Published online 5 February 2014 Nucleic Acids Research, 2014, Vol. 42, No. 7 4375–4390
doi:10.1093/nar/gku109

Direct activation of human and mouse *Oct4* genes using engineered TALE and Cas9 transcription factors

Jiabiao Hu[1,2], Yong Lei[1], Wing-Ki Wong[1], Senquan Liu[1,3], Kai-Chuen Lee[1,2], Xiangjun He[1,2], Wenxing You[1,2], Rui Zhou[1], Jun-Tao Guo[4], Xiongfong Chen[5], Xianlu Peng[6,7], Hao Sun[6,7], He Huang[3], Hui Zhao[1,2] and Bo Feng[1,2,*]

[1]Key Laboratory for Regenerative Medicine, Ministry of Education, School of Biomedical Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China, [2]SBS Core Laboratory, CUHK Shenzhen Research Institute, Shenzhen, China, [3]Bone Marrow Transplantation Centre, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, China, [4]Department of Bioinformatics and Genomics, The University of North Carolina at Charlotte, Charlotte, NC 28223, USA, [5]Advanced Biomedical Computing Center, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA, [6]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR, China and [7]Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong SAR, China

Received August 29, 2013; Revised January 11, 2014; Accepted January 13, 2014

ABSTRACT

The newly developed transcription activator-like effector protein (TALE) and clustered regularly interspaced short palindromic repeats/Cas9 transcription factors (TF) offered a powerful and precise approach for modulating gene expression. In this article, we systematically investigated the potential of these new tools in activating the stringently silenced pluripotency gene *Oct4* (*Pou5f1*) in mouse and human somatic cells. First, with a number of TALEs and sgRNAs targeting various regions in the mouse and human *Oct4* promoters, we found that the most efficient TALE-VP64s bound around −120 to −80 bp, while highly effective sgRNAs targeted from −147 to −89-bp upstream of the transcription start sites to induce high activity of luciferase reporters. In addition, we observed significant transcriptional synergy when multiple TFs were applied simultaneously. Although individual TFs exhibited marginal activity to up-regulate endogenous gene expression, optimized combinations of TALE-VP64s could enhance endogenous *Oct4* transcription up to 30-fold in mouse NIH3T3 cells and 20-fold in human HEK293T cells. More importantly, the enhancement of *OCT4* transcription ultimately generated OCT4 proteins. Furthermore, examination of different epigenetic modifiers showed that histone acetyltransferase p300 could enhance both TALE-VP64 and sgRNA/dCas9-VP64 induced transcription of endogenous *OCT4*. Taken together, our study suggested that engineered TALE-TF and dCas9-TF are useful tools for modulating gene expression in mammalian cells.

INTRODUCTION

Engineered transcription factors (TFs) have wide-ranging potential in modulating desired gene expression through targeting their promoters (1). Natural DNA-binding proteins, such as zinc finger proteins Gal4 and tetracycline repressors have been employed to modulate gene expression via fusion to transcriptional activators or repressors (2,3). However, lack of simple correlation between amino acid sequence and DNA recognition has made it difficult and costly to engineer these recombinant proteins specifically for an interested gene (4).

A recent breakthrough with transcription activator-like effector proteins (TALEs) makes it possible to establish universal types of engineered TFs that can potentially target any selected gene. TALEs originated in plant pathogen *Xanthomonas sp*., and have demonstrated a simple protein–DNA-binding principle (5,6). The TALE domain contains a highly conserved central region which is composed of a series of 33–35 amino acids repetitive elements. The highly variable di-residues at the 12th and

*To whom correspondence should be addressed. Tel: +852 3943 1455; Fax: +852 2603 5123; Email: fengbo@cuhk.edu.hk

13th positions in each element are referred to as repeat variable di-residues (RVDs), dictating the specific binding preference between a single repeat and a nucleotide. The RVDs with high affinity for nucleotides A, C and T have been identified as NI, HD and NG, respectively (5,6). Several RVDs—including NN, NH and NK—have been reported to recognize G nucleotide, in which NN occurs more frequently in natural TALEs and is widely used for DNA recognition in mammalian cells (6–8). This simple coding principle has enabled assembly of TALE repeat arrays for targeting almost any given DNA sequence. Hence, fusion of TALEs to transcriptional activator or repressor domains, such as VP16 or KRAB (2,9), could generate TALE-TFs that can target selected promoter regions and modulate expression of corresponding genes (10,11).

Apart from TALEs, an RNA-guided DNA-targeting approach was recently developed from the Type II prokaryotic clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system (12,13). In this system, foreign DNAs from invading viruses/plasmids stimulated the synthesis of CRISPR RNAs (crRNA) and trans-activating crRNAs (tracrRNA). In turn, these short RNAs annealed with foreign DNA and recruited Cas9 endonuclease to mediate the foreign DNA degradation (12,13). A simplified two-component CRISPR/Cas system was later established by replacing crRNA and tracrRNA with a single synthetic small guide RNA (sgRNA), which mimics the structure of annealed crRNA and tracrRNA (14). The sgRNAs were as efficient as crRNAs and tracrRNAs to direct Cas9 nuclease for introducing site-specific DNA cleavage, which subsequently resulted in targeted mutation/deletions (14). Hence, the CRISPR–Cas9 system has been regarded as a superior tool for genomic engineering both in human and mouse cells (15–17). Interestingly, transcriptional activators/repressors fused to a mutated Cas9 protein lacking endonuclease activity (dCas9–TF) could also be guided to desired DNA by sgRNAs, thus establishing a new platform for modulating gene expression (18–20). The requirement for single protein component dCas9, high fidelity of RNA–DNA binding as well as the simplicity of generating a new sgRNA to target a selected DNA sequence have made the CRISPR/Cas9 system a desirable tool for altering gene expression.

Intensive studies on stem-cell maintenance and differentiation substantiated that cellular identity is often determined by the activation or repression of key TFs. Octamer-binding TF 4 (Oct4) is a master TF that governs pluripotency in stem cells. Studies have demonstrated that Oct4 is essential for the formation of inner cell mass during embryogenesis (21), as well as the maintenance of embryonic stem cells (ESC) in culture (21,22). Moreover, Oct4 was shown to play a pivotal role in reinstating cellular pluripotency (23,24) and it alone could reprogramme somatic neural progenitor cells into induced pluripotent stem cells (iPSC) (25). The *Oct4* gene (also named *Pou5f1*) is driven by a TATA-less promoter, a proximal enhancer (PE) and a distal enhancer (DE) (26). Comparative analysis of *Oct4* regulatory elements in different species identified four conservative regions (CRs): CR1 (in proximal promoter), CR2 and CR3 (PE), as well as CR4 (DE) (27,28). The DE/CR4 is essential for regulating *Oct4* expression in morula, inner cell mass of blastocysts and primordial germ cells; while the PE/CR2 activates *Oct4* in the epiblast stage (26,29). The expression of *Oct4* is stringently silenced in differentiated cells. Upon iPSC induction, even though a couple of key TFs were simultaneously over-expressed in somatic cells—including those that can bind CR4 and activate *Oct4* expression in ESCs—the endogenous *Oct4* gene remained silent and its activation was observed only at the very late phase of reprogramming when true iPSCs started to emerge (30). Activation of silenced *Oct4* gene has become a hallmark event during epigenetic reprogramming into iPSCs, but the mechanism underlying its activation still remains elusive. In this context, highly specific TALE-TFs and sgRNA-guided dCas9-TFs offer new avenues for manipulating endogenous *Oct4* gene expression. It would be interesting to investigate whether direct activation of silenced *Oct4* gene by TALE-TFs or sgRNA/dCas9-TFs in somatic cells could promote reprogramming and facilitate iPSC generation. Moreover, TALE-TFs or dCas9-TFs could potentially be used for investigating the complex epigenetic regulations and chromatin architectures involved in the stringent suppression of *Oct4* gene in somatic cells.

Recent studies have attempted to use TALE- or dCas9-TFs to activate pluripotency genes in somatic cells. Zhang *et al.* showed that *SOX2* and *KIF4*—but not *OCT4* and *c-MYC* in human 293FT cells—could be activated by TALEs fused to VP64 (VP64 is a tetrameric repeat of VP16) (10). Bultmann *et al.* later demonstrated that TALE-VP16 could activate silenced *Oct4* gene in mouse ESC-derived neural stem cells with the assistance of epigenetic modifier inhibitors 5′-AzaC and valproic acid (VPA) (31); and Gao *et al.* showed that endogenous *Oct4* transcription could be induced by TALE-VP64s targeting the *Oct4* enhancer, which thus facilitated epigenetic reprogramming and enhanced iPSC generation in the presence of other reprogramming factors (32). More recently, concurrent application of multiple TALE- or dCas9-TFs was found to have a synergistic effect on activation of target genes (33–36).Perez-Pinera *et al.* detected around 10-fold increase of endogenous *NANOG* transcripts induced by simultaneous usage of five sgRNAs and dCas9-VP64 (35). Using a similar strategy, Mali *et al.* reported that transcription of endogenous *REX1* and *OCT4* could be activated to a high level in human HEK293T cells (37), and Cheng *et al.* showed that *SOX2* and *OCT4* mRNAs could be upregulated by 8- and 9-fold, respectively (38). These reports indicated that TALE- and sgRNA/dCas9-TFs could activate endogenous *Oct4* gene expression by targeting its promoter or enhancer. However, the detailed mechanism underlying this process has not been addressed. It remains unclear whether these engineered TFs could directly alter stringent epigenetic repression, or whether additional factors are needed to establish stable expression of *Oct4*. Furthermore, although generation of TALEs or sgRNAs

with specific DNA targets has become simple and fast, it is still difficult to predict the effectiveness of a single design. Therefore, more effort is needed to understand these new tools and explore their full potential.

