

# Generate TALE/TALEN as Easily and Rapidly as Generating CRISPR

Shuyan Zhang,<sup>1</sup> Huiting Chen,<sup>1</sup> and Jinke Wang<sup>1</sup>

<sup>1</sup>State Key Laboratory of Bioelectronics, Southeast University, Nanjing 210096, China

**TALE has always had potential as a gene-editing and regulatory tool. However, with the advent of CRISPR/Cas9, an easier to use tool with the same function, TALE has recently been abandoned because of the time-consuming and low-efficiency process required for its construction. The off-target activity of CRISPR/Cas9 has been a challenge to its *in vivo* application. By contrast, TALE has been applied *in vivo* for gene editing and therapy because of its high targeting capability. To overcome the key limitation of the TALE technique, we developed a high-efficiency method for constructing custom TALEs. We created 62 new monomers and developed a new pipeline that enabled assembly of custom TALEs in just 1 day. We verified the new method by assembling nine TALEs targeting the promoters of two transcription factor genes: HNF4 $\alpha$  and E47. The expression of the two endogenous genes in two cancer cells, HepG2 and PANC1, was activated by the constructed TALEs, which promoted differentiation of the two cancer cells. Using the new method, custom TALEs can be generated as easily and rapidly as CRISPR, thus promoting the wide application of TALE-based techniques.**

## INTRODUCTION

Transcription activator-like effectors (TALEs) are type III effector proteins from plant-pathogenic bacteria of the genus *Xanthomonas*. TALE is a programmable DNA-binding domain that can theoretically target any sequence. All TALEs are composed of an N-terminal translocation domain, a C-terminal nuclear localization signal (NLS) with an acidic transcriptional activation domain, and a central tandem repeat DNA binding domain (DBD).<sup>1</sup> The TALE DBD contains tandem repeats of 34 amino acid sequences (termed monomers) that are required for DNA recognition and binding.<sup>2–4</sup> The naturally occurring TALEs have been found to have a variable number of monomers, ranging from 1.5 to 33.5.<sup>1</sup> Although the sequence of each monomer is highly conserved, they differ primarily in two positions (the 12th and 13th) named as repeat variable diresidues (RVDs). Recent reports have found that the identity of these two residues determines the nucleotide binding specificity of each TALE repeat in a simple cipher that specifies the target base of each RVD (NI = A, HD = C, NG = T, NN = G).<sup>5,6</sup> Thus, each monomer targets one nucleotide, and the linear sequence of monomers in a TALE specifies the target DNA sequence in the 5'-to-3' orientation.

According to these TALE-DNA molecular codes, biologists have assembled a TALE DBD and fused it to other functional protein do-

main to manipulate the genome, including gene regulation<sup>7–10</sup> and gene editing.<sup>10–17</sup> The TALE system, including TALE and TALE nuclease (TALEN), thus became powerful tools for genome operations. Especially, TALEN has been a leading technology for biomedical treatment. The TALE system has high targeting capability. For instance, in a study in which an iPS cell line was edited with TALEN, no mutagenic activity was detected at other genome sites homologous to the target site.<sup>18</sup> Due to its high targeting effect, TALEN has been being employed to produce universal chimeric antigen receptor T (UCAR-T) cells by Collectis<sup>19</sup> and has been approved by the United States Food and Drug Administration (FDA) to be used in clinical cancer immunotherapy. The first patients to receive TALEN genetically engineered products have shown significant benefits.<sup>19</sup> These patients, treated in 2015, are still completely free of cancer to this day.<sup>20</sup>

With the advent of CRISPR in 2013, CRISPR/Cas9 rapidly became the most widely used genetic manipulation technique because of its simplicity. Since then, the TALE system has been almost completely replaced by the CRISPR/Cas9 system. As compared with the CRISPR/Cas9 system, a key limitation of the TALE system is its cumbersome and tedious vector construction process, a time-consuming (about 1 week) procedure, even for skilled experimenters. To address the problem, many vector assembly methods have been developed, including those derived from Golden Gate cloning,<sup>11</sup> regular cloning methods,<sup>21</sup> and high-throughput methods.<sup>22</sup> Because of the repetitive nature of TALEs, construction of the DNA-binding monomers is troublesome. Many groups used a hierarchical ligation strategy to overcome the difficulty of assembling monomers into ordered multimer arrays, taking advantage of degeneracy in codons surrounding the monomer junction and type II restriction enzymes.<sup>9–12,23,24</sup> However, these strategies are still limited by their relative low efficiency in constructing TALE expression vectors. Therefore, new TALE preparation methods as simple and rapid as CRISPR/Cas9 system are still in demand.

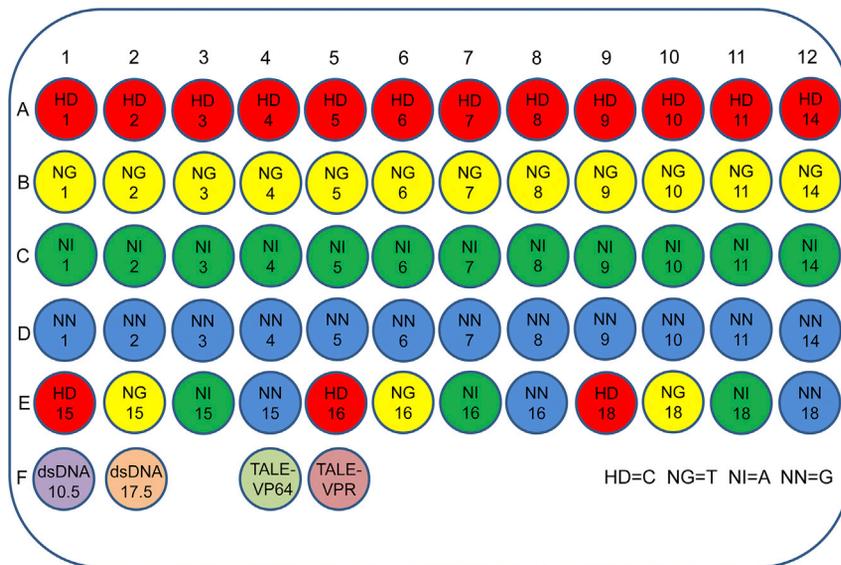
Currently, the most widely used TALE assembly kit is the Golden Gate TALEN and TAL Effector Kit 2.0 (Addgene). The monomers of this kit are plasmids contained in *E. coli*. In order to construct a custom TALE, the bacteria have to be amplified in culture, and

Received 29 August 2018; accepted 11 February 2019;  
<https://doi.org/10.1016/j.omtm.2019.02.004>.

**Correspondence:** Jinke Wang, State Key Laboratory of Bioelectronics, Southeast University, Nanjing 210096, China.

**E-mail:** [wangjinke@seu.edu.cn](mailto:wangjinke@seu.edu.cn)





**Figure 1. Schematic of Monomers in a 96-Well PCR Plate**

Each monomer was prepared by PCR amplification by using monomer plasmids included in an Addgene kit as a template and different primers designed to amplify variant monomers (Table S1). In total, there were 60 base-determinant monomers and 2 linker monomers (dsDNA<sub>10.5</sub> and dsDNA<sub>17.5</sub>). The position of each monomer is indicated in the well. The monomer plate can be easily regenerated by 96-well PCR amplification. This plate also contains two TALE backbone vectors, TALE-VP64 and TALE-VPR, which can be produced by *E. coli* DH5 $\alpha$  transformation and extraction and added to the wells.

## RESULTS

### Preparation of TALE Monomer and Backbone

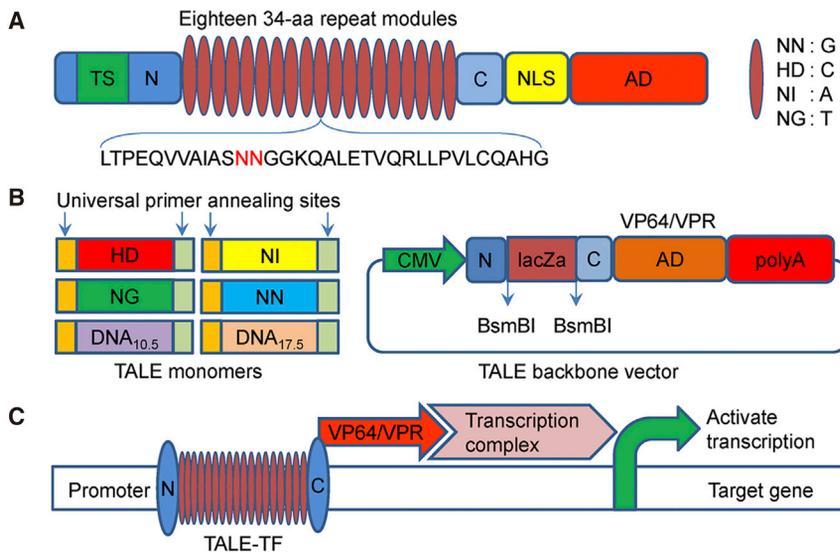
To overcome the limitations of the Addgene kit 2.0, we aimed to develop a new TALE construction strategy with plasmid-free monomers

and a bacteria-free assembly pipeline. To this end, a set of PCR primers (Table S1) was designed, and 60 linear base-determinant monomers (Figure 1) were amplified from Addgene plasmid monomers (Figure S1A). Importantly, all PCR-produced linear monomers could then be reproduced by a pair of universal primers (Table S1) in a 96-well PCR plate (Figure 2). To eliminate intermediate bacteria transformations and receptor plasmids (pFUSA, pFUSB, and pLR) used in TALE assembly with the Addgene kit 2.0 (Figure S1), this study created two linking monomers, dsDNA<sub>10.5</sub> and dsDNA<sub>17.5</sub> (Figure 1), which could also be produced by PCR amplification, using universal primers in a 96-well PCR plate (Table S2). The linker monomer dsDNA<sub>10.5</sub> was used to link monomers 10 and 11, and the linker monomer dsDNA<sub>17.5</sub> was used to link monomers 17 and 18. In addition, the two linker monomers provided partial coding sequences that were originally harbored by pFUSA and pFUSB. Once the monomer 96-well PCR plate (Figure 1) was constructed, it could be kept for constructing any TALEs that could bind various 18-bp targets. The monomer plate (Figure 1) can be easily replicated by 96-well high-fidelity PCR amplification. The quality of replicates was easily checked by direct sequencing.

various plasmids have to be purified. The purified plasmids are used to construct a custom TALE by using a Golden Gate digestion-ligation process that takes as long as 5 days (Figure S1). Moreover, the whole TALE assembly process has to undergo three bacterial transfections and colony selection and cultivation, which are all rate-limiting steps. Moreover, the efficiency of the digestion-ligation reaction with plasmids is low because a limited number of monomer molecules can be contained in a digestion-ligation reaction. In plasmid monomers, a large amount of plasmid backbone DNA is useless to TALE assembly, which prevents more effective monomer molecules from being added to the digestion-ligation reaction. In addition, the cutting efficiency of plasmid DNAs is lower than linear DNAs because of the compact supercoiled structure.<sup>25</sup>

Using the Addgene 2.0 assembly kit, we developed a new monomer library and developed a pipeline for rapidly assembly of custom TALEs that can constantly bind 18-bp targets. The new monomers can be used to assemble TALEs by the Golden Gate method.<sup>26–28</sup> The monomer library consists of 60 base-determinant monomers and 2 novel linker monomers. All monomers are net linear dsDNA fragment ended with two constant sequences. All monomers can be easily produced and reproduced by high-fidelity PCR amplification in a 96-well PCR plate, using a pair of universal primers. To construct TALEs with these monomers, we designed a rapid assembly pipeline that can finish a custom TALE assembly in just 1 day. To verify new monomers and TALE assembly pipeline, we constructed nine TALEs targeting the promoters of two transcription factor genes: HNF4 $\alpha$  and E47. The constructed TALEs could activate the expression of both exogenous reporter genes and endogenous genes. Moreover, the activation of endogenous HNF4 $\alpha$  and E47 genes in HepG2 and PANC1 led to the differentiation of these cancer cells into normal liver or pancreas cells. Therefore, the monomer library and TALE assembly pipeline developed in this study can promote the wide application of TALE-based techniques.

To produce the final functional TALEs, we also constructed two TALE backbones, pTALE-VP64 and pTALE-VPR (fusion of VP64, p65/RelA, and Rta), which still harbored a LacZ expression cassette for easily screening the final positive TALE colonies with blue-white spot screening (Figure 2). In the construction of TALE backbones, the NLS and acidic transcription activation domain of wild-type hax3 in the wild-type TALE protein were replaced by a mammalian NLS derived from the simian virus 40 large T antigen and the synthetic transcription activation domain VP64 or VPR (VP64-p65-Rta) (Figure 2B), respectively. To enhance the expression of the final TALE in mammalian cells, the TALE backbones contained a strong cytomegalovirus (CMV) promoter (Figure 2).



**Figure 2. Schematics of TALE Structure, Monomers, and TALE Backbone Vector for Assembling Custom TALEs**

(A) Natural structure of TALEs derived from *Xanthomonas* sp. Each DNA binding module consists of 34 amino acids, where the repeat variable diresidues (RVDs) in the 12th and 13th amino acid positions of each repeat specify the DNA bases being targeted according to the cipher NG = T, HD = C, NI = A, and NN = G. (B) Monomers and TALE backbones for constructing custom TALEs. Each monomer ends with two constant sequences that provide the annealing sites of a pair of universal primers that can be used to regenerate the monomer library. CMV, cytomegalovirus promoter; N, non-repetitive N-terminus of TALE; C, non-repetitive C-terminus of TALE; *BsmBI*, type II restriction sites used for the insertion of a custom TALE DNA binding domain; *LacZa*, *LacZa* expression cassette for blue-white spot screening; NLS, nuclear localization signal; VP64, synthetic transcriptional activator derived from VP16 protein of herpes simplex virus; VPR, fusion of VP64, p65/RelA, and Rta. (C) Schematic of activating gene expression with TALEs. TALE-TF, TALE transcription factor.

### Pipeline of TALE Construction

With the newly designed and produced monomers (Figures 1 and 2) and TALE backbones (Figure 2), a new pipeline was designed to construct a custom TALE in 1 day (Figure 3). This pipeline enables construction of a custom 18-bp TALE in two steps. First, four circular pentamers are prepared by Golden-Gate digestion-ligation reactions using new linear monomers, *BsaI*, and T4 DNA ligase (Figure 3). Second, the final TALE expression plasmid is prepared by a Golden-Gate digestion-ligation reaction using four pentamers, a TALE-backbone vector (TALE-VP64/VPR), *BsmBI*, and T4 DNA ligase. The final TALE plasmid can be obtained by one bacteria transformation, colony PCR identification, and plasmid extraction. The positive final TALE plasmid is further rapidly confirmed by *EcoRI* digestion identification. In comparison with the pipeline of the Addgene kit 2.0 (Figure S1B), two time-consuming bacteria transformations, overnight cultivation, and colony selections were removed. Moreover, we found that it was difficult to get white-positive colonies in these two steps when preparing TALEs with the Addgene kit. It was thus concluded that these two steps of the kit were error prone, which may be related to adding too many monomers in a Golden-Gate digestion-ligation reaction. For example, in the first two Golden-Gate digestion-ligation reactions, 10 and 7 monomers were added, respectively (Figure S1B). In addition to two receptor plasmids (pFUSA and pFUSB7), as many as 11 and 8 plasmids have to be included in two Golden-Gate digestion-ligation reactions. In contrast, a new pipeline was designed to contain only five short (about 100 bp) linear monomers in each of the first Golden-Gate digestion-ligation reactions, which is beneficial for enhancing digestion-ligation efficiency and fidelity.

