Site-Specific Integration of Large DNA Fragments - Evaluating and Redesigning Genome Editing Systems

Abstract

A major bottleneck of genome editing today is the tradeoff between the size of the integrated fragment and the ability to control where it is integrated. Currently, there are no techniques to reliably integrate sequences greater than few thousand base pairs in precise locations. PiggyBac transposase (PBase) is an enzyme capable of transposing mobile genetic elements of up to 200 kb. However, PBase integrates randomly into the genome, resulting in a lack of specificity. To address this issue, we introduce a novel technology for genomic insertion, achieving site-specific integration of fragments larger than previously possible. First, we established PBase mediated integration as a safe and viable technique by demonstrating no cytotoxicity upon insertion. Then, we fused PBase to a deactivated Cas9 (CRISPR-based programmable DNA-binding) to specifically direct PBase integration. This fusion was successful at targeting and displacing a red fluorescent protein gene with another genetic cassette. We show 38% increased on-target integrations as compared to WT PBase. We believe the ability to specifically integrate large DNA fragments will have a transformative impact on multiple fields, with applications such as the bioproduction of pharmaceuticals and fuels, correcting mutated genes, creating vaccine libraries, and storing data on DNA.

Site-Specific Integration of Large DNA Fragments - Evaluating and Redesigning Genome Editing Systems

Introduction

Genetic engineering technologies have revolutionized our ability to probe and edit the genome at a single nucleotide resolution. These technologies have been used to correct genetic disorders in vitro and in vivo. [1,2] Arguably, the next generation of genome editing lies with integrating large payloads to tackle a greater variety of genetic diseases. However, currently we can only integrate around 6-7 kb of DNA precisely [3]. Achieving mutation free insertion of large fragments of DNA is a valuable endeavor not only for medical applications but also for biomanufacturing and biopharmaceuticals.

HR is limited in both its payload capacity and feasibility, as HR usually occurs only during the S phase of the cell cycle. Error-prone non-homologous end joining is more efficient than HR throughout all stages of the cell cycle [3]. Furthermore, as the payload size increases, the efficiency of HR decreases. These limitations reveal the need for another method to integrate large sequences reliably.

Transposable elements appear to be a promising solution for large fragment stable integration, but there has been limited research to date. Transposable elements (transposons) are endogenous systems capable of excising, transporting, and integrating DNA sequences. They have multiple advantages over viral vector integration integration including larger payload capacity, reduced insertional mutagenesis, reduced immunogenicity, reduced affinity for coding regions, and reduced cytotoxicity [4]. However, their insertion sites are random, which can cause unknown off-target effects. Currently, transposons are typically used for mutagenesis, but we hypothesized that the

piggyBac transposon system can be used to integrate payloads over 10 kb at user-defined sequences.

piggyBac (PB) is a cut and paste transposon system that has been shown to mobilize >150 kb fragments in human embryonic stem cells[5],[6]. The PB system (**Figure 1**) contains two parts: a transposon and transposase. piggyBac transposase (PBase) recognizes ‘TTAA’ nucleotide sites in the genome and integrates the piggyBac transposon into the TTAA locus [7]. These sights are abundant in the genome, and therefore PBase integration is virtually random.

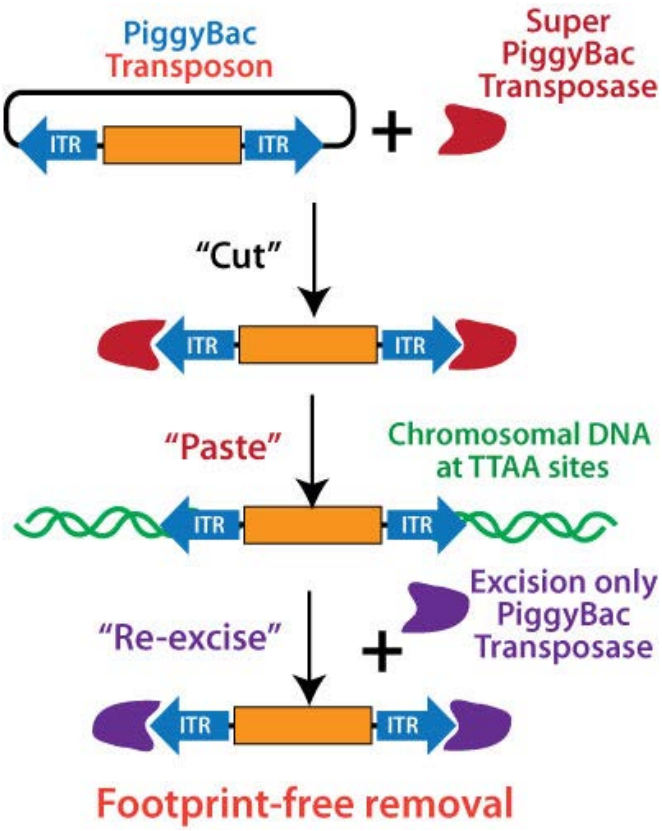


Figure 1 – PBase Mechanism
(a) The PiggyBac transposon enables the integration of large DNA fragments in the form of a transposon. The piggyBac recognizes inverse terminal repeats (ITRs). Upon ITR recognition, PBase excises the transposon from a donor vector. PBase then integrates into genomic TTAA sites. The transposon is permanently integrated with the ITRs. Figure adopted from Systems Bio International.

In order to bias targeting of transposable elements, researchers have designed DNA Binding Domain (DBD) -Transposase fusions [8,9,10,11,12]. Zinc Fingers Nucleases (ZFN) and Transposable Activator Like Effectors (TALEs) are both DBDs that have the capability to target user-defined sequences, but their difficulty to design and deliver make them less than ideal for genomic therapy. Fusions have been created using transposons such as *sleeping beauty*, but have shown successful integrations with efficiencies <1% [8],[13]. A PBase fusion may hold more promise considering that PBase has high efficiency of insertion, low local hopping, footprint-free excision, lack of overproduction inhibition, and larger cargo size compared to Tol2 and *sleeping beauty [14]*,[15]. However, previous PBase fusions have been unsuccessful at targeted integration as well. More recently discovered DBDs, such as dCas9, appear to be a promising alternative. To our knowledge, nobody has attempted to fuse dCas9 to PBase.

CRISPR-Cas9 is a powerful tool for targeted gene therapy. Using a guide RNA (gRNA), Cas9 is able to target user-specified regions of the genome. (**Figure 2)** The gRNA is complementary to a 20bp region in the genome that is user-specified and is supplied with Cas9 upon transfection [16,17]. To change the target locus, all that is required is a different gRNA. Deactivated Cas9 (dCas9) is a Cas9 protein without catalytic activity [18]. The interchangeability of Cas9 and piggyBac would yield a highly customizable system wherein both the target locus and payload can easily be exchanged.