In this article, we systematically investigated the potential of engineered TALE-VP64s and sgRNA/dCas9-VP64s for activation of silenced *Oct4* genes, both in mouse and human somatic cells. With a number of TALEs and sgRNAs generated to target various regions in the mouse and human *Oct4* promoters, we found that the most efficient TALE-VP64s bound around −120 to −80 bp, while highly effective sgRNAs targeted from −147 to −89-bp upstream of the transcription start site (TSS) to induce high activity of luciferase reporters; moreover, application of multiple TALE-VP64s or sgRNAs exhibited transcriptional synergy. For the activation of endogenous gene expression, we observed that individual activators often exhibited marginal or no activity; whereas optimized combinations of TALE-VP64s could up-regulate the transcription level of endogenous mouse *Oct4* genes to around 30-fold in NIH3T3, and activate human *OCT4* to approximate 20-fold in HEK293T cells. Interestingly, the activation of endogenous *OCT4* was also detected at protein level, whereas such expression was dependent on the exogenous TALE-VP64s or sgRNA/dCas9-VP64s and could not be sustained for a long period of time. Bisulfite sequencing analysis further showed that DNA methylation in *OCT4* promoter could not be reversed by transient transcriptional activation induced by either TALE-VP64s or sgRNA/dCas9-VP64s. Furthermore, examination of epigenetic modifiers showed that p300 could facilitate the activation of silenced *OCT4* genes mediated by TALE-VP64s or sgRNA/dCas9-VP64s, suggesting that these systems could also provide a platform for investigating the epigenetic repression of *Oct4* in somatic cells.

MATERIALS AND METHODS

DNA constructs

Luciferase reporter plasmids. Mouse or human *Oct4* promoters that cover 2.3 (mouse) or 2.5 kb (human) fragments upstream of ATG were cloned into pGL3 vector (Promega); 10 repeats of upstream activation sequence (UAS) element were inserted in the 5′-end of the *Oct4* promoter to serve as a positive control (39). Human *NANOG* luciferase reporter pNANOG-Luc was obtained from Addgene (Addgene #25900).

Mutated mOct4-Luc reporter plasmids. Site-directed mutagenesis was performed using PCR approach described previously (40). Briefly, the wild-type *mOct4-Luc* reporter was first amplified with primers that introduced the T to C mutation. The amplified fragment was then treated with DpnI to remove the original template plasmids; and the remaining product was used for transformation to obtain mutated reporter plasmids M-6, M-11, M-17, M-20 and M-25. Next, these mutated reporters were used as template for further amplification; primers used in this step were designed to anneal to DNA adjacent to the −120 to −104-bp region while carrying the target sequences of selected TALE-VP64s at 5′. The amplified fragments were then treated with DpnI as described above, yielding reporter plasmids that carried the relocated target sequences from −120 to −104 bp in mouse Oct4 promoter.

Fuw-tetO-TALE-VP64 plasmids. An nuclear localization signal (NLS)-VP64-HA fragment was assembled using PCR approach and inserted into the PspXI and HpaI sites in the Fuw-tetO vector, which was modified from the plasmid FUW-tetO-hOCT4 (Addgene #20726) by destroying the BsmBI site and introducing BsrGI, PspXI and HpaI sites enclosed by the EcoRI site (41). The DNA fragment coding LacZ flanked by TALE N- and C-terminals was released from the plasmid pTAL1 (Addgene #31031) (42) and inserted into the BsrGI and PspXI sites of the Fuw-tetO-NLS-VP64-HA to generate a Fuw-tetO-TALE(LacZ)-VP64 scaffold vector. Various 17-bp sequences preceded by a T were selected to be the candidate TALE target sequences (42). Corresponding TALE repeat arrays were then assembled using the Fuw-tetO-TALE(LacZ)-VP64 scaffold vector using the previously described Golden Gate cloning method (Supplementary Figure S1) (42,43).

Fuw-tetO-TALE-KRAB plasmids. The KRAB domain was amplified from pLV-tTRKRAB (Addgene #12249) and inserted into the PspXI and HpaI sites of Fuw-tetO-TALE-VP64s to replace VP64s.

dCas9-VP64/KRAB plasmids. The H840A mutation was introduced into the hCas9_D10A plasmid (Addgene # 41816) to produce a catalytically inactive Cas9 (dCas9) (14,19). The full-length dCas9, dCas9 with deletion of N-terminal RuvC1 domain (ΔN) or C-terminal HNH domain (ΔC) were amplified using PCR approach and inserted into BsrGI and PspXI sites of the modified Fuw-tetO vector, followed by insertion of VP64/KRAB.

sgRNA constructs. The 20-bp sequences that precede NGG, the protospacer adjacent motif (PAM) required for sgRNA targeting (44), were selected as candidate sgRNA target sequences. To generate an sgRNA, a pair of 26-mer oligos containing sgRNA target sequences were synthesized. They were annealed and then inserted into the BsmBI site in the sgRNA expression vector MLM3636 (Addgene # 43860) (Supplementary Figure S5A) (44).

shRNA constructs. Two shRNAs were designed to target separated 19-bp sequences in p300 coding sequence using WI siRNA selection program http://jura.wi.mit.edu/bioc/siRNAext/. They were p300 shRNA1: 5′-GCACGA CTTACCAGATGAA and p300 shRNA2: 5′-GCCTCAA ACTACAATAAAT. Synthesized oligonucleotides were cloned into pSUPER.puro (BglII and HindIII sites; Oligoengine).

Vectors containing p300, JMJD3 and JMJD2B were obtained from Addgene (Addgene #10717, #24167 and #24181).

Cell culture

HEK293T and NIH3T3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal bovine serum (FBS) (Life Technologies).

Mouse ESCs (E14) were cultured as previously described (45). Briefly, cells were cultured on gelatin-coated dishes in DMEM medium, supplemented with 15% heat-inactivated FBS (Life Technologies), 0.1 mM β-mercaptoethanol (Life Technologies), 2 mM L-glutamine, 0.1 mM MEM non-essential amino acid (Life Technologies) and 1000 U/ml of leukemia inhibitory factor (LIF) (Life Technologies). The culture medium was refreshed every day and the cells were passaged every two days.

Human ESCs (H1) were cultured as previously described (46). Briefly, they were cultured feeder-free on Matrigel (BD Biosciences). Medium containing 20% knockout serum replacement, 1 mM L-glutamine, 1% non-essential amino acids, 0.1 mM β-mercaptoethanol and 4 ng/ml basic fibroblast growth factor (bFGF) (Life Technologies) was conditioned by mouse embryonic fibroblast. Additional 8 ng/ml bFGF was added freshly to conditioned medium for human ESC culture. Medium was changed daily and cells were subcultured with 1 mg/ml collagenase IV (Life technologies) every 5–7 days.

Transfection

HEK293T cells were seeded into 12-well plates at a density of 2.4×10^5 cells/well one day before transfection. Unless otherwise stated, 1.6 μg TALE-VP64 plasmids (or 0.8 μg dCas9-VP64 + 0.8 μg of sgRNA) were used for Lipofectamine 2000 (Life Technologies) transfection in each well following the manufacturer's instruction. When more than one TALE-VP64 or sgRNA was examined, the amount for each TALE-VP64 or sgRNA plasmid equaled to 1.6 μg (TALE-VP64) or 0.8 μg (sgRNA) divided equally by the numbers of plasmids. Cells were grown for an additional 48 h before harvesting for qRT-PCR analysis or immunoblotting. The estimated transfection efficiency was around 83.1% using 1.6 μg pEGFP-N1 plasmid.

For transfection in NIH3T3, 1×10^5 cells were seeded into each well of 12-well plates one day before transfection. A mixture of 0.8 μg of FUW-M2rtTA and 0.8 μg TALE-VP64 (or 0.4 μg dCas9-VP64 + 0.4 μg of sgRNA) plasmids were used for transfection in each well. Similarly, when more than one TALE-VP64 or sgRNA was used, the amount of individual plasmids was divided equally as described above. Cells were grown in the presence of 1 μg/ml doxycycline (Dox) for an additional 48 h before harvesting for qRT-PCR analysis. The estimated transfection efficiency was ~78.8% using the plasmid pEGFP-N1.

RNA extraction, reverse transcription and quantitative real-time PCR

Total RNA was extracted from cell samples using TRIzol reagent (Life Technologies) and reverse-transcribed into cDNA using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Quantitative real-time PCR was performed with Power SYBR Green PCR Master Mix (Applied Biosystems) in an ABI7900HT Real Time PCR system. Measured transcript was normalized to Gapdh and samples were run in triplicate. Primers used for qRT-PCR analysis were provided in Supplementary Table S1.

