Molecular Therapy
Nucleic Acids
Original Article

One-Step *piggyBac* Transposon-Based CRISPR/Cas9 Activation of Multiple Genes

Shenglan Li,[1,2] Anqi Zhang,[1,2,3] Haipeng Xue,[1,2] Dali Li,[1,2] and Ying Liu[1,2]

[1]The Vivian L. Smith Department of Neurosurgery, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [2]Center for Stem Cell and Regenerative Medicine, The Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [3]Central Laboratory, Liaocheng People's Hospital, 252000 Shandong, China

Neural cell fate is determined by a tightly controlled transcription regulatory network during development. The ability to manipulate the expression of multiple transcription factors simultaneously is required to delineate the complex picture of neural cell development. Because of the limited carrying capacity of the commonly used viral vectors, such as lentiviral or retroviral vectors, it is often challenging to perform perturbation experiments on multiple transcription factors. Here we have developed a *piggyBac* (PB) transposon-based CRISPR activation (CRISPRa) all-in-one system, which allows for simultaneous and stable endogenous transactivation of multiple transcription factors and long non-coding RNAs. As a proof of principle, we showed that the PB-CRISPRa system could accelerate the differentiation of human induced pluripotent stem cells into neurons and astrocytes by triggering endogenous expression of different sets of transcription factors. The PB-CRISPRa system has the potential to become a convenient and robust tool in neuroscience, which can meet the needs of a variety of in vitro and in vivo gain-of-function applications.

INTRODUCTION

Neural cell fate determination is orchestrated by a tightly controlled transcription regulatory network. In the network, master regulatory genes or hub genes act in a predetermined sequence to govern proliferation and differentiation toward different lineages. The ability to manipulate expression of these genes simultaneously is a prerequisite for identifying important gene function and further delineating the complex picture of neural development.

Gene activation and suppression have been achieved by a variety of methods. Single gene perturbation is usually straightforward and relatively easy to carry out, while switching on or off the expression of several genes concurrently often requires multiple rounds of manipulation, which is frequently time consuming, labor intensive, and not always as efficient as one would anticipate. The CRISPR/Cas9 system has been developed and widely used as a tool for gene manipulation recently.[1–7] In the system, Crispr RNA (crRNA) and *trans*-activating crRNA (tracrRNA) work together to bind to Cas9, a CRISPR-associated RNA-guided type II endonuclease. crRNA and tracrRNA (collectively called single-guide RNA, or sgRNA) then bring Cas9 endonuclease to the targeting site in the genome. According to crystallography, Cas9 contains two lobes, a target recognition lobe and a nuclease lobe. The target recognition lobe is responsible for binding sgRNA and target DNA by recognizing the protospacer adjacent motif (PAM) sequence, while the nuclease lobe executes cleavage function on the target DNA.[8,9] Taking advantage of the properties of recognition and binding of Cas9, a nuclease-deficient Cas9 (dCas9) was developed to mute its catalytic activity and maintain only the binding activity. dCas9 has been subsequently fused with effector domains such as transcription activators, allowing manipulation of gene expression at the transcription level.[10–15]

The initially designed system with dCas9-VP64-transactivating domain often requires multiple sgRNAs to activate just one gene. To enhance activation efficiency, several versions of CRISPRa platforms have been developed by recruiting various copies of VP16 or by hybridizing multiple transcription activators to dCas9. Examples include the repeating peptide array SunTag system;[15] the three-component synergistic activation mediator (SAM) system;[16] and the tripartite activator VP64, p65, and Rta (VPR) fusion protein system.[17] Although these systems showed higher activation efficiency, they often would require the concurrent delivery of several transactivation components and sgRNAs to the same cells to achieve multiplex gene activation. A single all-in-one vector will expand potential applications of the transactivation platform, which will facilitate the evaluation of collaboration and interaction among multiple targeted genes.

One technical hurdle that hampers the construction of an all-in-one vector for stable activation is the limited payload capacity of the commonly used lentiviral or retroviral vectors, which hold a maximum cargo size of 7–8 kb, making it difficult to package multiple guide RNA-expressing fragments in one vector. To overcome this limitation, we decided to test whether the *piggyBac* (PB) transposon system, which could carry up to 100–200 kb of transgenes,[18] would be able to fulfill such requirement. The PB system features transposons that can cut and paste transgenes into the genome where a TTAA sequence is

Received 5 March 2017; accepted 10 June 2017;
http://dx.doi.org/10.1016/j.omtn.2017.06.007.

Correspondence: Ying Liu, Department of Neurosurgery, The University of Texas Health Science Center at Houston, 1825 Pressler Street, SRB 630G, Houston, TX 77030, USA.
E-mail: ying.liu@uth.tmc.edu

found. In addition, the PB system has been demonstrated as an efficient tool in various applications, including induced pluripotent stem cell (iPSC) reprogramming and genome-wide mutagenesis.[18–22]

In the current work, we constructed PB-based vectors that contained highly efficient CRISPRa SAM components, as well as modular cassettes that allow for simultaneous expression of multiple sgRNAs. This platform allowed us to test a total of 72 sgRNAs for 24 genes of interest and identify the optimal sgRNAs for gain-of-function experiments. In addition, we have provided proof-of-principle applications by plugging in different sets of sgRNAs to the all-in-one vectors, and we were able to accelerate neural cell differentiation and cell type conversion in human iPSCs.

RESULTS

Identification of the Optimal sgRNAs for Gene Activation in the PB-CRISPRa System

To identify the best sgRNAs in the PB-CRISPRa system for each gene, we first created a human HEK293FT stable cell line that overexpressed MS2-p65-HSF1 (MSPH) and dCas9-VP64. MSPH and dCas9-VP64 were cloned into the PB backbone under the control of EF1α and CAG promoter, respectively (Figure 1A, renamed as PB-SAM). The HEK293FT cell line was transfected with PB-SAM and PBase, followed by blasticidin and hygromycin selection. The stable cell line obtained after selection was then renamed as 293FT-SAM. In 293FT-SAM, compared to untransfected cells (background control), the expression of MSPH and dCas9-VP64 was elevated by more than 200-fold for at least 3 months (the longest time point tested) without silencing (Figure 1B; Figure S1), indicating that 293FT-SAM cells could be used to test activation efficiency of sgRNAs in our PB-CRISPRa system.

Next we sought to evaluate the efficiency of the PB-CRISPRa system for activating endogenous gene transcription using the 293FT-SAM cell line. We chose 21 transcription factors (TFs) and three long non-coding RNAs (lncRNAs) that have been reported to be involved in the process of neural development (Table S1). We synthesized three sgRNAs per gene and tested 72 sgRNAs in total. For each gene, all three sgRNAs were designed to target −200–+1 bp upstream of its transcription start site (TSS).[16] To quickly test the activation efficiency, we first co-transfected the 293FT-SAM cell line with all three sgRNAs. For 19 of 24 genes (79.2%), the mRNA expression level was increased by at least 2-fold (2- to 12,333-fold, mean fold increase was 763-fold; Figures 1C–1E and 1G), indicating that these sgRNAs could efficiently activate most genes of our interest respectively, although five genes (20.8%) failed to be activated in 293FT-SAM.

As the SAM system was reported to efficiently stimulate endogenous gene expression with only one sgRNA,[16] we next evaluated the activation efficiency of individual sgRNAs for each gene, aiming to identify the best sgRNA for endogenous activation for each gene on our list. Compared to basal transcript level, an average of 130-fold increase was observed upon transfection of single sgRNAs, as shown by qPCR results. Specifically, mRNA expression level was increased by at least 2-fold (2- to 3,846-fold) for 43 of the total 72 sgRNAs (60%) tested, of which two sgRNAs (ASCL1-gRNA1 and ASCL1-gRNA2) caused mRNA expression level to increase by more than 1,000-fold and nine sgRNAs caused the corresponding mRNA expression to increase by 100- to 1,000-fold. Additionally, 17 sgRNAs were able to activate the gene expression by 10- to 100-fold and 15 sgRNAs by 2- to 10-fold (Figures 1C–1F). The sgRNA that was able to maximally activate endogenous gene expression in 293FT-SAM was identified to be the best sgRNA for subsequent experiments using our PB-CRISPRa system.

The PB-CRISPRa System Activates TFs and lncRNAs in Human iPSCs

We asked whether the optimal sgRNAs obtained from our screening in 293FT-SAM were able to activate the endogenous expression of TFs or lncRNAs in human cell types other than 293FT. We decided to test human iPSCs, cells that would be able to give rise to neural lineage cells among a variety of cell types of all three germ layers. Following the success in 293FT-SAM cells, we created an iPSC-SAM line by transfecting Cy2 iPSCs (NIH) with PB-SAM and PBase plasmids, followed by hygromycin and blasticidin selection. Similar to 293FT-SAM, the expression of PB-SAM mRNAs in iPSC-SAM was about 260-fold and 150-fold higher than in untransfected Cy2 iPSCs (Figure 2A; Figure S2). Hence, the PB-SAM-transfected Cy2 iPSC line was renamed as iPSC-SAM. Next, the optimal sgRNA that elicited the most robust gene activation in 293FT-SAM was introduced to iPSC-SAM cells. Consistent with the results obtained from 293FT-SAM, all genes were activated by the optimal sgRNAs, as endogenous gene expression level was elevated by at least 2-fold (2- to 2,470-fold) (Figures 2B–2D). Moreover, for 17 of the 24 genes (70.8%), the expression level after sgRNA transfection was increased by at least 10-fold (Figure 2B). These data suggest that, using the PB-CRISPRa system, the sgRNAs identified from the 293FT-SAM screening allowed for robust endogenous gene activation in human iPSCs as well.