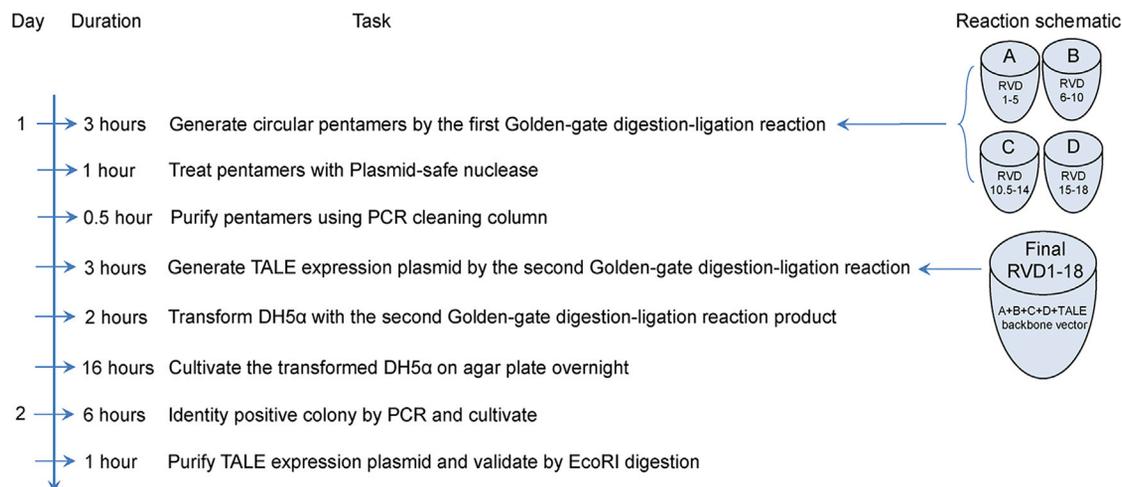
### Construction of Custom TALEs

To verify the newly produced monomers and TALE assembly pipeline, we assembled nine TALEs targeting promoters of human HNF4 $\alpha$  and E47 by following the assembly steps described in

Materials and Methods. The assembled TALEs were named HNF4 $\alpha$ -TALE1/2/3-VP64, HNF4 $\alpha$ -TALE1/2/3-VPR, and E47-TALE1/2/3-VPR. *E. coli* was transformed by the final TALEs and cultivated on solid agar. The results showed that many white spots were produced (Figures 4A and 4B). No blue spots were seen on the agar plate. The colony PCR detection of randomly picked colonies confirmed that the insert size was correct, and typically over 80% of colonies were *bona fide* positive colonies (Figures 4C and 4D; Figure S2). These data indicate the high efficiency of constructing custom TALEs with new monomers (Figure 1), backbone vectors (Figure 2), and pipeline (Figure 3). The subsequent *EcoRI* digestion of extracted plasmids further confirmed the results of colony PCR, which indicated that typically 80% of white spots were correctly and successfully assembled in the final TALEs (Figures 4E and 4F). We also assembled another TALE, HNF4 $\alpha$ -TALE-VP64, targeted to the same site as HNF4 $\alpha$ -TALE-VPR, in order to compare the activity of two transactivation domains: VP64 and VPR.

### Activation of Exogenous Gene with TALEs

To check if the constructed TALEs would function in mammalian cells, a fluorescence reporter construct was prepared by cloning the HNF4 $\alpha$ /E47 promoter that contained the TALE-binding sites and a minimal CMV promoter upstream of a reporter gene, ZsGreen (HNF4 $\alpha$ /E47-TALE reporter). Three TALEs targeting three different sites in the HNF4 $\alpha$  and E47 promoters were assembled (Figure 5A). The 293T cells were co-transfected by an assembled TALE and corresponding TALE reporters. The results revealed that the expression of corresponding TALE reporters was activated by all TALEs (Figures 5B–5D); however, the activation level was not the same (Figures 5B–5D). It was revealed that the TALEs targeting the sites near the transcription start sites (TSSs) showed the highest transcriptional activation activity (HNF4 $\alpha$ -TALE1-VP64/VPR and E47-TALE1-VPR; Figures 5B–5D). In comparison, the TALE-VPR activated



**Figure 3. Pipeline for Constructing a TALE with New Monomers**

Steps for the construction of a TALE are outlined. Time spent on each step is given. It can be seen that the whole TALE assembly procedure can be finished in 1 day. After a colony PCR screening, plasmid extraction, and *EcoRI* digestion confirmation, the usable final TALE can be obtained on the second day. The time taken by the process is identical to that taken by the current CRISPR.

much higher expression of ZsGreen than did TALE-VP64 (Figures 5B and 5C). In contrast, the control transfection did not activate reporter gene expression (Figure S3). These data indicate that the TALEs constructed with new monomers, backbone vectors, and pipeline were functional.

#### Activation of Endogenous Gene with TALEs

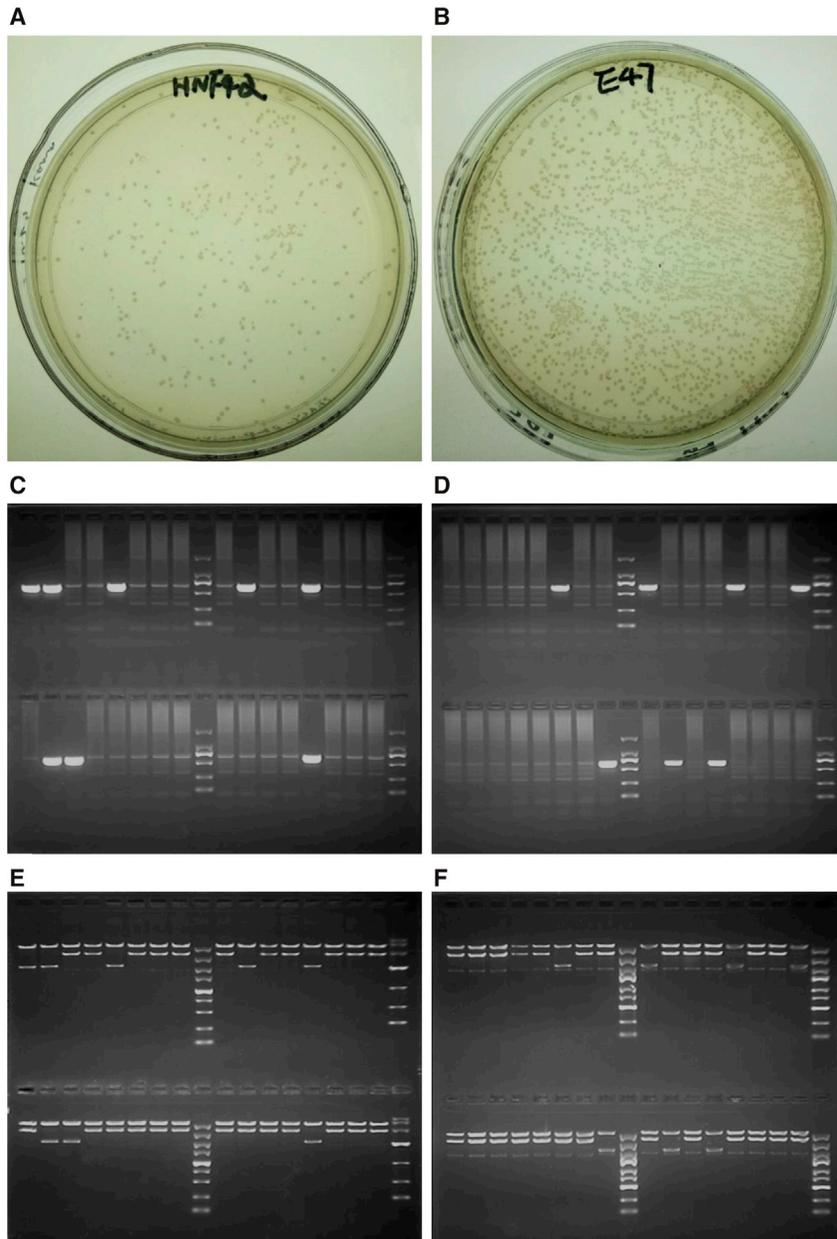
To find whether the constructed TALEs could activate the endogenous gene expression in mammalian cells, we first transfected HepG2 cells with HNF4α-TALE-VPR. The qPCR detection of HNF4α expression revealed that its expression was highly activated by HNF4α-TALE-VPR (Figure 6A). In addition, the expression of several characteristic hepatocyte markers was induced, including APOCIII, CYP1A2, G-6-P, GYS2, APOAI, and HPD (Figure 6A). At the same time, the expression of multiple stemness genes was significantly downregulated, including CD133, OCT3/4, BMI, SOX2, KLF4, and LIN28 (Figure 6B). Therefore, activation of endogenous HNF4α gene by the constructed TALE-VPR induced differentiation of hepatoma cells.

Similarly, the PANC1 cells were transfected with E47-TALE-VPR. The qPCR detection of E47 expression revealed that its expression was highly activated by E47-TALE-VPR (Figure 6C). Moreover, the expression of the cyclin-dependent kinase inhibitor p21 and the stress response protein TP53INP1, which play critical roles in cell-cycle arrest, was significantly upregulated. E47 expression also induced upregulation of the gap junction protein connexin 32 (CX32) and the ductal genes SOX9 and HNF6 in PANC1 cells. Meanwhile, E47 induced MIST1, which can regulate CX32 expression in PANC1 cells. E47 also downregulated the expression of the pancreatic ductal adenocarcinoma-associated cell cycle activators topoisomerase2A (Top2A) and aurora kinase A (AURKA).

#### Differentiation of Cancer Cells with TALEs

To explore whether the TALE-activated gene expression would result in phenotypic changes of the cells, cell viability, cycle, and migration were determined. The dynamic detection of cell viability revealed that HNF4α-TALE-VPR transfection reduced the cell viability of HepG2 cells on the third day after transfection (Figure 7A). Remarkably, on the fifth day after transfection, cell viability reduced more than 50%. The measurement of the cell cycle indicated that HNF4α-TALE-VPR transfection resulted in significant cell cycle arrest in the HepG2 cells (Figure 7B). On the third day after infection, only 9.66% of HNF4α-TALE-VPR-transfected HepG2 cells were in the S phase, whereas as many as 40.36% of the control cells were in the S phase (Figure 7B; Figure S4). The cell migration assay revealed that HNF4α-TALE-VPR transfection led to significant inhibition of HepG2 cell migration (Figure 7C). These phenotypic changes suggest that HNF4α-TALE-VPR-activated HNF4α expression induced differentiation of HepG2 cells into normal hepatocytes, which is in agreement with the significant upregulation of characteristic hepatocyte markers (Figure 6A) and the significant downregulation of stemness genes (Figure 6B).

The cell viability, cycle, and migration of PANC1 cells were also examined. Similarly, the results indicated that E47-TALE-VPR transfection resulted in significant reduction of cell viability (Figure 7C), cycle proliferation (Figure 7D; Figure S3), and migration (Figure 7E) in the PANC1 cells. These phenotypic changes were in agreement with the significant upregulation of P21, TP53INP1, and MIST1 genes, indicating that E47 reprogrammed aggressive PANC1 cells to a quiescent acinar state by restoring the expression of P21, TP53INP1, and MIST1. Mist1 is a key bHLH transcription factor that controls the acinar maturation program.



**Figure 4. Constructing Custom TALEs with New Monomers and Pipeline**

(A and B) *E. coli* colonies on solid agar that were transformed with HNF4 $\alpha$ -TALE1-VPR (A) and E47-TALE1-VPR (B). (C and D) Colony PCR detection of four groups of randomly selected white spots. HNF4 $\alpha$ -TALE1-VPR (C) and E47-TALE1-VPR (D) groups contained eight colonies. Full-length PCR products should be 2,051 bp long. However, PCR products were often less prominent, whereas the “ladder effect” represented a robust indicator of successful assembly. (E and F) Colony identification by *EcoRI* digestion. All colonies picked from HNF4 $\alpha$ -TALE1-VPR (E) and E47-TALE1-VPR (F) were cultivated for extracting plasmid. The extracted plasmids were digested with *EcoRI*, which cut out the full-length TALE of 3,537 bp. The incorrectly assembled TALEs produced a 2,143-bp band in *EcoRI* digestion. It can be seen that the identification results of colony PCR were in agreement with those of *EcoRI* digestion.

(Figure 8A). Similarly, the expression of E47-ZsGreen protein was significantly activated by E47-TALE1-VPR in the edited PANC1 cells (Figure 8B). To edit and activate endogenous genes simultaneously, HepG2 and PANC1 cells were also respectively co-transfected with a pair of TALEN vectors: a homologous donor and TALE1-VPR. Similar gene activation was observed in two cells (Figure 8).

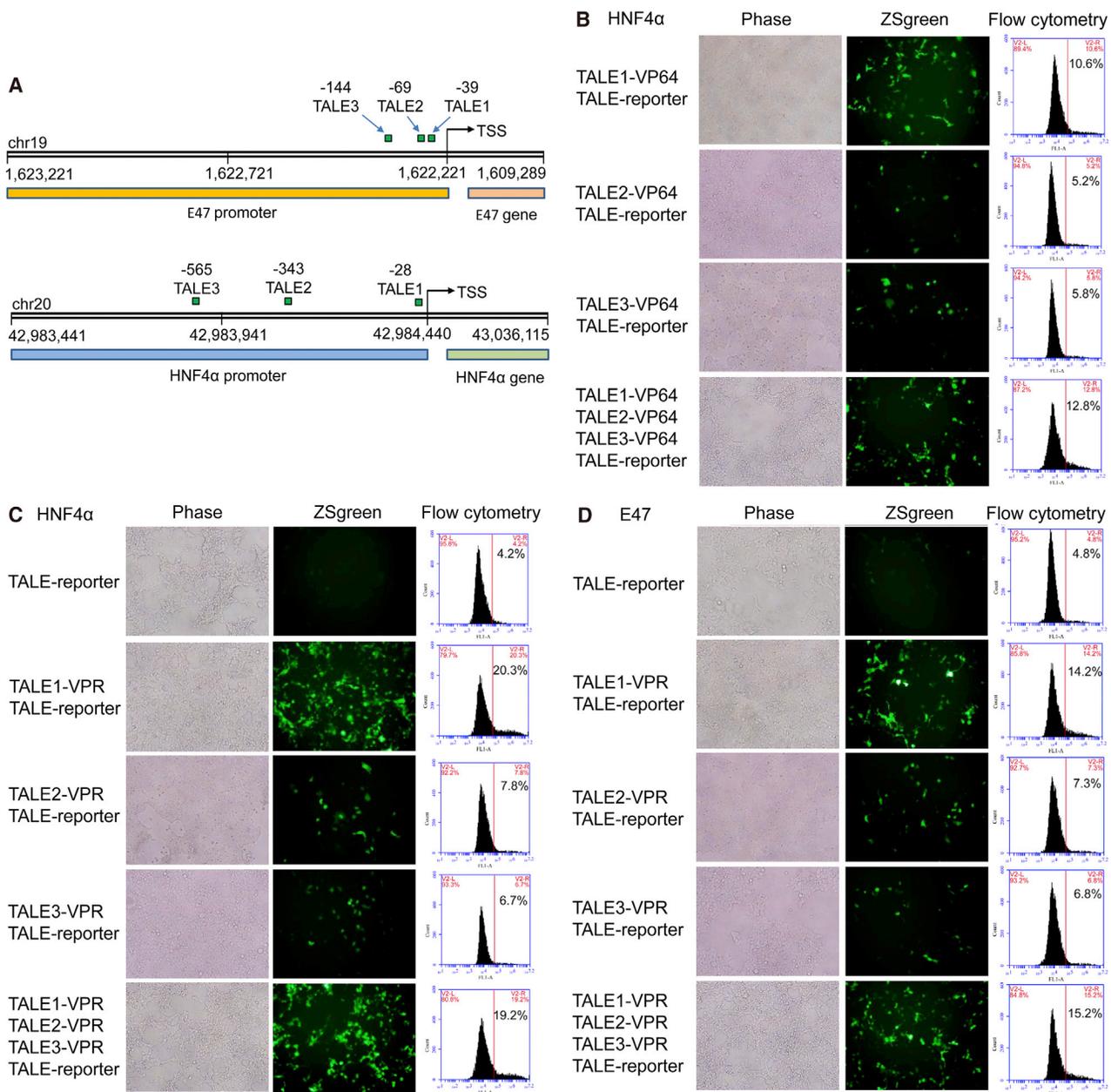
## DISCUSSION

The TALE system has been a promising gene operation tool that has high DNA-binding specificity. However, the low-efficiency and time-consuming vector construction process has always prevented its wide application. For overcoming this limitation, we greatly improved both monomers and the assembly pipeline of the Addgene kit. First, we made a new monomer library that consisted of 60 base-determinant monomers and 2 linker monomers. All base-determinant monomers were prepared by PCR