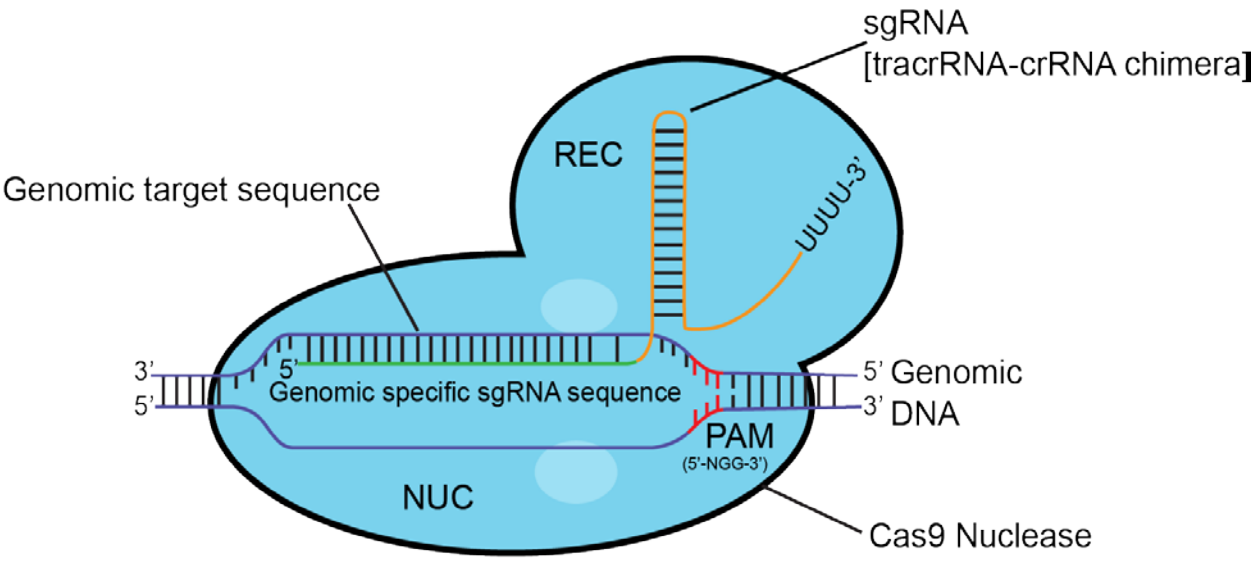


Figure 2 – dCas9 Mechanism
(a) Mechanism behind dCas9 DNA binding. A guide RNA (gRNA) is bound to the REC (recognition) domain of dCas9 because of its electromagnetic properties. The duplex is then guided by the gRNA to a target site, which has 20 complementary base pairs to the gRNA strand. A protospacer adjacent motif (PAM) (NGG) is required for Cas9 recognition of a gRNA binding site.

We created the first PBase-Cas9 fusion and designed it to target the AAVS1 safe harbor site in HEK-293 Cells. We optimized several factors including fusion orientation and linker length/type with the assistance of protein modelling software. The inherent flexibility of Cas9 and piggyBac makes our fusion a powerful tool that enables high value genomic modifications ranging from the insertion of chimeric genes to the correction of mutated ones.

Materials and Methods

Transfection: Tet-On HEK293 were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 9% Fetal Bovine Serum (FBS), 0.9% MEM NEAA (MEM Non-Essential Amino Acids) and 0.4% Penicillin Streptomycin. Cells were

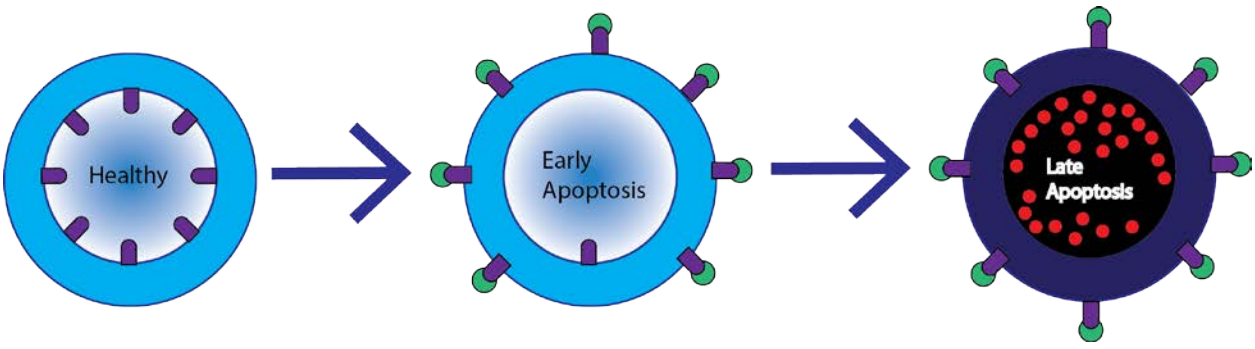
incubated at 37℃ at 5% CO_2 and 95% humidity. Lipofection was performed with Jetprime. The cells were plated in a 12 well format with 200,000 cells per well.

Microscopy: A Hamamatsu camera attached to an Olympus IX81 microscope with a 10X objective was used to capture images with the following filters (Chroma): ET 560/40X (excitation) and ET630/75m (emission) for mKate2 and ET 436/20X (excitation) and ET 480/40m (emission) for Tag-CFP. The brightfield images were captured at 10 ms exposure and Tag-CFP at 100 ms exposure. For all experiments, image normalization and processing were consistent. All raw images were captured with SlideBook 5.0, and ImageJ was used to analyze the pictures after capture.

Cell Cycle: Cell cycle analysis was performed using to assess the effect of PBase activity on the cell. Cells were fixed with ethanol and subsequently stained with propidium iodide (PI) according to the manufacturer's protocols [19]. **(Figure 3)**

Apoptosis: We used the Alexa Fluor® 488 annexin V/Dead Cell Apoptosis Kit with Alexa® Fluor 488 annexin V and PI was used for flow cytometry [19,20]. Cells were harvested two days after transfection and then washed in Phosphate-Buffered Saline (PBS). Then, they were diluted in 1X Annexin binding buffer to $1x10^6$ cells/mL, and 5 µL Alexa Fluor Annexin was added. **(Figure 3)**

a)



b)

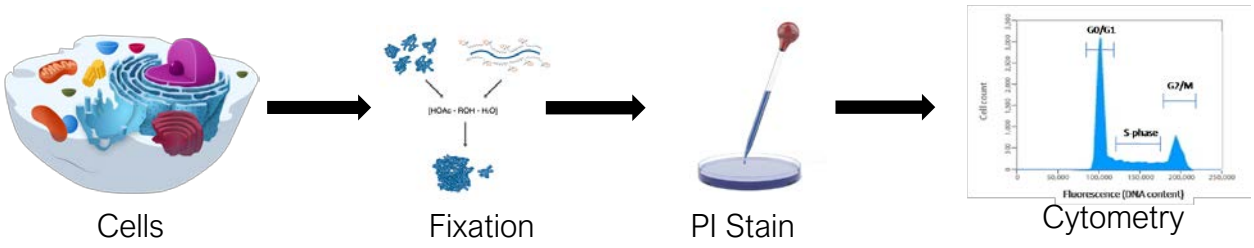


Figure 3 – Apoptosis and Cell Cycle Analysis
(a) Apoptosis results in the death of cells, and is indicative of the deleterious effects of an integrated sequence. Phosphatidylserine is a cellular signaling agent in the phospholipid cell membrane. Upon initiation of apoptosis, phosphatidylserine serine flips to the outside of the membrane, and can be stained with Annexin V. During late apoptosis, the cell ruptures, and intercellular DNA can be stained with propidium iodide (PI). PI and Annexin V signals are quantified using flow cytometry. **(b)** The cell cycle is the series of events that preface cell division. A PI stain is used to see changes in the level of DNA as each step in the cell cycle progresses.

Results

Investigating Cytotoxicity of piggyBac integration

We initially assessed the activity of a hyperactive piggyBac variant gifted from the Moisyadi Lab. The Moisyadi plasmid is a self-inactivating PBase containing an introduced intron in the piggyBac open reading frame (ORF) [21]. We removed this introduced intron to preserve complete activity of PBase. System Biosciences recommends a 1:5 or 2:5 ratio of transposase to transposon [22], and we tested to optimize this transfection ratio for future experiments. Each PBase variant was cotransfected with a transposon containing EF1α-Puro-pCMV-CFP flanked by ITRs. After two weeks of Puromycin stable cell selection, we quantified CFP expression using flow cytometry. The data was gated to remove apoptotic populations and multiple cells registered as a single event. A 2:5 ratio of transposase to transposon was shown to have greatest CFP fluorescence. **(Figure 4,5)**.