As we briefly mentioned already, four TFs (NFIA, NFIB, NFIX, and SOX9) and an lncRNA (HAR1A) failed to be activated in 293FT-SAM cells. These genes, however, were successfully activated in iPSC-SAM. As shown in Figures 2C and 2D, these genes achieved an increase in expression of 8- to 110-fold, as determined by qPCR. As basal transcription level has been reported to contribute to endogenous gene activation by CRISPRa,[16] we compared the basal expression level of all five genes, and we found that NFIA, NFIB, NFIX, and SOX9 had relatively higher basal transcription levels in 293FT than in iPSCs (Figure S3). To examine the potential correlation between the magnitude (fold increase in gene expression) of activation and basal transcription level in human iPSCs and 293FT cells, we performed a collective comparison of all 24 genes. As shown in Figures 2G, 2H, S3, and S4, iPSCs had a much higher fold increase in gene activation with relatively lower basal transcript level compared to 293FT cells. Linear regression analysis of the 24 genes showed a significant negative correlation between fold activation and basal transcript level in 293FT-SAM cells and human iPSC-SAM cells (Figures 2E and 2F; $r = 0.66$, $p = 0.0004$ in 293FT; $r = 0.76$, $p < 0.0001$ in iPSCs). A similar

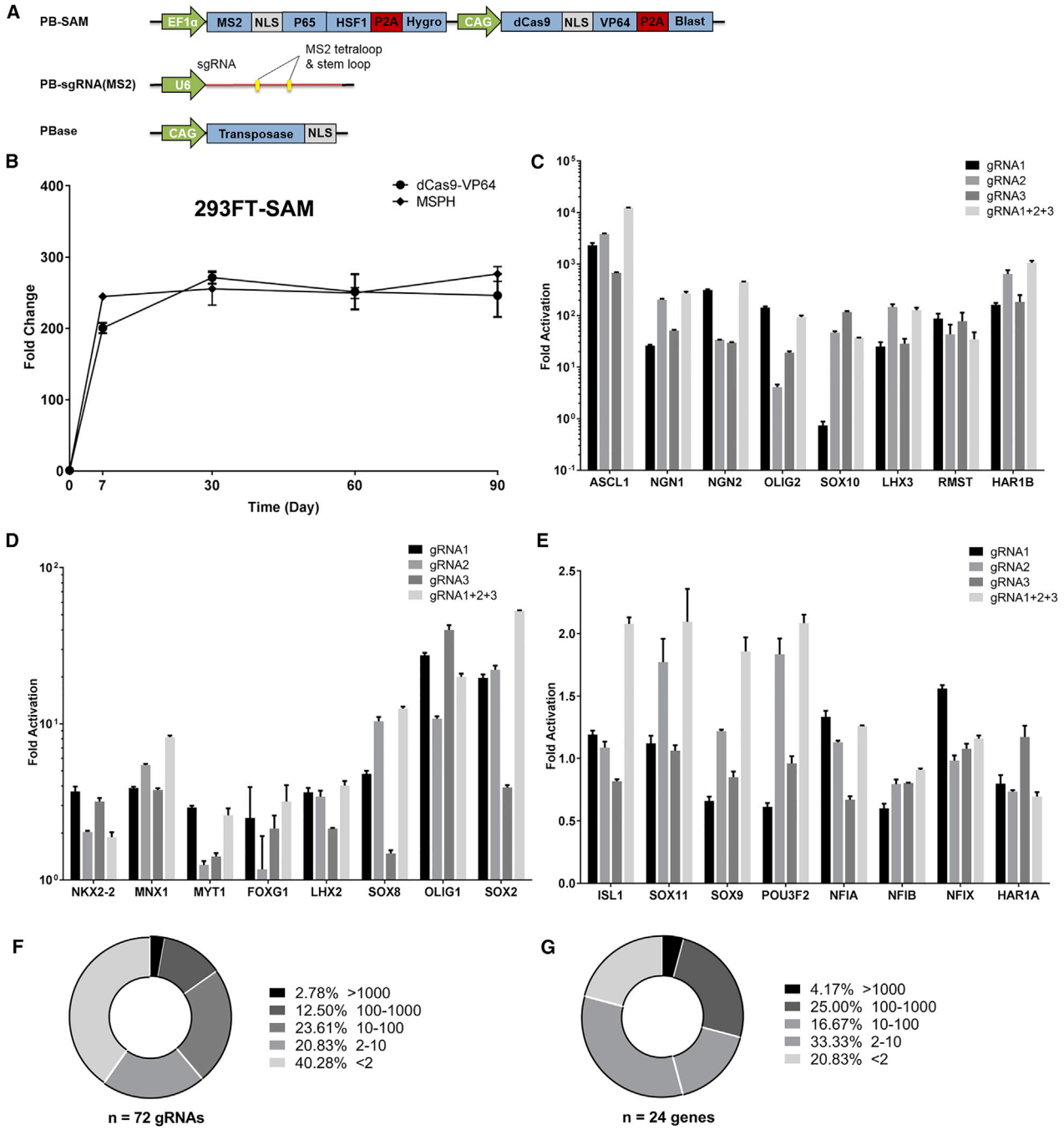
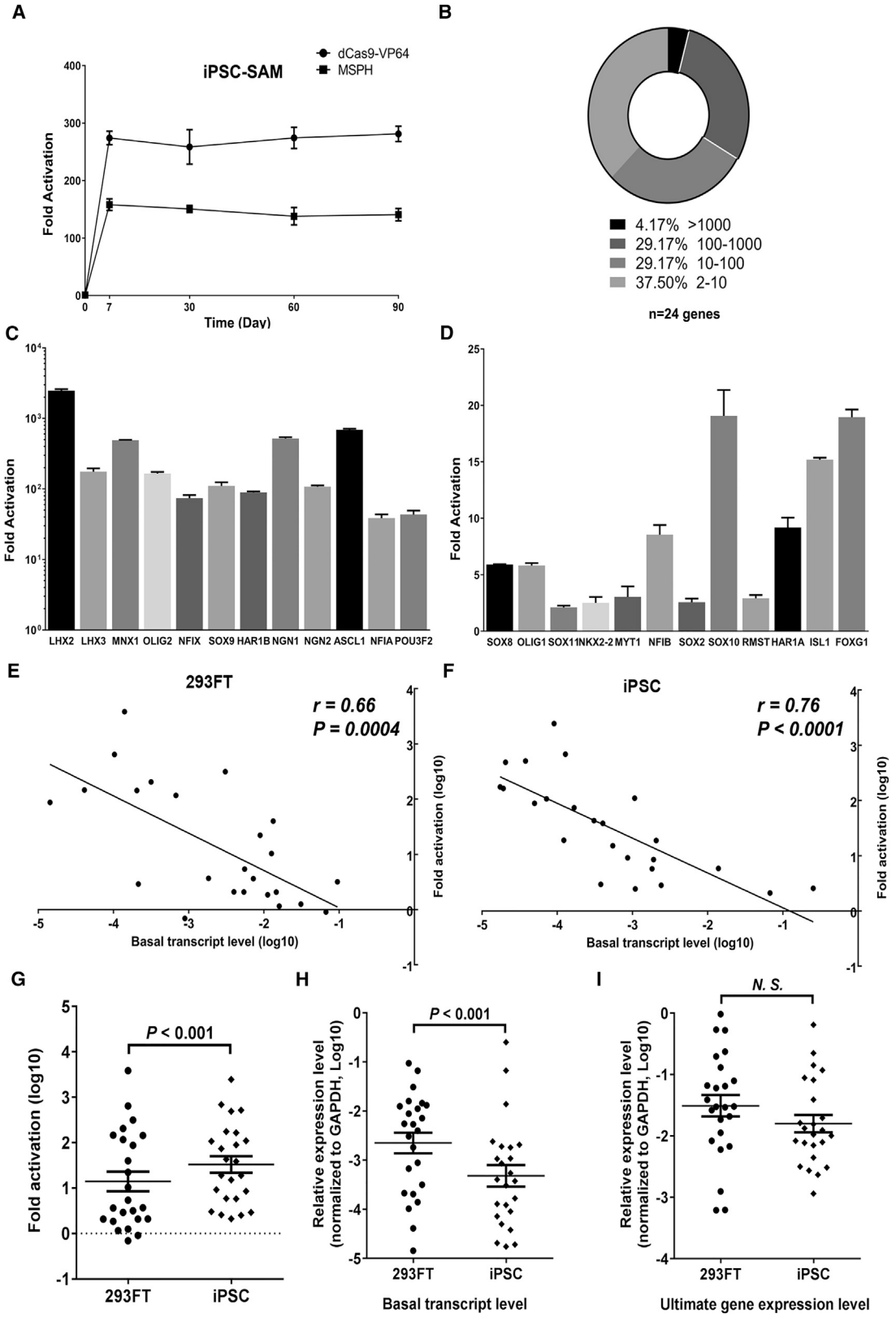


