occurs during the proper experimental window, and not, for example, during the propagation of cell lines or mice. For other applications, such as the mapping of TF binding to genomic DNA by PB transposition (14–16), the constitutive activity of the PBase prevents the determination of the precise timing of the binding events, a feature that is particularly useful when studying developmental processes. Therefore, a method for induction that can tightly regulate the temporal activity of the PB transposon system is desirable. Furthermore, post-translational control is appealing because protein activity can be switched on or off considerably faster than is possible with transcriptional control schemes (17,18).

Prior to the work presented here, the only existing inducible PB transposon system that operated post-

*To whom correspondence should be addressed. Tel: +1 314 362 2751; Email: rmitra@genetics.wustl.edu
Present Address: Robi David Mitra, Room 4301, Washington University School of Medicine, 4515 McKinley Ave, St. Louis, MO 63108, USA.

translationally utilized the mutated ligand-binding domain of the estrogen receptor (ERT2) (9). The PBase-ERT2 fusion is constitutively expressed but is sequestered outside of the nucleus by the cytoplasmic heat shock protein 90 (HSP90), which binds the ERT2 domain (Figure 1A). In the presence of a small molecule inducer (4-Hydroxy tamoxifen; 4OHT), HSP90 binding is abolished and the PBase-ERT2 fusion rapidly relocates to the nucleus where it directs transposition. The PBase-ERT2 fusion works well in several contexts (14), but our early experiences with this system revealed that dynamic range of induction varies widely between different cell types, possibly due to differences in HSP90 levels, which is not constant across cell types (19–22). We further found that ERT2-PBase fusions are not highly active when induced (relative to the unfused transposase, see results). Finally, we found that the chemical inducer 4OHT inhibited the *in vitro* differentiation of mouse embryonic stem cells, an observation consistent with similar reports that 4OHT adversely affects neurogenesis (23), myelinogenesis (23), myometrial differentiation (24) and sexual maturation (25).

In light of these observations, we sought to develop a PB transposon induction system that provides a tightly regulated transposase enzyme that displays a large difference in activity between the induced and un-induced state, that, when induced, deposits transposons with an efficiency equal to that of the native protein, that is highly active across a wide variety of cell types, and that is induced with a small molecule that has minimal effects on general cellular functions. To do so, we characterized and optimized fusions of the PB transposase with two different destabilized domain (DD) proteins, FK506- and rapamycin binding protein (FKBP) (17) and dihydrofolate reductase (DHFR) (18), and compared these to the ERT2 based induction system (9) in four different cell lines. Destabilized domains are small, inherently unstable proteins that bind small molecules and have been destabilized by mutation. When fused to PBase, the mutated domain unfolds and the fusion protein is degraded by the proteosome (Figure 1B). However, if the cognate small molecule ligand is provided, the DD is stabilized and the PBase fusion protein is rescued from degradation and transposase activity is restored in a rapid, dose-dependent and reversible manner. Unlike the ERT2-based system where induction is mediated by a specific cytoplasmic protein HSP90, DD degradation is mediated by the ubiquitin–proteasome pathway, a common protein degradation mechanism that is active in all mammalian cells (26,27).

To identify the optimal configuration for each of the three induction systems, we investigated 15 different fusions of the PB transposase with FKBP, DHFR or ERT and measured the fold-induction of activity between the uninduced and induced states, the maximum transposition rate of the activated constructs as compared to the unfused *piggyBac* transposase, and the effects of two of the inducer molecules on cellular transcription and differentiation.

## MATERIALS AND METHODS

### Plasmid construction

The plasmids that contain mutated destabilized domains, FKBP and DHFR, were purchased from Addgene with ID 31763 and 29326, respectively. The PBase-ERT2 plasmid (mPBase-L3-ERT2) was a kind gift from Dr Bradley (9). The coding sequence of mPBase-L3-ERT2 was sub-cloned into a yeast shuttle vector pRS314 containing CEN6, ARS and TRP1 to use 'gap repair' cloning technique in yeast cells (Strathern and Higgins 1991). This engineered yeast shuttle vector was used as a 'parental' plasmid (pRM1056) to derive other variants of PBase constructs by gap repair method (15). Briefly, PCR-generated sequences were cloned into linearized vectors by recognizing a 40 bp overlap at their ends. This 40 bp overlap can be engineered by designing primers for amplification of the desired sequences. For example, to replace the ERT2 sequence with the FKBP sequence, the pRM1056 plasmid was linearized by restriction digestion that cut the plasmid within the ERT2 sequence. The FKBP sequence was amplified with a pair of primers that have a 40 bp sequence that is homologous to pRM1056. The FKBP PCR products and linearized pBM1056 were co-transformed into yeast cells and the yeast cells were selected for Trp+ colonies. DNA extracted from Trp+ yeast colonies was introduced into *E. coli*. Finally, the plasmid was isolated by QIAprep Spin Miniprep Kit (QIAGEN) and was confirmed by Sanger sequencing. The engineered constructs used in this study are depicted in Figure 2.

### Cell culture and neural differentiation

Human embryonic kidneys cell lines (HEK293 and HEK293T) and human colon adenocarcinoma cell line HCT116 were maintained in Dubecco's Modified Eagle Media (DMEM; Gibco) supplemented with 10% fetal bovine serum. RW4 mouse embryonic stem cells (ESCs) were cultured in complete media consisting of DMEM supplemented with 10% new born calf serum, 10% fetal bovine serum (Gibco) and 0.3 mM of each of the following nucleosides: adenosine, guanosine, cytosine, thymidine and uridine (Sigma-Aldrich). To maintain their undifferentiated state, cells were also cultured in the presence of 1000 U/ml leukemia inhibitory factor (LIF; Chemicon) and 20 mM β-mercaptoethanol (BME; Invitrogen) in flasks coated with a 0.1% gelatin solution (Sigma-Aldrich).

Mouse ESCs were differentiated into ventral spinal neural cells using a retinoic acid (RA) and smoothened agonist induction protocol as described (28). Briefly, ESCs were cultured in suspension on low attachment plates (Corning) in modified DFK5 media consisting of DMEM/F12 base media (Gibco) containing 5% knockout serum replacement, 1x insulin transferrin selenium, 50 μM of non-essential amino acids, 100 mM of BME, 5 mM of thymidine and 15 mM of the following nucleosides: adenosine, cytosine, guanosine and uridine. During this process, ESCs aggregate into multicellular embryoid bodies (EBs). After the first 2 days, the EBs were moved to a 15 ml of conical and allowed to settle for 5 min. The media was aspirated and replaced with 10 ml of fresh DFK5 containing 2 mM of RA and 600 nM of smoothened agonist (Millipore). EBs were then cultured

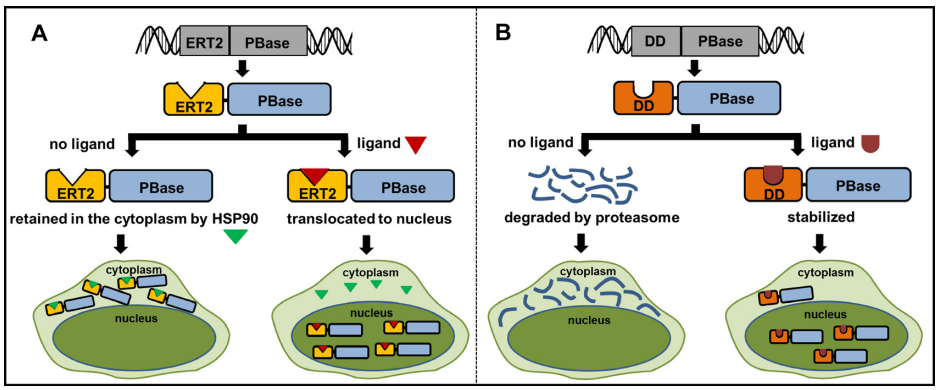


**Figure 1.** Schematic illustration of the ERT2 and destabilized domain (DD) based *PiggyBac* (PB) transposon induction systems. (**A**) For the ERT2-based PB transposon induction system, the PBase-ERT2 fusion is constitutively expressed but sequestered outside of the nucleus by HSP90, which binds the ERT2 domain. In the presence of ER antagonist 4-OHT, HSP90 dissociates and the PBase-ERT2 fusion rapidly relocates to the nucleus where it directs transposition. (**B**) For the DD-based PB transposon induction system, a DD (either FKBP or DHFR) was fused to the PBase. The DD confers the instability to the fused protein such that the PBase fusion protein was constitutively degraded. However, binding of a small molecule ligand (Shld1 for FKBP; TMP for DHFR) to the DD prevents PBase from degradation and stabilizes it.