### Activation of Edited Endogenous Genes with TALEs

To further characterize the activation of endogenous genes at the protein level by TALE-TF, four TALEN vectors (Table S5) were constructed with a new method for fusing the ZsGreen gene to endogenous genes by homology-directed repair (HDR). The HNF4 $\alpha$  and E47 genes in the HepG2 and PANC1 cells were edited by co-transfecting cells with a pair of TALEN vectors and a homologous donor, respectively (Figure 8). The edited HNF4 $\alpha$  and E47 genes were respectively activated by transfecting the edited HepG2 and PANC1 cells with HNF4 $\alpha$ -TALE1-VPR and E47-TALE1-VPR. The results indicate that the expression of HNF4 $\alpha$ -ZsGreen protein was significantly activated by HNF4 $\alpha$ -TALE1-VPR in the edited HepG2 cells

amplification by using Addgene kit plasmid monomers as templates. Two linker monomers were chemically synthesized. Importantly, all new monomers ended in two constant sequences and could be amplified by high-fidelity PCR using a pair of universal primers. One advantage of this plasmid-free monomer library is that it can be easily reproduced by high-fidelity PCR amplification in a 96-well PCR plate in a high-throughput format, eliminating the need to produce a plasmid monomer library by time-consuming bacteria cultivation and plasmid DNA extraction with the Addgene kit. Another advantage is that the linear monomers have no redundant plasmid sequences and thus allow more monomer molecules to be included in the first digestion-ligation reaction, which is beneficial for forming

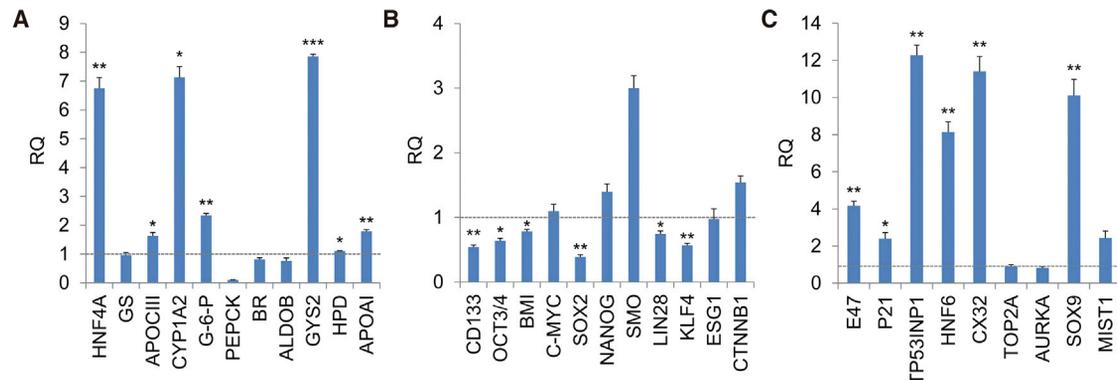


**Figure 5. Activating Exogenous Reporter Gene with TALEs**

293T cells were co-transfected with a TALE-VPR/VP64 plasmid and its corresponding reporter plasmid expressing ZsGreen. (A) Positions of TALE targets in the promoters of HNF4α and E47 genes. (B) Cells cotransfected HNF4α-TALE1/2/3-VP64 and HNF4α-TALE-reporter. (C) Cells cotransfected with HNF4α-TALE1/2/3-VPR and HNF4α-TALE-reporter. (D) Cells cotransfected with E47-TALE1/2/3-VPR and E47-TALE-reporter.

more circular pentamers in high efficiency. Second, this study made a new pipeline for assembling custom TALEs using new monomers. Following this pipeline, custom TALEs can be rapidly and easily assembled in just one day. In fact, with the off-the-shelf monomer plate, starting from a given target sequence, TALE preparation can be finished in as few as 2 days. These improvements should promote new wide application of TALE.

A limitation of the TALE assembly method developed in this study is that it can only be used to assemble TALEs with 18-bp binding sites. This study focused on TALEs of this length because TALEs were typically built with 18 repeats of 34 amino acids. With the constant T base, TALEs constructed by the new method can bind 19-bp target sequences in genomes. This is a frequently-used length of TALEs that carries high specificity and low cost. However, with the same strategy



**Figure 6. Activating Endogenous Genes with TALEs**

HepG2 and PANC1 cells were transfected with HNF4 $\alpha$ /E47-TALE1-VPR. (A) Expression of HNF4 $\alpha$  and characteristic hepatocyte markers in HepG2 cells. (B) Expression of “stemness” genes in HepG2 cells. (C) Expression of E47, cell cycle arrest-related genes (p21 and TP53INP1), and E47 target genes (HNF6, SOX9, CX32, and MIST1) in PANC1 cells. Data analysis was performed with the Applied Biosystems StepOne software v2.3, and  $C_t$  values of detected genes were normalized with that of  $\beta$ -actin. The relative expression level of target mRNAs was calculated as relative quantity (RQ) according to the equation:  $RQ = 2^{-\Delta\Delta C_t}$ . Student’s t test was used to check statistical significance. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . The dash lines refer to RQ values of negative controls.

as was used to make linear monomers in this study, other monomers can be similarly manufactured for preparing longer TALEs if needed.

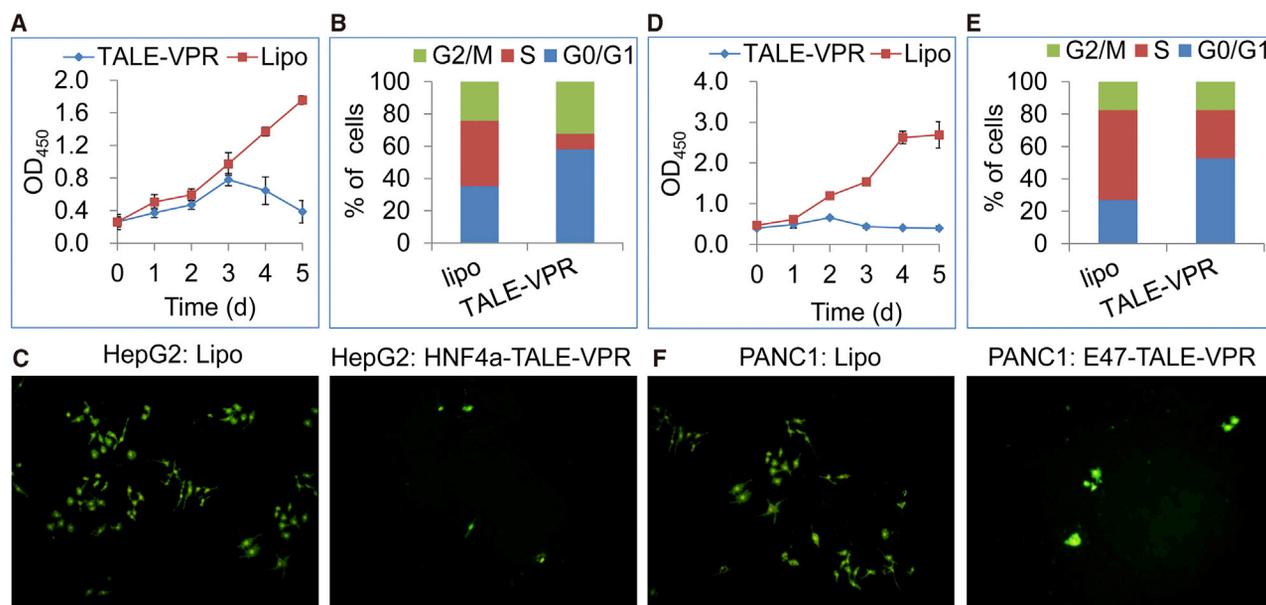
The gene transcriptional activation level is related to transactivation domains and TALE-targeting position. We used two different transcriptional activation domains: VP64 and VPR.<sup>9,29</sup> The results indicate that VPR always had higher transcriptional activation activity than VP64, in agreement with results obtained with CRISPR.<sup>29</sup> We therefore activated endogenous genes with TALE-VPR. In the future, more potent activation systems such as Suntag can be tried for use in TALEs.<sup>29</sup> To knock down gene expression, the transcriptional repression domain (RD) such as the Krüppel-associated box transcriptional repression domain (KRAB) can be fused to the TALE.<sup>30</sup> In addition, other epigenetic functional domains, such as DNMT3a (DNA methyltransferase), EZH2 (histone 3 lysine 27 methyltransferase), and LSD1 (histone demethylase) can be fused to the TALE for making variant gene regulators (GRs).<sup>31</sup> These GRs have wide application in basic research and biomedicine. The relative position of TALE targets to the TSS has a significant effect on TALE activation effect. The TALE-VPR regulation to endogenous genes is also dependent on the chromatin state.<sup>32,33</sup> For these reasons, this study typically built several TALE-VPRs for each target gene. It was found that well-functioning TALEs were always positioned in the proximal promoter region.

In this study, we selected HNF4 $\alpha$  and E47 as target genes to activate using the constructed TALEs because it has been reported that activation of these two transcription factors leads to differentiation of the cancer cells HepG2 and PANC1.<sup>34,35</sup> To see the functional results of the constructed TALEs, we transfected HepG2 and PANC1 cells with HNF4 $\alpha$ -TALE-VPR and E47-TALE-VPR, respectively. We found that the two genes were significantly activated by the transfected TALE-VPRs. Correspondingly, the expression of some target genes regulated by these two transcription factors was also regulated,

indicating the regulatory functions of these two activated transcription factors. Measurements of cell viability, cycle, and migration revealed that the TALEs-transfected cells underwent significant phenotypic changes, including cycle arrest and reduced viability and migration. HNF4 $\alpha$  is a central regulator of the differentiated hepatocyte phenotype, and forced re-expression of HNF4 $\alpha$  was sufficient to overcome repression of the hepatic phenotype in dedifferentiated hepatoma cells.<sup>35</sup> TALE-VPR targeting the HNF4 $\alpha$  promoter significantly upregulated HNF4 $\alpha$  expression in HepG2 cells, which led to the significant upregulation of hepatocyte marker genes and downregulation of stemness genes. For example, CD133 is currently thought to serve as a target for identifying cancer stem or progenitor cells in a portion of hepatocellular carcinoma (HCC).<sup>36–38</sup> Upregulation of HNF4 $\alpha$  expression induces downregulation of CD133 and some genes that are highly expressed in human embryonic stem cells and are involved in the establishment or maintenance of pluripotency,<sup>38–40</sup> including OCT3/4, BMI, SOX2, KLF4, LIN28, and ESG1. The current study revealed that E47 regulated its target genes, such as p21 and TP53INP1 in PANC1 cells,<sup>41,42</sup> in agreement with a previous report that E47 upregulated expression of p21 and TP53INP1 genes in all tested PANC1 cell lines associated with G0/G1 arrest.<sup>34</sup> These data suggest that the TALE-VPRs targeting HNF4 $\alpha$  and E47 promoters have a potential clinical application as gene therapy reagents.

## Conclusions

To overcome key technical limitations of TALE application, we developed a fast and efficient strategy for preparing TALE expression plasmids in just 1 day. We manufactured a new set of linear monomers that could be easily reproduced by high-fidelity PCR amplification in a 96-well PCR plate in a high-throughput format. With new monomers, any 18-bp TALE expression plasmids could be rapidly assembled using a newly designed pipeline. This method can promote wide application of TALE-based tools.



**Figure 7. Assays of Cell Viability, Cycle, and Migration**

HepG2 and PANC1 cells were transfected with HNF4 $\alpha$ /E47-TALE1-VPR. (A) HepG2 cell viability. (B) HepG2 cell cycle. (C) HepG2 cell migration. (D) PANC1 cell viability. (E) PANC1 cell cycle. (F) PANC1 cell migration. CCK-8 was used to determine the viability of cells at different time points. Flow cytometry was used to detect the cell cycle. The percentage of cells in individual cell cycle phases was determined. A Transwell assay was used to detect cell migration.

## MATERIALS AND METHODS

### Materials and Reagents

The chemical and biochemical reagents used in this study and their manufacturers are listed in detail in the [Supplemental Information](#).

### Preparation of Plasmids

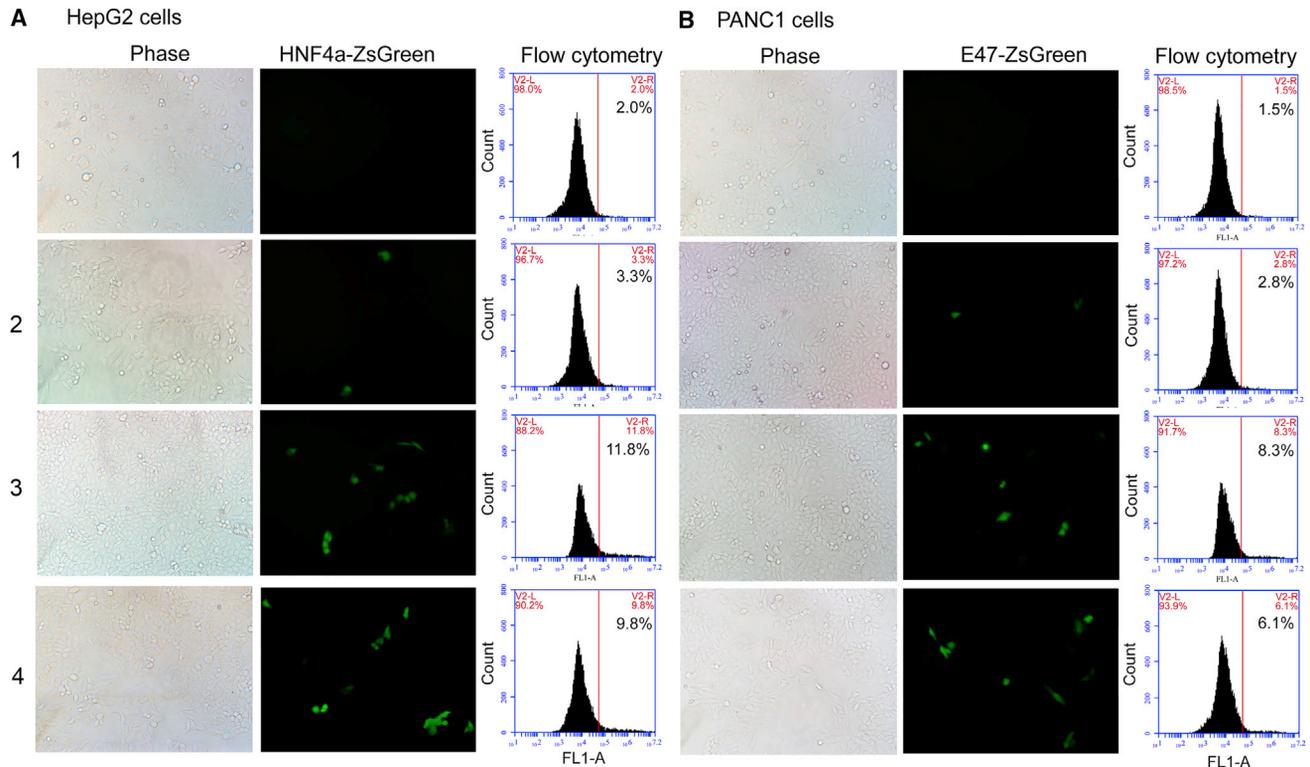
The VP64 and VPR fragments were amplified by PCR from pcDNA-dCas9-VP64 and pcDNA-dCas9-VPR, using PrimeSTAR HS DNA Polymerase (Takara Bio), according to the manufacturer's instructions, in which the *EcoRI* and *NotI* sites were introduced at the two ends. The PCR primers were VP64-F/R and VPR-F/R (Table S3). The fragments were ligated into pPIRES2-EGFP to produce pPIE-VP64 and pPIE-VPR. The TAL(LacZa) fragment was recovered from pTAL2 (Addgene kit) with *EcoRI*, which was ligated into pPIE-VP64 and pPIE-VPR to produce the TALE backbone vector pTALE-VP64 and pTALE-VPR. The 1,000-bp promoter fragments of HNF4 $\alpha$  and E47 genes were amplified by PCR from HepG2 and PANC1 cell genomic DNA (gDNA) using PrimeSTAR HS DNA Polymerase, in which the *XhoI* and *HindIII* sites were introduced at the two ends. The PCR primers were HNF4 $\alpha$ -F/R and E47-F/R (Table S3). The fragments were ligated into pEZX-miniCMV-ZsGreen to produce the reporter vectors pHNF4 $\alpha$ -TALE-reporter and pE47-TALE-reporter. The maps of the plasmids are shown in the [Supplemental Information](#). All PCR programs were run on a MasterCycler Pro (Eppendorf). For preparing high-quality plasmids, all plasmids were amplified in *E. coli* DH5 $\alpha$  and extracted with the EndoFree Plasmid kits (QIAGEN), according to the manufacturer's instructions. All plasmids were further verified by DNA sequencing.