Figure 1. Identification of the Best sgRNAs in Gene Activation by 293FT-SAM

(A) PB vectors used for creating 293FT-SAM. (B) Generation of 293FT-SAM cell line by co-transfecting 293FT with PB-SAM and PBase expression vectors. qPCR results showed that SAM components MS2-p65-HSF1 (MSPH) and dCas9-VP64 were stably expressed for at least 90 days. (C–E) qPCR results of gene expression activation of individual TFs (ASCL1, NEUROG1, NEUROG2, OLIG2, SOX10, LHX3, NKX2-2, MNX1, MYT1, FOXG1, LHX2, SOX8, OLIG1, SOX2, ISL1, SOX11, SOX9, POU3F2, NFIA, NFIB, and NFIX) and lncRNAs (RMST, HAR1B, and HAR1A) after transfecting 293FT-SAM cell line with corresponding sgRNA vectors. A total of three sgRNAs per gene was tested separately to evaluate their activation efficiency. The rightmost bars for each gene represent the extent of gene activation when all three sgRNA vectors for that gene were co-transfected. (F) A pie chart showing the range of fold change of gene expression after each sgRNA vector was transfected into the 293FT-SAM line. For each gene, three sgRNAs were tested separately. A total of 72 sgRNAs was tested. Basically, 43 sgRNAs were able to augment target gene expression by >2-fold. (G) For the 24 genes targeted, 19 genes were activated to be expressed at >2-fold. The qPCR results were normalized to GAPDH mRNA level. P2A, self-cleaving peptide P2A sequence; Hygro, hygromycin resistance cassette; Blast, blasticidin resistance cassette; NLS, nuclear localization signal. In (B)–(E), data are presented as mean ± SEM (n = 3).



(legend on next page)

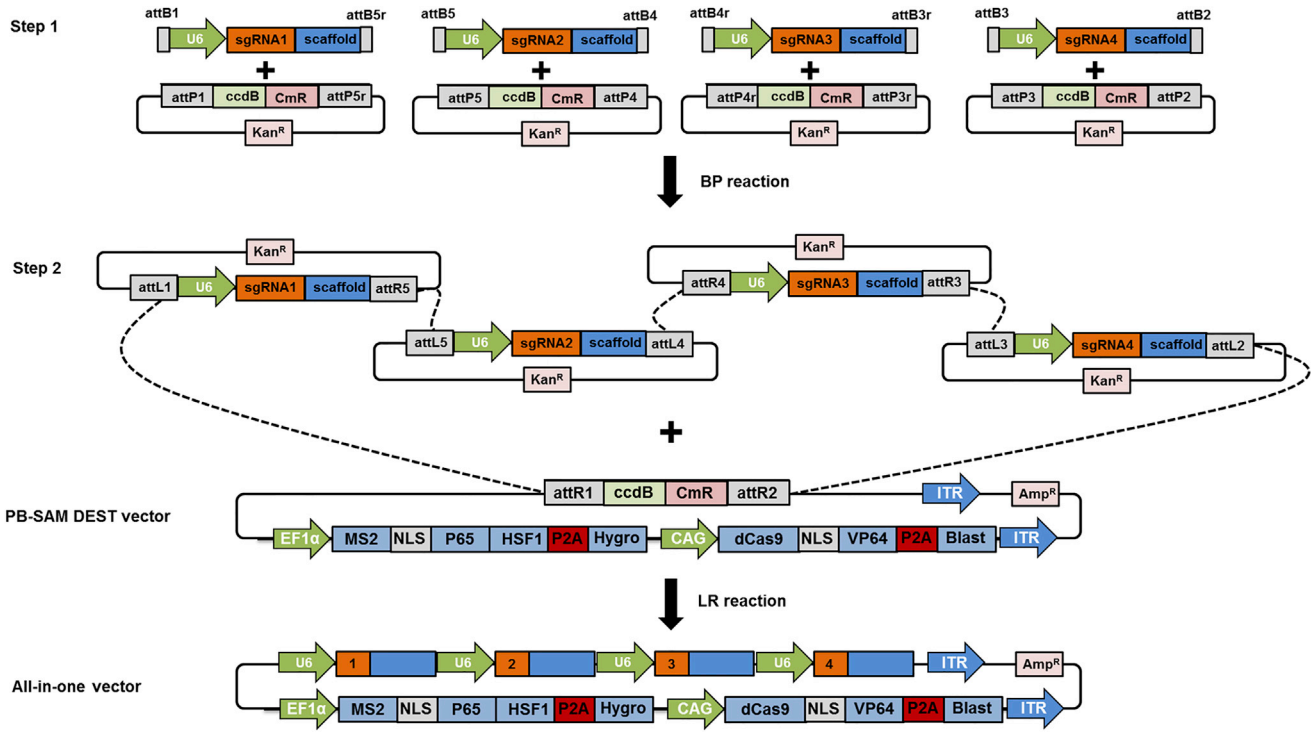


Figure 3. A Schematic Representation of Constructing PB-CRISPRa All-in-One Vectors with Multisite Gateway Cloning Strategy

The PB-SAM vector was amplified to become PB-SAM R1-R2 DEST vector containing attR1 and attR2 sites, a ccdB cassette, and a chloramphenicol resistance cassette (CmR). PB-sgRNA (MS2) vectors containing different sgRNA inserts were amplified to be attached to appropriate attB sites. All-in-one vectors were assembled by four-way LR reactions. P2A, self-cleaving peptide P2A sequence; Hygro, hygromycin resistance cassette; Blast, blasticidin resistance cassette; Amp^R, ampicillin resistance cassette; Kan^R, kanamycin resistance cassette; NLS, nuclear localization signal; ITR, inverted terminal repeat sequence.

correlation was also observed when data were pooled for all 24 genes in both cell types (Figure S5; $r = 0.73$, $p < 0.0001$). Collectively, these data indicate that relatively lower basal transcript levels of the genes tested in iPSCs is one of the major factors that might contribute to higher activation magnitude induced by optimal sgRNAs. Interestingly, although iPSCs achieved a greater fold increase (i.e., ratio of gene expression level of post- versus pre-transactivation), there was no significant difference in the ultimate gene expression level between 293FT cells and iPSCs (Figure 2I), indicating that endogenous gene expression probably has an upper limit.

All-in-One Vectors for Stable Activation of Multiple Genes in Single Cells

To strengthen the application of the PB-CRISPRa system, we attempted to incorporate all components into one PB vector to enable simultaneous activation of multiple genes in single cells. As shown in Figure 3, the all-in-one PB vector contains EF1α promoter-driven MSPH, connected by CAG promoter-driven dCas9-VP64 and followed by multiple U6 promoter-driven sgRNAs. Additional U6-sgRNA expression cassettes could be incorporated by Multisite Gateway cloning as desired.

We first tested whether the all-in-one PB-CRISPRa vector was able to express all SAM components and stably activate endogenous expression of a single gene. To facilitate the test, we used an OLIG2-EGFP human iPSC knockin reporter, in which an EGFP cassette was knocked into one of the two alleles of OLIG2 by gene targeting and has been shown to faithfully reflect the endogenous expression of OLIG2[23–25] (S.L. and Y.L., unpublished data). The all-in-one vector containing OLIG2-sgRNA1 was transfected to the OLIG2-EGFP

Figure 2. The Optimal sgRNAs Determined by 293FT-SAM Are Able to Activate TFs and lncRNAs in Human iPSCs

(A) Generation of the human iPSC-SAM cell line by transfecting human iPSC line Cy2 with PB-SAM and PBase expression vectors. Similar to the 293FT-SAM line, the SAM components MSPH and dCas9-VP64 were stably expressed starting from day 7 post-transfection for at least 90 days. (B) A pie chart showing the gene activation magnitude for all 24 genes. All optimal sgRNAs were able to augment target gene expression by >2-fold. (C and D) Activation of individual TFs and lncRNAs in human iPSC-SAM line by the optimal sgRNAs that were identified from 293FT-SAM. (E and F) A significant negative correlation was found between fold activation (log10) and basal transcript level (log10, normalized to GAPDH) in both 293FT-SAM (E) and iPSC-SAM (F). (G and H) Compared to 293FT, most genes tested had a relatively lower basal transcript level (log10, normalized to GAPDH, H) in human iPSCs while a much higher fold increase (log10, G) when activated in iPSC-SAM. (I) No significant difference between 293FT-SAM and iPSC-SAM in ultimate gene expression levels was found for our genes of interest upon CRISPR activation. In (C), (D), and (G)–(I), data are presented as mean ± SEM (n = 3).