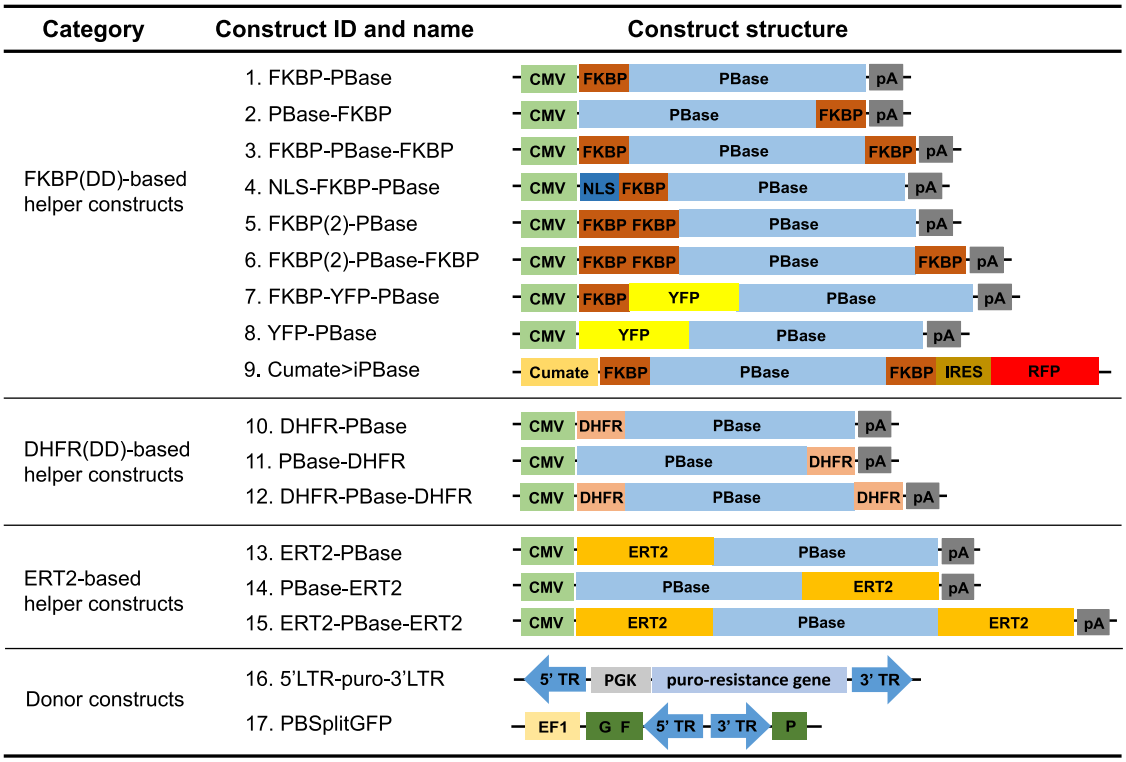


**Figure 2.** Schematic illustration of the constructs used in this study. For the helper constructs: PBase, *PiggyBac* transposase; CMV, cytomegalovirus promoter; pA, bovine growth hormone polyA signal. FKBP, a destabilizing domain derived from FK506 and rapamycin binding protein. DHFR, a destabilizing domain derived from *Escherichia coli* dihydrofolate reductase; YFP, yellow fluorescent protein; ERT2, mutated ligand-binding domain of the estrogen receptor; Cumate, cumate operator (mutated cmv promoter). For the donor construct: 5′TR-puro-3′TR contains the minimal PB terminal repeats (5′LR and 3′TR) flanking a PGK promoter driven puromycin resistance cassette; EF1, elongation factor 1 promoter; GFP, green fluorescent protein.

on the adhesive plates (Corning) for an additional 4 days for further differentiation, and media was replaced every 2 days.

### Cell transduction and transgenic cell line generation

All lentiviruses used for cell transduction were produced by the Hope Center Viral Vectors Core at Washington University School of Medicine. A CymR-expressing HEK293 cell line was made by transducing the cell with lentivirus containing the cumate operator repressor driven by an EF1a promoter at a multiplicity of infectivity of 20 to ensure that CymR is overexpressed. Two days after transduction, neomycin was added to the culture medium (500 ng/ml) and maintained for 7 days. Then, the CymR-expressing

HEK293 cell line was further transduced with lentivirus containing an FKBP-PBase-FKBP-IRES-RFP fusion gene driven by a cumate inducible promoter at multiplicity of infectivity of 0.5 to favor single-copy integration. To obtain a pure population, transduced cells were then sorted for red fluorescent protein (RFP) positive cells.

### Cell transfection and drug administration

All plasmids used for the transfection of cells were prepared using EndoFree Plasmid Maxi Kit (Qiagen) following the manufacturer's protocol. About $10^5$ cells were electroporated with a total of 0.6 μg of DNA (0.1 μg helper plasmid and 0.5 μg donor plasmid) by the Neon transfection system (Invitrogen) and plated to one well of a 6-well plate. Immediately upon transfection, cells were treated with 1 μM 4-hydroxy tamoxifen (4OHT, Sigma) for the ERT2 based induction system, 1 μM Shld1 (Clontech) for FKBP based induction system, 10 μM trimethoprim (TMP, Sigma) for the DHFR based induction system. Negative controls were mock-treated with 95% ethanol as vehicle. For the transgenic cell line expressing inducible PBase driven by a cumate promoter, only donor plasmid was used in transfection. Cumate inducer (System Biosciences) was added at concentrations ranging from 0 μg/ml to 70 μg/ml to regulate the strength of cumate promoter. For the purpose of reproducibility, experiments were done in triplicates. To select cells in which the PB transposon transposed, cells were trypsinized 2 days after transfection and cultured on 10 cm dish (Corning) with 10 ml fresh media containing puromycin (1 μg/ml). The puromycin selection normally takes 7 days before visible cell colonies are formed.

For experiments testing the reversibility of the FKBP-based PB transposon system, Shld1 (1 μM) was added to the culture medium immediately after transfection. One day after transfection, a small aliquot of cells were used for imaging and flow cytometry analyses and the rest of the cells were passaged to a 12-well plate and provided fresh Shld1-free medium. The passaged cells were grown to about 100% confluence before another passage to a 12-well plate. One well of cells were used for imaging and flow cytometry analyses every day. After 3 days of destabilization, the cells were treated with fresh medium containing 1 μM Shld1 to re-stabilize the FKBP-YFP-PBase fusion protein. The re-stabilization lasted for another 3 days during which the cells were analyzed by flow cytometry every day.

### Cell colony staining and counting

The visible drug-resistant cell colonies were fixed with phosphate buffered saline (PBS) containing 4% paraformaldehyde for 1 h and then stained with 1% methylene blue in 70% EtOH for 30 min, washed in distilled water and air-dried overnight. Colonies with diameters more than 0.3 mm were counted by ImageJ software (National Institutes of Health). The number of puromycin-resistant colonies formed from the cells transfected with both donor and helper plasmids was normalized to that with donor and helper backbone plasmids (yeast shuttle vector pRS314) prior to any further calculations. The transposition activity of a PBase fusion protein was calculated as the normalized number of colonies from the inducible domain (i.e. FKBP, DHFR or ERT2) fused PBase divided by that from 'wild type' unfused PBase. For experiments that used the transgenic cell line expressing FKBP-PBase-FKBP driven by a cumate promoter, we used normalized number of puromycin-resistant colonies to estimate the transposition activity due to the unavailability of the cell line expressing 'wild-type' unfused PBase. Fold-induction was calculated as the normalized number of colonies from the induced samples divided by that from untreated samples.

### Imaging and flow cytometry

Fluorescent images were taken on Zeiss Axioskop fluorescence microscope equipped with a QICAM FAST 1394 digital CCD camera. Cells were grown to 90% confluence, trypsinized from the plate and suspended in PBS, washed once with PBS and resuspended in Hank's balanced salt solution supplemented with 2 mM ethylenediaminetetraacetic acid. Cellular fluorescence was analyzed on an iCyt Reflection HAPS2 cell sorter at the Washington University Siteman Flow Cytometry Core. The gate was set relative to the cells transfected with non-fluorescent control plasmids to eliminate background. Cells transfected with a positive control fluorescent reporter plasmid were also used to eliminate false positive singles. About 10 000 cells were analyzed from each FACS and post-sort analysis was performed with FloJo software to obtain the percentage of fluorescent positive cells.