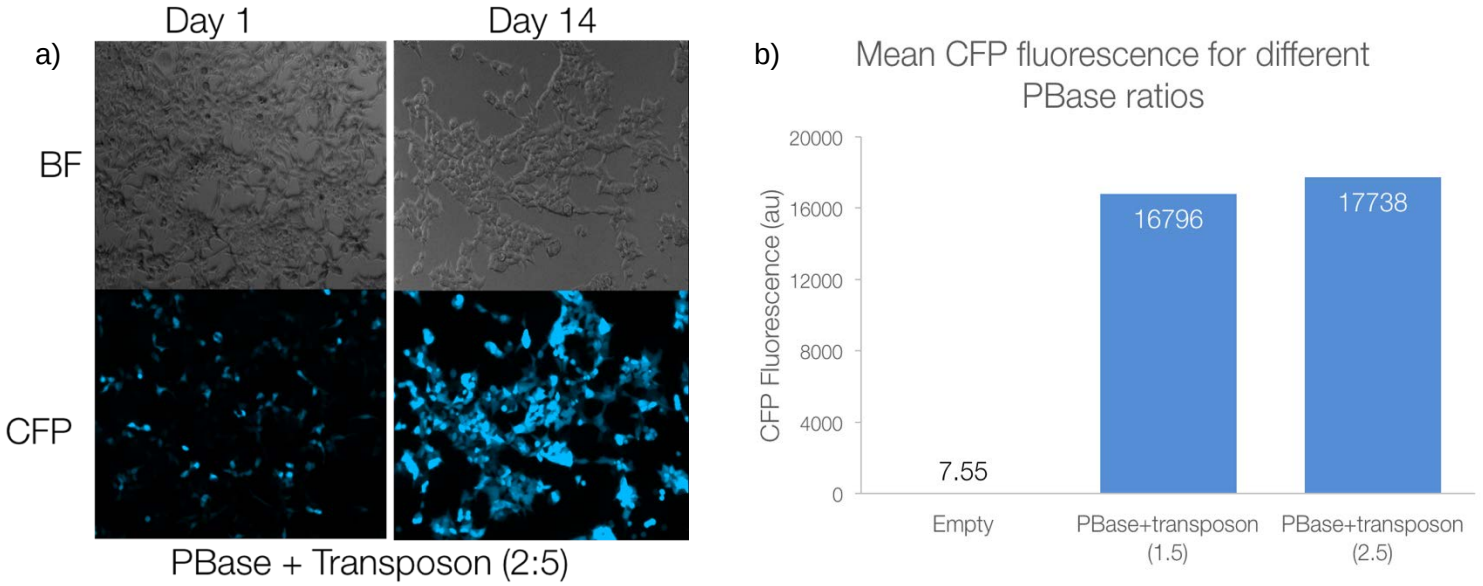


Figure 4 – Cell Microscopy and Flow Cytometry Results
(a) Fluorescent microscopy images of PB stably integrating CFP into the genome during Puromycin stable cell selection. By day 14, the only surviving cells were those transfected with PBase. There was no significant change in cell morphology from day 1 to day 14. **(b)** Flow Cytometry data of Puromycin stable cell selection 14 days post transfection was gated to remove noise. 2:5 ratio of transposase has the highest integration rate (*P<.0001, one tailed-Student's t test; n=5).

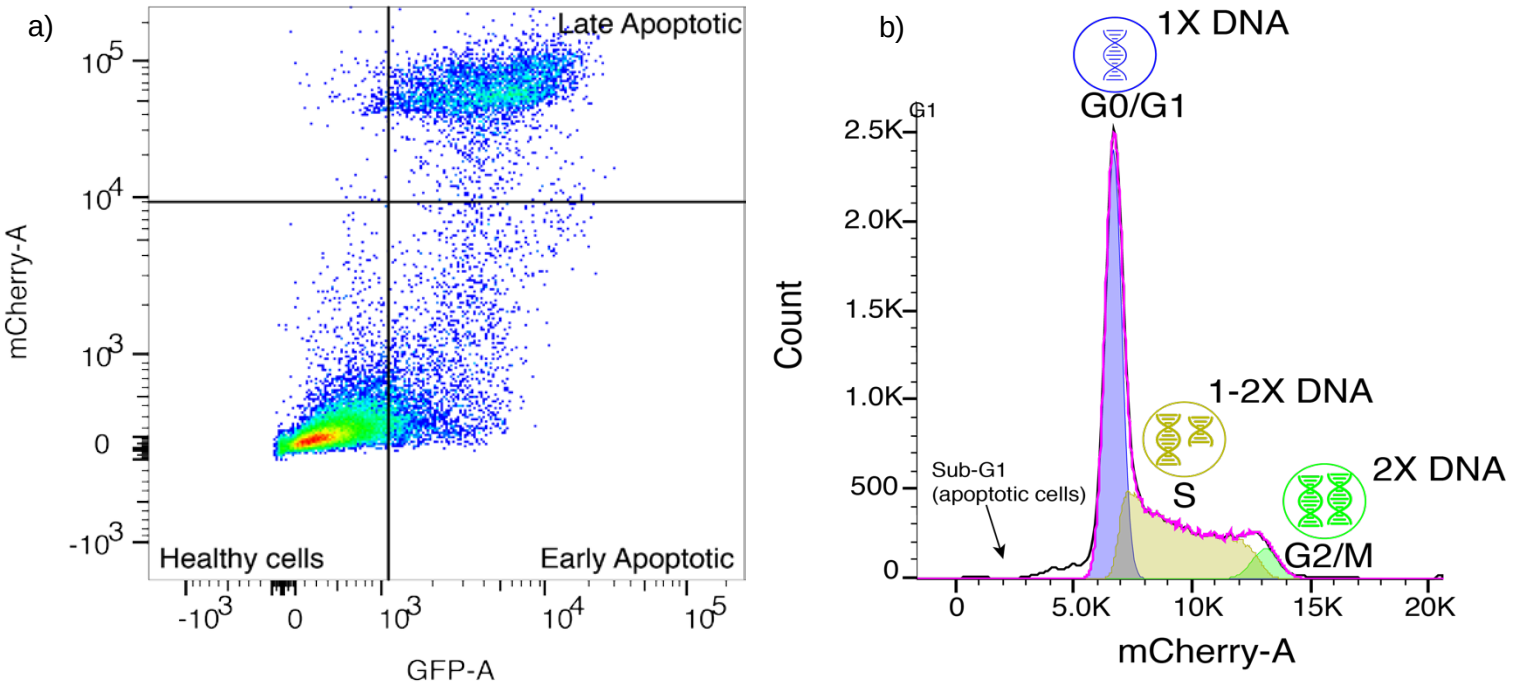
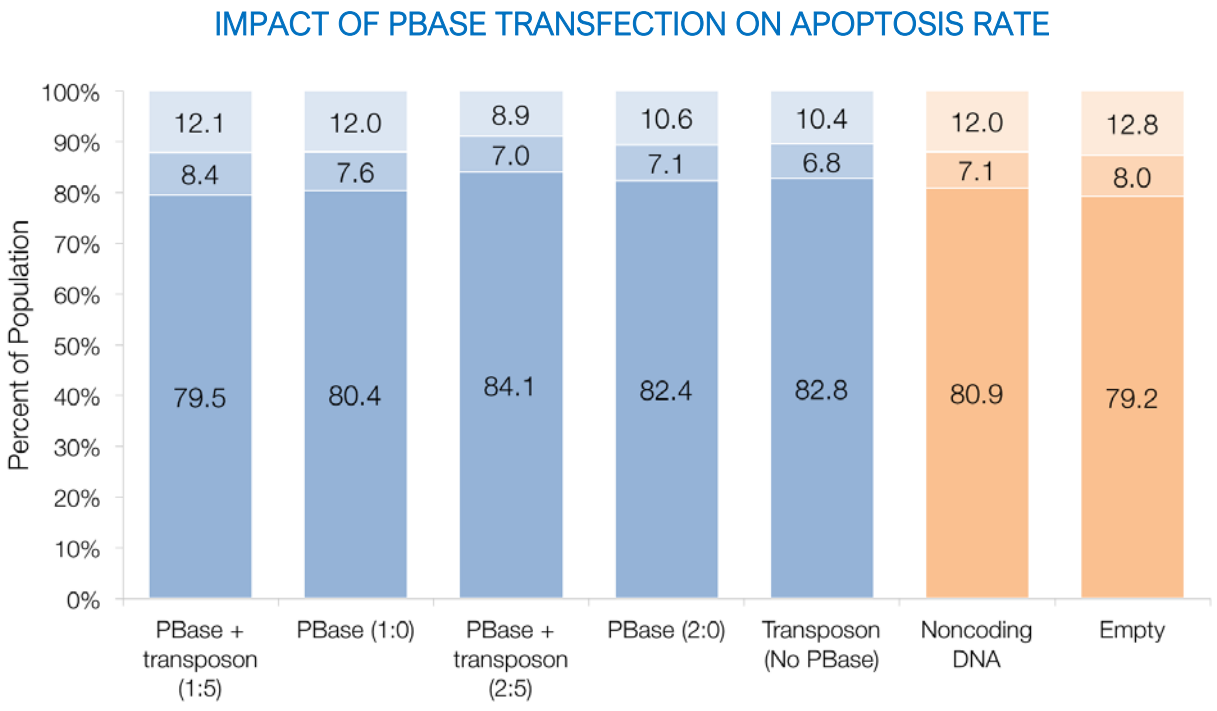