Western blot

Cells were collected by trypsinization and washed with phosphate buffered saline (PBS). Samples were then incubated in lysis buffer [50 mM Tris, 0.5% NP40, 1 mM EDTA, 10% glycerol, 400 mM sodium chloride and Protease Inhibitor Cocktail (Roche)] and cleared by centrifuging at 20 000 × g, 4°C for 30 min. Protein concentration was determined with BCA Protein Assay Reagent (Thermo Scientific). From each sample, 10 μg protein were resolved by SDS/PAGE and subsequently transferred to polyvinylidene difluoride membranes (Bio-Rad). Membranes were blocked with 5% non-fat dry milk in PBST buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.46 mM KH_2PO_4, 0.05% Tween-20) for 1 h at room temperature and incubated with Oct3/4 or GAPDH antibodies (Santa Cruz) overnight. Membrane were then washed three times with PBST buffer and incubated with rabbit anti-goat-HRP or goat anti-mouse-HRP (GE technology) at a dilution of 1:20 000 in PBST for 1 h at room temperature. Signals were detected using the Amersham ECL Select Western Blotting Detection Kit (GE Health Care Life Sciences) and exposed to Super RX-N film (Fuji).

Dual luciferase reporter assay

HEK293T cells were seeded into 96-well plates at a density of 2×10^4 cells/well one day before transfection. Cells were transfected with 100 ng TALE-VP64 plasmid (or 50 ng dCas9-VP64 + 50 ng sgRNA), 100 ng corresponding reporter plasmids and 10 ng Renilla (Promega) using Lipofectamine 2000 according to the manufacturer's instructions. When more than one TALE-VP64 or sgRNA was used, the amount of individual plasmids was divided equally as described above.

To examine the effect of Dox-inducible expression, 50 ng FUW-M2rtTA and 50 ng pLV-tTRKRAB were co-transfected with 12.5 ng TALE-VP64, 50 ng of the corresponding reporter plasmids and 10 ng Renilla into the HEK293T cells in each well of 96-well plates. Cells were then maintained in the presence or absence of 1 μg/ml Dox for 2 days before analysis.

To examine the transcriptional repression by TALE-KRAB or dCas9-KRAB, mouse ESCs E14 were seeded into 24-well plates at a density of 8×10^4 cells/well 6 h before transfection. The cells were then transfected with 200 ng TALE-KRAB plasmid (or 100 ng dCas9-KRAB + 100 ng sgRNA), 200 ng FUW-M2rtTA, 200 ng corresponding reporter plasmids and 20 ng Renilla (Promega). The estimated transfection efficiency was around 74.9% using the plasmid pEGFP-N1.

Two days after transfection, luciferase reporter assays were carried out using the Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's instructions. Relative luciferase activity was measured using a GLOMA20/20 Luminometer (Promega). The activity of the firefly luciferase was normalized with that of Renilla luciferase.

Statistical analysis

Statistical significance was determined using unpaired two sample Student's *T*-test. $P < 0.05$ was considered as significant. Data were shown as mean ± SEM ($n = 3$).

In vitro methylation assay

In vitro methylation of *mOct4-Luc* and *hOCT4-Luc* reporter plasmids was performed using CpG methyltransferase M.SssI (New England Biolabs) as previously described (31). Briefly, 45 μg of reporter plasmid were incubated overnight with 45 units of M.SssI enzyme. The methylation status of the plasmid DNA was then confirmed by digestion with MspI and HpaII (New England Biolabs). The effect of TALE-VP64s on the *in vitro* methylated reporter was then examined by luciferase assay as described above.

Bisulfite sequencing analysis

The genomic DNA from cell samples was extracted using PureLink™ Genomic DNA Mini Kit (Life Technologies). For each sample, 300 ng of genomic DNA was used for bisulfite conversion using the EZ DNA Methylation-Gold™ Kit (Zymo Research Corporation) in accordance with the manufacturer's instructions. The promoter sequence was amplified using HotStarTaq Plus DNA Polymerase (QIAGEN) with the following primers: mouse *Oct4*: 5′-ATGGGTTGAAATATTGGGTTTATT TA and 5′-CCACCCTCTAACCTTAACCTCTAAC; human *OCT4* (region 1): 5′-AAGTTTTTGTGGGGGA TTTGTAT and 5′-CCACCCACTAACCTTAACCTCT A; human *OCT4* (region 2): 5′-GAGAGAGGGGTTGA GTAGTTTT and 5′-CACTAACCCCACTCCAACCTA AA. Amplified PCR products were then ligated into the pGEM-T Easy vector (Promega) and sequenced with M13 forward primer.

RESULTS

TALE-VP64s activated mouse and human *Oct4* promoters by targeting a non-conserved proximal region

We constructed a scaffold vector to assemble TALEs, NLS, VP64 and HA-tag under the control of Dox-inducible tetO-CMV fusion promoter (Figure 1A, upper panel) (41). Based on this vector, we assembled different TALE DNA-binding domains using previously described Golden Gate cloning method (Supplementary Figure S1) (42,43). In total, we generated 26 TALE-VP64s, each recognizing a 17-bp sequence on various regions across 2.3-kb upstream genomic region of mouse *Oct4* gene (Figure 1B; Supplementary Figure S2) (27). These TALE-VP64s hereinafter were referred to as mO1–mO26. Four of the mO TALE-VP64s targeted the conserved DE region CR4, eight recognized the PE CR2, four bound to the proximal promoter CR1 region and the remaining 10 located in the non-conserved regions upstream of CR4 (mO1–mO4) or between CR1 and CR2 (mO17–mO22) (Figure 1B; Supplementary Figure S2).

The effects of individual mO TALE-VP64s were examined by luciferase assay in HEK293T cells. Each mO was co-transfected with a luciferase reporter under the control of 2.3 kb mouse *Oct4* promoter (*mOct4-Luc*) (Figure 1A, lower panel). Ten repeats of UAS elements were also inserted into the reporter plasmid (Figure 1A, lower panel) (39); and Gal4-VP64, which could target the UAS region, was included to serve as a positive control. As expected, mO TALE-VP64s targeting different regions of *Oct4* promoter exhibited various effects in activating the *mOct4-Luc* reporter (Figure 1B). Yet it was surprising that mO5–mO16, which targeted the conserved enhancer regions (CR2 and CR4) that were known to be bound by ESC TFs such as OCT4 itself, SOX2, NANOG and KLF4 for maintaining the transcription of *Oct4* gene (47), showed only moderate activity (Figure 1B). In contrast, mO21, which targeted a sequence outside of the conserved regions, at −120 to −104-bp upstream of the TSS (27), exhibited exceptionally high activity in boosting the transcription (Figure 1B). This transcriptional activation by mO21 was well correlated with the presence of Dox (Figure 1C). To examine whether these TALEs could mediate gene repression, we fused the TALE domains in mO21 and mO22 to a transcriptional repressor KRAB (9,48). These $mO21^{KRAB}$ and $mO22^{KRAB}$ fusion proteins were then examined in mouse ESCs, where the transcription of *mOct4-Luc* reporter could be maintained at a high level. Indeed, luciferase assay showed that $mO21^{KRAB}$ and $mO22^{KRAB}$ repressed the transcription of *mOct4-Luc* reporter by ∼80% and 30%, respectively (Figure 1D). Collectively, these indicated that optimized TALE-VP64s could modulate the transcription of mouse *Oct4* promoter efficiently by targeting the non-conserved −120 to −104-bp region.

To examine whether the activity of a TALE-VP64 depends on its target position, we selected five TALE-VP64s mO6, mO11, mO17, mO20 and mO25 and analyzed the positional effect by moving their target sequences from the original locations into the observed optimal region at −120 to −104 bp. First, 5′ T in the target sequences of mO6, mO11, mO17, mO20 and mO25 were mutated into C to abolish the recognition by TALEs to their original target sites. The mutated reporter constructs were named M-6, M-11, M-17, M-20 and M-25. Next, the original mO21 target sequence at the −120 to −104-bp region in the plasmids M-6, M-11, M-17, M-20 and M-25 was replaced with the target sequences of mO6, mO11, mO17, mO20 and mO25. This generated the reporter constructs MR-6, MR-11, MR-17, MR-20 and MR-25 which carried the re-located target sequences. Relocation of the target sequences into the proposed optimal region dramatically increased the transcriptional activation induced by these TALE-VP64s (Figure 1E), suggesting that the target positions on the promoter region influenced the activity of TALE-VP64s.