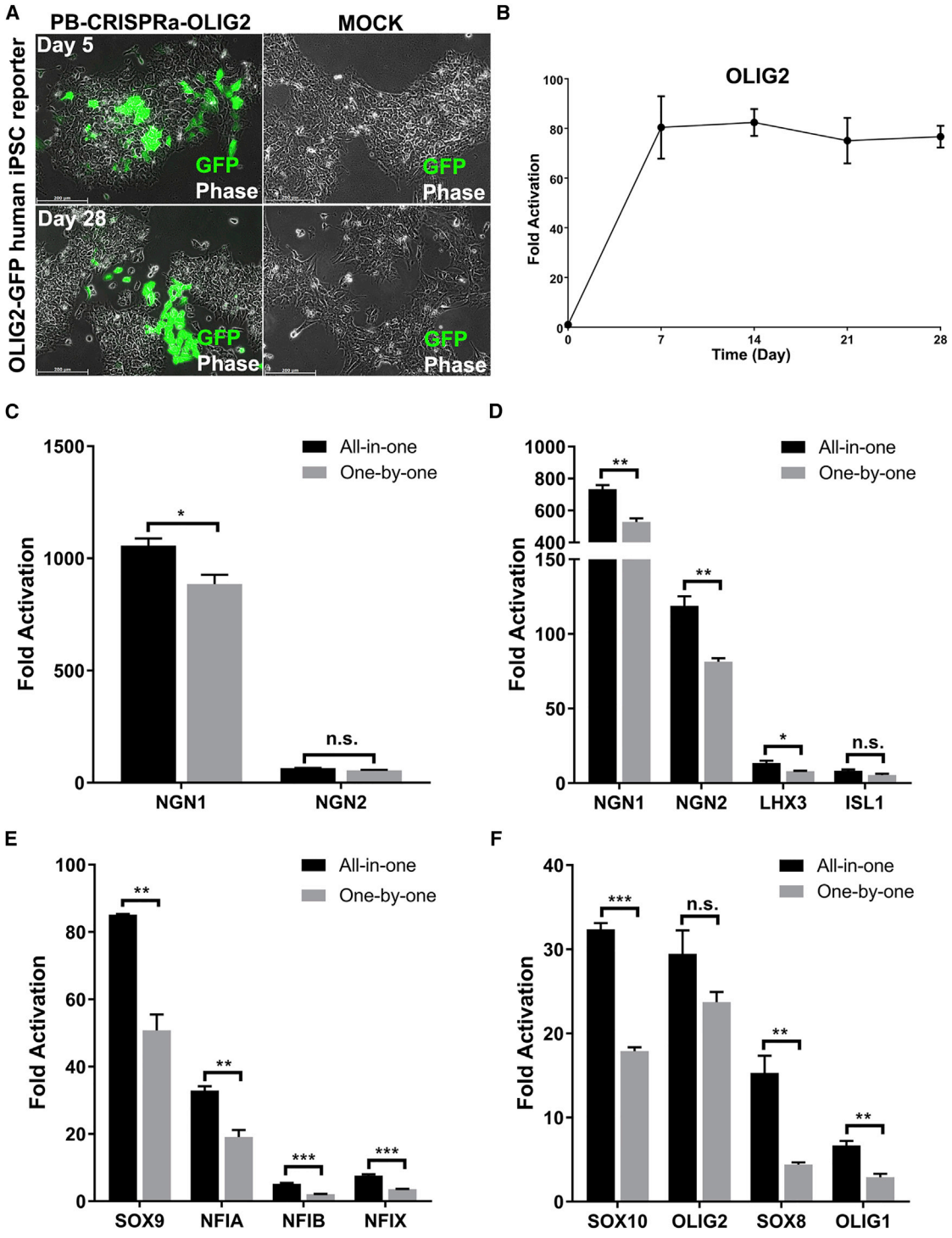


Figure 4. Stable Activation of Multiple Genes with All-in-One PB-CRISPRa Vectors

(A and B) An all-in-one PB-CRISPRa vector containing OLIG2-sgRNA1 was able to continuously activate EGFP (surrogate marker for endogenous OLIG2) expression in OLIG2-EGFP human iPSC knockin reporter cell line, as indicated by the robust, native GFP signal from day 5 and day 28 post-transfection (A). The elevation of endogenous OLIG2 expression was validated by qPCR, which showed a 70- to 80-fold increase in OLIG2 gene expression starting from day 7 post-transfection that was maintained for at least 28 days (B). (C–F) Comparison of gene activation efficiency between all-in-one vectors and multiple single sgRNA vectors (one-by-one). The qPCR results showed

(legend continued on next page)

reporter iPSC line, and EGFP expression was observed as early as day 1 after transfection. Starting from day 7 post-transfection, OLIG2 (EGFP) was activated in 35%–40% cells, and the expression level was elevated by about 80-fold as measured by qPCR. The activation was maintained for at least 4 weeks after transfection, indicating that the PB-CRISPRa system was able to stably activate expression of single genes in human iPSCs (Figures 4A and 4B; Figure S6).

To evaluate the efficiency of the all-in-one PB-CRISPRa vectors for concurrently activating multiple neural lineage-specific TFs, we generated a two-in-one NEUROG1 and NEUROG2 vector; a four-in-one NEUROG1, NEUROG2, LHX3, and ISL1 vector; a four-in-one SOX10, OLIG2, SOX8, and OLIG1 vector; and a four-in-one SOX9, NFIA, NFIB, and NFIX vector. To compare activation efficiency, 293FT cells were transfected with all-in-one PB CRISPRa vectors or co-transfected with two or four sgRNA vectors designed to activate single TFs (designated as one by one). Two days after transfection, expression of all genes was elevated. In addition, cells transfected with the all-in-one plasmids, especially the four-in-one vectors, had significantly higher activation efficiency than the one-by-one groups, with an upregulation of 5- to 1,050-fold increase in gene expression (Figures 4C–4F). Thus, with the all-in-one strategy, we were able to activate multiple TFs in the same cells more efficiently than with the separate one-by-one system.

Rapid Generation of Neurons and Astrocytes from Human iPSCs via Endogenous Activation with All-in-One PB-CRISPRa Vectors

NEUROG1 and NEUROG2 play important roles in neuronal development and cell fate determination.[26–40] Exogenous overexpression of NEUROG2 alone or in combination with other TFs was able to direct or skew the fate of pluripotent stem cells, even mature somatic cells, toward neuronal lineages.[41,42] To test whether the all-in-one system was able to convert human iPSCs to neurons by endogenous activation of NEUROG1 and NEUROG2, we transfected two human iPSC knockin reporter lines NEUROG2-mCherry and Doublecortin (DCX)-ZsGreen with an all-in-one PB-SAM-NEUROG1/2 gRNA vector and PBase. At 24 hr post-transfection, cells were treated with blasticidin and hygromycin B. Then 3 days after transfection, 30%–40% of cells started to show a bipolar morphology and expressed NEUROG2 (mCherry, red) or DCX (ZsGreen, green), as observed under the fluorescence microscope (Figures 5A–5F). Within an additional 4 days, the morphology of the transfected cells highly resembled that of more mature neurons with longer neurites and extensive arborization. More than 80% of the transfected cells were confirmed by positive immunocytochemistry staining of NeuN, NF160, and MAP2, markers of mature neurons (Figures 5G–5O).

To expand the potential applications of the PB-CRISPRa all-in-one vectors, we also attempted the induction of human iPSCs toward the astrocyte lineage. iPSCs were transfected with an all-in-one PB-SAM-SOX9-NFIA-NFIB-NFIX gRNA vector and PBase. Treatment with blasticidin and hygromycin B was carried out 24 hr post-transfection to select cells that were successfully transfected with the all-in-one vector. At 14 days post-transfection, the majority of cells expressed multiple well-accepted astrocyte markers, CD44 (80%), S100B (95%), and GFAP (65%), indicating that the transfected cells were committed to the astrocyte lineage. Consistently, these cells did not express molecules of other lineages, such as A2B5, a glial progenitor marker, or β3 tubulin, a pan neuronal marker (Figure 6). In addition, concurrent activation of the four TFs seemed to be necessary for rapid induction toward the astrocyte lineage, as activation of only one TF (SOX9, NFIA, NFIB, or NFIX) failed to induce GFAP immunoreactivity from human iPSCs during the same induction period (data not shown).

DISCUSSION

In this work, we developed a PB-based CRISPRa system that allows for rapid sgRNA screening and stable activation of multiple TFs and lncRNAs. We tested the system in human iPSC lines and 293FT cells, and we compared the upregulation magnitude of gene expression level and basal transcript level, across genes and cell types. In addition, we showed that the PB-CRISPRa system could accelerate human iPSC differentiation into neurons and astrocytes by simultaneously augmenting endogenous expression of specific sets of TFs.