Figure 5 – Flow Cytometry Gating
(a) Apoptosis assay is gated using mKate and GFP to isolate live, early apoptotic, and late apoptotic cells. Cells expressing high levels of both mKate and GFP are considered late apoptotic**. (b)** A cell cycle assay was conducted to verify that piggyBac has low toxicity in cells. A Watson (pragmatic) model was used to determine which cells are in which stage of the cell cycle.

The next step was to investigate the cytoxicity of PBase integrations. There is no clear consensus on the effect of PBase integrations on the cell as random integrations can cause unkown positional effects. Preliminary microscopy brightfield images showed that PBase integrations had no effect on cell morphology, (fig.). We performed apoptosis and cell cycle assays to ensure that PBase was safe to use before designing a fusion protein.

The apoptosis assay verified that PBase transposition did not induce cell apoptosis. We compared varying concentrations of transposon to transposase to a control group (untransfected cells and cells transfected with noncoding DNA). All populations had similar proportions of live, early apoptotic, and late apoptotic cells. This suggests that the addition of PBase had a marginal effect on cell death. (**Figure 6)**

a)



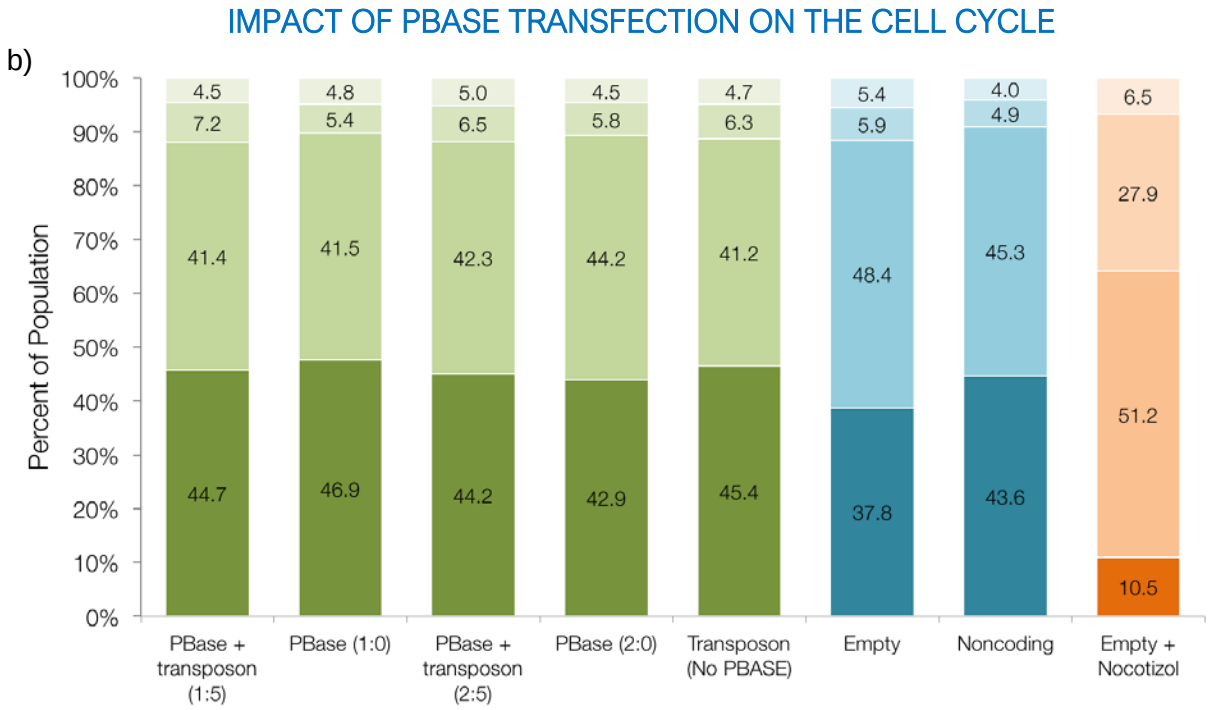


Figure 6 – Apoptosis and Cell Cycle Results
(a) Apoptosis assay of HEK293 cells 48 hours post transfection demonstrates low cytotoxicity. Control cells and treated cells have similar populations of cells in each phase of the cell cycle. (Chi squared independence; *P < .01). **(b)** Cell Cycle assay flow cytometry data shows that piggyBac has marginal effect on cell cycle arrest. Nocodazole (2 μg/mL) induced cell death was used as a benchmark for this assay (Chi squared independence; *P < .01).

Furthermore, we performed a cell cycle assay to show the effects of PBase integration on cell cycle arrest. Once again, the same ratios of transposon to transposase were compared to a negative control group of untransfected cells. We also set up a positive control of cells transfected with Nocodazole (2 μg/ml) to induce G2 arrest. The percentage of cells in the G1, S, and G2 phases of the cell cycle for the experimental group were more similar to the negative control than the positive control. Both of the assays confirmed that PBase integrations did not induce significant levels of apoptosis or cell cycle arrest. With these promising results we established PBase as a safe tool for integration of foreign DNA fragments. **(Figure 6)**

Design and Mechanism of Cas9-Pbase Fusion

The first consideration when designing a Cas9-PBase fusion was the orientation of both proteins. Previous fusions to PBase have had the most success when fusing to the N terminus [12], while Cas9 has been successfully fused at either terminus [23–26]. Therefore, we fused the C terminus of Cas9 to the N terminus of PBase. We also considered other factors including the type of linker, linker length, and the addition of additional nuclear localization sequences. Owens et al. created a TALE-PBase fusion using either a linker or directly attached to the plasmid backbone. The plasmid tethering strategy did not work, so we used a simple GGGGS flexible linker to design our fusion. Since PBase already contains a sequence for nuclear localization, an additional nuclear localization signal was not necessary .