### RNA extraction and sequencing

Total RNA was isolated from the RW4 mouse ESCs using the PureLink RNA Mini kit (Ambion) according to the manufacturer's instructions. The quantity of RNA was measured using a spectrophotometer (NanoDrop 2000c; Thermo Scientific). Samples with a RNA concentration (A260/A280 $\geq$1.8 ng/μl) and purity (A230/A260 $\geq$ 2.0 ng/μl) were selected. The Agilent 2100 Bioanalyser was used to determine the RNA integrity number. The degradation level was identified using the RNA 6000 Nano LabChip kit (Agilent). Samples with RNA integrity number $>$ 9.8 were further processed using TruSeq mRNA Library Preparation Kit (Illumina) and then sequenced by Illumina MiSeq platform at the Genome Technology Access Center at Washington University in St. Louis. The gene expression data generated for this study can be found under the NCBI Gene Expression Omnibus accession number GSE78857. The expression data is also publicly available at the Center for Genome Sciences by request.

### Reads mapping and statistical analysis

Trimmomatic (v0.32) (29) was employed on RNA-Seq FASTQ files to clip the illumina adaptors and remove the reads of low quality. The cleaned reads were mapped back to mm10 genome reference from UCSC database using STAR (v2) (30). We used the HTSeq package (31) to estimate the count of uniquely mapped reads for each of the annotated genes in the mm10 gene transfer format (.GTF) file. Differential expressed genes were analyzed with R (v 2.13.0)

using DESeq (v1.4.1) (32) available in Bioconductor (v 2.8). The resulting *P*-values were adjusted using the Benjamini–Hochberg correction and only genes that were significant at a false discovery rate of 0.05 were considered as expressed. For other comparisons between different experimental conditions, the statistical significance was assessed by paired Student's *t*-test and a $P$-value $< 0.05$ was considered significant.

## RESULTS

### Overview

We evaluated and optimized three induction systems for their abilities to provide tight post-translation control over the PB transposon. We used one degradation domain (DD) derived from human FKBP12 (FKBP) (17) and another from *E. coli* DHFR (18) and compared these against the ERT2-based system in four different cell lines: human embryonic kidney cells (HEK293 and HEK293T), human colorectal cancer cells (HCT116) and RW4 mouse ESCs. To obtain the most efficient induction possible, a number of different fusions constructs were investigated for each system (Figure 2). We fused DD or ERT2 induction domains in-frame at the N terminus, C terminus or at both termini of the PBase. For N terminal fusion proteins, an 18 amino acid linker sequence was included between the PBase and the induction domain (33). For the C terminal fusion protein, a 24 amino acid linker was added between the PBase and induction domain (9).

In order to evaluate these different fusion proteins, we needed a simple, robust way to measure PB transposition. We adopted the commonly used method of co-transfecting cells with two plasmids: a 'helper' plasmid expressing the PBase and a 'donor' plasmid carrying the PB transposon with a drug-resistance marker (15). Once transfected into the host cells, the PBase mobilizes the transposon from the donor plasmid and integrates it into the host genome, conferring the cell with drug resistance. The number of drug-resistant colonies is closely correlated with the number of transposon integration events and is a key indicator of the efficiency of the PB transposon system (9,15,34). We used this system to measure, for each post-translational control system, the absolute transposition efficiency of the activated construct relative to the unfused *piggyBac* transposase, and the fold-induction achieved by each method in four different cell lines. We then went on to optimize the most promising system, characterize the dynamics of the temporal control achieved by the optimal system and evaluate the effects of two of the inducer molecules on cellular transcription and function.

### Evaluating the transposition efficiencies of inducible PBase

We first measured the transposition activity of the induced PBase fusion proteins relative to that of unfused PBase. The transposition activity for each variant of helper constructs was obtained by transfecting the above-mentioned four cell lines with an equimolar ratio of donor and helper plasmids in the presence of chemical inducers (Shld1 for FKBP fusions; TMP for DHFR fusions; 4OHT for ERT2 fusions). Typical images of stained colonies from negative control, inducer-treated and non-treated co-transfected experiments, and positive control with unfused PBase are shown in Figure 3A. The induced transposition activities relative to unfused PBase are shown in Figure 3B, and raw colony counts are provide in Supplementary Figures S1 and S2. In our assays, the DD-based induction systems had significantly higher activities than the ERT2-based induction system in all cell lines. The highest PBase activity was observed for the constructs of FKBP-PBase and DHFR-PBase, both of which had activities that were ∼95% of 'wild-type' levels. In contrast, all of the ERT2 fusion proteins displayed activities less than 25% that of the unfused PBase. Thus, we conclude that, under inducing conditions, the FKBP and DHFR DDs have little effect on PBase transposition activity while ERT2 appears to significantly interfere with the PBase activity. For the two DD-based induction systems, the transposition activities between N and C terminally tagged PBase fusions are not significantly different ($P > 0.05$). However, we observed a significant decrease of PBase activity when both termini were attached to either DD, suggesting that at least one terminus should be exposed for near-optimal PBase transposition activity.

### The FKBP DD system achieved the highest fold-change between induced and uninduced states

We next sought to quantify how tightly the different induction systems regulated the PBase protein. To do so, we again transfected the four cell lines with donor and helper plasmids; however, in this assay, we omitted the small-molecule inducers. We found that the DHFR-based PB induction system had a much higher background than did the FKBP and ERT2-based ones (Figure 3C). To determine which PBase fusion protein provided the largest dynamic range, we used the corresponding induced and non-induced transposition rates to calculate the fold-induction for each protein (Figure 3D). This analysis revealed that the FKBP-based induction system displayed a significantly higher fold-induction than both the DHFR and ERT2-based systems in all cell lines ($P < 0.05$). The FKBP-PBase-FKBP fusion protein showed the highest fold-induction while the ERT2-PBase-ERT2 fusion protein showed the lowest. Taking fold-induction and the absolute activity of induced PBase into consideration, the FKBP-based inducible PB transposon system outperformed its DHFR and ERT2-based counterparts.

### HSP90 levels correlate with the dynamic range of the ERT2 system

The dynamic ranges of the ERT2 controlled PBase fusion proteins were highly variable across the different cell lines (Figure 3D). For example, the PBase-ERT2 fusion displayed a 17-fold change in transposase activity between induced and uninduced conditions in mouse RW4 embryonic stem cell, but a significantly lower fold-induction in the other cell lines ($P < 0.05$). One possible explanation for this observation is that the HSP90 protein is typically highly expressed in mouse ES cells. HSP90 is the molecule that sequesters the fusion protein in the absence of inducer, and it has been previously reported the HSP90 RNA (35) and protein (36) levels are significantly higher in mouse ES