Based on this result, we designed eight TALE-VP64s targeting the non-conserved region as well as CR1 in human *OCT4* promoter (hO1–hO8) (Figure 1F; Supplementary Figure S2). Similarly, an *hOCT4-Luc* reporter plasmid was constructed to examine the activity of individual hO TALE-VP64s. Interestingly, hO3 and hO4 that targeted the −113 to −80-bp region upstream of the TSS exhibited significantly higher

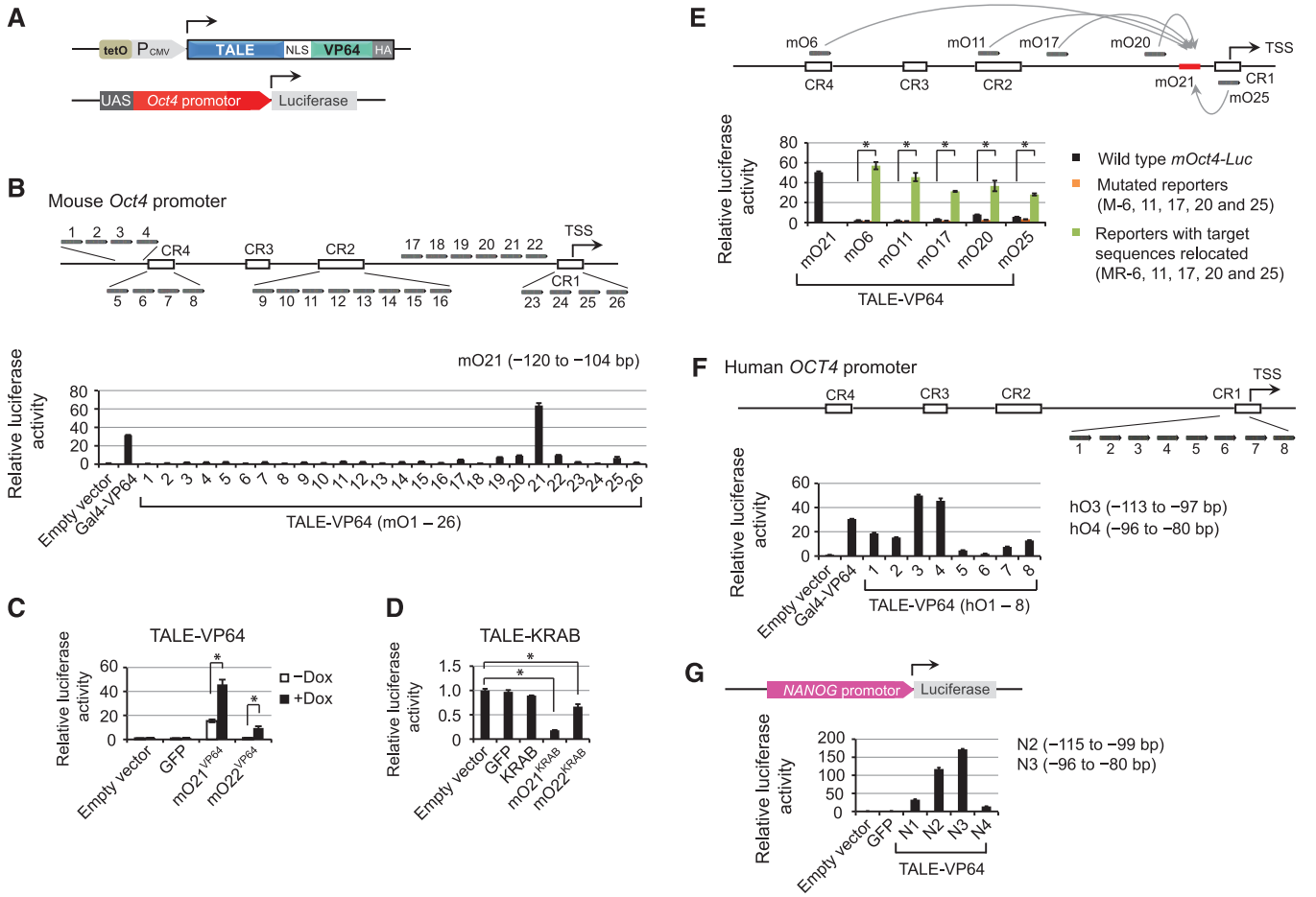

Figure 1. Analysis of TALE-VP64s targeting human and mouse *Oct4* promoters. (A) Schematic representation of TALE-VP64 fusion proteins and *mOct4*/*hOCT4-Luciferase* (*Luc*) reporters. TALE domain was fused with NLS, VP64 and HA tag under the control of tetO-CMV promoter (upper panel). Luciferase reporters contained UAS element and *mOct4*/*hOCT4* promoter (lower panel). (B) Luciferase activity of mouse *Oct4* promoter induced by various TALE-VP64s (mO1–mO26). Schematic diagram showed the relative target locations of mO1–mO26 within the 2.3-kb regulatory region upstream of the TSS in mouse *Oct4* gene. Effects of individual mOs were examined via co-transfection with *mOct4-Luc* reporter plasmid into HEK293T cells and relative luciferase activity was measured at 48 h after transfection. (C) Dox-dependent activity of mO21 and mO22 TALE-VP64s. Individual mOs were co-transfected with *mOct4-Luc* reporter plasmid into HEK293T cells, in the presence or absence of Dox. Relative luciferase activity was measured at 48 h after transfection. (D) TALE-KRAB fusion proteins $mO21^{KRAB}$ or $mO22^{KRAB}$ were co-transfected with *mOct4-Luc* reporter plasmid into E14 mouse ESCs, and relative luciferase activity was measured at 48 h after transfection. (E) Effect of positions on the activity of TALE-VP64s. TALE-VP64s mO6, mO11, mO17, mO20 and mO25 were co-transfected with wild-type *mOct4-Luc* reporter, reporters with point-mutation on the original target sequences of corresponding TALE-VP64s (constructs M-6, M-11, M-17, M-20 and M-25), and reporters with these target sequences relocated to the −120- to −104-bp region (constructs MR-6, MR-11, MR-17, MR-20 and MR-25), to assess their ability for activating the luciferase reporter. Relative luciferase activity was measured at 48 h after co-transfection. (F) Luciferase activity of human *OCT4* promoter induced by TALE-VP64s targeting its CR1 and non-conserved region. TALE-VP64 hO1–hO8 and their target sites were illustrated. Relative luciferase activity was measured at 48 h after co-transfection with *hOCT4-Luc* reporter plasmid into HEK293T cells. (G) Schematic diagram of *hNANOG-Luc* reporter and its relative luciferase activity induced by TALE-VP64s (N1–N4). Data were shown as mean ± SEM ($n = 3$). *$p < 0.05$.

activity than other TALE-VP64s in the luciferase assay (Figure 1F; Supplementary Figure S2). Besides *OCT4*, we also generated four TALE-VP64s (N1–N4) targeting DNA regions close to the TSS in human *NANOG* promoter (Supplementary Figure S2). We found that TALE-VP64 (N3) which targeted from −96 to −80-bp upstream of the *NANOG* TSS showed the highest activity in the luciferase assay using *hNANOG-Luc* reporter (Figure 1G). Together with the result obtained with mO1–mO26, these data suggested that a positional effect of the engineered TFs TALE-VP64s might exist independent of genomic and epigenetic contexts.

Furthermore, we generated TALE-VP64s targeting several other genes, including human *SOX2*, *KLF4*, *c-MYC* and *CDH1* (or *E-Cadherin*) (Supplementary Figure S2). Individual TALE-VP64s were transfected into HEK293T cells and the expression of targeted genes was examined using qRT-PCR. We found that the activation of endogenous genes using individual TALE-VP64s was relatively inefficient. The mRNA level induced by TALE-VP64s targeting *KLF4* and *CDH1* was around 6-fold and 11-fold; while that in *c-MYC* and *SOX2* genes was only around 3-fold (Supplementary Figure S3). Although the most efficient TALE-VP64s

against *KLF4*, *c-MYC* and *CDH1* lied within a similar region as that in mO21, hO3, hO4 and N3 (Supplementary Figure S3A–C), TALE-VP64 targeting a similar region in *CDH1* (E2) and *SOX2* promoter (S2) exhibited low activity or no obvious gene activation (Supplementary Figure S3C, D). These suggested that the ability of TALE-VP64s to activate endogenous genes might be also determined by other factors, such as intrinsic epigenetic modification.

TALE-VP64s activated silenced *Oct4* gene expression in mouse and human somatic cells

In order to activate silenced *Oct4* genes in somatic cells, which are constantly masked by DNA methylation on CpG di-nucleotides in their promoter regions (45), we examined whether TALE-VP64s could activate *mOct4-Luc* reporter methylated *in vitro* by *Escherichia coli* CpG methyltransferase M.SssI. Recent studies have reported that TALE repeats use different codes to discriminate between cytosine (C) and methylated cytosine (mC) (49,50). Compared to HD, which specifically recognizes unmethylated cytosine (C), the code NG (previously known to bind to thymidine) or N* (lacking the residue at position 13 and recognizing both cytosine and thymidine) were found to bind mC with higher affinity (49,50). With the current Golden Gate cloning system (42), we generated a series of mO21 variants, which carried NG, instead of the original HD, to target the potential mC at the fourth, eighth or both nucleotides (Figure 2A). The abilities of these mO21 variants to activate either unmethylated (Figure 2B, upper panel) or *in vitro* methylated *mOct4-Luc* reporters (Figure 2B, lower panel; Supplementary Figure S4A) were examined using luciferase assays. We found that mO21*—which carried NG targeting mC at both the fourth and eighth positions—exhibited higher activity in initiating transcription from the methylated *mOct4-Luc* reporter (Figure 2B), compared with the original mO21 as well as other variants which carried one NG or irrelevant TALE code NI(A) at the fourth and eighth positions (Figure 2A and B). The overall transcriptional activity of mO21 variants was lowered using methylated reporter (Figure 2B), which was probably due to the transcriptional repression associated with DNA methylation. Non-optimal mO21 variants could still activate the reporters, although with reduced activity (Figure 2B). This observation might be explained by the nature that TALE domains could tolerate one or two mismatches (37).

Afterwards, we transfected mO21 and its variants into mouse fibroblast NIH3T3 cells, and examined the level of endogenous *Oct4* transcription using qRT-PCR. Three-fold up-regulation of *Oct4* mRNA induced by mO21* was observed, whereas mO21 showed only marginal activity (Figure 2C, upper panel). In support of this observation, bisulfite sequencing analysis showed that two CpG within the mO21 target sequence were indeed methylated (Figure 2C, lower panel). This suggested that using NG instead of HD for targeting mC is more suitable for activating a silenced gene.