The cloned Cas9-PBase fusion requires a TTAA target site for PBase to bind to. Upon transfection, the dCas9/gRNA ribonucleoprotein complex localizes to its complementary target sequence upstream of the desired TTAA integration locus. Upon detection of inverted terminal repeats (ITRs) on the plasmid bearing a transposon, PBase excises the transposon and integrates it into a site-directed genomic locus. **(Figure 7)**

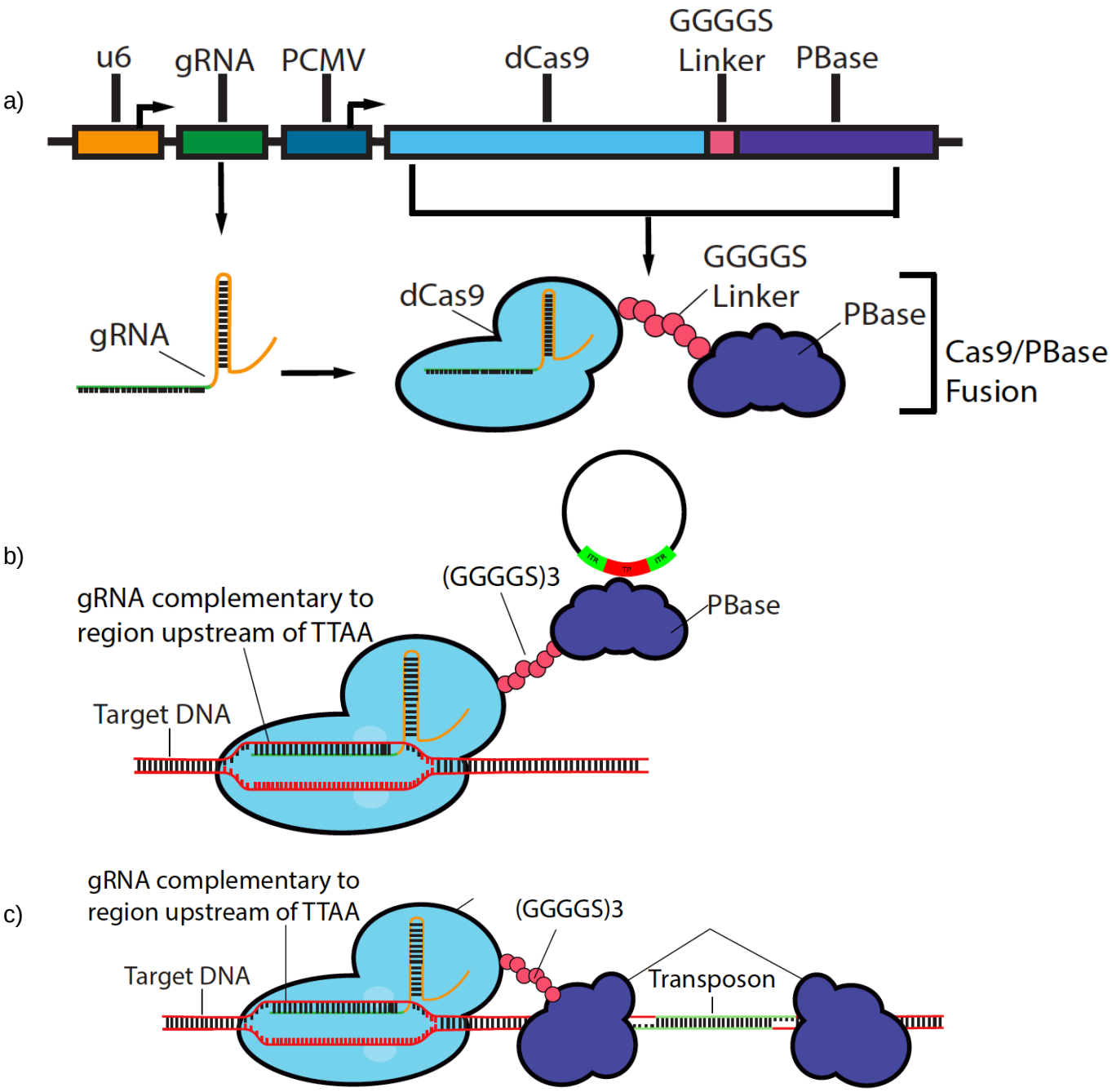


Figure 7 – PBase/dCas9 Fusion Mechanism

(a) DNA construct with PBase fused to the C-terminus of Cas9 using a GGGGS flexible linker. Once translated, a gRNA is loaded into the dCas9 for specific targeting of the genome. **(b**) The dCas9/gRNA heteroduplex locates its complementary target sequence, upstream of the desired TTAA integration locus. The fusion protein is localized to this region, where Upon detection of ITRs on a transposon, PBase separates into subunits and flanks the transposon on either side of each ITR. **(c)** PBase integrates the transposon into a site-directed genomic locus. This transposon can be scarcely re-exercised.

Reporter Integration Assay

We hypothesized that fusing dCas9 to PBase would potentially decrease the integration frequency of PBase due to the overall size of the protein being larger (~6kb). HEK-293 cells were cotransfected with the PBase-dCas9 fusion and a transposon containing ITR-EF1α-Puro-pCMV-CFP-ITR in a 2:5 ratio. After two weeks of puromycin selection, flow cytometry was used to quantify CFP levels in 100,000 cells. To normalize for cells receiving either more than one copy or no copies of the reporter during transfection, the mean CFP production was recorded. Surprisingly, rather than attenuating PBase activity, fusion to Cas9 had ~61% increased integration events than wild type PBase alone (P<.001, one tailed Student's t-test). This showed that a fusion to dCas9 did not compromise PBase activity, and, in fact, increased integration frequency. These results are consistent with Kettlun et. al who also demonstrated increased integration events when fusing a transposon to ZFN. **(Figure 8)**

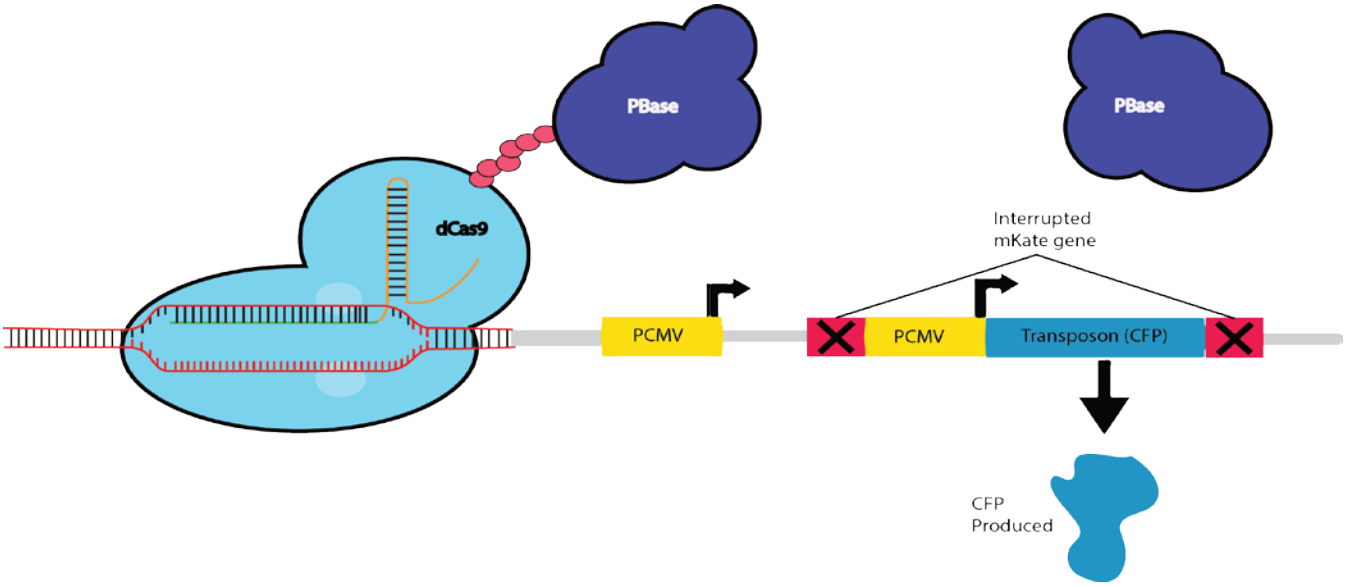


Figure 8 – Transcript Interruption Assay

(a) While puro selection identifies the successful integration of PBase, transcription interruption/NextGen assays give data on the specific site of integration. Integration of a transposon into the ORF of an expressed gene will cease its translation. In this assay,

mKate is interrupted with CFP, which is translated. The ratio of mKate to CFP quantifies the site-specific integration efficiency.