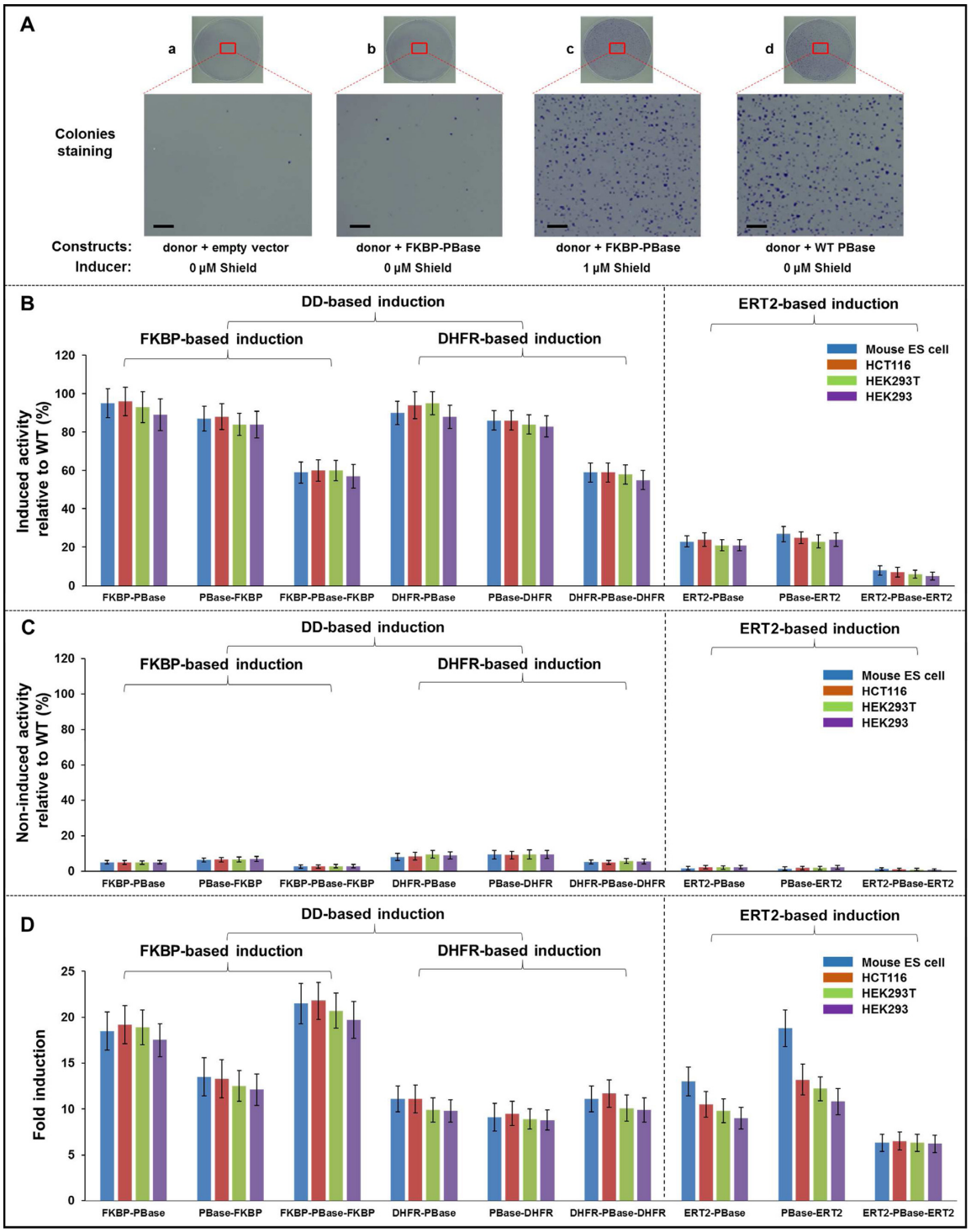


**Figure 3.** The performance of different helper constructs in DD and ERT2 based PB transposon induction systems. (**A**) Typical images of colony forming and staining assays to evaluate the transposition efficiency. The scale bar in the enlarged image equals 0.1 mM. The construct used is FKBP-PBase. (a) Cells transfected with donor and helper backbone plasmids (yeast shuttle vector pRS314) were used to estimate the background or random insertions. (b and c) Cells transfected with FKBP-PBase and donor plasmids either in the absence or presence of Shld1 were used to evaluate the transposition efficiency. (d) Cells transfected with unfused PBase and donor plasmid were used to estimate the maximum transposition efficiency. (**B**) Induced PBase activity of different PBase fusions relative to wild-type PBase across four cell lines. The number of puromycin-resistant colonies formed from the cells transfected with both donor and helper plasmids was normalized to that with donor and helper backbone plasmids prior to any further calculations (background subtraction). The induced transposition activity of an inducible domain (i.e. FKBP, DHFR or ERT2) fused PBase was calculated as the normalized number of colonies from the PBase fusion divided by that from 'wild type' unfused PBase. Experiments were done in triplicate. (**C**) Non-induced PBase activity of different PBase fusions across four cell lines. Experimental conditions were the same as in B except that no drug was added for the non-induced samples. (**D**) Fold induction of different PBase fusions across four cell lines. The induction fold was calculated as the normalized number of colonies from induced samples in B divided by that from untreated samples in C.

cells than in differentiated cell lines or embryoid bodies. To confirm HSP90 is indeed highly expressed in our mESCs, we performed RNA sequencing on the RW4 and HCT116 lines. We found that the beta-actin normalized HSP90 expression level in the RW4 mouse ES cell line was about 4-fold higher than that in HCT116 cell line (Supplementary Data 1 and 2), which is consistent with the theory that HSP90 levels explain the differences observed between RW4 mouse ESCs and HCT116 cells.

### There is a trade-off between maximal PBase activity and fold-induction

In the FKBP-based induction system, the PBase fusion protein shuttles between the cytoplasm where FKBP-directed degradation occurs and nucleus where PB transposition occurs. Therefore, we hypothesized that this nucleocytoplasmic transport could be fine-tuned to create an induction system with both efficient transposition and rapid degradation. To try to further improve transposition efficiency, we fused the nuclear localization sequence (NLS) from the SV40 large T-antigen upstream of, and in-frame with, the FKBP-PBase fusion, which had the highest transposition activity in our previous tests (Figure 3B). We reasoned that the NLS should direct more fusion protein to the nucleus, increasing the transposition rate upon administration of Shld1. We used transient transfection to introduce an equimolar ratio of NLS-FKBP-PBase fusion-carrying plasmid and donor plasmid across four different cell lines, and compared the results to that of FKBP-PBase plasmid without the NLS. After induction, the engineered fusion protein (NLS-FKBP-PBase) displayed an increase in transposition activity relative to the FKBP-PBase fusion protein (Figure 4A, $P < 0.05$). However, we also observed higher background transposition in the uninduced sample for this construct. As a result, the fold-induction for the NLS-FKBP-PBase fusion was lower than that of FKBP-PBase (Figure 4B, $P < 0.05$). This suggests that the NLS sequence sequesters some protein in the nucleus and prevents efficient degradation in the absence of Shld1. Based on this observation, we concluded that adding NLS to PBase fusion protein is not optimal for an inducible PB transposon system.

To test the possibility of further reducing the background transposition that occurs in the absence of chemical inducer, we made two more fusion constructs. One contains two tandem FKBP domains at the N terminus of PBase (Figure 2, construct 5) since our previous results showed that an N-terminally tagged PBase was more tightly regulated than the corresponding C-terminally tagged PBase. The second construct utilized three FKBP domains: two in tandem on the N-terminus and one on the C-terminus (Figure 2, construct 6). We tested the transposition efficiency and dynamic range of these constructs as before, and found that the triple FKBPs fusion protein (FKBP-FKBP-PBase-FKBP) showed the highest fold-induction in all cell lines (Figure 4A, right panel), followed by the double FKBPs fusion proteins (FKBP-PBase-FKBP and FKBP-FKBP-PBase). All three constructs achieved a significantly higher fold induction than did FKBP-PBase ($P < 0.05$).

We next examined the maximal induced transposition efficiency of these constructs relative to the unfused PBase (Figure 4B, right panel). We found that the transposition efficiencies of FKBP-FKBP-PBase-FKBP, FKBP-FKBP-PBase and FKBP-PBase-FKBP were 60%, 79% and 80% of the unfused PBase, which is significantly lower than the 95% relative activity we observed for FKBP-PBase ($P < 0.05$), indicating that transposition activity gradually decreased when more FKBP domains were added to PBase. Taken together, these results suggest that there is a trade-off between the efficient degradation of the PBase fusions under non-inducing conditions and the enzymatic activity of these fusions under fully inducing conditions (Figure 4C). Specifically, if high activity upon induction is the objective, then the FKBP-PBase fusion is optimal, while if a high fold-change in induction is required, the FKBP-FKBP-PBase-FKBP fusion is the best choice.

### Fusion protein levels can be tuned to achieve low background transposition and high inducibility

For some applications, it is not important to obtain the maximum possible PBase activity, but it is instead preferable to have an assay with low background levels and to achieve a large fold-change in activity upon induction. We reasoned that the high background levels observed in Figures 3 and 4 might be due to the fact that the PBase-DD fusion proteins were all highly expressed and therefore the proteasome was getting overloaded. To test this hypothesis, we created lentivirus containing an FKBP-PBase-FKBP-IRES-RFP fusion gene under the control of a cumate inducible promoter (37) (Figure 2, construct 9), which allows gene expression to be tuned by varying the cumate concentration in the culture media. Next, we infected cymR expressing HEK 293 cells, and isolated transduced cells by performing FACS to purify RFP positive cells. We transfected these cells with donor plasmid (Figure 2, construct 16), titrated FKBP-PBase-FKBP protein levels by culturing the cell with varying amounts of cumate, in the presence or absence of Shld1, and determined PBase activity by counting puromycin resistant colonies as previously described. The results are shown in Figure 5. As lower concentrations of cumate were added to the media, much lower background and larger fold-changes in induction were achieved. For example, when 10 μg/ml cumate was added to the media, no background transposition was observed, and we observed a more than 100-fold change in PBase activity after induction with Shld1. These results demonstrate that, when the FKBP-PBase-FKBP protein is expressed at moderate levels, this system has essentially no background transposition yet is still robustly inducible. This level of control comes at some cost, however, as the maximum level of PBase activity is roughly one-fourth the maximum rate we observed (i.e. compare the 10 μg/ml cumate and Shld1 culture condition to the PBase activity of cells grown with 70 μg/ml cumate and Shld1). For experiments that do not require rapid, post-transcriptional induction, it is possible to have the best of both worlds, namely zero background and high absolute activity, by growing cells in the absence of cumate and Shld1 and then adding both chemicals to induce transposition.