### Preparation of Monomers

Monomers were prepared by high-fidelity PCR amplification. The plasmid monomers of the Addgene kit were used as templates. The primers for amplifying each linear monomer are shown in Table S1. High-fidelity PCR amplification was performed using PrimeSTAR HS DNA Polymerase, according to the manufacturer's instructions. Each PCR reaction (240  $\mu$ L) contained 48 ng plasmid monomer, 10  $\mu$ M forward and reverse primers (Table S1), and 1  $\times$  PrimeSTAR HS DNA Polymerase. The PCR products were purified with the MinElute Gel Extraction kit, according to the manufacturer's instructions. The PCR templates for amplifying the two linker monomers dsDNA<sub>10.5</sub> and dsDNA<sub>17.5</sub> were chemically synthesized (oligo DNA<sub>10.5</sub> and DNA<sub>17.5</sub> in Table S2). The double-stranded DNA<sub>10.5</sub> and DNA<sub>17.5</sub> (dsDNA<sub>10.5</sub> and dsDNA<sub>17.5</sub>) were amplified by high-fidelity PCR with universal primers (Table S1), using PrimeSTAR HS DNA Polymerase. Once the first set of linear monomers was prepared by high-fidelity PCR amplification, as described above, all linear monomers could be reproduced by high-throughput in 96-well plates by high-fidelity PCR amplification, using a pair of universal primers (Table S1).

### Construction of Custom TALEs

Custom TALEs were prepared according to new pipelines schematically shown in Figure 3. First, according to a 19-bp target sequence of 5'-T<sub>0</sub>N<sub>1</sub>N<sub>2</sub>N<sub>3</sub>N<sub>4</sub>N<sub>5</sub>N<sub>6</sub>N<sub>7</sub>N<sub>8</sub>N<sub>9</sub>N<sub>10</sub>N<sub>11</sub>N<sub>12</sub>N<sub>13</sub>N<sub>14</sub>N<sub>15</sub>N<sub>16</sub>N<sub>17</sub>N<sub>18</sub>-3' (where N = A, G, T, or C) and the amino acid-coding rules of the monomers (NI = A, HD = C, NG = T, and NN = G), four digestion-ligation reactions for producing pentamers (N<sub>1</sub>-N<sub>2</sub>-N<sub>3</sub>-N<sub>4</sub>-N<sub>5</sub>, N<sub>6</sub>-N<sub>7</sub>-N<sub>8</sub>-N<sub>9</sub>-N<sub>10</sub>, dsDNA<sub>10.5</sub>-N<sub>11</sub>-N<sub>12</sub>-N<sub>13</sub>-N<sub>14</sub>, and N<sub>15</sub>-N<sub>16</sub>-N<sub>17</sub>-dsDNA<sub>17.5</sub>-N<sub>18</sub>)



**Figure 8. Activating Exogenous Reporter Gene with TALEs**

(A) Editing and activating the endogenous HNF4 $\alpha$  gene in HepG2 cells. (B) Editing and activating the endogenous E47 gene in PANC1 cells. (1) Cells transfected only by Lipofectamine as the control. (2) HepG2/PANC1 cells co-transfected by HNF4 $\alpha$ /E47-TALEN and HOM-HNF4 $\alpha$ /E47-ZsGreen for editing HNF4 $\alpha$ /E47 gene to form fused HNF4 $\alpha$ /E47-ZsGreen gene in genome. (3) Edited HepG2/PANC1 cells transfected by HNF4 $\alpha$ /E47-TALE1-VPR for activating the fused HNF4 $\alpha$ /E47-ZsGreen gene in genome. (4) HepG2/PANC1 cells co-transfected by HNF4 $\alpha$ /E47-TALEN, HOM-HNF4 $\alpha$ /E47-ZsGreen and HNF4 $\alpha$ /E47-TALE1-VPR.

were set up. The 20  $\mu$ L reactions contained 10 U *Bsa*I, 400 U T4 DNA ligase, 1 $\times$  T4 DNA ligase buffer, 2  $\mu$ g BSA, and 200 ng of each of five monomers. The reactions were incubated in a thermocycler for 10 cycles of 37 $^{\circ}$ C for 5 min and 16 $^{\circ}$ C for 10 min. The enzymes were inactivated at 50 $^{\circ}$ C for 5 min and 80 $^{\circ}$ C for 5 min. The remained linear DNA was removed by adding 1  $\mu$ L Plasmid-Safe nuclease (10 U/ $\mu$ L; Epibio) and 1  $\mu$ L ATP (10 mM) to each pentamer reaction and incubating at 37 $^{\circ}$ C for 30 min. The Plasmid-Safe nuclease was inactivated at 70 $^{\circ}$ C for 30 min. The pentamers were purified on a PCR cleaning column and collected in water at a concentration of 100 ng/ $\mu$ L. Second, four pentamers were cloned into TALE backbones by another digestion-ligation reaction. A 20  $\mu$ L reaction contained 75 ng TALE backbone (pTALE-VP64 or pTALE-VPR), 10 U *Bsm*BI, 400 U T4 DNA ligase, 1 $\times$  T4 DNA ligase buffer, and 200 ng of each of four pentamers. The reaction was incubated in a thermocycler for 10 cycles of 37 $^{\circ}$ C for 5 min and 16 $^{\circ}$ C for 10 min. The reactions were then incubated at 50 $^{\circ}$ C for 5 min and 80 $^{\circ}$ C for 5 min to inactivate enzymes. Third, *E. coli* DH5 $\alpha$  was transformed with the TALE products and cultivated overnight on lysogeny broth (LB) agar containing kanamycin, X-gal, and IPTG at 37 $^{\circ}$ C. The positive colonies were identified by colony PCR with primers TAL-F and TAL-R (Table S2) using 1 $\times$  premix Taq. Finally, the selected colonies were cultivated overnight in liquid

LB medium containing 100  $\mu$ g/mL kanamycin at 37 $^{\circ}$ C. A small number of plasmids was prepared with the QIAprep Spin Miniprep Kit (QIAGEN) and verified by *Eco*RI digestion. Large quantities of plasmids were extracted with the EndoFree Plasmid kit (QIAGEN), according to the manufacturer's instructions.

#### Cell Culture and Transfection

Cells were cultured with DMEM containing 10% fetal bovine serum (FBS), 100 U/mL penicillin, and 100  $\mu$ g/mL streptomycin in a humidified incubator with 5% CO $_2$  at 37 $^{\circ}$ C. Cells were seeded in a 96-well plate at a density of 3,000 cells/well, a 24-well plate at a density of 5  $\times$  10 $^4$  cells/well, or a 6-well plate at a density of 2  $\times$  10 $^5$  cells/well and cultivated for 24 h. Cells were then transfected with plasmids by using Lipofectamine 2000 (Thermo Fisher Scientific), according to the manufacturer's instruction. All vectors used for transfection were isolated using EndoFree Plasmid kits. The transfected cells were cultured for an additional 48 or 72 h.

#### Activation of Reporter Construct with TALEs

293T cells cultivated in 24-well plates were co-transfected with 400 ng of TALE vector (HNF4 $\alpha$ -TALE-VP64, HNF4 $\alpha$ -TALE-VPR, or E47-TALE-VPR) and 400 ng of ZsGreen reporter vector

(HNF4 $\alpha$ -TALE-reporter or E47-TALE-reporter). Cells were photographed with a fluorescence microscope (IX51 with a DP71 camera; Olympus) and quantitatively analyzed with a flow cytometer (FACSCalibur; BD).

#### Detection of Gene Expression

Cells cultivated in 24-well plates were transfected by TALE-VPR. The elution buffer of the EndoFree Plasmid kit was used as blank control. After 48 h, the total RNA was extracted by Trizol. The cDNA was prepared using the FastKing RT Kit (with gDNase; Tian-gen Biotech), according to the manufacturer's instructions. The gene expression was detected by qPCR using 2 $\times$  Fast SYBR Green Master Mix (Applied Biosystems), according to the manufacturer's instructions.  $\beta$ -Actin was used as an internal reference to analyze the relative mRNA expression of different genes. PCR primers are shown in Table S4. The qPCR programs were run on a real-time PCR machine, StepOne Plus (Applied Biosystems). Each qPCR detection was performed in at least three technical replicates. Melting curve analysis was performed. Data analysis was performed using the Applied Biosystems StepOne software v2.3, and  $C_t$  values were normalized with that of  $\beta$ -actin. The relative expression level of target mRNAs was calculated as relative quantity (RQ) according to the equation:  $RQ = 2^{-\Delta\Delta C_t}$ . Student's t test was used to check the statistical significance.

#### Detection of Cell Viability and Cycle

Cells cultivated in 96-well plates were transfected with 200 ng TALE-VPR plasmid, using Lipofectamine 2000. The transfected cells were cultivated for variant time. Cell viability was detected with CCK8 (Biosharp), according to the manufacturer's instruction. The optical density (OD) value was read by a BioTek plate reader at the wavelength of 450 nm. The cell cycle was analyzed by flow cytometry. Cells were collected by trypsinization, resuspended with 40  $\mu$ L 4% paraformaldehyde, cells were added with 1  $\mu$ L 100  $\mu$ g/mL DAPI, and incubated at room temperature for 30 min. The cells were precipitated by centrifugation, washed once with PBS, then resuspended in PBS, and analyzed with a flow cytometer (FACSCalibur; BD). G0/G1, S, and G2/M phase were generated by modeling data with ModFitLT software (Verity Software House).

#### Detection of Cell Migration

Cells cultivated in 24-well plates were transfected with 800 ng of TALE-VPR plasmid using Lipofectamine 2000. The transfected cells were cultivated for 48 h, then collected by trypsinization and resuspended in serum-free DMEM at a final concentration of  $1 \times 10^5$  cells/mL. The cells were added to the upper chamber at the volume of 200  $\mu$ L per well. DMEM containing 10% FBS and 50  $\mu$ g/mL fibronectin was added to the lower chamber at the volume of 600  $\mu$ L per well. Cells were incubated at 37°C for 48 h. The Transwell was removed. The lower chamber was washed once with PBS and 100  $\mu$ L of PBS was added, containing 10 mg/mL acridine orange. After staining for 30 min at room temperature, the cells were washed three times with PBS, photographed with fluorescence microscope IX51 or DP71 (Olympus).

#### Gene Editing and Activation by TALEs

A pair of TALEN vectors was respectively designed for editing endogenous HNF4 $\alpha$  and E47 genes. The target sequences of these TALENs were shown in Table S5. The TALEN vectors were constructed using the simplified TALE vector assembly method developed in this study. To fuse a ZsGreen coding sequence to the endogenous HNF4 $\alpha$  and E47 genes by HDR, the ZsGreen coding sequence with two homologous arms (HOM-HNF4 $\alpha$ -ZsGreen and HOM-E47-ZsGreen) were prepared by PCR amplification, using the primers listed in Table S6. The HNF4 $\alpha$  gene in HepG2 cell was edited by co-transfecting HepG2 cell with a pair of TALEN vectors (Left-HNF4 $\alpha$ -TALEN and Right-HNF4 $\alpha$ -TALEN) and homologous donor (HOM-HNF4 $\alpha$ -ZsGreen). The E47 gene in PANC1 cell was edited by co-transfecting PANC1 cells with a pair of TALEN vectors (Left-E47-TALEN and Right-E47-TALEN) and homologous donor (HOM-E47-ZsGreen). To activate expression of endogenous HNF4 $\alpha$ -ZsGreen and E47-ZsGreen genes, HepG2 and PANC1 cells were respectively transfected with HNF4 $\alpha$ -TALE1-VPR and E47-TALE1-VPR. To edit and activate endogenous genes simultaneously, HepG2 and PANC1 cells were also respectively co-transfected with a pair of TALEN vectors, a homologous donor, and TALE1-VPR. The cells were photographed with a fluorescence microscope and quantified with flow cytometry.

#### SUPPLEMENTAL INFORMATION

Supplemental Information can be found with this article online at <https://doi.org/10.1016/j.omtm.2019.02.004>.

#### AUTHOR CONTRIBUTIONS

S.Z. performed most of the experiments. H.C. performed PCR amplification. J.W. conceived of and directed the study. J.W. and S.Z. wrote the manuscript.

#### CONFLICTS OF INTEREST

The authors declare no competing interests.

#### ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (61571119).

#### REFERENCES

- Boch, J., and Bonas, U. (2010). Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* 48, 419–436.
- Römer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U., and Lahaye, T. (2007). Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* 318, 645–648.
- Kay, S., Hahn, S., Marois, E., Wieduwild, R., and Bonas, U. (2009). Detailed analysis of the DNA recognition motifs of the Xanthomonas type III effectors AvrBs3 and AvrBs3Deltarep16. *Plant J.* 59, 859–871.
- Römer, P., Strauss, T., Hahn, S., Scholze, H., Morbitzer, R., Grau, J., Bonas, U., and Lahaye, T. (2009). Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper Bs3 alleles. *Plant Physiol.* 150, 1697–1712.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512.

6. Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501.
7. Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148.
8. Morbitzer, R., Römer, P., Boch, J., and Lahaye, T. (2010). Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc. Natl. Acad. Sci. USA* 107, 21617–21622.
9. Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., and Arlotta, P. (2011). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* 29, 149–153.
10. Geissler, R., Scholze, H., Hahn, S., Streubel, J., Bonas, U., Behrens, S.E., and Boch, J. (2011). Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* 6, e19509.
11. Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39, e82.
12. Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P., and Yang, B. (2011). Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.* 39, 6315–6325.
13. Wood, A.J., Lo, T.W., Zeitler, B., Pickle, C.S., Ralston, E.J., Lee, A.H., Amora, R., Miller, J.C., Leung, E., Meng, X., et al. (2011). Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307.
14. Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761.
15. Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 29, 731–734.
16. Tesson, L., Usal, C., Ménotet, S., Leung, E., Niles, B.J., Remy, S., Santiago, Y., Vincent, A.I., Meng, X., Zhang, L., et al. (2011). Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.* 29, 695–696.
17. Sun, N., Liang, J., Abil, Z., and Zhao, H. (2012). Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol. Biosyst.* 8, 1255–1263.
18. Park, C.Y., Kim, J., Kweon, J., Son, J.S., Lee, J.S., Yoo, J.E., Cho, S.R., Kim, J.H., Kim, J.S., and Kim, D.W. (2014). Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPSCs using TALENs. *Proc. Natl. Acad. Sci. USA* 111, 9253–9258.
19. Qasim, W., Amrolia, P.J., Samarasinghe, S., Ghorashian, S., Zhan, H., Stafford, S., Butler, K., Ahsan, G., Gilmour, K., Adams, S., et al. (2015). First Clinical Application of Talen Engineered Universal CAR19 T Cells in B-ALL. *Blood* 126, 2046.
20. Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., Butler, K., Rivat, C., Wright, G., Somana, K., et al. (2017). Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* 9, eaaj2013.
21. Sander, J.D., Cade, L., Khayter, C., Reyon, D., Peterson, R.T., Joung, J.K., and Yeh, J.R.J. (2011). Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.* 29, 697–698.
22. Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., and Joung, J.K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* 30, 460–465.
23. Weber, E., Gruetzner, R., Werner, S., Engler, C., and Marillonnet, S. (2011). Assembly of designer TAL effectors by Golden Gate cloning. *PLoS ONE* 6, e19722.
24. Morbitzer, R., Elsaesser, J., Hausner, J., and Lahaye, T. (2011). Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res.* 39, 5790–5799.
25. New England Biolabs. Cleavage of Supercoiled DNA. <https://www.neb.com/tools-and-resources/selection-charts/cleavage-of-supercoiled-dna>.
26. Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* 3, e3647.
27. Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009). Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS ONE* 4, e5553.
28. Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* 6, e16765.
29. Chavez, A., Tuttle, M., Pruitt, B.W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S.J., Cecchi, R.J., Kowal, E.J.K., Buchthal, J., et al. (2016). Comparison of Cas9 activators in multiple species. *Nat. Methods* 13, 563–567.
30. Zhang, Z., Wu, E., Qian, Z., and Wu, W.S. (2014). A multicolor panel of TALE-KRAB based transcriptional repressor vectors enabling knockdown of multiple gene targets. *Sci. Rep.* 4, 7338.
31. Wang, H., Guo, R., Du, Z., Bai, L., Li, L., Cui, J., Li, W., Hoffman, A.R., and Hu, J.-F. (2018). Epigenetic Targeting of Granulin in Hepatoma Cells by Synthetic CRISPR dCas9 Epi-suppressors. *Mol. Ther. Nucleic Acids* 11, 23–33.
32. Murakami, M.T., Sforça, M.L., Neves, J.L., Paiva, J.H., Domingues, M.N., Pereira, A.L., Zeri, A.C., and Benedetti, C.E. (2010). The repeat domain of the type III effector protein PthA shows a TPR-like structure and undergoes conformational changes upon DNA interaction. *Proteins* 78, 3386–3395.
33. Scholze, H., and Boch, J. (2011). TAL effectors are remote controls for gene activation. *Curr. Opin. Microbiol.* 14, 47–53.
34. Kim, S., Lahmy, R., Riha, C., Yang, C., Jakubison, B.L., van Niekerk, J., Staub, C., Wu, Y., Gates, K., Dong, D.S., et al. (2015). The basic helix-loop-helix transcription factor E47 reprograms human pancreatic cancer cells to a quiescent acinar state with reduced tumorigenic potential. *Pancreas* 44, 718–727.
35. Yin, C., Lin, Y., Zhang, X., Chen, Y.X., Zeng, X., Yue, H.Y., Hou, J.L., Deng, X., Zhang, J.P., Han, Z.G., and Xie, W.F. (2008). Differentiation therapy of hepatocellular carcinoma in mice with recombinant adenovirus carrying hepatocyte nuclear factor-4alpha gene. *Hepatology* 48, 1528–1539.
36. Yin, S., Li, J., Hu, C., Chen, X., Yao, M., Yan, M., Jiang, G., Ge, C., Xie, H., Wan, D., et al. (2007). CD133 positive hepatocellular carcinoma cells possess high capacity for tumorigenicity. *Int. J. Cancer* 120, 1444–1450.
37. Suetsugu, A., Nagaki, M., Aoki, H., Motohashi, T., Kunisada, T., and Moriwaki, H. (2006). Characterization of CD133+ hepatocellular carcinoma cells as cancer stem/progenitor cells. *Biochem. Biophys. Res. Commun.* 351, 820–824.
38. Ma, S., Chan, K.W., Hu, L., Lee, T.K., Wo, J.Y., Ng, I.O., Zheng, B.J., and Guan, X.Y. (2007). Identification and characterization of tumorigenic liver cancer stem/progenitor cells. *Gastroenterology* 132, 2542–2556.
39. Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920.
40. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872.
41. Joshi, U.S., Dergahm, S.T., Chen, Y.Q., Dugan, M.C., Crissman, J.D., Vaitkevicius, V.K., and Sarkar, F.H. (1998). Inhibition of pancreatic tumor cell growth in culture by p21WAF1 recombinant adenovirus. *Pancreas* 16, 107–113.
42. Gironella, M., Seux, M., Xie, M.J., Cano, C., Tomasini, R., Gommeaux, J., Garcia, S., Nowak, J., Yeung, M.L., Jeang, K.T., et al. (2007). Tumor protein 53-induced nuclear protein 1 expression is repressed by miR-155, and its restoration inhibits pancreatic tumor development. *Proc. Natl. Acad. Sci. USA* 104, 16170–16175.

**OMTM, Volume 13**

**Supplemental Information**

**Generate TALE/TALEN as Easily and Rapidly  
as Generating CRISPR**

**Shuyan Zhang, Huiting Chen, and Jinke Wang**

## Supplementary information

### Reagents used in this study and their manufacturers

The plasmids pIRES2-EGFP, pcDNA-dCas9-VP64, pcDNA-dCas9-VPR, pEGFP-C1, and pEZX-miniCMV-ZsGreen were kept by our laboratory. The Golden Gate TALEN and TAL Effector Kit 2.0 was purchased from Addgene. PrimeSTAR® HS (Premix), Premix Taq™ (TaKaRa Taq™ Version 2.0), DNA Markers, cDNA reverse transcription kit, and some restriction endonucleases, including EcoRI, XhoI, HindIII, and NotI, were purchased from Takara. BsaI, BsmBI, and T4 DNA Ligase were purchased from New England Biolabs (NEB). The plasmid-safe nuclease was purchased from Epicentre Biotechnologies. Lipofectamine™ 2000, Opti-MEM, DAPI, and Trizol were purchased from Invitrogen. The Endo-Free Plasmid kit was purchased from CWBio. Diethylpyrocarbonate (DEPC) was purchased from Sigma. 293T, HepG2, and PANC1 cells were purchased from the Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences (CAS). All oligonucleotides were manufactured by Sangon Biotech (Shanghai). The MinElute Gel Extraction kit and QIAprep Spin Miniprep Kit were purchased from Qiagen. The FastKing RT Kit (with gDNase), the Cell Counting Kit-8 (CCK8) and the competent *E. coli* DH5α were purchased from Tiangen. The fetal bovine serum (FBS) was purchased from HyClone. The Fast SYBR® Green Master Mix was purchased from Applied Biosystems (ABI). The transwell was purchased from Corning. Fibronectin was purchased from Becton Dickinson (BD).

Table S1. Primers for TALE monomers amplification and clony PCR

Name	Sequence (5'-3')	Usage
PCR-TAL-F	TATCATCATGCCTCCTCTAGAG	monomer amplification (universal primers)
PCR-TAL-R	TTGGTCATGGGTGGCTCGAGG	
PCR-TAL-F1	TATCATCATGCCTCCTCTAGAGGTCTCCCTATCTTAAAC CGGCCAACATAACCGTCTCCCCCTGAACCTGACCCCG GACCAAGTGGTGGCTATCGCCAGC	First monomer amplification
PCR-TAL-R5	TTGGTCATGGGTGGCTCGAGGGTCTCCATA GAGTCTGTCTTTCCCCTTTCCCGTCTCCTGCACCG	Fifth monomer amplification
PCR-TAL-F6	TATCATCATGCCTCCTCTAGAGGTCTCCCTATCTTAAAC CGGCCAACATAACCGTCTCGTGCAGCGGC	Sixth monomer amplification
PCR-TAL-R10	TTGGTCATGG GTGGCTCGAGGGTCTCCATA GAGTCTGTCTTTCCCCTTTCCCGTCTCCCGCC	Tenth monomer amplification
PCR-TAL-R14	TTGGTCATGG GTGGCTCGAGGGTCTCCTCG AAGTCTGTCTTTCCCCTTTCCCGTCTCCAACAGCCG	Fourteenth monomer amplification
PCR-TAL-F15	TATCATCATGCCTCCTCTAGAGGTCTCCTCGACTTAAAC CGGCCAACATAACCGTCTCCTGTTGCCGG	Fifteenth monomer amplification
PCR-TAL-F18	TATCATCATGCCTCCTCTAGAGGTCTCGCTATCGCCAGC	Eighteenth monomer amplification
PCR-TAL-R18	TTGGTCATGGGTGGCTCGAGGGTCTCTCGAAGTCTGT CTTTCCCCTTTCCCGTCTCTCGTTGGTCAAC	Eighteenth monomer amplification
TAL-F	TTGGCGTCGGCAAACAGTGG	Colony PCR amplification and TALE sequencing
TAL-R	GGCGACGAGGTGGTCGTTGG	Colony PCR amplification and TALE sequencing

Table S2. Oligonucleotides used as templates for amplifying two linker monomers

Name	Encoded chain sequence (5'-3')
DNA <sub>10.5</sub>	TATCATCATGCCTCCTCTAGAGGTCTCCTCGACTTAAACCGGCCAACATACCCGTCTCTGGCG GCAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCG AGACCCTCGAGCCACCCATGACCAA
DNA <sub>17.5</sub>	TATCATCATGCCTCCTCTAGAGGTCTCCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCC AGGACCATGGCCTGACCCCGGACCAAGTGGTGGCTATCGAGACCCTCGAGCCACCCATGAC CAA

Table S3. Primers for constructing TALE backbones and fluorescence reporter construct

Name	Sequence (5'-3')	Usage
VP64-F	CCGGAATTCAGCGCTGGAGGAGGTGGAAG	VP64 amplification
VP64-R	AAGGAAAAAAGCGGCCGCTCAGTTAATCAGCATGTC	
VPR-F	CCGGAATTCGACGCATTGGACGATTTTGATCTGGATAT GCTGGGAAGTG	VPR amplification
VPR-R	AAGGAAAAAAGCGGCCGCAAACAGAGATGTGTGCGAA GATGGACAGTCCTGTGCTG	
HNF4 $\alpha$ -F	CCGCTCGAGTGAGATCCAAAACCTGAGAC	HNF4 $\alpha$ promoter amplification
HNF4 $\alpha$ -R	CCCAAGCTTAAGCCCACCCAGCCGGAGAG	
E47-F	CCGCTCGAGGCTCAGTAGC CACCAACCACC	E47 promoter amplification
E47-R	CCCAAGCTTTCTGTGGAG GGGAGCTGGTAAG	

Table S4. Primers for qPCR detection of gene expression

Name	Forward primer (5'-3')	Reverse primer (5'-3')
GS	CCTGCTTGATGCTGGAGTC	GAAAAGTCGTTGATGTTGGA
APOCIII	GGGTACTCCTTGTGTTC	AAATCCCAGAAGCTCAGAGAAC
CYP1 $\alpha$ 2	CTGGCCTCTGCCATCTTCTG	TTAGCCTCCTTGCTCACATGC
G-6-P	GGCTCCATGACTGTGGGATC	TTCAGCTGCACAGCCCAGAA
PEPCK	GTGTCCCTCTAGTCTATGAAGC	ATTGACTTGATCCTCCAGATAC
BR	ACAAGGTGCTGCGGGAATCA	ACTGGTGGGAGGGGTAGGTG
ALDOB	AGGAGGACTCTTCTCTCCCAA	GATTCATCTGCAGCCAGGAT
APOAI	GCCTTGGGAAAACAGCTAAA	AGGCCCTCTGTCTCCTTTTC
GYS2	CCAGTGGGAAGTCGAAGAAC	TTCTCTCCCCATTCATCTGC
HPD	TTGGGAAGGTGAAGTTTGCT	GCATTTGGGCAGTTTAGGAA
CD133	ACATGAAAAGACCTGGGGG	GATCTGGTGT CCCAGCATG
CD90	CGGAAGACCCCAGTCCA	ACGAAGGCTCTGGTCCACTA
Oct3/4	CGACCATCTGCCGCTTTGAG	CCCCCTGTCCCCCATTCCTA
SMO	ATCTCCACAGGAGAGACTGGTTCCGG	AAAGTGGGGCCTTGGAACATG
$\beta$ -catenin	AGGTCTGAGGAGCAGCTTCA	ATTGTCCACGCTGGATTTTC
Bmi	GGAGACCAGCAAGTATTGTCTATTT	CATTGTCGCTGGGCATCGTAAG
LIN28	TGTAAGTGGTTCAACGTGCG	CCTCACCTCCTTCAAGCTC
c-Myc	GCTGCTTAGACGCTGGATTT	CTCCTCCTCGTCGCAGTAGA
Klf4	GCGGCAAACCTACACAAAAG	CCCCGTGTGTTTACGGTAGT
Sox2	GCGAACCATCTCTGTGGTCT	GGAAAGTTGGGATCGAACAA
ESG1	GCGCAGTATCACAGCCTTAAA	TCAATCTCTTGGCGATTCA
NANOG	GATTTGTGGGCCTGAAGAAA	TTGGGACTGGTGAAGAATC
$\beta$ -actin	CCTGGCACCCAGCACAAAT	GGGCCGGACTCGTCATACT
P21	GGATGTCCGTCAGAACCCAT	CCCTCCAGTGGTGTCTCGGTG
TP53INP1	AAACCTTCTCATTGAACATCCC	CCATTGTGCTTGACTTGCC
HNF6	AACCCTGGAGCAAACCTCAA	AAGACCAACCTGGGCTTTTT
Cx32	CTGCAGACATTCTCTGGGAAA	GCACCATGATTCTGAAGATGA
Top2a	GGGTAGCAATAATCTAAACCT	CCAGTTCTTCAATAGTACCCT
Aurka	TCTAGTCTCCTTAACCACTTATCT	GACACATGGCCTCTTCTGTATC
MIST1	CCAGCACTACCAGCAGCA	AGGACTGGGCGCTAGGTG
Sox9	CACCCGGATTACAAGTACCAG	AAGGTCTCGATGTTGGAGATG

Table S5. Target sequences of TALENs

Name	Sequence (5'-3')	Usage
Left-HNF4a-TALEN	AATGCTCCGGGCTGGCA	edit HNF4a
Right-HNF4a-TALEN	GGGGCGGGGAGGGCCGAG	
Left-E47-TALEN	AACCCCGCCGGGCACATG	edit E47
Right-E47-TALEN	GAAAGCGGGTGGCTCGTC	

Table S6. Primers amplifying HOM-HNF4a-ZsGreen and HOM-E47-ZsGreen

Name	Sequence (5'-3')	Usage
HOM-HNF4a-F	CTACTGCAGGCTCAAGAAATGCTTCCGGGCTGG CAGGAGGAGGTTCCGGTGGAGGTGGTTCTGGA ATGGCCCAGTCCAAGCACGGCCTGA	HOM-HNF4a-ZsGreen amplification
HOM-HNF4a-R	GGTGGGGCAG TGGTGGTGGG GCGGGGAGGG CCGAG TTAGGGCAAGGCGGAGCCGGAGGCG	
HOM-E47-F	GCCTGAGCGAAGCCCACAACCCCGCCGGGCAC ATGGGAGGAGGTTCCGGTGGAGGTGGTTCTGG A ATGGCCCAGTCCAAGCACGGCCTGA	HOM-E47-ZsGreen amplification
HOM-E47-R	GGCCAGAGCA CAGGGCTGAA AGCGGGTGGC TCGTCTTAGGGCAAGGCGGAGCCGGAGGCG	