Transcript Interruption Assay

While the integration assay established that our PBase-dCas9 fusion was still catalytically active, no indication was given as to the targeting capacity of the fusion. We designed a transcript interruption assay wherein a HEK-293 cell line with constitutive expression of mKate was targeted by a gRNA. Integration of pCMV-CFP into the ORF of mKate would disrupt mKate transcription, and produce CFP. Therefore, the ratio of mKate to CFP gives the ratio of on-target displacement events to total integration events. A smaller ratio indicates a greater proportion of displacement events since it indicates that the total amount of mKate is lower, while the total amount of CFP is higher. The PBase-dCas9 fusion was co-transfected with a reporter plasmid containing ITR-EF1α-Puro-pCMV-CFP-ITR in a 2:5 ratio and a gRNA targeting stably integrated mKate in the cells. After two weeks of selection, the ratio between mKate to CFP in the fusion was ~38% less than wild type PBase alone. This confirmed that instances of on target integration were occurring. **(Figure 8,9)**

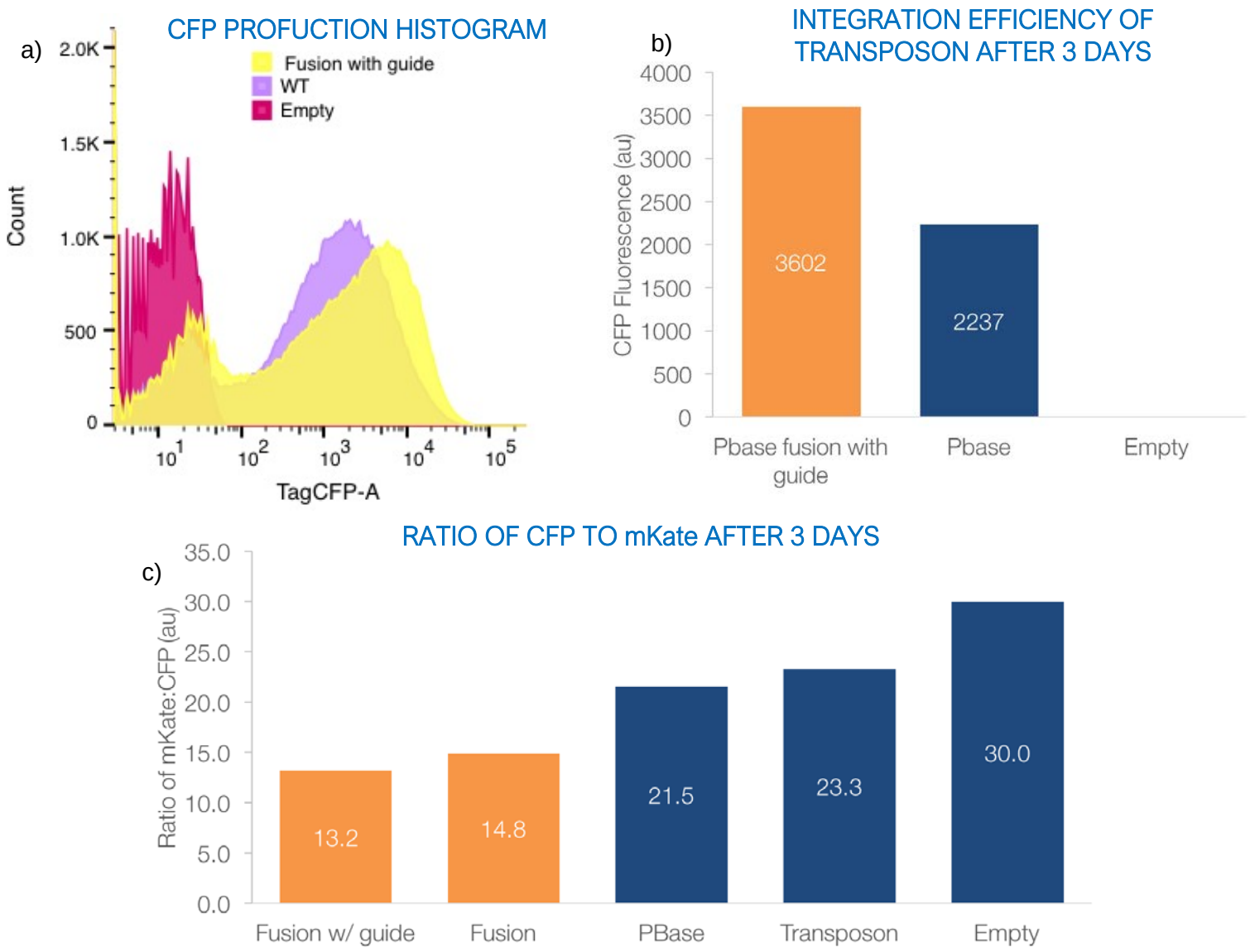


Figure 9 – Integration and Transcript Interruption Assay Results
(a) A CFP-Puro reporter plasmid was transfected at at 1:5 ratio between transposon and transposase. Transfection and puromycin selection demonstrates that the total number of cells affected by transposition of a CFP reporter with a dCas9-PBase fusion are higher than the number affected by WT PBase alone. **(b)** The mean production of CFP indicates the frequency of integration. The fusion seems to have marginally increased the integration frequency while proving that fusion to dCas9 has not compromised the catalytic activity of PBase. (P<.001, one-tailed Student's t-test) **(c)** Lower levels of mKate production indicate increased amounts of site-directed integration. Cotransfection of dCas9-PBase with a gRNA targeting mKate should result in less mKate production than dCas9-PBase and PBase alone. The ratio of mKate to CFP gives the ratio of displacement events to integration events. A smaller ratio indicates more CFP has been integrated and less mKate is being produced. Transfection of the dCas9-PBase fusion resulted in smaller ratios than PBase alone, indicating site-directed integration.

Genomic PCR Assay

To further confirm that site directed integration was achieved, we co-transfected HEK-293 cells with the PBase-dCas9 fusion supplied with and without a gRNA to target the AAVS1 safe harbor site. Two weeks post transfection, genomic DNA was harvested. When PBase integrates, it can either integrate in the forward or reverse direction. To account for this, we performed a nested PCR of the region starting 5 kb upstream and 5kb downstream of the integration locus. Integration within 3.4 kb upstream of the target site was observed with the wild type PBase (which is expected due to its random integration), as well as the fusion with the gRNA. The fusion without the gRNA did not have any integrations in the region, demonstrating that gRNA was necessary for targeted integrations, and evidenced successful targeted integration.