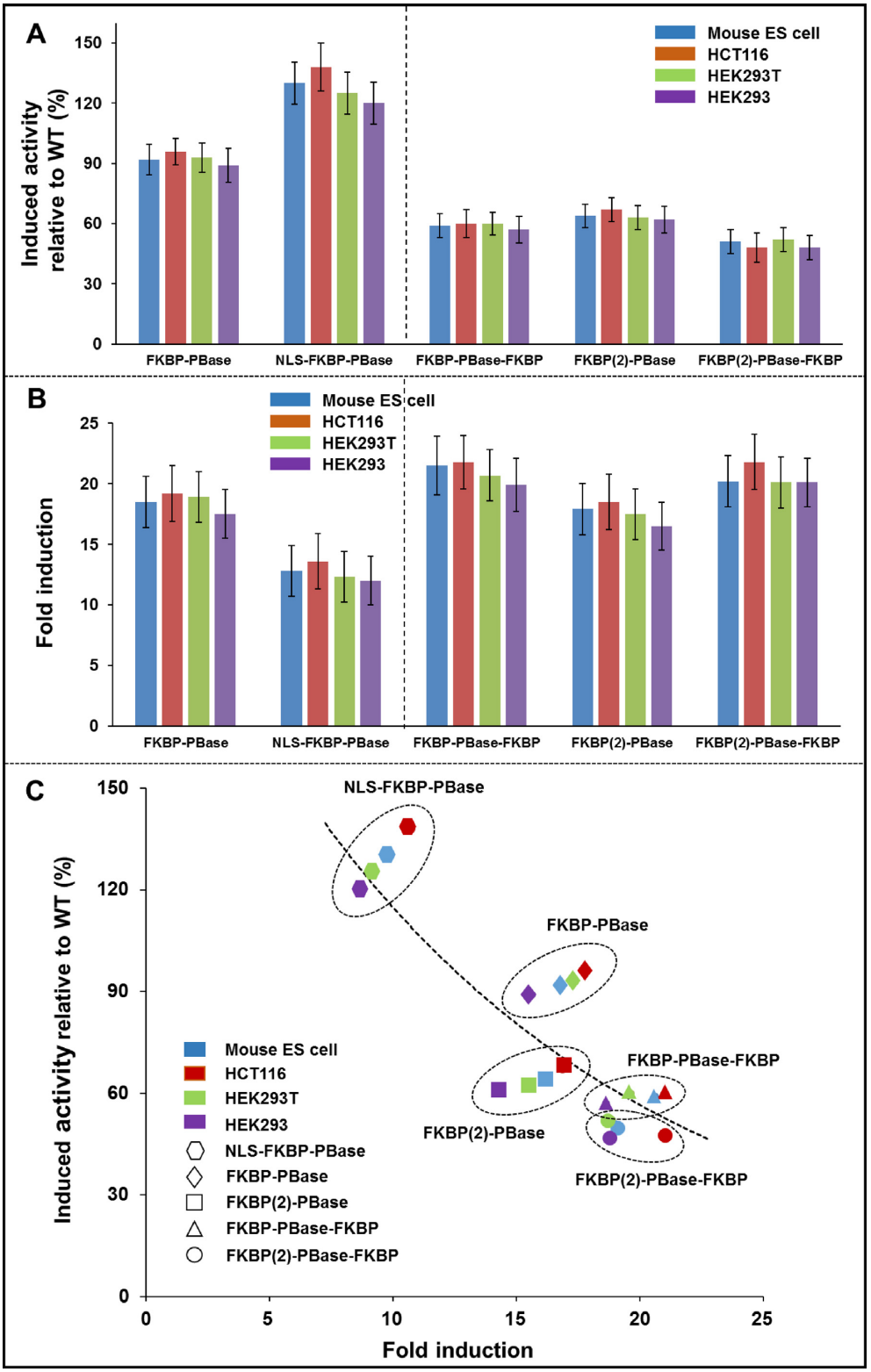


**Figure 4.** Optimization of the induced transposition activity and fold induction for FKBP-based PB transposon induction system. (**A**) Induced transposition activity of different PBase fusion proteins relative to wild-type PBase across four cell lines. Experiments were done in triplicates. The induced transposition activity of a PBase fusion protein was calculated as the normalized number of colonies from the PBase fusion in the presence of corresponding chemical inducer divided by that from 'wild type' unfused PBase. (**B**) Fold induction of different PBase fusion proteins across four cell lines. The induction fold was calculated as the normalized number of colonies from chemical inducer treated samples divided by that from untreated samples. (**C**) Comparison of the induced transposition activity with the fold induction for different PBase fusion proteins across four cell lines.

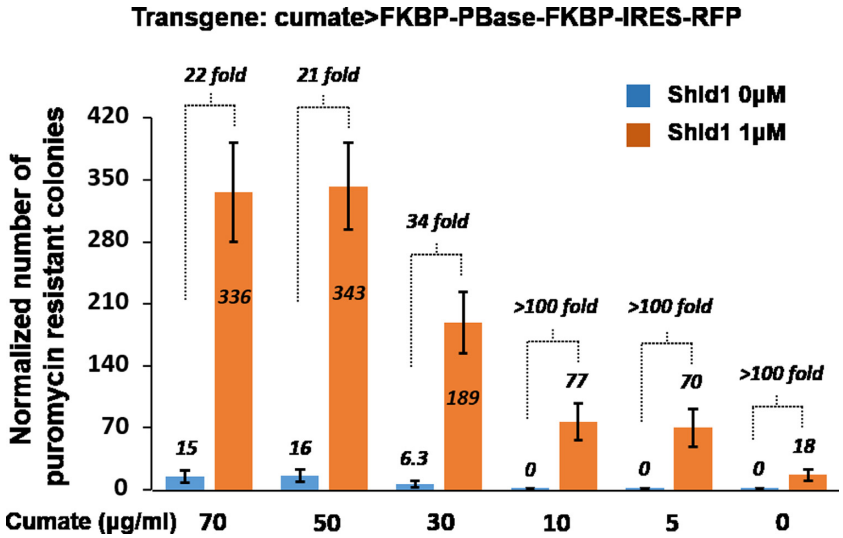


**Figure 5.** The performance of the cumate-regulated FKBP-based PB transposon induction system. Cumate was included in culture media at concentrations ranging from 70 μg/ml (full induction of the cumate promotor) to 0 μg/ml (no induction). For each cumate concentration, Shld1 (yellow) or vehicle (70% ethanol, blue) was added to the medium to activate the PB transposon induction system. The normalized number of puromycin-resistant colonies and the fold-activation are plotted at different cumate/Shld-1 concentrations.

### The FKBP-PBase-FKBP protein levels can be tuned post-transcriptionally

We next sought to determine if PB activity could be post-transcriptionally tuned by adjusting Shld1 concentration. To do so, we transfected four cell lines with donor and helper plasmids (Figure 2, construct 3 and 16) and subjected these transfected cells to various concentrations of Shld1: 0 nM, 8 nM, 40 nM, 200 nM and 1 μM. We calculated the normalized number of puromycin resistant cell colonies for each Shld1 concentration. We found that the colony number increased almost linearly with the increase of the concentration across all cell lines ($r = 0.81$, Figure 6A), indicating the FKBP-based PB transposon induction is tunable and dose-dependent.

Counting puromycin-resistant colonies reveals the number of cells with at least one integrated transposon; however, it does not provide information about the average number of transposition events that have occurred per cell. To address this, we also measured PBase activity at various concentrations of Shld1 using an alternative readout for transposition. We created a donor vector (Figure 2, construct 17) in which a GFP gene is split by a PB transposon, rendering it inactive (Supplementary Figure S3A). Nuclear PBase will excise the transposon creating a functional GFP gene that is then expressed, and so the average fluorescence of the cell population is proportional to the number of transposon excision events. The results of these experiments are shown in Supplementary Figure S3B and S3C. The mean cellular fluorescence increases monotonically with Shld1 in a near-linear fashion, validating the results obtained with the puromycin donor, and supporting the thesis that PB induction is tunable and dose-dependent.