**a** Golden Gate TALEN and TAL Effector Kit 2.0 Plate Map

	1	2	3	4	5	6	7	8	9	10	11	12
A	pHD1	pHD2	pHD3	pHD4	pHD5	pHD6	pHD7	pHD8	pHD9	pHD10	pNG1	pNG2
B	pNG3	pNG4	pNG5	pNG6	pNG7	pNG8	pNG9	pNG10	pNI1	pNI2	pNI3	pNI4
C	pNI5	pNI6	pNI7	pNI8	pNI9	pNI10	pNN1	pNN2	pNN3	pNN4	pNN5	pNN6
D	pNN8	pNN8	pNN9	pNN10	pNK1	pNK2	pNK3	pNK4	pNK5	pNK6	pNK7	pNK8
E	pNK9	pNK10	pNH1	pNH2	pNH3	pNH4	pNH5	pNH6	pNH7	pNH8	pNH9	pNH10
F	pLR-HD	pLR-NG	pLR-NI	pLR-NN	pFUS_B1	pFUS_B2	pFUS_B3	pFUS_B4	pFUS_B5	pFUS_B6	pFUS_B7	pFUS_B8
G	pFUS_B9	pFUS_B10	pFUS_A	pFUS_A30A	pFUS_A30B	pLR-NH	pTAL1	pTAL2	pTAL3	pTAL4	pZHY500	pZHY501
H	pCP5b	pKEB31										

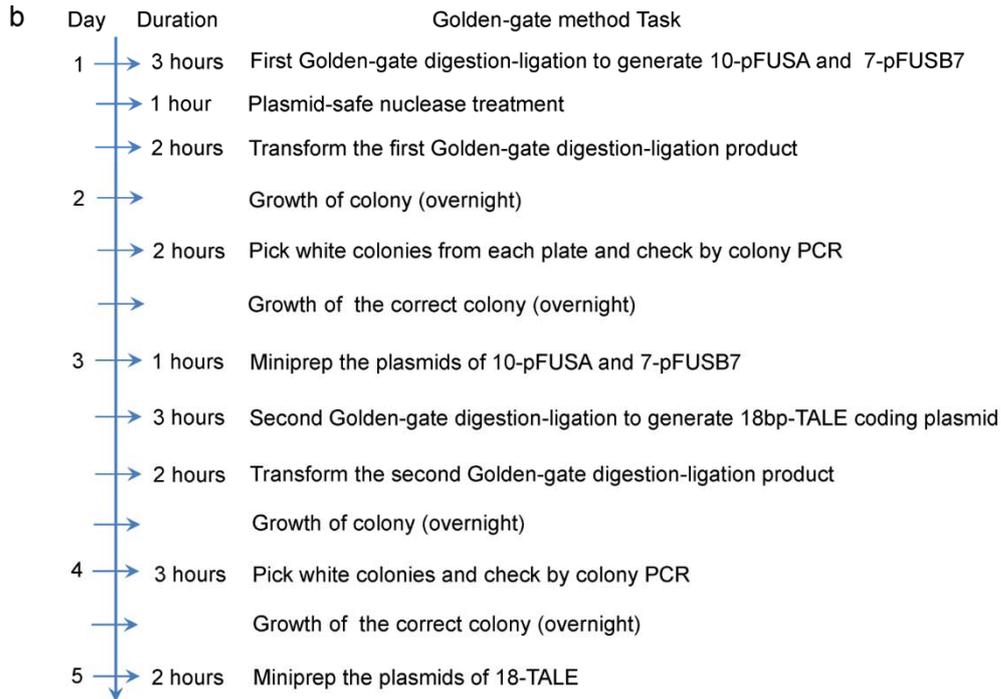


Figure S1. Golden Gate TALEN and TAL Effector Kit 2.0 (Addgene). (a) Plasmid 96-well plate of Addgene kit 2.0. Plasmids were provided as transformed *E.coli* DH5 $\alpha$ . (b) Pipeline for constructing TALE with Addgene kit 2.0. This scheme shows an example for constructing a TALE that can bind a target site at the length of 18 bp. 10-pFUSA, a plasmid prepared with 10 monomer plasmids and one receptor plasmid pFUSA in a Golden-gate digestion-ligation reaction; 7-pFUSB7, a plasmid prepared with 7 monomer plasmids and one receptor plasmid pFUSB7 in a Golden-gate digestion-ligation reaction. As a result, 10-pFUSA codes the amino acid binding the N1–N10 nucleotides and 7-pFUSB7 codes the amino acid binding the N11–N17 nucleotides. The nucleotides coding the amino acid binding the last nucleotide (N18) are contained in another monomer plasmid (pLR). The second Golden-gate digestion-ligation reaction contains four plasmids including 10-pFUSA, 7-pFUSB7, pLR, and a TALE backborn vector (pTAL1).

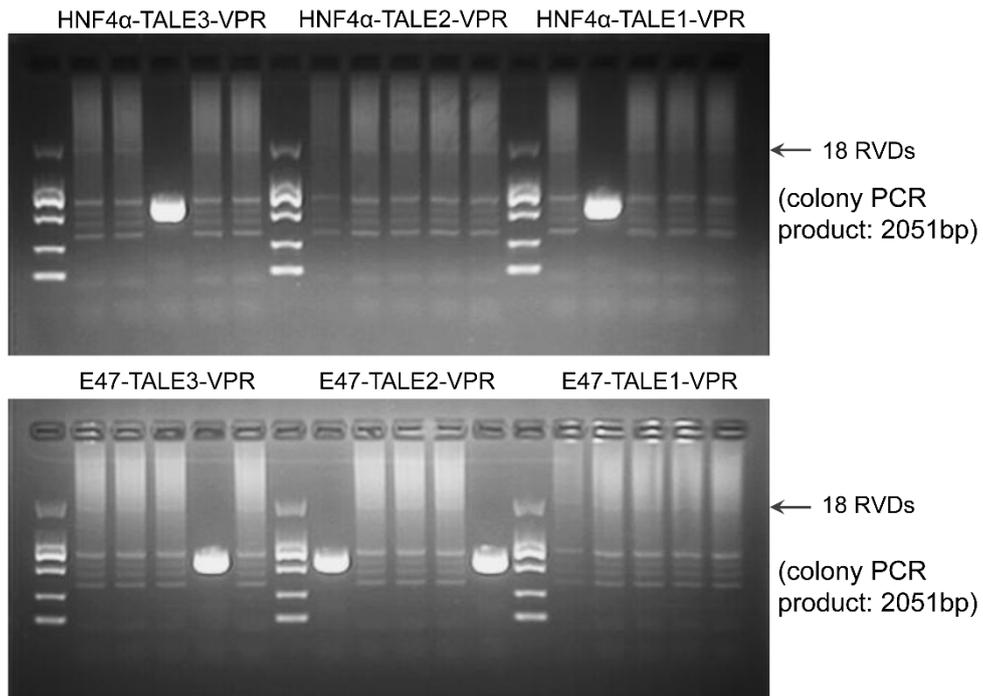


Figure S2. Detection of final TALE-VPR colonies with colony PCR. Full-length PCR products are often less prominent while the "ladder effect" represents a robust indicator of successful assembly. DNA Marker: DL2000 mark. Randomly picked white spots were detected with by colony PCR colony PCR, which indicated that most white spots are correctly assembled TALEs.

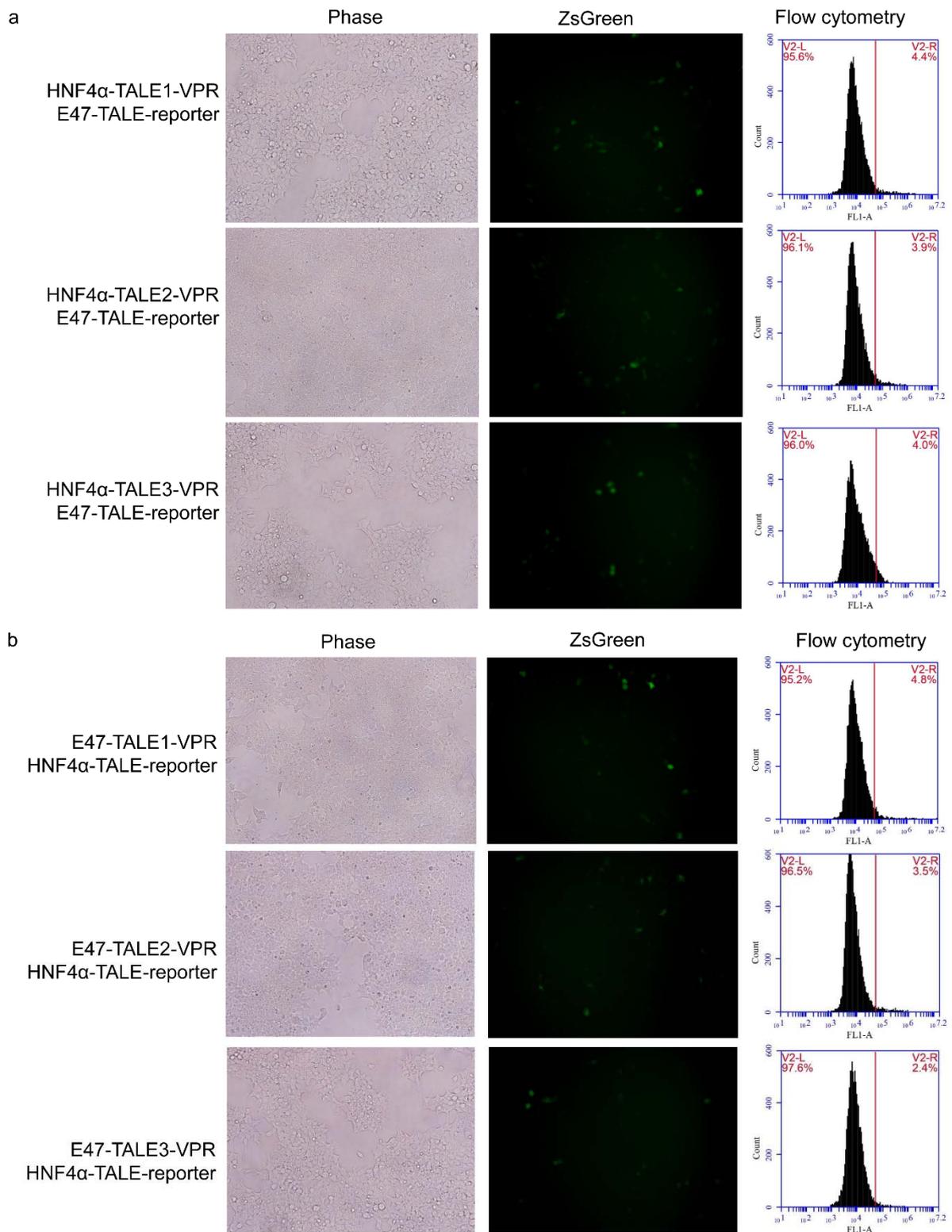


Figure S3. Activating exogenous reporter gene with TALEs. 293T cells were co-transfected with a TALE-VPR plasmid and its unrelated corresponding reporter plasmid expressing ZsGreen. (a) Cells co-transfected with HNF4 $\alpha$ -TALE1/2/3-VPR and E47-TALE-reporter. (b) Cells cotransfected with E47-TALE1/2/3-VPR and HNF4 $\alpha$ -TALE-reporter.