Two excellent examples of CRISPR-mediated gene activation systems, including the SAM[16] and the VPR fusion protein systems,[17] were reported to be able to regulate endogenous gene expression robustly. The SAM system was predicted to activate the endogenous expression of any given gene with only one sgRNA, as long as it targeted a PAM within 200 bp upstream of a TSS.[16] To simplify vector design and reduce experimental variability, we decided to adapt the SAM platform to our multiplex transactivation system. We packaged all SAM system activation components into a PB vector, and we generated SAM overexpression stable cell lines for sgRNA screening and verification. Of the sgRNAs tested here, 60% (43 of 72) were able to increase mRNA expression by at least 2-fold in 293FT-SAM cells. When tested in human iPSCs, all of the optimal sgRNAs identified from 293FT-SAM achieved at least 2-fold activation as well (Figure 2). Thus, our strategy offers a convenient strategy for quick optimization of best sgRNAs for the activation of multiple endogenous genes, which provides additional feasibility and flexibility for gain-of-function experiments.

Several elegant tools have been developed to achieve multiplex sgRNA expression based on RNA processing mechanisms, including Csy4-cleavable cassettes,[43,44] tRNA-sgRNA cassettes,[45] and sgRNA-small hairpin RNA (shRNA) structure.[46] Endonuclease Csy4 from

significant increases in gene expression in human iPSC line Cy2 after being transfected with all-in-one vectors compared to those that were co-transfected with multiple single sgRNA vectors. The all-in-one vectors used here are as follows: a two-in-one NEUROG1 and NEUROG2 vector (C); a four-in-one NEUROG1, NEUROG2, LHX3, and ISL1 vector (D); a four-in-one SOX9, NFIA, NFIB, and NFIX vector (E); and a four-in-one SOX10, OLIG2, SOX8, and OLIG1 vector (F). In (B)–(F), all data are presented as mean ± SEM (n = 3; *$p \leq 0.05$, Student's t test).

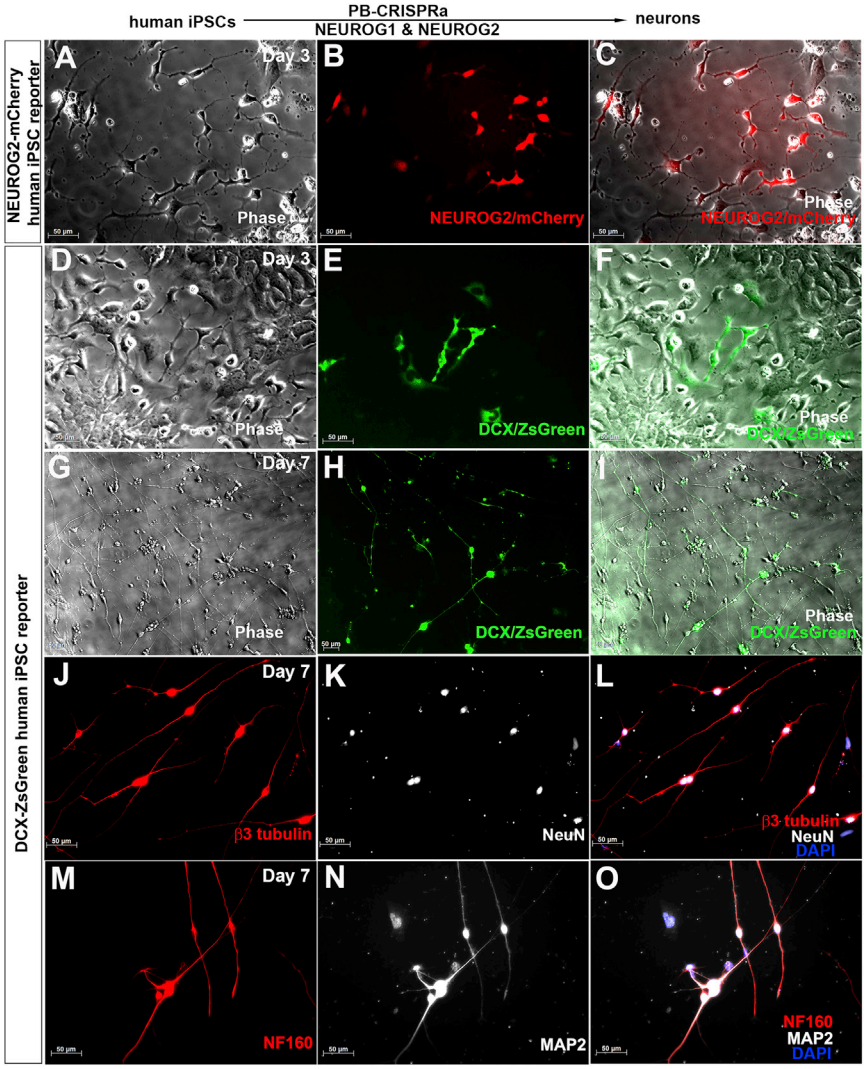


Figure 5. Rapid Generation of Neurons from Human iPSCs by Endogenous Activation of Neuronal Lineage-Specific TFs with the PB-CRISPRa System

(A–C) Experiments were performed in NEUROG2-mCherry knockin human iPSC reporter line. Three days after transfection with PB-SAM-NEUROG1/2 all-in-one vector, iPSCs underwent morphological changes and showed a bipolar morphology, typical of neural progenitors. Importantly, these cells also expressed mCherry (red), the surrogate marker for NEUROG2. (D–F) Experiments were performed in Doublecortin (DCX)-ZsGreen human iPSC reporter line. Three days after transfection with PB-SAM-NEUROG1/2 all-in-one vector, iPSCs started to express ZsGreen (green), the surrogate marker for DCX, indicating that the activation of NEUROG1 and NEUROG2 had promoted the iPSCs to differentiate into DCX-expressing young neurons. (G–I) Seven days after transfection, the DCX-ZsGreen reporter cells presented longer neurites and extensive arborization, typical of more mature neurons. (J–O) Immunocytochemistry revealed that these cells also expressed pan neuronal marker β3 tubulin (J) and mature neuronal markers NeuN (K), NF160 (M), and MAP2 (N). (L) and (O) are overlapped images with DAPI. Scale bar, 50 μm.

Pseudomonas aeruginosa recognizes a 28-nt RNA sequence and cleaves the target RNA. For multiplex sgRNA expression, sgRNAs are fused with the 28-nt Csy4-cleavable RNA sequence, and the transcript is processed into multiple sgRNAs with the addition of Csy4 protein. One limitation of this strategy is that the endonuclease Csy4 might be toxic to cells. The tRNA-sgRNA cassette approach takes advantage of the endogenous tRNA-processing system. Multiple tRNA-sgRNA units are assembled in a synthetic tRNA-sgRNA cassette, with each sgRNA containing a target-specific spacer, which can be cleaved by endogenous RNase to release mature sgRNAs and tRNA.[45] The sgRNA-shRNA structure hijacks the endogenous shRNA processing system, and multiple sgRNAs and shRNAs are assembled in an sgRNA-shRNA structure format with interval sequences of the Drosha cutting site.[46] Thus, the sgRNA-shRNA transcript is processed into functional sgRNAs and shRNAs by endogenous Drosha. These approaches, together with the present work that uses multiple Pol III U6 promoters in a single vector, provide a wide range of selections for simultaneous sgRNA expression for the purpose of CRISPR-mediated endogenous gene activation.

During development, master TFs dictate the gene expression regulatory network to drive cell fate determination, which provides a basis for direct conversion experiments across cell types, even different germ layers.[47] Recently, several groups reported successful conversion of fibroblasts to neurons, myocytes, or iPSCs by endogenous activation of TFs using the CRISPRa systems.[48–50] The CRISPRa system also allows for fast differentiation of pluripotent stem cells to endodermal lineages, trophoblast stem cells, extraembryonic endoderm cells, and neurons.[17,50,51] However, there are at least two drawbacks in these systems: (1) most experiments required multiple sgRNAs to activate a single TF, and (2) dCas9-activator fragments were delivered separately. These limitations might decrease conversion efficiency. An all-in-one vector could help overcome the technical barriers and boost reprogramming efficiency.[52] Our PB-CRISPRa system is able to stably activate multiple TFs by expressing corresponding optimal sgRNAs for multiple genes at the same time. Our data indicated that the all-in-one vectors achieved higher activation magnitude than the one-by-one vectors that relied on co-transfection of activators and multiple sgRNA plasmids (Figure 4). Using this system, we were able to accelerate directed neural lineage differentiation from human iPSCs (Figures 5 and 6). Thus, our PB-CRISPRa system is a necessary addition to the collection of gene activation tools.