Next, we generated a variant of hO3, named hO3*, by targeting the potential mC in its target sequence with NG (Figure 2D). Similar to mO21*, the hO3* was more effective in activating the transcription of *in vitro* methylated *hOCT4-Luc* reporter, compared to the original hO3 (Figure 2E). Moreover, qRT-PCR analysis showed that hO3* could up-regulate the endogenous *OCT4* transcription 2.5-fold, whereas hO3 showed no obvious enhancement (Figure 2F). Another TALE-VP64 hO4, which targeted a region containing no CpG sites, could also induce endogenous *OCT4* transcription, but to a lesser extent than hO3* (Figure 2F).

To achieve a higher up-regulation of endogenous *Oct4* genes, we further examined the combinatory effect of multiple TALE-VP64s that targeted different regions in the same promoter. TALE-VP64 mOs that targeted CR1, CR2, CR4 and non-conserved regions were combined separately and transfected into NIH3T3 cells (Figure 2G). In addition, two bigger pools of mOs were examined, targeting multiple regions of CR1, CR2, CR4 and non-conserved regions. The highest activity (around 30-fold up-regulation of endogenous *Oct4* transcription; Figure 2G) was observed with a combination of nine effective TALE-VP64s (mO10, 11, 16, 17, 19, 20, 21*, 22 and 25) (individual relative luciferase activities >3, Figure 1B). Next, the combinatory effects of hO TALE-VP64s were also examined. Indeed, the pool of all eight hOs showed a significant synergistic effect in activating the transcription of endogenous *OCT4* gene (Figure 2H). Removal of individual hOs from the combination showed that hO1 and hO3* contributed to the combinatory activity of the pool, whereas hO2 might have minor inhibitory effect (Figure 2H). The highest activity (~20-fold) was achieved using combinations of hOs with numbers 1, 3*, 4 and 6, or 1, 3*, 5, 7 and 8 (Figure 2H). More importantly, we detected OCT4 protein in these multiple TALE-VP64-transfected HEK293T cells using a specific antibody described previously (Figure 2I) (46). Collectively, TALE-VP64s did indeed activate endogenous *OCT4* genes and produce mature proteins.

sgRNA-guided dCas9-VP64s activated silenced *Oct4* genes

To investigate the potential of sgRNA-mediated dCas9-TFs for modulating expression of endogenous *Oct4* genes, we generated dCas9-TFs by fusing dCas9 protein containing two mutations in the RuvC1 and HNH nuclease domains (D10A and H840A) to VP64 or KRAB (Figure 3A) (14). Based on the aforementioned result with TALE-VP64s, we generated eight sgRNAs targeting different 20-bp DNA sequences on either template (T) or non-template (NT) strands in mouse *Oct4* promoter around CR1 and the nearby non-conserved region (Figure 3B; Supplementary Figure S5). These sgRNAs were cloned by annealing 26-nt oligo pairs, followed by insertion into an sgRNA scaffold vector (Supplementary Figure S5A) (44). Each sgRNA was then co-transfected with dCas9-VP64 and *mOct4-Luc* reporter for luciferase assay. We found that five out of the eight sgRNAs significantly activated the transcription of

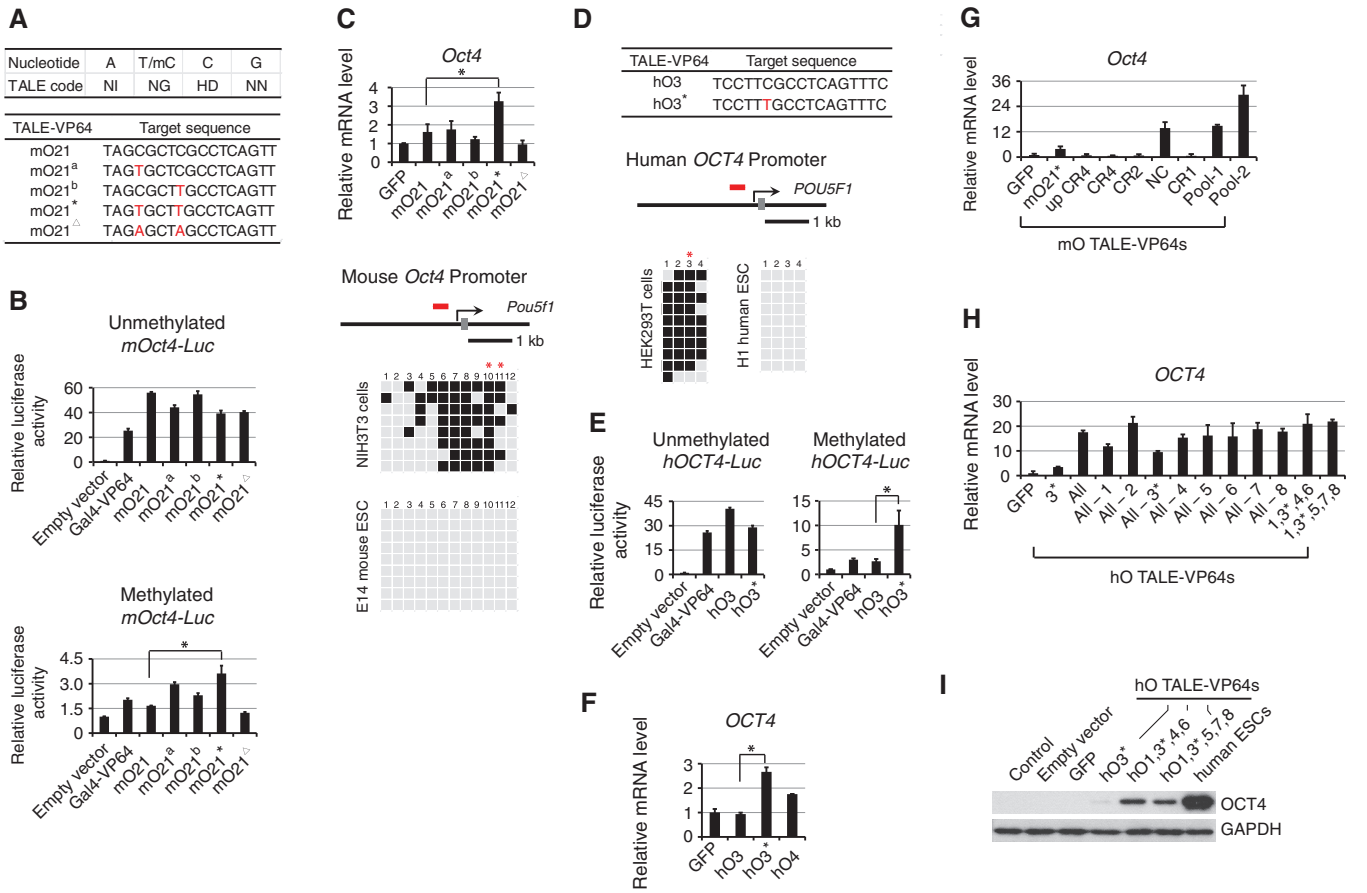

Figure 2. Activation of silenced *Oct4* genes in mouse and human somatic cells by TALE-VP64s. (A) TALE code and variants of mO21 targeting potential mC using NG at the fourth, eighth or both nucleotides in its target sequences. 'T' in bold indicates the use of TALE code NG (T/mC) in the place of original code HD (C). TALE-VP64 mO21$^{\triangle}$ containing TALE code NI (A) in the place of HD (C) at the fourth and eighth nucleotide positions was included as negative control. (B) Activity of TALE-VP64 mO21 and its variants for activating unmethylated (upper panel) and methylated (lower panel) *mOct4-Luc* reporter. Individual TALE-VP64s were co-transfected with *in vitro* methylated/unmethylated reporter plasmids into HEK293T cells. Relative luciferase activity was measured at 48 h after transfection. (C) Activation of endogenous mouse *Oct4* gene by TALE-VP64 mO21 and its variants (upper panel). NIH3T3 cells were transfected and harvested at 48 h after transfection. Relative *Oct4* mRNA level was examined using qRT-PCR. The methylation status of endogenous *Oct4* promoter in NIH3T3 cells was analyzed using bisulfite sequencing (lower panel). Mouse ESCs E14 were included as a control. The grey bar in the schematic diagram indicates the analyzed region (from −463 to −33 bp). Grey squares represent unmethylated CpG di-nucleotides and black squares represent methylated ones. The positions of the fourth and eighth nucleotide of the mO21 target sequence are indicated by asterisks. (D) TALE-VP64 hO3* was designed to target potential mC using NG at the sixth position in its target sequence (upper panel). The DNA methylation status of human *OCT4* promoter in HEK293T cells was analyzed using bisulfite sequencing (lower panel). Human ESCs H1 were included as control. Grey squares indicate unmethylated CpG dinucleotides and black squares indicate methylated CpG di-nucleotides. Cytosine in the sixth nucleotide of hO3 target sequence is indicated with a asterisk. The analyzed region (from −217 to −32 bp) is indicated as grey bar in the schematic diagram. (E) Effects of hO3 and hO3* on unmethylated (left panel) and methylated (right panel) *hOCT4-Luc* reporters. Individual TALE-VP64s were co-transfected with *in vitro* methylated/unmethylated *hOCT4-Luc* reporters into HEK293T cells. Relative luciferase activity for each sample was examined at 48 h after transfection. (F) Activation of endogenous human *OCT4* gene by hO3 and hO3*. qRT-PCR analysis was performed at 48 h after transfection in HEK293T cells. (G) Analysis of combinations of TALE-VP64 mOs for activating the endogenous mouse *Oct4* gene in NIH3T3 cells. Combinations of mOs targeting CR4 (mO5, 6, 7, 8), non-conserved regions upstream of CR4 (up CR4) (mO1, 2, 3, 4), CR2 (mO9, 10, 11, 12, 13, 14, 15, 16), non-conserved region between CR2 and CR1 (NC) (mO17, 18, 19, 20, 21, 22) as well as CR1 (mO23, 24, 25, 26) were analyzed separately as well as in combination. Pool-1: all mOs targeting CR1, CR2 NC and CR4; Pool-2: mO10, 11, 16, 17, 19, 20, 21*, 22 and 25. qRT-PCR were performed at 48 h after transfection. (H) Analysis of various combinations of TALE-VP64 hOs for activating the endogenous human *OCT4* gene in HEK293T cells. Significant transcriptional synergy was observed with multiple combinations. (I) OCT4 proteins were detected in HEK293T cells at 48 h after transfection with single or combinations of TALE-VP64 hOs. Western blot was performed with anti-OCT4. GAPDH was included as control for equal loading. Data were shown as mean ± SEM ($n = 3$).* $P < 0.05$.