### The FKBP based PBase system is reversible

To determine if the degradation of FKBP-PBase is reversible, we fused a yellow fluorescent protein (YFP) in-frame between FKBP and PBase by an 18 amino acid linker at N terminus of PBase (33) (Figure 2, construct 7), allowing the visualization and semi-quantitative analysis of the expression of the PBase fusion by monitoring fluorescence intensity. The four cell lines mentioned above were transfected with the helper plasmid (FKBP-YFP-PBase) and separately, a control plasmid (YFP-PBase). Shld1 was added to the culture medium immediately upon transfection and incubated for one day. Next, the cells were passaged and cultured in drug-free medium for three days. Finally, Shld1 was added back again to re-stabilize the FKBP-YFP-PBase fusion for another three days. The typical images from fluorescence microscopy at three turning points were shown in Figure 6B: (i) one day after treatment with Shld1, (ii) three days after removing Shld1 and (iii) three days after retreatment with Shld1. We also measured the percentage of YFP positive cells every day by flow cytometry. To control for plasmid dilution during the experiment, the percentages of YFP positive cells from the FKBP-YFP-PBase fusions were normalized by the percentage of YFP positive cells in the corresponding control YFP-PBase samples. The results are plotted in Figure 6C. Taking Figure 6B and C together, we observed a sharp drop in the YFP positive cell population at the second measurement time point, indicating that without the inducer Shld1, the fusion proteins were quickly degraded. The normalized percentage of YFP positive cells was nearly completely restored to the original value three days after Shld1 was reapplied, indicating the system is reversible. By fitting the curves in Figure 6C to an exponential decay functions, we estimated the half-life of the FKBP-PB fusion protein to be 30 h. Our results suggested FKBP-based inducible PB transposon system is tunable and reversible.

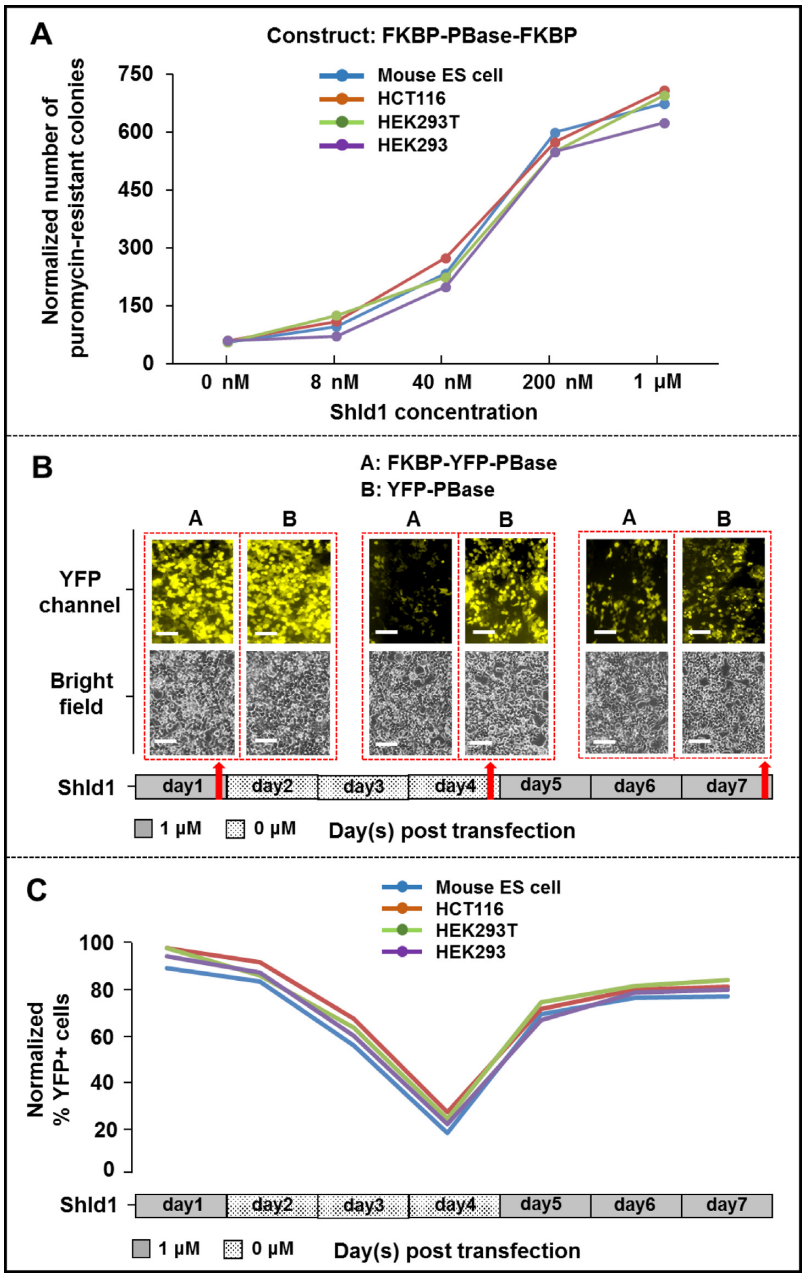


**Figure 6.** Tunability and reversibility of the FKBP-based PB transposon induction system. (**A**) FKBP-PBase-FKBP activity is tunable. Cells were transfected with donor and helper plasmids (FKBP-PBase-FKBP) and subjected to various concentrations of Shld1: 0 nM, 8 nM, 40 nM, 200 nM and 1 μM. The normalized number of puromycin-resistant colonies observed is plotted versus the Shld1 concentration for four cell lines. (**B**) Shld-1 induction is reversible. Fluorescent images are taken at various timepoints after induction with 1uM Shld1 or removal of Shld1 from transfected HEK293T cells. The white scale bar equals 50 μm. The red arrows indicate the time points of measurement. Bright field and fluorescent images were shown at top and bottom panel respectively. (**C**) Quantification of the reversibility of Shld-1 induction. The cells imaged in panel B were quantified by FACS. The normalized percentage of YFP positive cells is plotted at each time point for four cell lines.

### Unlike 4OHT, Shld1 does not interfere with general cellular functions

It has previously been reported that 4OHT adversely affects developmental processes such as neurogenesis (23), myelinogenesis (23), myometrial differentiation (24) and sexual maturity (25). Therefore, we next sought to test if cells treated with Shld1 or 4OHT display any developmental phenotypes. We added these chemicals to EBs generated from mouse ESCs, and differentiated them into ventral spinal neural cells with retinoic acid and smoothened agonist (28). We found that in the presence of 2 μM 4OHT, EB differentiation was inhibited and neuron-like cells were not generated. In contrast, EBs that were mock-treated or treated with 2 μM Shld1 differentiated normally (Figure 7A). These results suggest that 4OHT inhibits EB differentiation, while Shld1 does not.