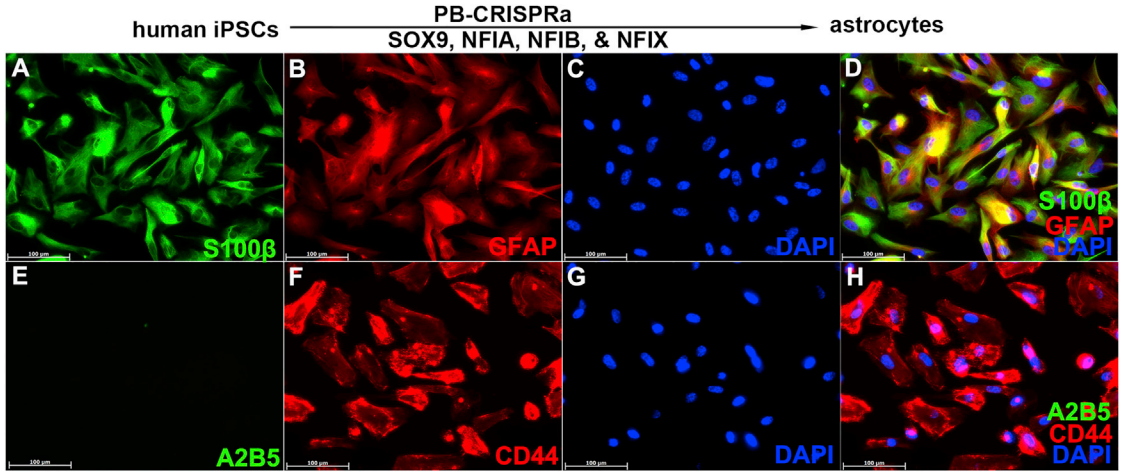


Figure 6. Rapid Generation of Astrocytes from Human iPSCs by Endogenous Activation of Astrocyte Lineage-Specific TFs with the PB-CRISPRa System
(A–H) Human iPSCs cell line Cy2 was transfected with all-in-one PB-SAM-SOX9-NFIA-NFIB-NFIX vector. At 14 days post-transfection, nearly all cells expressed astrocyte markers S100B (A), GFAP (B), and CD44 (F), while they did not express glial progenitor marker A2B5 (E). Nuclei are revealed by DAPI in (C) and (G). (D) and (H) are overlapped images. Scale bar, 100 μm.

Simultaneous perturbation of multiple TFs is a desired approach to unfold the collateral functions of TFs during development. Genetic manipulation of multiple genes in human cells has been more challenging to achieve than in animal models as it often requires multiplex targeting, which is difficult to perform even with multiple rounds of transfection. Indeed, technical limitations often restrict researchers going from single- or double-gene manipulation to multiple-gene perturbation. Our PB multiplex system has a large packaging capacity and is able to carry multiple components in a single plasmid. This is important as it brings extra flexibility in controlling the gene expression of groups of genes of interest.

One of the most appealing applications for the CRISPR system is the potential for gene therapy. Critical challenges for in vivo delivery of transgenes for gene therapy include limited cargo size and low efficiency. The PB transposon is a non-viral system with large cargo capacity. Previous reports showed that PB vectors were successfully delivered in vivo to the liver and the lung via mouse tail vein injections, and they maintained long-term gene expression.[7,18,53] A recent work on rapid in vivo CRISPR library screening of tumor suppressor genes using the PB system further confirmed the high efficiency and broadened potential applications of the cooperation of PB and CRISPR systems.[22] As a logical extension, our PB-CRISPRa system has the potential to be developed into a highly efficient gene delivery tool for in vivo gene activation in the future. When combined with previously validated reagents, such as nanoparticle, liposome, or electroporation, PB-CRISPRa also could activate genes in different target organs.

One limitation for PB vectors is that they are prone to integrate into transcription units of the genome. Integrated transgenes could function as oncogenes or may elicit immune responses in the target cells. Although re-expression of PBase could mediate the excision of integrated PB vectors, efficiency could be a challenge. Several procedures, such as inclusion of an inducible herpes simplex virus-1 thymidine kinase (hsv-TK) cassette in the PB vectors, allow for the selection of cells free of PB vector integration. In this case, cells with successful removal of PB vectors do not express TK-encoded thymidine kinase, and they will not convert Fialuridine (FIAU), the drug used for selection, to the toxic Fluorouracil (5-FU).[54,55] To further enhance the excision efficiency, a hyperactive PBase was developed that had an increased excising efficiency in the human iPSC reprogramming process.[56,57] The same PBase has also been effectively applied to gene therapy in vivo.[58]

As 293FT cells have served as a workhorse for screening sgRNAs for multiple genes previously, we started out to test all of our 72 sgRNAs in 293FT cells first. However, the ability of the SAM system to activate a gene is dependent upon the epigenetic properties of that locus, which vary in different cell types. Therefore, it is possible that results from 293FT cells would not be predictive of human iPSCs. To address this question, we compared the activation efficiency of individual sgRNAs across the two cell types for selected genes (Figure S7). The overall trends were that the sgRNAs that were able to activate a target in 293FT cells also could efficiently activate it in human iPSCs. It is also interesting to note that the rank-order activity of the three sgRNAs for a given locus was similar, but not entirely the same, for the two cell types. For instance (Figure S7), the best sgRNAs that had the highest gene activation in MTY1, FOXG1, or POU3F1 were the same in the two cell lines, while the second or third ranked sgRNAs switched positions. Future experiments comparing additional loci across different cell types are warranted to fully reveal the effects of epigenetic properties of different loci.

Our data showed that the transactivation efficiency was higher with four-in-one constructs compared to the corresponding four one-by-one

constructs. This is probably due to the following reasons. First, the transfection efficiency for individual PB vectors varies. Second, the four-in-one strategy guaranteed the delivery of sgRNAs for activating all four TFs as well as SAM components (Figure 4), while the one-by-one strategy could not ensure concurrent delivery of SAM components and sgRNAs into the same cells. It would be interesting to determine whether there was a potential for cross-regulation among those genes. In that case, a multiplex-inducible system that could control the activation of individual TFs is needed. Assays that measure gene expression in single cells will facilitate the determination of such potential regulation.

In summary, we have built PB-CRISPRa, a platform that could quickly and stably activate multiple genes in human cells simultaneously. Our system has the potential to be readily scaled up for dissection of genome-wide complex transcription regulatory networks.

MATERIALS AND METHODS

Vector Construction

A vector containing EF1α promoter-driven MSPH SAM components and a vector containing CAG promoter-driven dCas9-VP64 components were obtained from Addgene (61426 and 61425). The PB backbone and PBase expression vectors were gifts from Dr. Sen Wu. To make the PB SAM vector (Figure 1A), MSPH fragments with a nuclear localization signal (NLS) and hygromycin resistance cassette and the dCas9-VP64 components with NLS and blasticidin resistance cassette were digested and ligated into PB backbone vector via Gibson assembly (New England Biolabs)[59] (Figure 1A; Figure S8). The sgRNA (tracrRNA and crRNA) with MS2 stem loop fragment (Figure 1A) (Addgene, 61427) was digested and cloned into the PB vector via Gibson assembly. Sequence information of sgRNAs for 21 neural TFs and three lncRNAs is listed in Table S1. Sense and antisense strands of sgRNAs with appropriate overhangs for each gene were synthesized, incubated with T4 polynucleotide kinase (PNK) and T4 ligase buffer at 37°C for 30 min and 95°C for 5 min, and cooled down to room temperature within 2 hr. The resultant sgRNA insert of each gene was cloned into the PB-sgRNA (MS2) vector after six cycles of incubation at 37°C for 5 min and 21°C for 5 min, respectively, with FastDigest BbsI and T7 ligase in the presence of Tango buffer, DTT, and ATP.[60] Expression of sgRNA was driven by U6 promoter. All inserts were verified by sequencing.

All-in-one vectors containing expression cassettes of MSPH, dCas9-VP64, and sgRNAs for two to four genes desired to test were constructed using the Multisite Gateway strategy (Figure 3). Briefly, the PB-SAM vector described above was amplified to become PB-SAM R1-R2 DEST vector with attR1 and attR2 sites, ccdB cassette, and chloramphenicol resistance cassette (CmR). PB-sgRNA (MS2) vectors containing different sgRNA inserts were amplified to be attached with appropriate attB sites (Figure 3). All-in-one vectors were assembled by four-way LR reactions as described previously.[61] Primers used in Multisite Gateway cloning are listed in Table S2.

Cell Culture

HEK293FT cells were obtained from ATCC and cultured in high-glucose DMEM supplemented with 10% heat-inactivated fetal bovine serum, 1× GlutaMAX, and 1× non-essential amino acid (NEAA). A collection of human iPSC fluorescence protein reporters was created or obtained to facilitate the testing of the activation of neural lineage TFs or lncRNAs. Human NEUROG2-mCherry knockin reporter iPSC line was previously generated in our lab,[62] and OLIG2-EGFP iPSC knockin reporter line was created from human iPSC ND2-0 using a similar targeting strategy as previously described.[23] Human DCX-ZsGreen iPSC reporter was obtained from the NIH. A non-engineered human iPSC line Cy2 was also obtained from the NIH. All human iPSCs described herein (NEUROG2-mCherry knockin reporter, OLIG2-EGFP iPSC knockin reporter, DCX-ZsGreen reporter, and Cy2) were cultured in chemically defined mTeSR medium (STEMCELL Technologies) in feeder-free conditions and passaged every 4–5 days with Accutase (Innovative Cell Technologies) onto Matrigel- (Becton Dickinson) coated culture plates at a ratio of 1:4–1:8 with ROCK inhibitor Y-27632 (10 μM).