mOct4-Luc reporter (Figure 3B). Interestingly, the most effective sgRNAs T2 and NT3 targeted regions ranging from −147 to −89-bp upstream of the TSS (Figure 3B; Supplementary Figure S5B), which was similar to the highly effective TALE-VP64s. Deletion of either RuvC1 or HNH nuclease domains, each of which is responsible for binding to one strand of target DNA (14), abolished transcriptional activation mediated by dCas9-VP64 or transcriptional repression mediated by dCas9-KRAB (Figure 3C).

Similar to TALE-VP64s, when multiple sgRNAs were applied in combination, synergistic effects were observed (Figure 3D). The combinations of sgRNAs, NT1 to NT4 or T1, T2, NT2 and NT3, further boosted the transcription of

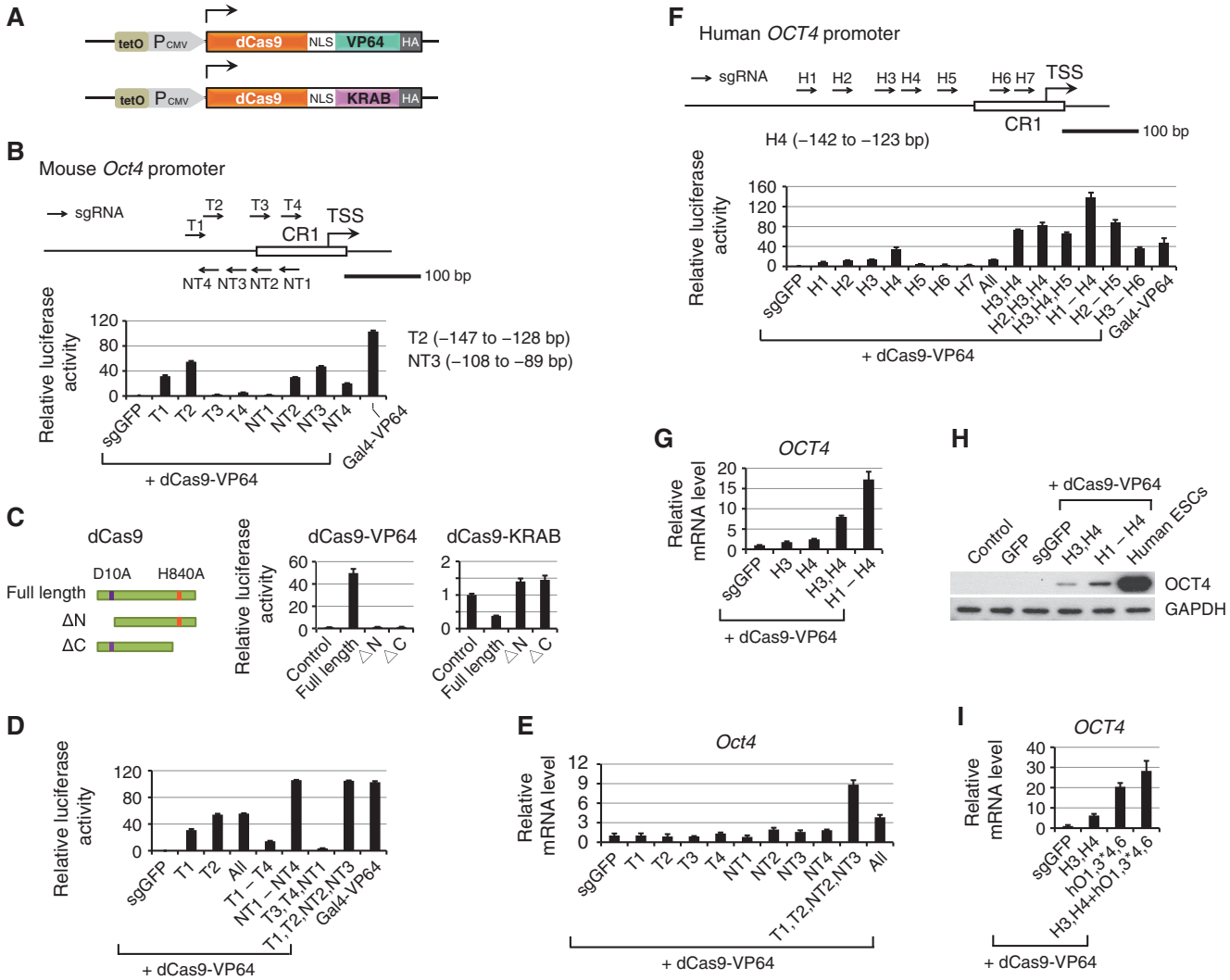

Figure 3. Activation of mouse and human *Oct4* promoters by sgRNA-guided dCas9-VP64s. (A) Schematic representation of dCas9-VP64 and dCas9-KRAB fusion proteins. (B) Schematic diagram of eight sgRNAs (T1–T4, NT1–NT4) and their target sites in mouse *Oct4* promoter. Effects of individual sgRNAs were examined via co-transfection of full-length dCas9-VP64 and *mOct4-Luc* reporter into HEK293T cells. sgRNA targeting GFP (sgGFP) was included as a negative control. Relative luciferase activity was measured at 48 h after transfection. (C) Analysis of dCas9 protein. Full-length dCas9 protein, dCas9 protein with deletion of N-terminal RuvC1 domain (ΔN) or deletion of C-terminal HNH domain (ΔC) were generated (left panel) and fused to VP64 or KRAB separately. Individual fusion proteins were co-transfected with sgRNA T2 and *mOct4-Luc* reporter. VP64 fusion proteins were examined for their ability to mediate sgRNA T2-guided transcriptional activation of *mOct4-Luc* reporter in HEK293T cells (middle panel) and KRAB fusion proteins were examined for their ability to mediate sgRNA T2-guided transcriptional repression of *mOct4-Luc* reporter in mouse ESCs E14 (right panel). Relative luciferase activity was measured at 48 h after transfection. (D) Analysis of combinatory effects of sgRNAs. Single or combined sgRNAs were co-transfected with dCas9-VP64 and *mOct4-Luc* reporter into HEK293T cells. Relative luciferase activity was measured at 48 h after transfection. Combination of sgRNA (NT1–NT4), as well as sgRNA (T1, T2, NT2 and NT3) showed significant synergistic effect in activating transcription of *mOct4-Luc* reporter. (E) Activation of endogenous mouse *Oct4* genes by sgRNA-guided dCas9-VP64s. Individual or combined sgRNAs were co-transfected with dCas9-VP64 into NIH3T3 cells. Up-regulation of endogenous *Oct4* genes was examined by qRT-PCR analysis at 48 h after transfection. (F) Schematic diagram of seven sgRNAs and their corresponding target sites in human *OCT4* promoter (upper panel). Effects of individual and combined sgRNA were examined by luciferase assay (lower panel). HEK293T cells were co-transfected with sgRNA(s), dCas9-VP64 and *hOCT4-Luc* reporter and relative luciferase activity was measured at 48 h after transfection. Combination of sgRNAs H3 and H4, as well as H1–H4 showed a significant synergistic effect in activating transcription of *hOCT4-Luc* reporter. (G) Activation of endogenous human *OCT4* gene by sgRNA-guided dCas9-VP64s. Individual or combined sgRNAs were transfected with dCas9-VP64s into HEK293T cells. Up-regulation of endogenous *OCT4* gene was examined using qRT-PCR analysis at 48 h after transfection. (H) OCT4 proteins were detected in HEK293T cells at 48 h after transfection with dCas9-VP64 and combinations of sgRNAs (H3 and H4, or H1–H4). Western blot was performed with anti-OCT4. GAPDH was included as control for equal loading. (I) Combinatory effect of sgRNA/dCas9-VP64s and hO TALE-VP64s. sgRNAs H3 and H4 and TALE-VP64 (hO1, 3*, 4 and 6) were transfected together with dCas9-VP64 into HEK293T cells. Expression of endogenous *OCT4* gene was examined using qRT-PCR at 48 h after transfection. Data were shown as mean ± SEM ($n = 3$).

mOct4-Luc reporter about 2-fold, compared with the most potent single sgRNA T2 (Figure 3D). Moreover, we found that the combination of multiple competent sgRNAs could induce endogenous *Oct4* mRNA ~9-fold in NIH3T3 cells, whereas individual sgRNAs showed marginal or no increase (Figure 3E). This indicated that combined sgRNAs could mediate dCas9-VP64 to activate silenced *Oct4* genes in mouse somatic cells.