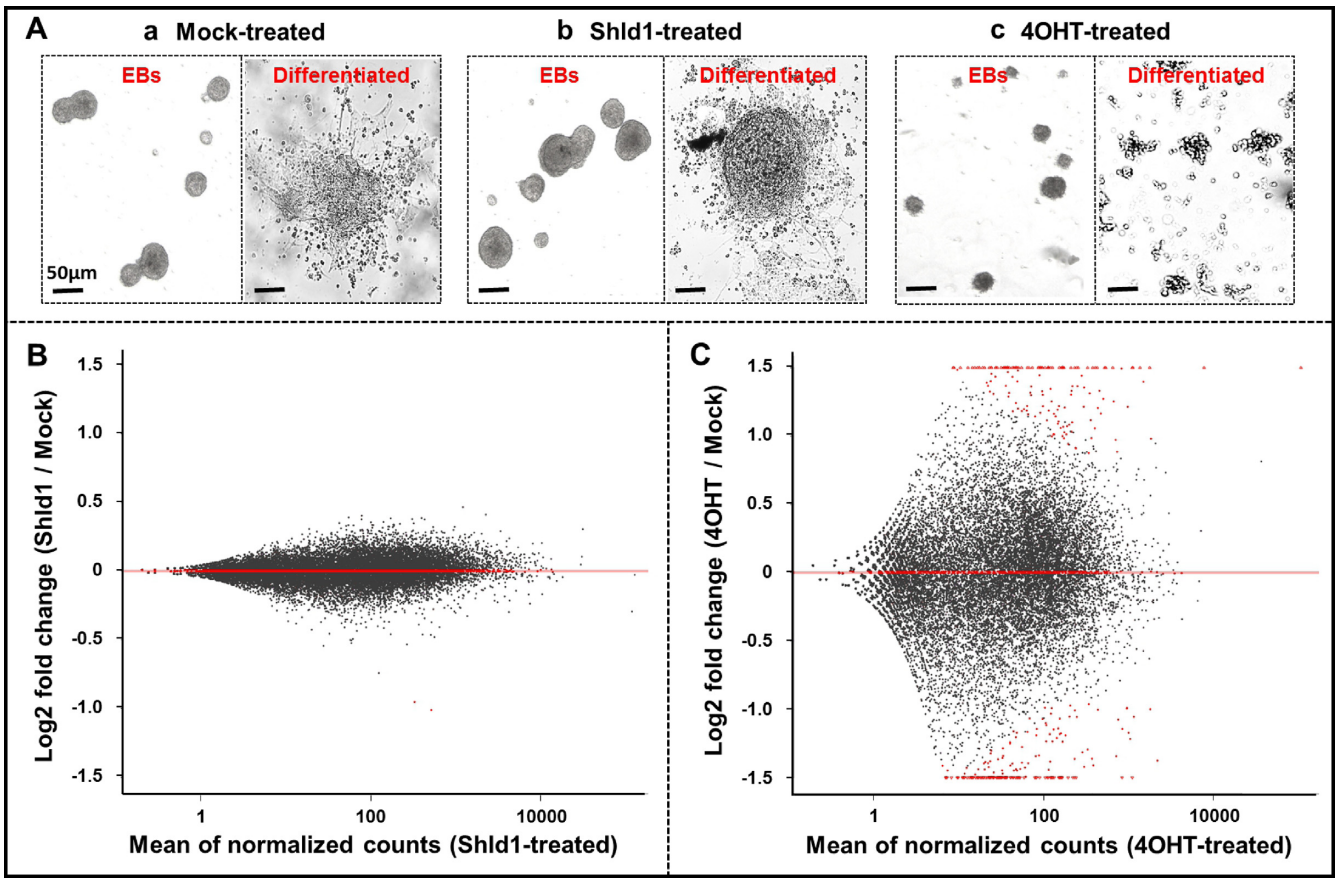


**Figure 7.** The effects of chemical inducer on RW4 mouse ES cells (ESCs). (**A**) Images of embryoid bodies (EBs) and differentiated EBs treated with (a) 95% ethanol as vehicle, (b) 2 μM Shld1 and (c) 2 μM 4-hydroxytamoxifen during EB formation and neural differentiation from mouse ESCs to neural lineages. (**B** and **C**) The comparison of gene expression profile of drug-treated samples with that of mock-treated ones. Experiments were done in duplicate. Mean of the normalized counts from drug-treated samples was plotted against log2 fold change of drug-treated samples to mock-treated ones under the category of (**B**) Shld1 versus Mock and (**C**) 4OHT versus Mock. Red dots indicate the differentially expressed genes.

To further explore the effects of the inducer molecules on differentiation, we sought to quantify the extent to which these Shld1 and 4OHT perturb the transcriptional network of EBs. EBs were treated with Shld1 and 4OHT, respectively, for 2 days, induced with retinoic acid and smoothened agonist for 2 more days, and then subjected to gene expression profiling by RNA-Seq. Cells mock-treated with 95% ethanol as vehicle were used as controls. For each condition, two biological replicates were performed and correlation co-efficiencies between them are above 0.96 (Supplementary Figure S5), demonstrating reproducibility. The mean of the normalized counts from drug-treated samples was plotted against log2 fold change of drug-treated samples to mock-treated ones under the category of Shld1 versus Mock and 4OHT versus Mock (Figure 7B and C). A significant change of gene expression profile was observed for 4OHT-treated samples but not in Shld1-treated ones. After Benjamini–Hochberg correction for multiple hypotheses (see methods), we identified 260 differentially expressed genes in the 4OHT-treated samples (Supplementary data 3). We performed gene ontology analysis on these genes using the DAVID Bioinformatics package (38,39) and functional annotation clustering revealed that 'heat shock' and 'stress response' were the most over-represented terms (adjusted $P < 0.01$), a result that supports our experimental observation that 4OHT is toxic to EBs. In contrast, only one gene (FOS, NM_005252) was differentially expressed in the Shld1 treated samples, suggesting that this chemical has little to no effect on the transcriptional network that controls development of EBs.

## DISCUSSION

We have systematically characterized, in four cell lines, 15 different PB transposase fusion proteins representing three different induction systems. We found that the FKBP-based system achieved the broadest dynamic range of induction across four cell lines. Remarkably, in the presence of chemical inducer, this system had transposition efficiencies that were almost as high (~95%) as the 'wild-type' PB transposon. Our results, coupled with the fact that Shld1 does not affect ESC development, suggest that this will be the preferred induction system for many types of experiments, especially those involving cellular differentiation or organismal development.

In our plasmid-based experiments (Figure 3), we observed a lower dynamic range for the ERT2-based PB transposon induction system that was previously observed by

Cadinanos *et al*. (∼20-fold versus ∼270-fold) (9). This discrepancy is likely explained by the fact that, in our donor plasmid, the puromycin resistance gene is driven by a PGK promoter whereas in Cadinanos *et al*., the puromycin resistance gene is promoterless and downstream of a splicing acceptor site. In that system, only insertions in active genes will produced colonies, but in our system, all insertions will produce resistant clones. This would explain why we observed more background puromycin-resistant clones when the transposase was not induced since both studies normalized to background and small absolute changes in the denominator can lead to large changes in the calculated fold change.

The moderate background transposition of the DD induction system observed in our plasmid experiments (Figure 3) was largely the result of the over-expressed fusion protein overloading the cells' proteasome system, since when the FKBP-fused PBase was placed under the control of a cumate-titratable promoter and delivered at lower copy number, we observed no background transposition while still achieving robust transposition after induction. Since no background hops were observed at several cumate concentrations, the fold-induction of the system is not defined; we conservatively estimate it as >100-fold.

An interesting difference between DD-based systems and the ERT2 based system is that the performance of the DD systems was more uniform across the difference cell lines. We hypothesize that this difference is due to the fact that the DD system uses the ubiquitin–proteasome system to constantly degrade FKBP-fused PBase. The ubiquitin–proteasome system is a general degradation machinery and uniformly expressed and works efficiently in all mammalian cell types. In contrast, the ERT2-based induction system is dependent on a specific cytoplasmic protein HSP90, a gene whose expression varies widely between cells. Our RNA sequencing data showed that the HSP90 expression level in the RW4 mouse ES cell line is about 4-fold higher than in the HCT116 cell line, suggesting that a low amount of HSP90 levels may lead to either a high basal activity without induction or a low induced activation.

In the FKBP-based PB transposon induction system, we observed that N-terminally tagged PBase (FKBP-PBase) has a higher fold induction than the C-terminally tagged counterpart (PBase-FKBP), which may suggest that FKBP fused at the N terminus is more likely to be exposed and thereby recognized by the proteasome than when fused to the C terminus. This terminal specific higher activity may relate to the distribution of functional components of PBase. It was reported that the PBase has a functional nuclear targeting signal in the 94 C-terminal residues (40). C-terminal fused PBase, therefore, is more likely to negatively affect the nuclear translocation of the PBase protein, which in turn increases the likelihood of being recognized and degraded by the proteasome in the cytoplasm. We also observed that fusing additional DDs to the PB transposase increases the fold-change in activation between the induced and uninduced conditions. However, the addition of multiple DDs causes the absolute activity of the induced protein to be significantly reduced, suggesting a trade off between how tightly a protein can be regulated, and its maximum activity in the induced state. Our results suggest that the FKBP-DD is optimal when a high level of PB transposition is preferred, while FKBP-FKBP-PBase-FKBP is the choice when tight regulation is the priority.

The FKBP system will be likely prove considerably more useful than then ERT2 system for the study developmental processes, since we found that the ERT2 inducer tamoxifen is a strong inhibitor of embryoid body differentiation. Our results are consistent with previous reports that tamoxifen has deleterious effects on a number of developmental systems (23–25). In contrast, we found that Shld1, the FKBP inducer, has no such phenotype, and furthermore, few transcriptional changes were observed in our RNA-Seq analysis of embryoid bodies treated with Shld1. Together, these results suggest that the FKBP DD system is prefereable to the ERT2 system for experiments involving cellular differentiation.