Generation of 293FT-SAM and Human iPSC-SAM Stable Cell Lines

To generate 293FT-SAM line, HEK293FT cells (1×10^6 cells in one well of a six-well plate) were co-transfected with PB-SAM (1.5 μg) and PBase (0.5 μg) vectors using Lipofectamine 3000 (Life Technologies), following the manufacturer's instructions. Then 48 hr after transfection, blasticidin (15 μg/mL) and hygromycin (200 μg/mL) were added to the culture medium, and a 293FT-SAM stable cell line was obtained after 2 weeks of selection. Similarly, to generate a human iPSC-SAM stable cell line, 1×10^6 human Cy2 iPSCs (NIH) were co-transfected with PB-SAM (3.75 μg) and PBase (1.25 μg) vectors using Nucleofector 2B with human stem cell nucleofector kit (Lonza). Then 48 hr after electroporation, blasticidin (10 μg/mL) and hygromycin (37.5 μg/mL) were added to the culture medium, and the Cy2-iPSC-SAM stable cell line was obtained 2 weeks after selection. For both 293FT-SAM and Cy2-iPSC-SAM cells, the expression of the two SAM components was examined by qRT-PCR as described below.

Transactivation in Stable Cell Lines

Approximately 5×10^5 293FT-SAM cells (in one well of a 12-well plate) were transfected with 1 μg PB-MS2-sgRNA or sgRNA mass using Lipofectamine 3000. For experiments in human iPSC-SAM stable cells, 1×10^6 cells were electroporated with 5 μg PB-sgRNA vector as described above. Then 48 hr after transfection, expression of the genes intended to activate was evaluated by qPCR. The sgRNAs that elicited the highest activation were selected for further experiments using normal human iPSCs (Cy2) or iPSC reporter lines.

Transactivation in Human iPSCs and iPSC Reporter Cell Lines with All-in-One Vectors

One million plain iPSCs or iPSC reporter cells were co-electroporated with all-in-one vectors (3.75 μg) and PBase vector (1.25 μg) as described above. OLIG2-EGFP human iPSC reporter line was used

to estimate OLIG2 activation. Similarly, NEUROG2-mCherry and DCX-ZsGreen human iPSC reporter lines were used for testing neuronal conversion, where NEUROG2 or DCX expression was estimated by mCherry or ZsGreen expression under a fluorescence microscope. One day after electroporation, iPSC mTeSR culture medium was changed to N2B27 medium containing DMEM/F12 and Neurobasal medium (1:1), with 1× N2 supplement, 1× B27 supplement, 1× NEAA, and 1× Glutamax, supplied with 20 ng/mL brain-derived neurotrophic factor (BDNF), 20 ng/mL glial cell-derived neurotrophic factor (GDNF), and 20 ng/mL neurotrophin-3 (NT3). Hygromycin (37.5 μg/mL) and blasticidin (8 μg/mL) were used for selection (Life Technologies). The Cy2 iPSC line was used for astrocyte induction, and astrocyte differentiation efficiency was estimated by GFAP immunocytochemistry. One day after electroporation, culture medium was switched to N2B27 medium supplied with 20 ng/mL recombinant bone morphogenetic protein 4 (BMP4), 8 ng/mL basic fibroblast growth factor (bFGF), and 10 ng/mL ciliary neurotrophic factor (CNTF, all from Life Technologies). Hygromycin (37.5 μg/mL) and blasticidin (8 μg/mL) were used for selection.

qRT-PCR

Total RNAs were extracted using Quick-RNA miniprep kits (Zymo Research). RNA (1 μg) was converted to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad). qRT-PCR was performed to determine mRNA levels using the iQ SYBR Green Supermix kit (Bio-Rad). GAPDH was used as an internal control. The relative fold change in gene expression was evaluated using the comparative threshold cycle ΔΔCt method.[63] The qPCR primers are listed in Table S3.

Immunocytochemistry

Immunocytochemistry was performed as previously described.[62,64] Briefly, cells grown on glass coverslips were fixed using 4% paraformaldehyde and incubated in blocking buffer (5% goat serum, 1% BSA, and 0.1% Triton X-100) for 60 min. Cells were then incubated in primary antibodies diluted in blocking buffer at 4°C overnight. Appropriate secondary antibodies were used for single and double labeling. All secondary antibodies were tested for cross-reactivity and nonspecific immunoreactivity. Bis-benzamide (Sigma) was used to identify the nuclei. For live staining (A2B5 and CD44), cells were incubated in primary antibodies diluted in culture medium at 37°C for 30 min. Appropriate secondary antibodies were then incubated at room temperature for 15 min. The following primary antibodies were used: A2B5 (1:20, Developmental Studies Hybridoma Bank [DSHB]), CD44 (1:200; EMD Millipore), β3 tubulin (1:4,000, Sigma), GFAP (1:2,000, Dako), S100β (1:200, R&D Systems), NeuN (1:1,000, Sigma), NF160 (1:500, Sigma), MAP2 (1:500, Millipore), and Cas9 (1:500, EpiGentek). Images were captured using a Zeiss Axiovision microscope with z stack split-view function. Note that all images involving EGFP, ZsGreen, and mCherry were captured directly under the fluorescence or confocal microscope without immunocytochemistry antibody staining. For each marker, the percentage of fluorescence protein-expressing cells or positively stained cells was calculated from at least ten low-magnification fields.

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.omtn.2017.06.007.

AUTHOR CONTRIBUTIONS

Conceptualization, S.L. and Y.L.; Methodology, S.L., H.X., and A.Z.; Investigation, S.L., H.X., A.Z., and D.L.; Writing – Original Draft, S.L. and Y.L.; Writing – Review & Editing, S.L. and Y.L.; Funding Acquisition, Y.L.; Resources, Y.L.; Supervision, Y.L.

ACKNOWLEDGMENTS

The authors thank Dr. Joanna S. OLeary for editing the manuscript. This work was supported by the start-up fund from The Vivian L. Smith Department of Neurosurgery, McGovern Medical School, The University of Texas Health Science Center at Houston, Memorial Hermann Foundation-Staman Ogilvie Fund (to Y.L.), the Bentsen Stroke Center (to Y.L.), the TIRR Foundation through Mission Connect (014-115 to Y.L.), and the Craig H. Neilsen Foundation (338617 to Y.L.).

REFERENCES

1. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.
2. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.
3. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. *31*, 227–229.
4. DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343.
5. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell *153*, 910–918.
6. Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. *31*, 833–838.
7. Jore, M.M., Brouns, S.J., and van der Oost, J. (2012). RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. Cold Spring Harb. Perspect. Biol. *4*, a003657.
8. Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA *109*, E2579–E2586.
9. Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949.
10. Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods *10*, 973–976.
11. Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647–661.
12. Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013).

CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

13. Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.
14. Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., and Lim, W.A. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell *160*, 339–350.
15. Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell *159*, 635–646.
16. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.
17. Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. Nat. Methods *12*, 326–328.
18. Woodard, L.E., and Wilson, M.H. (2015). piggyBac-ing models and new therapeutic strategies. Trends Biotechnol. *33*, 525–533.
19. Cary, L.C., Goebel, M., Corsaro, B.G., Wang, H.G., Rosen, E., and Fraser, M.J. (1989). Transposon mutagenesis of baculoviruses: analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. Virology *172*, 156–169.
20. Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., and Xu, T. (2005). Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell *122*, 473–483.
21. Wu, S., Ying, G., Wu, Q., and Capecchi, M.R. (2007). Toward simpler and faster genome-wide mutagenesis in mice. Nat. Genet. *39*, 922–930.
22. Xu, C., Qi, X., Du, X., Zou, H., Gao, F., Feng, T., Lu, H., Li, S., An, X., Zhang, L., et al. (2017). piggyBac mediates efficient in vivo CRISPR library screening for tumorigenesis in mice. Proc. Natl. Acad. Sci. USA *114*, 722–727.
23. Xue, H., Wu, S., Papadeas, S.T., Spusta, S., Swistowska, A.M., MacArthur, C.C., Mattson, M.P., Maragakis, N.J., Capecchi, M.R., Rao, M.S., et al. (2009). A targeted neuroglial reporter line generated by homologous recombination in human embryonic stem cells. Stem Cells *27*, 1836–1846.
24. Douvaras, P., Wang, J., Zimmer, M., Hanchuk, S., O'Bara, M.A., Sadiq, S., Sim, F.J., Goldman, J., and Fossati, V. (2014). Efficient generation of myelinating oligodendrocytes from primary progressive multiple sclerosis patients by induced pluripotent stem cells. Stem Cell Reports *3*, 250–259.
25. Liu, Y., Jiang, P., and Deng, W. (2011). OLIG gene targeting in human pluripotent stem cells for motor neuron and oligodendrocyte differentiation. Nat. Protoc. *6*, 640–655.
26. Dixit, R., Wilkinson, G., Cancino, G.I., Shaker, T., Adnani, L., Li, S., Dennis, D., Kurrasch, D., Chan, J.A., Olson, E.C., et al. (2014). Neurog1 and Neurog2 control two waves of neuronal differentiation in the piriform cortex. J. Neurosci. *34*, 539–553.
27. Fode, C., Gradwohl, G., Morin, X., Dierich, A., LeMeur, M., Goridis, C., and Guillemot, F. (1998). The bHLH protein NEUROGENIN 2 is a determination factor for epibranchial placode-derived sensory neurons. Neuron *20*, 483–494.
28. Fode, C., Ma, Q., Casarosa, S., Ang, S.L., Anderson, D.J., and Guillemot, F. (2000). A role for neural determination genes in specifying the dorsoventral identity of telencephalic neurons. Genes Dev. *14*, 67–80.
29. Hindley, C., Ali, F., McDowell, G., Cheng, K., Jones, A., Guillemot, F., and Philpott, A. (2012). Post-translational modification of Ngn2 differentially affects transcription of distinct targets to regulate the balance between progenitor maintenance and differentiation. Development *139*, 1718–1723.
30. Lee, S.K., Lee, B., Ruiz, E.C., and Pfaff, S.L. (2005). Olig2 and Ngn2 function in opposition to modulate gene expression in motor neuron progenitor cells. Genes Dev. *19*, 282–294.
31. Ma, Q., Chen, Z., del Barco Barrantes, I., de la Pompa, J.L., and Anderson, D.J. (1998). neurogenin1 is essential for the determination of neuronal precursors for proximal cranial sensory ganglia. Neuron *20*, 469–482.
32. Ma, Q., Kintner, C., and Anderson, D.J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. Cell *87*, 43–52.
33. Ma, Y.C., Song, M.R., Park, J.P., Henry Ho, H.Y., Hu, L., Kurtev, M.V., Zieg, J., Ma, Q., Pfaff, S.L., and Greenberg, M.E. (2008). Regulation of motor neuron specification by phosphorylation of neurogenin 2. Neuron *58*, 65–77.
34. Mizuguchi, R., Sugimori, M., Takebayashi, H., Kosako, H., Nagao, M., Yoshida, S., Nabeshima, Y., Shimamura, K., and Nakafuku, M. (2001). Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. Neuron *31*, 757–771.
35. Novitch, B.G., Chen, A.I., and Jessell, T.M. (2001). Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. Neuron *31*, 773–789.
36. Perrin, F.E., Boniface, G., Serguera, C., Lonjon, N., Serre, A., Prieto, M., Mallet, J., and Privat, A. (2010). Grafted human embryonic progenitors expressing neurogenin-2 stimulate axonal sprouting and improve motor recovery after severe spinal cord injury. PLoS ONE *5*, e15914.
37. Roybon, L., Deierborg, T., Brundin, P., and Li, J.Y. (2009). Involvement of Ngn2, Tbr and NeuroD proteins during postnatal olfactory bulb neurogenesis. Eur. J. Neurosci. *29*, 232–243.
38. Roybon, L., Hjalt, T., Stott, S., Guillemot, F., Li, J.Y., and Brundin, P. (2009). Neurogenin2 directs granule neuroblast production and amplification while NeuroD1 specifies neuronal fate during hippocampal neurogenesis. PLoS ONE *4*, e4779.
39. Scardigli, R., Schuurmans, C., Gradwohl, G., and Guillemot, F. (2001). Crossregulation between Neurogenin2 and pathways specifying neuronal identity in the spinal cord. Neuron *31*, 203–217.
40. Zirlinger, M., Lo, L., McMahon, J., McMahon, A.P., and Anderson, D.J. (2002). Transient expression of the bHLH factor neurogenin-2 marks a subpopulation of neural crest cells biased for a sensory but not a neuronal fate. Proc. Natl. Acad. Sci. USA *99*, 8084–8089.
41. Busskamp, V., Lewis, N.E., Guye, P., Ng, A.H., Shipman, S.L., Byrne, S.M., Sanjana, N.E., Murn, J., Li, Y., Li, S., et al. (2014). Rapid neurogenesis through transcriptional activation in human stem cells. Mol. Syst. Biol. *10*, 760.
42. Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron *78*, 785–798.
43. Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P., and Lu, T.K. (2014). Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. Mol. Cell *54*, 698–710.
44. Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat. Biotechnol. *32*, 569–576.
45. Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. USA *112*, 3570–3575.
46. Yan, Q., Xu, K., Xing, J., Zhang, T., Wang, X., Wei, Z., Ren, C., Liu, Z., Shao, S., and Zhang, Z. (2016). Multiplex CRISPR/Cas9-based genome engineering enhanced by Drosha-mediated sgRNA-shRNA structure. Sci. Rep. *6*, 38970.
47. Sadahiro, T., Yamanaka, S., and Ieda, M. (2015). Direct cardiac reprogramming: progress and challenges in basic biology and clinical applications. Circ. Res. *116*, 1378–1391.
48. Black, J.B., Adler, A.F., Wang, H.G., D'Ippolito, A.M., Hutchinson, H.A., Reddy, T.E., Pitt, G.S., Leong, K.W., and Gersbach, C.A. (2016). Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells. Cell Stem Cell *19*, 406–414.
49. Chakraborty, S., Ji, H., Kabadi, A.M., Gersbach, C.A., Christoforou, N., and Leong, K.W. (2014). A CRISPR/Cas9-based system for reprogramming cell lineage specification. Stem Cell Reports *3*, 940–947.

50. Balboa, D., Weltner, J., Eurola, S., Trokovic, R., Wartiovaara, K., and Otonkoski, T. (2015). Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation. Stem Cell Reports *5*, 448–459.

51. Wei, S., Zou, Q., Lai, S., Zhang, Q., Li, L., Yan, Q., Zhou, X., Zhong, H., and Lai, L. (2016). Conversion of embryonic stem cells into extraembryonic lineages by CRISPR-mediated activators. Sci. Rep. *6*, 19648.

52. Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. Nat. Rev. Mol. Cell Biol. *12*, 36–47.

53. Saridey, S.K., Liu, L., Doherty, J.E., Kaja, A., Galvan, D.L., Fletcher, B.S., and Wilson, M.H. (2009). PiggyBac transposon-based inducible gene expression in vivo after somatic cell gene transfer. Mol. Ther. *17*, 2115–2120.

54. Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., and Liu, P. (2008). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. Proc. Natl. Acad. Sci. USA *105*, 9290–9295.

55. Wu, S., Wu, Y., Zhang, X., and Capecchi, M.R. (2014). Efficient germ-line transmission obtained with transgene-free induced pluripotent stem cells. Proc. Natl. Acad. Sci. USA *111*, 10678–10683.

56. Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive piggyBac transposase for mammalian applications. Proc. Natl. Acad. Sci. USA *108*, 1531–1536.

57. Yusa, K. (2013). Seamless genome editing in human pluripotent stem cells using custom endonuclease-based gene targeting and the piggyBac transposon. Nat. Protoc. *8*, 2061–2078.

58. Li, M.A., Pettitt, S.J., Eckert, S., Ning, Z., Rice, S., Cadiñanos, J., Yusa, K., Conte, N., and Bradley, A. (2013). The piggyBac transposon displays local and distant reintegration preferences and can cause mutations at noncanonical integration sites. Mol. Cell. Biol. *33*, 1317–1330.

59. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

60. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

61. Macarthur, C.C., Xue, H., Van Hoof, D., Lieu, P.T., Dudas, M., Fontes, A., Swistowski, A., Touboul, T., Seerke, R., Laurent, L.C., et al. (2012). Chromatin insulator elements block transgene silencing in engineered human embryonic stem cell lines at a defined chromosome 13 locus. Stem Cells Dev. *21*, 191–205.

62. Li, S., Xue, H., Wu, J., Rao, M.S., Kim, D.H., Deng, W., and Liu, Y. (2015). Human Induced Pluripotent Stem Cell NEUROG2 Dual Knockin Reporter Lines Generated by the CRISPR/Cas9 System. Stem Cells Dev. *24*, 2925–2942.

63. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408.

64. Chen, C., Jiang, P., Xue, H., Peterson, S.E., Tran, H.T., McCann, A.E., Parast, M.M., Li, S., Pleasure, D.E., Laurent, L.C., et al. (2014). Role of astroglia in Down's syndrome revealed by patient-derived human-induced pluripotent stem cells. Nat. Commun. *5*, 4430.