Next, we generated seven sgRNAs targeting CR1 as well as the non-conserved region in human *OCT4* promoter (Figure 3F; Supplementary Figure S5B). Luciferase assay showed that four out of the seven sgRNAs (H1–H4) could activate the transcription of *hOCT4-Luc* reporter (Figure 3F). Of these four sgRNAs, H4—which targeted −142 to −123-bp-region upstream of the TSS—exhibited the greatest activity. On the other hand, sgRNAs H5, H6 and H7, which targeted the −85 to −14-bp region, induced marginal transcriptional activation (Figure 3F). This was consistent with the result using sgRNAs targeting the mouse *Oct4* promoter, where T3, T4 and NT1, targeting the −87 to −30-bp region, showed significantly lower activity, compared with T2 and NT3, which targeted the −147 to −128-bp and −108 to −89-bp regions (Figure 3B; Supplementary Figure S5B). These results suggested that dCas9-VP64 might function with optimal activity when being recruited to the region around −147 to −89-bp upstream of the TSS, thus it further clarified the proximal TSS region for sgRNA targeting reported in Mali *et al.*'s study (37). The minor difference in the optimal targeting locations between TALE-VP64s and sgRNA/dCas9-VP64s might be caused by their size and structural difference.

Consistently, combination of these sgRNAs showed a synergistic effect in activating the *hOCT4-Luc* reporter (Figure 3F). Among various combinations examined, the pool of H1–H4 showed the highest activity (Figure 3F). This pool of sgRNAs was also able to elevate endogenous *OCT4* transcription around 17-fold (Figure 3G) and induced OCT4 protein in transfected HEK293T cells, though at a lower level compared to that in human ESCs (Figure 3H). Interestingly, we also observed transcriptional synergy when sgRNAs (H3 and H4) and TALE-VP64s (hO1, 3*, 4 and 6) were applied simutanously (Figure 3I), suggesting that these engineered TFs could be used in combination for activating gene expression.

TALE-VP64-induced activation of endogenous *OCT4* was transient

Next, we examined whether the observed up-regulation of endogenous *Oct4* gene could be stably maintained. HEK293T cells transfected with single or combinations of hO TALE-VP64s were sub-cultured every 2 days and maintained for three passages. Specific primers amplifying a common fragment in all hO TALE-VP64s showed that the overall expression of TALE-VP64s decreased quickly after the second passage (Supplementary Figure S6). In accordance with the decrease of TALE-VP64 expression, we observed that the initial up-regulation of endogenous *OCT4* transcript induced by the combined hO TALE-VP64s also decreased rapidly and was undetectable after three passages (Figure 4A). This indicated that the up-regulation of endogenous *OCT4* transcription was dependent on the exogenous TALE-VP64s, and *OCT4* transcriptional activation could not be maintained when the ectopic expression of hO TALE-VP64s decreased due to instability of plasmids. Bisulfite sequencing analysis on these TALE-VP64-transfected HEK293T cells consistently showed that genomic DNA in *OCT4* promoter remained to be hypermethylated, even when its transcription was elevated by the exogenous TALE-VP64s (Figure 4B). Beside the region that contains hO3 target sequence (Figure 2D), analysis on another region covering the TSS showed a similar unchanged DNA hypermethylation pattern (Figure 4B, region 2). This result suggested that transient up-regulation of endogenous *OCT4* transcription did not reverse the DNA methylation status and thus could not establish sustainable *OCT4* expression. Similarly to TALE-VP64s, we found that the dCas9-VP64-mediated activation of *OCT4* gene also decreased rapidly in subsequent passages (Supplementary Figure S7). These results suggested that neither TALE-VP64 nor dCas9-VP64-induced activation of *OCT4* gene could reverse its epigenetic repression and achieve stable expression in the long term.

Furthermore, we examined whether TALE-VP64-induced activation of *OCT4* gene could activate other pluripotency genes, including downstream targets of OCT4 protein in ESCs. However, qRT-PCR analysis of the HEK293T cells transfected with single or combined hO TALE-VP64s showed that activation of *OCT4* had no effect on the transcriptional activation of *SOX2*, *KLF4*, *NANOG*, *c-MYC* and *CDH1* genes, which work in the same pluripotency network (Figure 4C). This is consistent with the previous notion that endogenous pluripotency genes could not be activated directly by forced expression of TFs during reprogramming (30,51); instead, they were probably up-regulated via an indirect path by multiple factors, after the global epigenetic status had been significantly modified.

Next we examined whether activation of other pluripotency genes has an effect on *OCT4* transcription. With the TALE-VP64s targeting *SOX2*, *KLF4*, *NANOG*, *c-MYC* and *CDH1,* which were generated previously (Supplementary Figure S2), we identified combinations that could increase the transcription level of endogenous *SOX2* by around 26-fold (S1, S2 and S3), *KLF4* by around 9-fold (K2, K3 and K4), and *CDH1* by around 27-fold (E2, E3 and E4) (Figure 4D and Supplementary Figure S8A). Similar to *OCT4* activation, activation of *SOX2*, *KLF4* and *CDH1* showed no obvious effect on transcription of other pluripotency genes (Figure 4D). Combinations of TALE-VP64s to *NANOG* and *c-MYC* induced relatively low transcriptional activation of their endogenous genes (∼2- and 4-fold, respectively) (Supplementary Figure S8), which might be explained by its stringent epigenetic repression (*NANOG*) or the high basal level of the endogenous gene (*c-MYC*).

p300 enhanced the activation of endogenous *OCT4* induced by TALE-VP64s and dCas9-VP64s

Histone H3 lysine 27 acetylation (H3K27ac) and histone H3 lysine 4 trimethylation (H3K4me3) are known epigenetic modifications for maintaining the *Oct4* locus in an active or permissive status in pluripotent cells; whereas histone H3 lysine 9 and 27 trimethylation (H3K9me3 and H3K27me3) mediate epigenetic repression