Plasmid integration Assay

We also investigated targeting a plasmid for gene integration.[1] This would provide unique applications in dynamic gene circuits, as genetic elements can be exchanged post transfection. Previous studies have attempted integration into plasmids using a ZFN-PBase fusion, and were successful in recovering plasmids containing integrated transposon sequences [2,8]. We co-transfected our dCas9-PBase fusion, a transposon containing ITR-EF1α-Puro-pCMV-CFP-ITR, and a reporter plasmid containing EF1α-Puro-pCMV-mKate. A gRNA targeting mKate was supplied. We performed a plasmid DNA isolation through Qiagen protocol, and performed a PCR of the transposon sequence. However, no transposon DNA was recovered, and we were unable to recreate the data from previous studies. (Data not shown).

Discussion

The piggyBac transposon shows therapeutic promise, and targeted insertions into safe harbor sites are the next step towards safer editing. The unanswered question is: how precise is precise enough? The challenge with creating a new protein fusion is that there is no direct comparison to be made. Certainly current transposon-DBD fusions are neither precise nor versatile enough [27]. At the same time, a comparison to Cas9 is invalid because our protein is designed for stable integration, not DNA cleavage. Perhaps the best comparison to assess how effective our fusion is would be to compare to the Tale-PBase fusion created by Owens et. al. This is, to our knowledge, the most recent DBD-protein fusion. They were able to demonstrate 0.019% targeting efficiency. In our research, we showed a 38% increase in site-directed integration. However, one of the limitations of our study is we were unable to perform Illumina sequencing to calculate exactly what percent of integrations were on target, and we relied on a less accurate displacement assay instead. Experimentation is ongoing, and we are planning on performing Illumina sequencing in the future to obtain a more precise metric on the increase in targeting efficiency.

To increase the relevance of our technology as a therapeutic treatment, the targeting ability must be further optimized. We hypothesize that the hyperactive PBase variant was increasing too frequently, increasing the probability of an off-target integration. In order to increase specificity and reduce off-target insertions, we are attenuating PBase using two different techniques: splitting PBase protein to separate catalytic residues and inserting point mutations.

We are designing self-complementing split PBase proteins that separate the catalytic residues needed for editing capability. The two subunits will auto-assemble in vivo. This reduces the overall size of the protein, which makes it both easier to package into a viral vector and more mobile within the cell. The PBase is catalytically inactive until both parts are rejoined inside the cell, further increasing the specificity. We split PBase to separate the catalytic residues of D268, D346, and D447. These amino acids are aspartic acids making up the catalytic core of PBase, and PBase cannot excise DNA without all three amino acids [28]. The crystal structure of PBase has not yet been experimentally determined, but with the assistance of I-TASSER protein crystallization software [29,30], we were able to simulate the protein folding of PBase. **(Figure 10)**

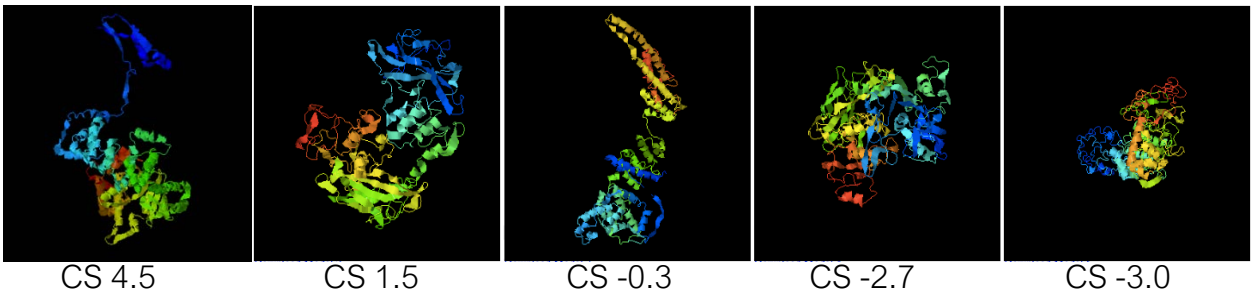


Figure 10 – I-TASSER Crystal Structure Prediction
(a) Prediction software assisted in designing the locus for splitting PBase. PBase was split between catalytic residues canonical for excision (D268, D346, D447). Secondary structure analysis led to the splitting of PBase at helix (H), strand (S), and coil (C) motifs.

In addition to separating catalytic residues, we opted to separate them between varying secondary structure motifs, namely: helix-helix-helix, helix-strand-coil, and strand-coil-strand. We cloned PBase into two subunits such that it was split at each of these three motifs, rendering each independent subunit non-functional. We plan to test the auto-assembly of each of these splits, and explore the use of inteins to increase the auto assembly rate.

The second approach we are exploring is attenuating PBase with point mutations. A set of mutations were created at the catalytic cores to create an inactive protein to act as a control. Keith et al identified specific mutations that significantly reduce PBase activity (D447E, D450E, D286N). We created mutations at these amino acid residues, and are currently fusing the mutated PBase to dCas9. Other factors we plan to optimize include linker length, type, and orientation of termini. Increasing the linker length may lead to increased mobility of PBase.

We introduce a new technology that enables high-value genetic modifications in human cells. While additional prototyping is necessary to ensure robust and reliable operation, here we demonstrate that large gene cassettes can be stably integrated in a customizable manner where the user can determine the size of the cassette, the locus of integration, and the cell type.

We envision targeted transposition becoming the standard approach for integrating large fragments, enabling myriad of applications. A direct application is integrating biosynthetic pathways into Chinese Hamster Ovary (CHO) cells to produce biopharmaceuticals. Upon further optimization our technology holds promise in a form of gene replacement therapy where payload size is not restricted and integration into safe harbor sites eliminates undesirable effects. For example, we foresee researchers integrating drug-vaccine libraries into cells, to create a programmable synthetic immune system. The future of genome editing is exciting and we hope that others build and expand upon our work.

Works Cited

1. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature. 2016;540: 144–149.

2. Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science. 2014;345: 1184–1188.

3. Mao Z, Bozzella M, Seluanov A, Gorbunova V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. Cell Cycle. 2008;7: 2902–2906.

4. Vargas JE, Chicaybam L, Stein RT, Tanuri A, Delgado-Cañedo A, Bonamino MH. Retroviral vectors and transposons for stable gene therapy: advances, current challenges and perspectives. J Transl Med. 2016;14. doi:10.1186/s12967-016-1047-x

5. Li Z, Michael IP, Zhou D, Nagy A, Rini JM. Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. Proc Natl Acad Sci U S A. 2013;110: 5004–5009.

6. Rostovskaya M, Fu J, Obst M, Baer I, Weidlich S, Wang H, et al. Transposon-mediated BAC transgenesis in human ES cells. Nucleic Acids Res. Oxford University Press; 2012;40: e150–e150.

7. Yusa K. piggyBac Transposon. Microbiology Spectrum. 2015;3. doi:10.1128/microbiolspec.mdna3-0028-2014

8. Voigt K, Gogol-Döring A, Miskey C, Chen W, Cathomen T, Izsvák Z, et al. Retargeting sleeping beauty transposon insertions by engineered zinc finger DNA-binding domains. Mol Ther. 2012;20: 1852–1862.

9. 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: 1636–1644.

10. Yant SR, Huang Y, Akache B, Kay MA. Site-directed transposon integration in human cells. Nucleic Acids Res. 2007;35: e50.

11. Ye L, You Z, Qian Q, Zhang Y, Che J, Song J, et al. TAL effectors mediate high-efficiency transposition of the piggyBac transposon in silkworm Bombyx mori L. Sci Rep. 2015;5: 17172.

12. Owens JB, Mauro D, Stoytchev I, Bhakta MS, Kim M-S, Segal DJ, et al. Transcription activator like effector (TALE)-directed piggyBac transposition in

human cells. Nucleic Acids Res. 2013;41: 9197–9207.

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

14. Munoz-Lopez M, Garcia-Perez J. DNA Transposons: Nature and Applications in Genomics. Curr Genomics. 2010;11: 115–128.