In summary, we have investigated the applicability of three different induction systems to the PB transposase by characterizing and optimizing 15 fusion constructs to develop a broadly applicable PB transposon induction system based on the FKBP degradation domain. This induction system satisfies five important criteria: (i) The system has a low basal transposition activity when 'off ' and maintains high transposition activity when 'on'; (ii) The PBase fusion protein shows high transposition activity almost equal to that of the unfused 'wild type' PBase; (iii) The system can be applied across different cell lines with high performances; (iv) The induction is reversible and responds in a dose-dependent manner; (v) The chemical inducer does not interfere with general cellular function. Taken together, our experiments suggest that this system can be readily applied to diverse research applications such as PB transposon-mediated genome engineering and technology development.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENT

The RW4 Mouse ESC line was generously provided by Dr David Gottlieb. The authors thank Dr Christian Shively and Dr Joe Dougherty for advice and useful discussions in preparation of the manuscript. The authors are grateful to Jessica Hoisington-Lopez and the Washington University Genome Technology Access Center for their expert assistance with the use of Illumina Genome Analyzer. The authors thank the Alvin J. Siteman Cancer Center at Washington University School of Medicine and Barnes-Jewish Hospital in St. Louis, MO, for the use of the Siteman Flow Cytometry Core, which provided the Flow cytometry analysis service. The Siteman Cancer Center is supported in part by an NCI Cancer Center Support Grant #P30 CA91842.

## FUNDING

National Institutes of Health [U01MH109133 and 5R01NS076993]; WM Keck Foundation Grant. Funding for open access charge: National Institutes of Health [U01MH109133 and 5R01NS076993].

*Conflict of interest statement.* None declared.

## REFERENCES


1. Carlson,C.M. and Largaespada,D.A. (2005) Insertional mutagenesis in mice: new perspectives and tools. *Nat. Rev. Genet.*, **6**, 568–580.
2. Ivics,Z., Hackett,P.B., Plasterk,R.H. and Izsvak,Z. (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*, **91**, 501–510.
3. Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H. Jr (1996) High frequency retrotransposition in cultured mammalian cells. *Cell*, **87**, 917–927.
4. Fraser,M.J., Brusca,J.S., Smith,G.E. and Summers,M.D. (1985) Transposon-mediated mutagenesis of a baculovirus. *Virology*, **145**, 356–361.
5. Wu,S.C., Meir,Y.J., Coates,C.J., Handler,A.M., Pelczar,P., Moisyadi,S. and Kaminski,J.M. (2006) piggyBac is a flexible and highly active transposon as compared to sleeping beauty, Tol2, and Mos1 in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 15008–15013.
6. Fraser,M.J., Ciszczon,T., Elick,T. and Bauser,C. (1996) 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.*, **5**, 141–151.
7. Wilson,M.H., Coates,C.J. and George,A.L. Jr (2007) PiggyBac transposon-mediated gene transfer in human cells. *Mol. Ther.*, **15**, 139–145.
8. Ding,S., Wu,X., Li,G., Han,M., Zhuang,Y. and Xu,T. (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*, **122**, 473–483.
9. Cadinanos,J. and Bradley,A. (2007) Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res.*, **35**, e87.
10. Wu,S., Ying,G., Wu,Q. and Capecchi,M.R. (2007) Toward simpler and faster genome-wide mutagenesis in mice. *Nat. Genet.*, **39**, 922–930.
11. Woltjen,K., Michael,I.P., Mohseni,P., Desai,R., Mileikovsky,M., Hamalainen,R., Cowling,R., Wang,W., Liu,P., Gertsenstein,M. *et al.* (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature*, **458**, 766–770.
12. Yusa,K., Rad,R., Takeda,J. and Bradley,A. (2009) Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nat. Methods*, **6**, 363–369.
13. Rad,R., Rad,L., Wang,W., Cadinanos,J., Vassiliou,G., Rice,S., Campos,L.S., Yusa,K., Banerjee,R., Li,M.A. *et al.* (2010) PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. *Science*, **330**, 1104–1107.
14. Wang,H., Mayhew,D., Chen,X., Johnston,M. and Mitra,R.D. (2011) Calling Cards enable multiplexed identification of the genomic targets of DNA-binding proteins. *Genome Res.*, **21**, 748–755.
15. Wang,H., Mayhew,D., Chen,X., Johnston,M. and Mitra,R.D. (2012) 'Calling cards' for DNA-binding proteins in mammalian cells. *Genetics*, **190**, 941–949.
16. Wang,H., Johnston,M. and Mitra,R.D. (2007) Calling cards for DNA-binding proteins. *Genome Res.*, **17**, 1202–1209.
17. Banaszynski,L.A., Chen,L.C., Maynard-Smith,L.A., Ooi,A.G. and Wandless,T.J. (2006) A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell*, **126**, 995–1004.
18. Iwamoto,M., Bjorklund,T., Lundberg,C., Kirik,D. and Wandless,T.J. (2010) A general chemical method to regulate protein stability in the mammalian central nervous system. *Chem. Biol.*, **17**, 981–988.
19. Pick,E., Kluger,Y., Giltnane,J.M., Moeder,C., Camp,R.L., Rimm,D.L. and Kluger,H.M. (2007) High HSP90 expression is associated with decreased survival in breast cancer. *Cancer Res.*, **67**, 2932–2937.
20. Zagouri,F., Sergentanis,T., Nonni,A., Papadimitriou,C., Pazaiti,A., Michalopoulos,N.V., Safioleas,P., Lazaris,A., Theodoropoulos,G., Patsouris,E. *et al.* (2010) Decreased Hsp90 expression in infiltrative lobular carcinoma: an immunohistochemical study. *BMC Cancer*, **10**, 409–414.
21. Zhao,H., Yang,H., Zhao,H., Chen,M. and Wang,T. (2011) The molecular characterization and expression of heat shock protein 90 (Hsp90) and 26 (Hsp26) cDNAs in sea cucumber (Apostichopus japonicus). *Cell Stress Chaperones*, **16**, 481–493.
22. Deb,R., Sajjanar,B., Singh,U., Kumar,S., Singh,R., Sengar,G. and Sharma,A. (2014) Effect of heat stress on the expression profile of Hsp90 among Sahiwal (Bos indicus) and Frieswal (Bos indicus × Bos taurus) breed of cattle: a comparative study. *Gene*, **536**, 435–440.
23. Nobakht,M., Najafzadeh,N. and Kordestani Shargh,B. (2009) Effects of tamoxifen on morphological and ultrastructural aspects of developing hippocampus of rat. *Iranian Biomed. J.*, **13**, 237–243.
24. Mehasseb,M.K., Bell,S.C. and Habiba,M.A. (2009) The effects of tamoxifen and estradiol on myometrial differentiation and organization during early uterine development in the CD1 mouse. *Reproduction*, **138**, 341–350.
25. Singh,R., Singh,A.K. and Tripathi,M. (2012) Effect of a non steroidal tamoxifen on the gonad and sex differentiation in Nile tilapia, Oreochromis niloticus. *J. Environ. Biol.*, **33**, 799–803.
26. Sellmyer,M.A., Chen,L.C., Egeler,E.L., Rakhit,R. and Wandless,T.J. (2012) Intracellular context affects levels of a chemically dependent destabilizing domain. *PloS One*, **7**, e43297.
27. Egeler,E.L., Urner,L.M., Rakhit,R., Liu,C.W. and Wandless,T.J. (2011) Ligand-switchable substrates for a ubiquitin-proteasome system. *J. Biol. Chem.*, **286**, 31328–31336.
28. Wichterle,H., Lieberam,I., Porter,J.A. and Jessell,T.M. (2002) Directeddifferentiation of embryonic stem cells into motor neurons. *Cell*, **110**, 385–397.
29. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
30. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
31. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
32. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
33. Wilson,M.H. and George,A.L. Jr (2010) Designing and testing chimeric zinc finger transposases. *Methods Mol. Biol.*, **649**, 353–363.
34. Wang,W., Bradley,A. and Huang,Y. (2009) A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. *Genome Res.*, **19**, 667–673.
35. Mansergh,F.C., Daly,C.S., Hurley,A.L., Wride,M.A., Hunter,S.M. and Evans,M.J. (2009) Gene expression profiles during early differentiation of mouse embryonic stem cells. *BMC Dev. Biol.*, **9**, 5–22.
36. Bradley,E., Bieberich,E., Mivechi,N.F., Tangpisuthipongsa,D. and Wang,G. (2012) Regulation of embryonic stem cell pluripotency by heat shock protein 90. *Stem Cells*, **30**, 1624–1633.
37. Mullick,A., Xu,Y., Warren,R., Koutroumanis,M., Guilbault,C., Broussau,S., Malenfant,F., Bourget,L., Lamoureux,L., Lo,R. *et al.* (2006) The cumate gene-switch: a system for regulated expression in mammalian cells. *BMC Biotechnol.*, **6**, 43–60.
38. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
39. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
40. Keith,J.H., Fraser,T.S. and Fraser,M.J. Jr (2008) Analysis of the piggyBac transposase reveals a functional nuclear targeting signal in the 94 c-terminal residues. *BMC Mol. Biol.*, **9**, 72–84.