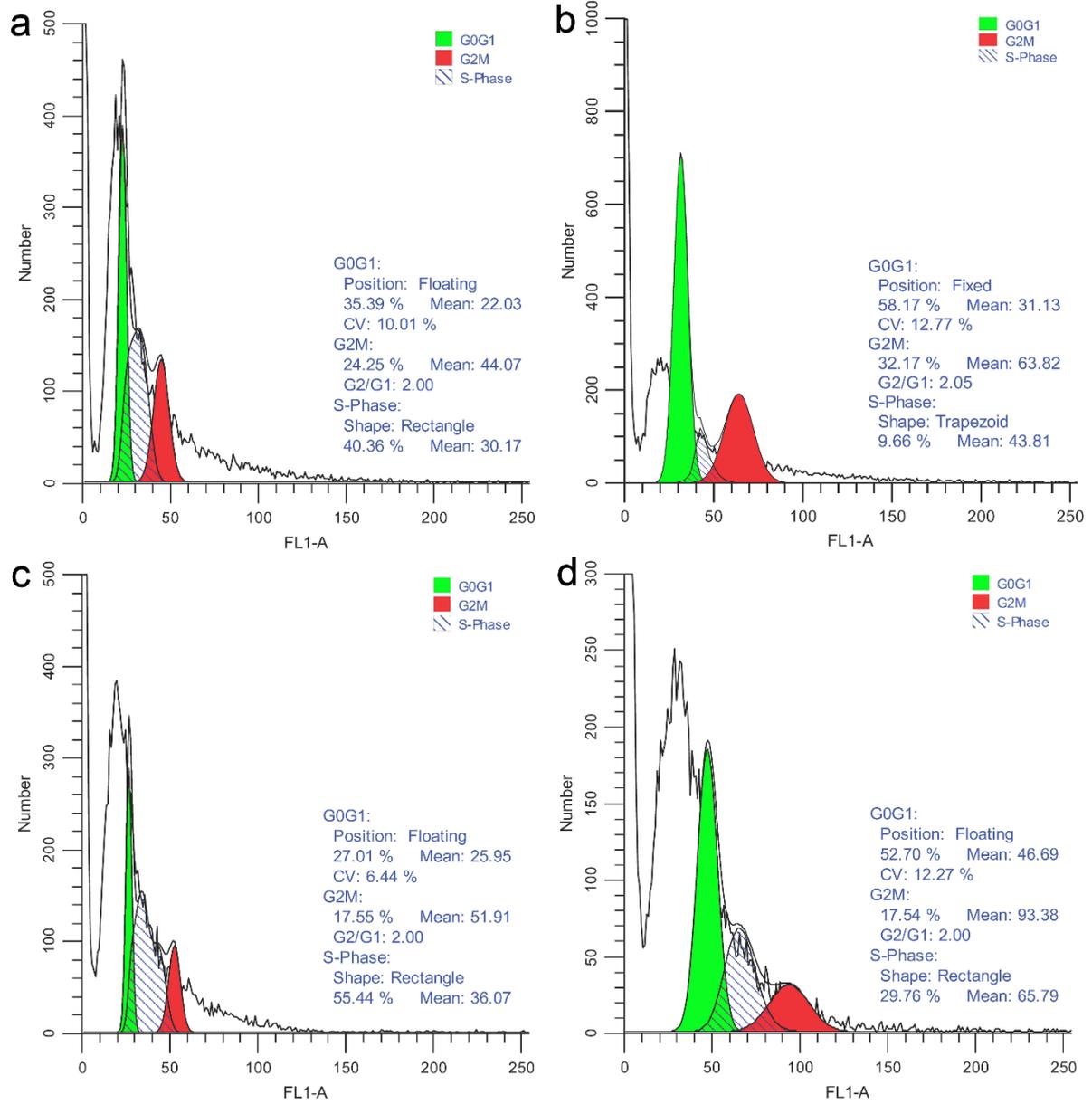
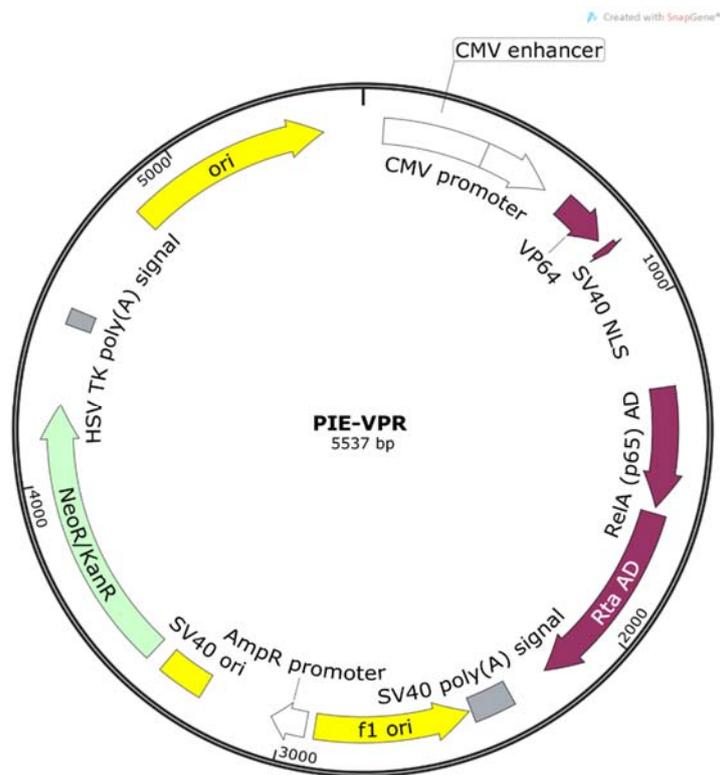
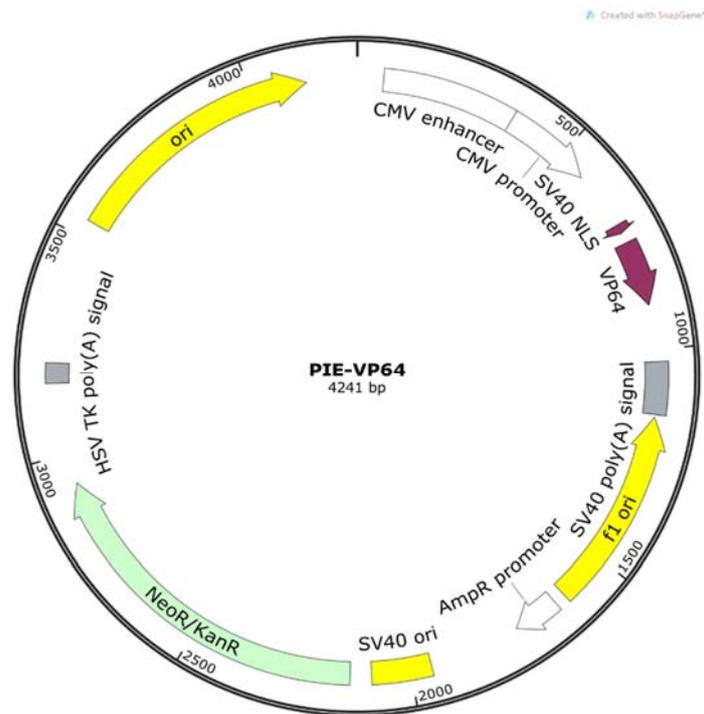
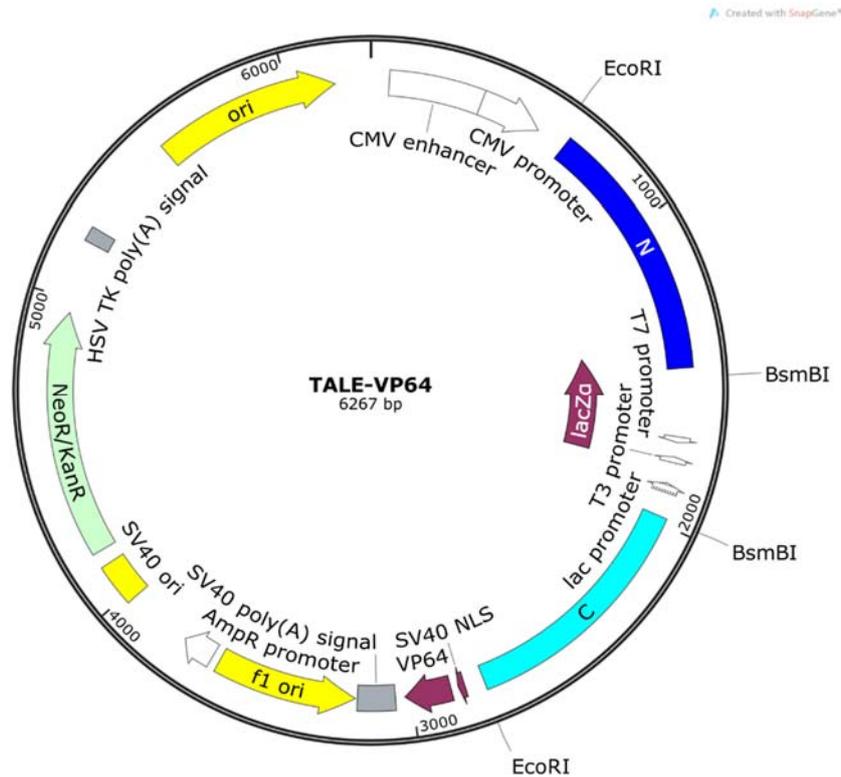


Figure S4. Detection of cell cycle with flow cytometry. (a) HepG2 treated by lipo. (b) HepG2 treated by lipo and HNF4 $\alpha$ -TALE1-VPR. (c) PANC1 treated by lipo. (d) PANC1 treated by lipo and E47-TALE1-VPR.

## Vectors and their elements and sequences

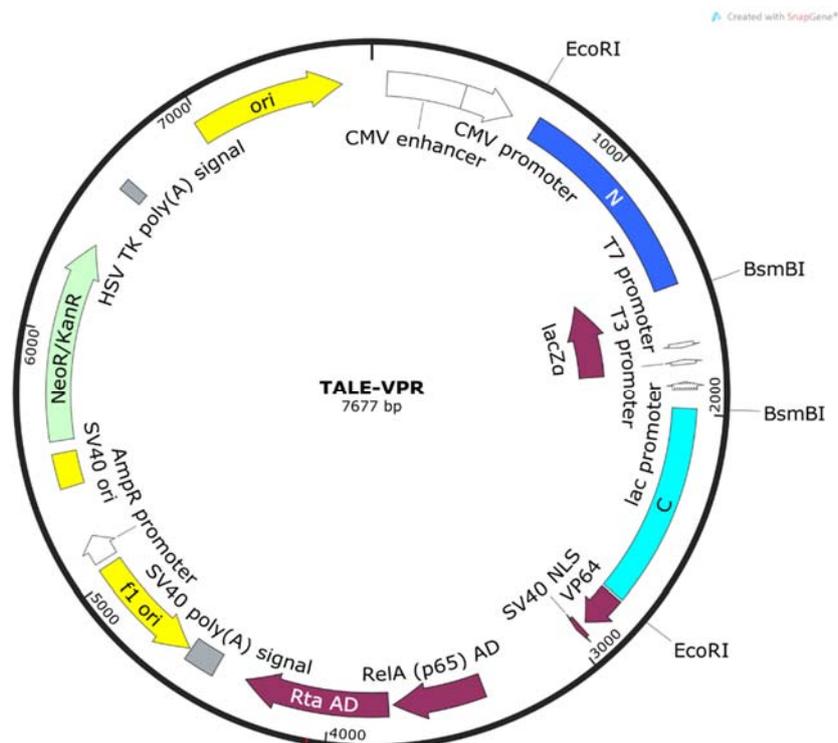




CMV enhancer + CMV promoter + N-term of TALE + Repeating sequence area of TALE instead of LacZα + C-term of TALE+VP64 (SV40NLS+VP64)

CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACG  
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG  
AGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCT  
ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACT  
TTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT  
ACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTCCAAGTCTCCACCCCATGACGT  
CAATGGGAGTTTTGTTTTGGCACAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGC  
CCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTATGGAT  
CCCATTTCGTCGCGCAGGCCAAGTCCTGCCCGGAGCTTCTGCCCGGACCCCAACCGGATAGG  
GTTCAGCCGACTGCAGATCGTGGGGTGTCTGCGCCTGCTGGCAGCCCTCTGGATGGCTTGCC  
GCTCGGCGGACGGTGTCCCGACCCGGCTGCCATCTCCCCCTGCGCCCTCACCTGCGTTCTCG  
GCGGGCAGCTTCAGCGATCTGCTCCGTCCGTTTCGATCCGTCGTTCTTGATAACATCGCTTCTTGA  
TTCGATGCCTGCCGTCGGCACGCCGATACAGCGGCTGCCCCAGCAGAGTGGGATGAGGCGCA  
ATCGGCTCTGCGTGCAGCCGATGACCCGCCACCCACCGTGCCTGTCGCTGTCACTGCCGCGCG  
GCCGCCGCGCGCAAGCCGGCCCCGCGACGGCGTGTGCGCAACCCTCCGACGCTTCGCCGG  
CCGCGCAGGTGGATCTACGCAGCTCGGCTACAGTCAGCAGCAGCAAGAGAAGATCAAACCG  
AAGGTGCGTTCGACAGTGGCGCAGCACCACGAGGCACTGGTGGGCCATGGGTTTACACACGC  
GCACATCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCACGTATCAGCA  
CATAATCACGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGT  
CCGGCGCACGCGCCCTGGAGGCCTTGCTCACGGATGCGGGGAGTTGAGAGGTCCGCCGTTA  
CAGTTGGACACAGGCCAACTTGTGAAGATTGCAAACCGTGGCGGCGTGACCGCAATGGAGGC  
AGTGCATGCCCTGGAGACGGGCGCCGCTACAGGGCGCGTCCATTTCGCCATTCAGGCTGCGCA  
ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGAT

GTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAACGA  
CGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACGGGCCCCCCTCG  
AGGTCCCTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATA  
GCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATA  
AAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC  
CCGTTTTCCACGGTCTCTCCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGACGTCC  
TGCCATGGATGCAGTGAAAAAGGGATTGCCGCACGCGCCGGAATTGATCAGAAGAGTCAATCG  
CCGTATTGGCGAACGCACGTCCCATCGCGTTGCCGACTACGCGCAAGTGGTTCGCGTGCTGGA  
GTTTTTCCAGTGCCACTCCCACCCAGCGTACGCATTTGATGAGGCCATGACGCAGTTCGGGATG  
AGCAGGAACGGGTTGGTACAGCTCTTTCGCAGAGTGGGCGTCACCGAACTCGAAGCCCGCGG  
TGGAACGCTCCCCCAGCCTCGCAGCGTTGGGACCGTATCCTCCAGGCATCAGGGATGAAAAG  
GGCCAAACCGTCCCCTACTTCAGCTCAAACACCGGATCAGGCGTCTTTGCATGCATTCGCCGAT  
TCGCTGGAGCGTGACCTTGATGCGCCCAGCCCAATGCACGAGGGAGATCAGACGCGGGCAAG  
CAGCCGTAAACGGTCCCGATCGGATCGTGCTGTCACCGGCCCTCCGCACAGCAGGCTGTCGA  
GGTGC GCGTTC CCGAACAGCGCGATGCGCTGCATTTGCCCTCAGCTGGAGGGTAAAACGCC  
GCGTACCAGGATCTGGGGCGGCCTCCCGGATCCTGGTACGCCATGGCTGCCGACCTGGCAGC  
GTCCAGCACCGTGATGTGGGAACAAGATGCGGACCCCTTCGCAGGGGCAGCGGATGATTTCCC  
GGCATTCAACGAAGAGGAACTCGCATGGTTGATGGAGCTATTGCCTCAGAAGAGCGCTGGAGG  
AGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCGGACCTAAGAAAAAGAGGAA  
GGTGGCCGCTGCTGGCTCTGGACGGGCTGACGCATTGGACGATTTTGATCTGGATATGCTGGGA  
AGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCCTTGATGACTTTGACC  
TCGACATGCTCGGCAGTGACGCCCTTGATGATTTGACCTGGACATGCTGATTA ACTGA

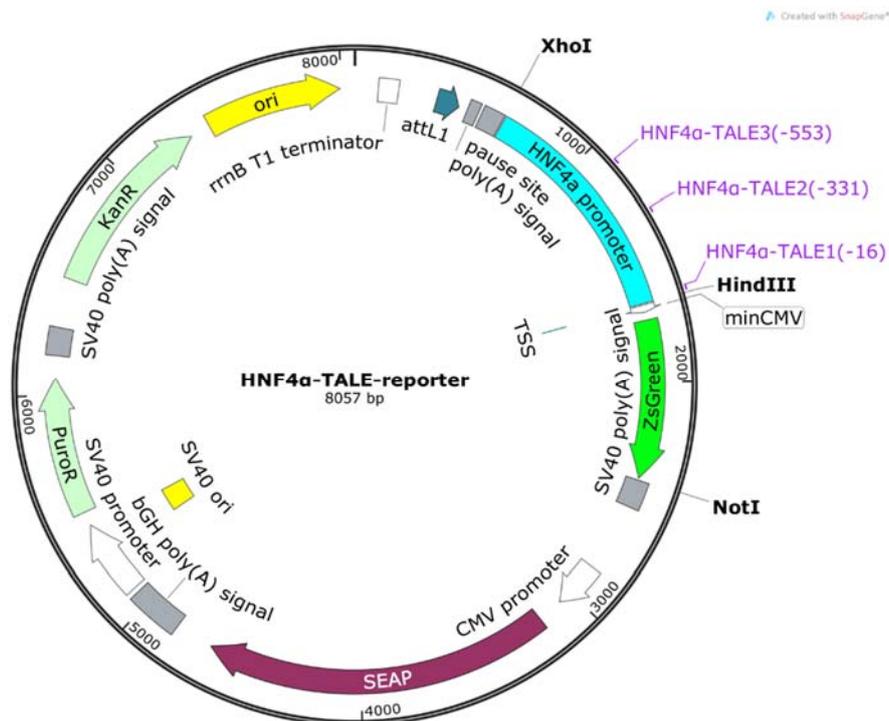


CMV enhancer + CMV promoter + N-term of TALE + Repeating sequence area of TALE instead of LacZα + C-term of TALE + VPR (VP64+SV40 NLS + Rel A(p65)AD + RtaAD)

CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACG  
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG  
AGTATTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCT  
ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACT  
TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT  
ACATCAATGGGCGTGGATAGCGGTTTACTCACGGGGATTTCOAAGTCTCCACCCCATGACGT  
CAATGGGAGTTTTGTTTTGGCACAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGC  
CCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTATGGAT  
CCCATTTCGTCGCGCAGGCCAAGTCCTGCCCGGAGCTTCTGCCCGGACCCCAACCGGATAGG  
GTTGACCCGACTGCAGATCGTGGGGTGTCTGCGCCTGCTGGCAGCCCTCTGGATGGCTTGCCC  
GCTCGGCGGACGGTGTCCCGGACCCGGCTGCCATCTCCCCTGCGCCCTCACCTGCGTTCTCG  
GCGGGCAGCTTCAGCGATCTGCTCCGTCCGTTTCGATCCGTCGTTCTTGATAACATCGCTTCTTGA  
TTCGATGCCTGCCGTCGGCACGCCGCATACAGCGGCTGCCCCAGCAGAGTGGGATGAGGCGCA  
ATCGGCTCTGCGTGCAGCCGATGACCCGCCACCCACCGTGCCTGTCGCTGTCACTGCCGCGCG  
GCCCGCGCGCCAAGCCGGCCCCGCGACGGCGTGTGCGCAACCCTCCGACGCTTCGCCCGG  
CCGCGCAGGTGGATCTACGCACGCTCGGCTACAGTCAGCAGCAGCAAGAGAAGATCAAACCG  
AAGGTGCGTTCGACAGTGGCGCAGCACCACGAGGCACTGGTGGGCCATGGGTTTACACACGC  
GCACATCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCACGTATCAGCA  
CATAATCACGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGT  
CCGCGCACGCGCCCTGGAGGCCTTGCTCACGGATGCGGGGAGTTGAGAGGTCCGCCGTTA  
CAGTTGGACACAGGCCAACTTGTGAAGATTGAAAACGTGGCGGCGTGACCGCAATGGAGGC  
AGTGCATGCCCTGGAGACGGGCGCGCTACAGGGCGCGTCCATTGCCATTCAGGCTGCGCA

ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGAT  
GTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA AACGA  
CGGCCAGTGAGCGCGCTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCTCG  
AGGTCTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATA  
GCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATA  
AAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC  
CCGTTTTCCACCGGTCGTCTCCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGACGTCC  
TGCCATGGATGCAGTGAAAAAGGGATTGCCGCACGCGCCGGAATTGATCAGAAGAGTCAATCG  
CCGTATTGGCGAACGCACGTCCCATCGCGTTGCCGACTACGCGCAAGTGGTTCGCGTGCTGGA  
GTTTTTCCAGTGCCACTCCCACCCAGCGTACGCATTTGATGAGGCCATGACGCAGTTCGGGATG  
AGCAGGAACGGGTTGGTACAGCTCTTTCGCAGAGTGGGCGTCACCGAACTCGAAGCCCGCGG  
TGGAACGCTCCCCCAGCCTCGCAGCGTTGGGACCGTATCCTCCAGGCATCAGGGATGAAAAG  
GGCCAAACCGTCCCCTACTTCAGCTCAAACACCGGATCAGGCGTCTTTGCATGCATTCGCCGAT  
TCGCTGGAGCGTGACCTTGATGCGCCCAGCCCAATGCACGAGGGAGATCAGACGCGGGCAAG  
CAGCCGTAAACGGTCCCGATCGGATCGTGCTGTCACCGGCCCTCCGCACAGCAGGCTGTCGA  
GGTGCGCGTTCCCGAACAGCGCGATGCGCTGCATTTGCCCTCAGCTGGAGGGTAAAACGCC  
GCGTACCAGGATCTGGGGCGGCCTCCCGATCCTGGTACGCCATGGCTGCCGACCTGGCAGC  
GTCCAGCACCGTGATGTGGGAACAAGATGCGGACCCCTTCGCAGGGGCAGCGGATGATTTCC  
GGCATTCAACGAAGAGGAACTCGCATGGTTGATGGAGCTATTGCCTCAGAAGGACGCATTGGA  
CGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTGACCTTGACATGCTTGGTT  
CGGATGCCCTTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTGACCT  
GGACATGCTGATTA ACTCTAGAAGTTCCGGATCTCCGAAAAGAAAACGCAAAGTTGGTAGCCA  
GTACCTGCCGACACCGACGACCGGCACCGGATCGAGGAAAAGCGGAAGCGGACCTACGAGA  
CATTCAAGAGCATCATGAAGAAGTCCCCCTCAGCGGCCCCACCGACCTAGACCTCCACCTA  
GAAGAATCGCCGTGCCAGCAGATCCAGCGCCAGCGTGCCAAAACCTGCCCCCAGCCTTACC  
CCTTACCAGCAGCCTGAGCACCATCAACTACGACGAGTTCCCTACCATGGTGTTCCCCAGCG  
GCCAGATCTCTCAGGCCTCTGCTCTGGCTCCAGCCCCTCCTCAGGTGCTGCCTCAGGCTCCTGC  
TCCTGCACCAGCTCCAGCCATGGTGTCTGCACTGGCTCAGGCACCAGCACCCGTGCCTGTGCT  
GGCTCCTGGACCTCCACAGGCTGTGGCTCCACCAGCCCCTAACCTACACAGGCCGGCGAGG  
GCACACTGTCTGAAGCTCTGCTGCAGCTGCAGTTCGACGACGAGGATCTGGGAGCCCTGCTGG  
GAAACAGCACCGATCCTGCCGTGTTACCGACCTGGCCAGCGTGGACAACAGCGAGTTCCAG  
CAGCTGCTGAACCAGGGCATCCCTGTGGCCCCTCACACCACCGAGCCATGCTGATGGAATAC  
CCCGAGGCCATCACCCGGCTCGTGACAGGCGCTCAGAGGCCTCCTGATCCAGCTCCTGCCCT  
CTGGGAGCACAGGCCTGCCTAATGGACTGCTGTCTGGCGACGAGGACTTCAGCTCTATCGCC  
GATATGGATTTCTCAGCCTTGCTGGGCTCTGGCAGCGGCAGCCGGGATTCCAGGGAAGGGATG  
TTTTTGCCGAAGCCTGAGGCCGGCTCCGCTATTAGTGACGTGTTGAGGGCCGCGAGGTGTGC  
CAGCCAAAACGAATCCGGCCATTTATCCTCCAGGAAGTCCATGGGCCAACCGCCCACTCCCC  
GCCAGCCTCGCACCAACACCAACCGGTCCAGTACATGAGCCAGTCGGGTCACTGACCCCGGC  
ACCAGTCCCTCAGCCACTGGATCCAGCGCCCGCAGTGACTCCCGAGGCCAGTCACCTGTTGGA  
GGATCCCGATGAGGAGACTAGCCAGGCTGTCAAAGCCCTTCGGGAGATGGCCGATACTGTGAT  
TCCCAGAAGGAAGAGGCTGCAATCTGTGGCCAAATGGACCTTTCCATCCGCCCCCAAGGG  
CCATCTGGATGAGCTGACAACCACACTTGAGTCCATGACCGAGGATCTGAACCTGGACTCACC  
CCTGACCCCGGAATTGAACGAGATTCTGGATACTTCTGAACGACGAGTGCCTCTTGATGCC  
ATGCATATCAGCACAGGACTGTCCATCTTCGACACATCTCTGTTTGCGGCCGCGACTCTAGATC

ATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCT  
 GAACCTGAAACATAA

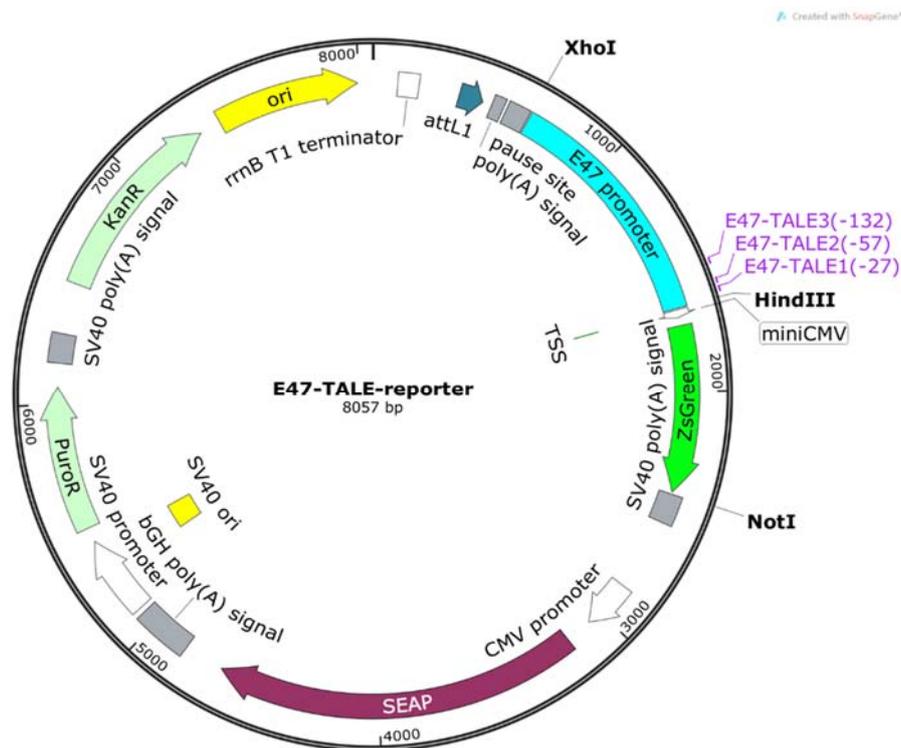


HNF4α promoter+miniCMV+ZsGreen

TALE-binding sites are shown in highlight: TALE1, TALE2, and TALE3

TGAGATCCAAAAGTGGACAAAAGAAACGGGGCTGTTCCAAAAAAGCTAGGTGGCAGG  
 TGTCTAACATGCCAGGGAGCTAAAACAGAGTGTGTGAGTTTCAGCAGCAGGTTGAATTTAGA  
 ATGGGGAAGGAGACCAGAGGAGACGCCAGACAGGATGACTTTGTCCCATTGGCCTGGAGGC  
 AGCCCCATGTTTCTCCACCCCTCATATCACTCACCAGTTTGTAAATAGTATCTTTGAATGACGAT  
 CTGATTAAGGTCCGTCTCCTCCATTAGTCCACAAGTTTCGGGGGTACATCTACTTTGCTCATT  
 CCATATCCCAGAGTCTAGCACAAGGCCTGGTACATAGTAGGTGCTCAATAAATATGTTAGAT  
 GAAAGGAAGATAACACCTCTATGTACTAGCAGTGAGACTCCAGGCATGCAATTTCTCTCCTC  
 CTTCAGTCCCTTCATCTCAAGGTTTAATTTAAATATGGTAACGCCTGTATGCAACTCCAGCAT  
 CCAGTAGGCACTCACTAAACACAGTTCTCCACCCTCCTTTTTCTCTGCCCTCCCTCGGTTT  
 TCCCCTACTTCTGTCATGGTGACACACCCATAGTTTGGAGCCATAAAACCAACCCAGGTTG  
 GACTCTCACCTCTCAGCCCTTCTGCTCCGGCCCTGTCTCAAATTTGGGGGGCTGATGTCCCC  
 ATACACCTGGCTCTGGGTTCCCCTAACCCAGAGTGCAGGACTAGGACCCGAGTGGACCTCA  
 GGTCTGGCCAGGTCGCCATTGCCATGGAGACAGCAACAGTCCCCAGCCGCGGGTCCCTAAG  
 TGACTGGTTACTCTTTAACGTATCCACCACCTTGGGTGATTAGAAGAATCAATAAGATAACC  
 GGGCGGTGGCAGCTGGCCGCACTCACCCTTCTGTTGGACGGGCTCCTGGTGGCTGTGCTG  
 CTGCTGTGAGCGGGCCCCTGCTCCTCCAAGCCCCAGCTCTCCGGCTGGGTGGGCTTTAGAGG  
 GTATATAATGGAAGCTCGACTTCCAGATGGCCAGTCCAAGCACGGCCTGACCAAGGAGATG  
 ACCATGAAGTACCGCATGGAGGGCTGCGTGGACGGCCACAAGTTCGTGATCACCGGCGAGGG  
 CATCGGCTACCCCTTCAAGGGCAAGCAGGCCATCAACCTGTGCGTGGTGGAGGGCGGCCCT  
 TGCCCTTCGCCGAGGACATCTTGTCCGCCCTTCATGTACGGCAACCGCGTGTTCACCGAGT  
 ACCCCAGGACATCGTCGACTACTTCAAGAACTCTGCCCCGCCGGCTACACCTGGGACCGCT

CCTTCCTGTTCGAGGACGGCGCCGTGTGCATCTGCAACGCCGACATCACCGTGAGCGTGGAG  
 GAGAACTGCATGTACCACGAGTCCAAGTTCTACGGCGTGAACCTCCCCGCCGACGGCCCCGT  
 GATGAAGAAGATGACCGACAACCTGGGAGCCCTCCTGCGAGAAGATCATCCCCGTGCCCAAGC  
 AGGGCATCTTGAAGGGCGACGTGAGCATGTACCTGCTGCTGAAGGACGGTGGCCGCTTGCGC  
 TGCCAGTTCGACACCGTGTACAAGGCCAAGTCCGTGCCCGCAAGATGCCCGACTGGCACTT  
 CATCCAGCACAAGCTGACCCGCGAGGACCGCAGCGACGCCAAGAACCAGAAGTGGCACCTG  
 ACCGAGCACGCCATCGCCTCCGGTCCGCCTTGCCTAA



E47 promoter + miniCMV promoter + ZsGreen

TALE-binding sites are shown in highlight: TALE1, TALE2, and TALE3

GCTCAGTAGCCACCAACCACCTCTGCCAGTGCCTGTGCTGGCCCCAAGAGCCCTGGGCATA  
 CATCTCCTCACCCCTCCAGCCGACCAGCTCCCCACAGGCCTGGGTTTCGGGGTTAGAGTCAGGT  
 TGACTTCCCAGGCCAAAGGCGTGGTCAGGCTCTGCTCCCAACAGTATCACTCACTCCCCCTCT  
 TCCCAGGGTGCACCTTCTCTGCCCTGAAACCGCCCCCAGTTTATTACAACATCAGGCCAC  
 CCAGCATACTTCCCGCTCCAGCACAGCAGCAGACGTTACCAACAGGCCCCAGCCCCACAGTC  
 CAGGCAGTGCCAACCAGTCAGAGCTGGCGTACTCACAGACCCCTCTTCTGTTACTTTAGTG  
 AAGGACGAATGTTGCTCTGGGACCAAGGGAAGCTCAGCACTGGAAGATTGATTCCAGCTGGG  
 GAACAAGGGAGAAGGATGGGATGGGGTGTGGAGCTGGGGTTTTGATGGGTCGATAGGAGTT  
 TTCTGGATTAAGGAGGAAAAAGGTTTAGTAGGCAGAGCCAGATCCAAGTCTCAAGTACAAA  
 GGTCAGAAGTTCAGGGAGGCGGGAGGTGGGGGCTGGAAAAGGAGTAGGAGAGAGAGACCA  
 GGACATCATAAGAACCTCAAACCTGGCCAGGGGCTGGGATTTACCTAAAGGGGGGTGGCTTT  
 ACAGGGAAGGGAACGGGGTGGTAGATGTGCTACCAGGAGGGGACAGGAGGCCGCAAGGT  
 GCATGGGCTGATCGTGGTCCCTCCGTCTGACTGCACCCCCACCGCCCCACCCGCCCCGAGGT  
 GTACCCACCCAGCTCAGGTGAGGACTACGGCAGGGATGCCACCGCCACCCGTCGCCAAGA

CCCCAGCAGCACCTATCCCGCCCCCTTCTACGTGGCAGGTACATGGCAGGGCGGGGGC CG  
CCAGGGACGGTAGGGCAGGGCTGGGGT CCCCCTCGGGCCCCACCTTACCAGCTCCCCTCC  
ACAGATAGAGGGTATATAATGGAAGCTCGACTTCCAGATGGCCCAGTCCAAGCACGGCCTGA  
CCAAGGAGATGACCATGAAGTACCGCATGGAGGGCTGCGTGGACGGCCACAAGTTCGTGATC  
ACGGCGAGGGCATCGGCTACCCCTTCAAGGGCAAGCAGGCCATCAACCTGTGCGTGGTGGA  
GGGCGGCCCTTGCCCTTCGCCGAGGACATCTTGTCCGCCCTTCATGTACGGCAACCGCGT  
GTTACCGAGTACCCCAGGACATCGTCGACTACTTCAAGAACTCCTGCCCCGCCGGCTACAC  
CTGGGACCGCTCCTTCTGTTCGAGGACGGCGCCGTGTGCATCTGCAACGCCGACATCACCGT  
GAGCGTGGAGGAGAACTGCATGTACCACGAGTCCAAGTCTACGGCGTGA ACTTCCCCGCCG  
ACGGCCCCGTGATGAAGAAGATGACCGACA ACTGGGAGCCCTCCTGCGAGAAGATCATCCCC  
GTGCCAAGCAGGGCATCTTGAAGGGCGACGTGAGCATGTACCTGCTGCTGAAGGACGGTGG  
CCGCTTGCCTGCCAGTTCGACACCGTGTACAAGGCCAAGTCCGTGCCCCGCAAGATGCCCG  
ACTGGCACTTCATCCAGCACAAAGCTGACCCGCGAGGACCGCAGCGACGCCAAGAACCAGAA  
GTGGCACCTGACCGAGCACGCCATCGCCTCCGGCTCCGCCTTGCCCTAA