15. Hackett PB, Largaespada DA, Switzer KC, Cooper LJN. Evaluating risks of insertional mutagenesis by DNA transposons in gene therapy. Transl Res. 2013;161: 265–283.

16. de la Fuente-Núñez C, Lu TK. CRISPR-Cas9 technology: applications in genome engineering, development of sequence-specific antimicrobials, and future prospects. Integr Biol . 2017;9: 109–122.

17. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science. 2016;351: 84–88.

18. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346: 1258096.

19. ThermoFisher. Cell Cycle Assays for Flow Cytometry. In: ThermoFisher Scientific [Internet]. [cited 7 Aug 2017]. Available: https://www.thermofisher.com/us/en/home/life-science/cell-analysis/flow-cytometry/cell-health-and-viability-assays-for-flow-cytometry/cell-cycle-assays-for-flow-cytometry.html

20. ThermoFisher. Alexa Fluor 488 Dye. In: ThermoFisher Scientific [Internet]. [cited 7 Aug 2017]. Available: https://www.thermofisher.com/us/en/home/life-science/cell-analysis/fluorophores/alexa-fluor-488.html?gclid=CjwKCAjw0qLOBRBUEiwAMG5xMLtzhyEWejhIpBchOSwCr8rjx3lLMYA0WzE1mm1bxEoPKISfmOr5jxoCbm8QAvD_BwE&s_kwcid=AL!3652!3!222058295958!e!!g!!alexa%20fluor%20488&ef_id=VxrzUQAAAIUDXEu3:20170926020905:s

21. Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, et al. piggyBac transposase tools for genome engineering. Proc Natl Acad Sci U S A. 2013;110: E2279–87.

22. Biosciences S. PiggyBacTM Transposon Vector System. 2014\.

23. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. 2013;41: 7429–7437.

24. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol. 2014;32:

577–582.

25. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014;32: 569–576.

26. Oakes BL, Nadler DC, Savage DF. Protein Engineering of Cas9 for Enhanced Function. Methods in Enzymology. 2014. pp. 491–511.

27. Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, et al. CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. Protein Cell. 2015;6: 363–372.

28. Keith JH, Schaeper CA, Fraser TS, Fraser MJ Jr. Mutational analysis of highly conserved aspartate residues essential to the catalytic core of the piggyBac transposase. BMC Mol Biol. 2008;9: 73.

29. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9: 40.

30. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12: 7–8.

Appendix

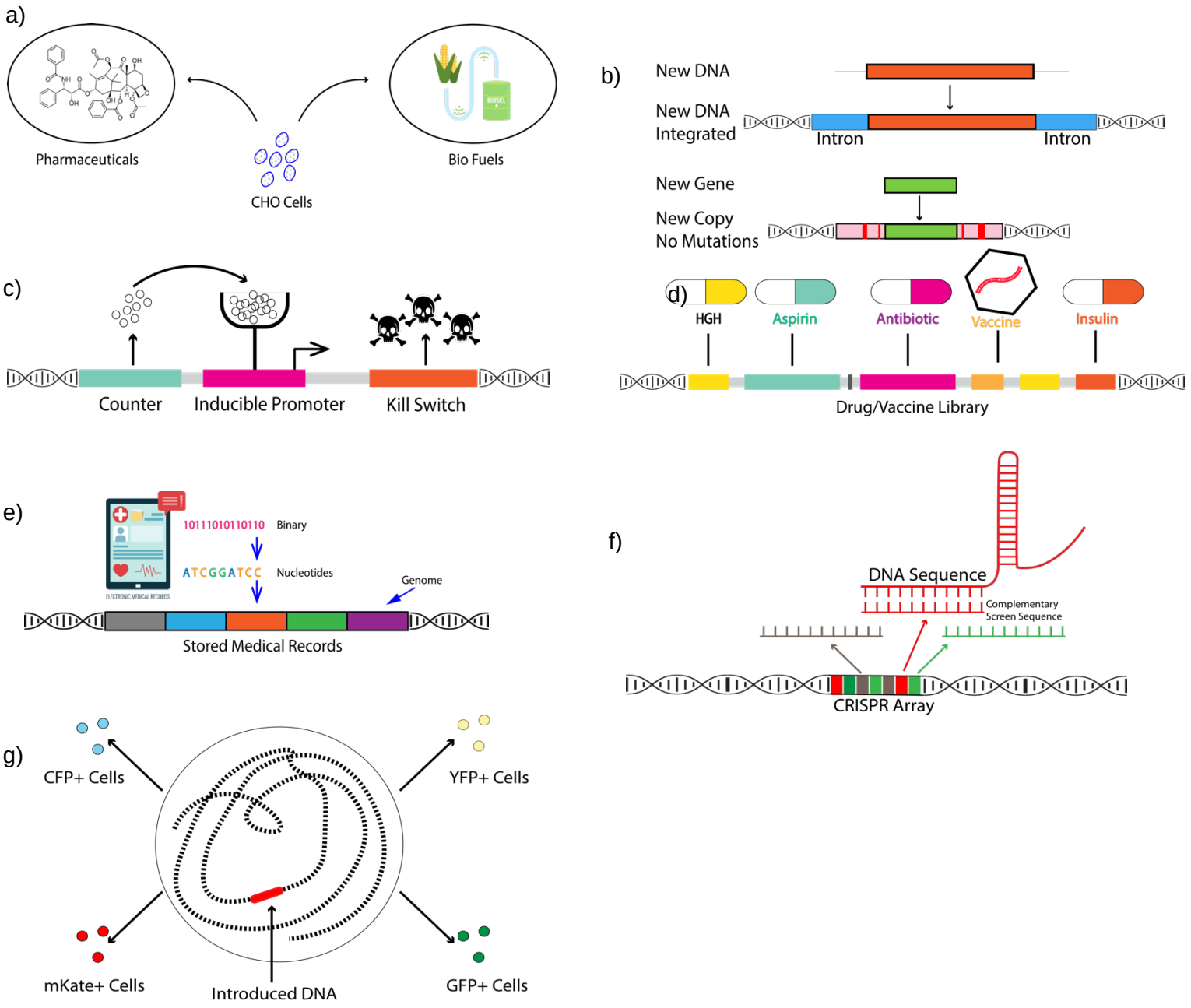


Figure 12 – Potential Long-Term Applications of Cas9-PBase Fusion
(a) CHO Cell mediated bioproduction - Chinese Hamster Ovary cells can be modified to produce biofuels and biopharmaceuticals using biosynthetic pathways. **(b) DNA Integration/Correction** - Leptin is known as the “satiety hormone” which increases fullness as food is consumed and regulated body weight. Leptin binds to LEPR, a 217-kb receptor protein that has be related to obesity and pituitary dysfunction when mutated. Using a PB-Cas9 fusion, the entire gene could be replaced with a new, functional gene. **(c) Cell Division Counter** - Cancer is

characterized by cells replicating at uncontrollable rates. The integration of a DNA copy number "counter" along with a sequence that triggers apoptotic activity in the event of excessive cell division serves as a mechanism for cancer control**(d) Drug/Vaccine Library** - Integrating a large library of genes coding for drugs and attenuated viruses increases the bioavailability of medications while allowing cell-specific reactions to induction of drug/vaccine release. **(e) Medical Record/Information Storage** - In the age of technology, privacy is a luxury. Medical records can be stored within a patient's genome that prevents data from being lost while maintaining patient privacy.**(f) CRISPR Array Integration** - CRISPR Arrays are repeated sequences of DNA that can store and screen against specific fragments of DNA. Integrating CRISPR arrays would be extremely useful for genetic screening. **(g) Model Cell Lines** - The development of model cell lines for cancer and diabetes studies can be streamlined by integrating massive genetic pathways all at once.