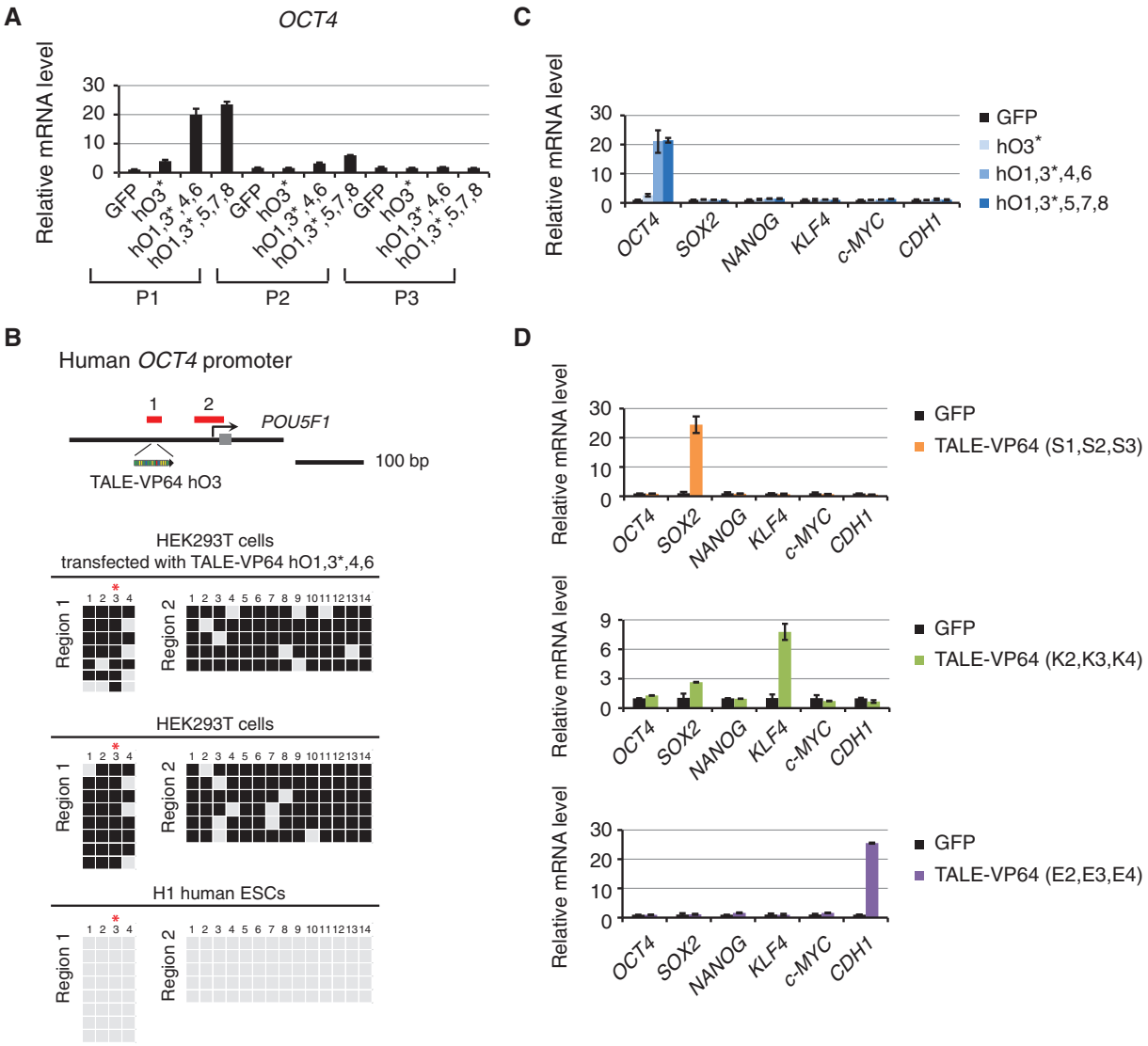

Figure 4. Transient activation of endogenous *OCT4* gene by TALE-VP64s. (A) Expression of endogenous *OCT4* gene was examined in three consecutive passages after transfection with single or combined TALE-VP64 hOs in HEK293T cells. qRT-PCR analysis was performed after each passage, i.e. at Day 2 (P1), Day 4 (P2) and Day 6 (P3) after transfection. Rapid down-regulation of *OCT4* was observed after second passage. (B) Methylation status of *OCT4* promoter in the HEK293T cells transfected with the combination of TALE-VP64 hO1, 3*, 4 and 6. Genomic DNA was collected at 48 h (P1) after transfection and analyzed using bisulfite sequencing. Grey squares indicate unmethylated CpG di-nucleotides and black squares indicate methylated ones. Grey bars indicate region 1 (from −217 to −32 bp) and region 2 (from −21 to 309 bp) being analyzed by bisulfite sequencing. Region 1 is identical to that in Figure 2D. Human ESCs and untransfected HEK293T cells were included as controls. (C) qRT-PCR analysis of the expression of pluripotency genes interconnected with *OCT4*. HEK293T cells transfected with single or combined TALE-VP64 hOs were collected at 48 h after transfection. Expression of endogenous *OCT4*, *SOX2*, *NANOG*, *KLF4*, *c-MYC* and *CDH1* were analysed using qRT-PCR. (D) Activation of endogenous *SOX2*, *KLF4* and *CDH1* by corresponding TALE-VP64s. HEK293T cells transfected with different TALE-VP64 combinations were collected at 48 h after transfection. Expression of endogenous *OCT4*, *SOX2*, *NANOG*, *KLF4*, *c-MYC* and *CDH1* were analyzed using qRT-PCR. Data were shown as mean ± SEM ($n = 3$).

and are associated with *Oct4* gene silencing in somatic cells. Hence, we examined whether introducing epigenetic modification enzymes that catalyse active epigenetic marks could facilitate the activation of silenced *OCT4* genes in the presence of corresponding TALE-VP64s or sgRNA/dCas9-VP64s. We transfected H3K27 demethylase JMJD3 (52), H3K9 demethylase JMJD2B (53) or histone acetyltransferase p300 (54) individually with either hO3* or the combination of hO1, 3*, 4 and 6 into HEK293T cells. Analysis by qRT-PCR showed that p300 significantly enhanced the endogenous *OCT4* transcription induced by these TALE-VP64s, either in a single form (Figure 5A) or in a combination (Figure 5B). Further analysis of these epigenetic modifiers using *hOCT4-Luc* reporter plasmid showed that they did not enhance the TALE-VP64-induced activation of *OCT4* reporter (Supplementary Figure S9). This suggested that the enhancement effect of p300 on activating transcription was dependent on its epigenetic modification function in a chromatin context.

Next, we examined whether these active epigenetic modifiers could also enhance the transcription activation

mediated by dCas9-VP64s. sgRNAs H3 and H4, which targeted the human *OCT4* promoter, were co-transfected with dCas9-VP64 and JMJD3, or JMJD2B, or p300 in HEK293T cells. Analysis by qRT-PCR at 48 h post-transfection showed that p300 also enhanced the transcriptional activation induced by sgRNA-guided dCas9-VP64 (Figure 5C). Importantly, when the endogenous *p300* was depleted using specific shRNAs, we observed a reduction of *OCT4* mRNA in the presence of TALE-VP64s or sgRNAs/dCas-VP64s (Figure 5D), suggesting that p300 was indeed involved in the activation of endogenous *OCT4* transcription induced by the engineered TFs. We also observed that the alteration of *p300* level did not influence the expression of TALE-VP64s (Figure 5E), eliminating the possibility that p300 enhanced the activation of *OCT4* transcription indirectly through modulating the level of engineered TFs. Taken together, these data indicated that p300 could facilitate the activation of endogenous *OCT4* gene mediated by TALE-VP64s or sgRNA/dCas9-VP64s.

Since TALEs and sgRNAs can tolerate one or two mismatches (37), they may target undesired DNA sequences due to close similarity of the sequences and cause off-target effect. To assess if p300 also enhanced activation of undesired genes due to off-target effect, we carried out genome wide search and identified six genes which carry two-mismatch, and 29 genes carrying three-mismatch of hO3* target sequence in their proximal promoters (maximum 250-bp upstream of the TSSs) (Supplementary Table S2). Complete match or one-mismatch sequences could not be found within proximal region in any promoters. We selected 20 potential off-target genes and analyzed their expression in the HEK293T cells that had been transfected with hO3* or combination of hO1, 3*, 4, 6, in the presence or absence of p300. With the valid expression data of thirteen genes out of the twenty, we found that other than *OCT4*, all the mRNA levels remained unchanged in cells transfected with these TALE-VP64s and p300 (Supplementary Figure S10). This suggested that off target effect is not a primary concern for the TALE-VP64s-induced activation of endogenous *OCT4*.

DISCUSSION

Direct activation of pluripotency gene *Oct4* has great potential to facilitate the reinstatement of pluripotency in somatic cells, opening up an opportunity to improve the current technology for iPSC generation. However, due to the complexity of the epigenetic control that stringently represses the *Oct4* gene in somatic cells, direct activation of silenced *Oct4* genes has not been achieved. The recent advent of engineered TFs TALE-TF and CRISPR/Cas9-TF has opened up new avenues for manipulating the transcription of *Oct4* gene by directly targeting its promoter and thus has the potential to bypass epigenetic repression and activate silenced *Oct4* genes directly.

In this article, we systematically examined a number of TALE-VP64s that target a wide range of loci in mouse and human *Oct4* promoters. Using luciferase assay, we identified the highly effective TALE-VP64 mO21, hO3 and hO4, which targeted from −120 to −104-bp upstream of the TSS in mouse *Oct4* promoter and from −113 to −80 bp in human *OCT4* promoter, respectively. Interestingly, moving the target sequences of several inefficient TALE-VP64s from their original locations into the −120 to −104-bp region of mouse *Oct4* promoter significantly increased the transcriptional activation induced by the same TALE-VP64s (Figure 1E), suggesting that the target position was greatly implicated in the activity of TALE-VP64s. A similar positional influence was also observed in dCas9-VP64s. Individual sgRNAs (T2, NT3 and H4) targeting mouse and human *Oct4* promoters around the −147 to −89-bp region could effectively activate the corresponding luciferase reporters in the presence of dCas9-VP64s; whereas sgRNAs (T3, T4, NT1, H5, H6 and H7) targeting regions closer to the TSS (from −87 to −14 bp) showed lower activity. This is consistent with previous notions that wild-type VP16 complex could bring strong activity from a proximal position (55) and engineered TFs targeting the proximal region of promoters tend to exhibit high activity (35–37). The observed positional influence is also supported by the Bultmann *et al.*'s and Gao *et al.*'s studies, which ruled out the potential affinity variance between specific TALE domains and their targets by fluorescence polarization assay and chromatin immunoprecipitation (31,32).

However, detailed comparison with recent studies showed a rather diverse distribution of the promoter regions targeted by the most effective TALE or dCas9 activators (Supplementary Figures S11 and S12) (31,32,34,37,56). Moreover, the positional influence observed in luciferase assay was not apparent for endogenous gene activation (Supplementary Figures S3 and S13) (33–35,38). Therefore, besides the position of target regions, other factors are likely to play a role in determining the activity of a TALE- or dCas9-activator. These include the size and structure of the activators, the epigenetic repression in a local genomic context, as well as factors that have not been identified yet. Currently, there is still a lack of mechanistic studies in this phenomenon. Based on the previous investigations on VP16, the positional influence observed is possibly introduced through its interactions with components in the transcription machinery (57).

Using an optimized TALE code NG for mC and combinations of multiple TALE-VP64s, we observed that transcriptions of endogenous *Oct4* genes were increased by around 30-fold in mouse NIH3T3 cells and 20-fold in human HEK293T cells. Similarly to TALE-VP64s, combination of multiple sgRNAs showed a synergistic effect. The combined application of effective sgRNAs could activate the silenced *Oct4* genes in both NIH3T3 and HEK293T cells, whereas individual sgRNAs showed no obvious activity. Importantly, ours is the first report to show that the activation of endogenous *OCT4* genes by multiple TALE- or dCas9-VP64s could produce mature proteins, providing additional evidence that simultaneous application of engineered TFs is an effective approach for modulating

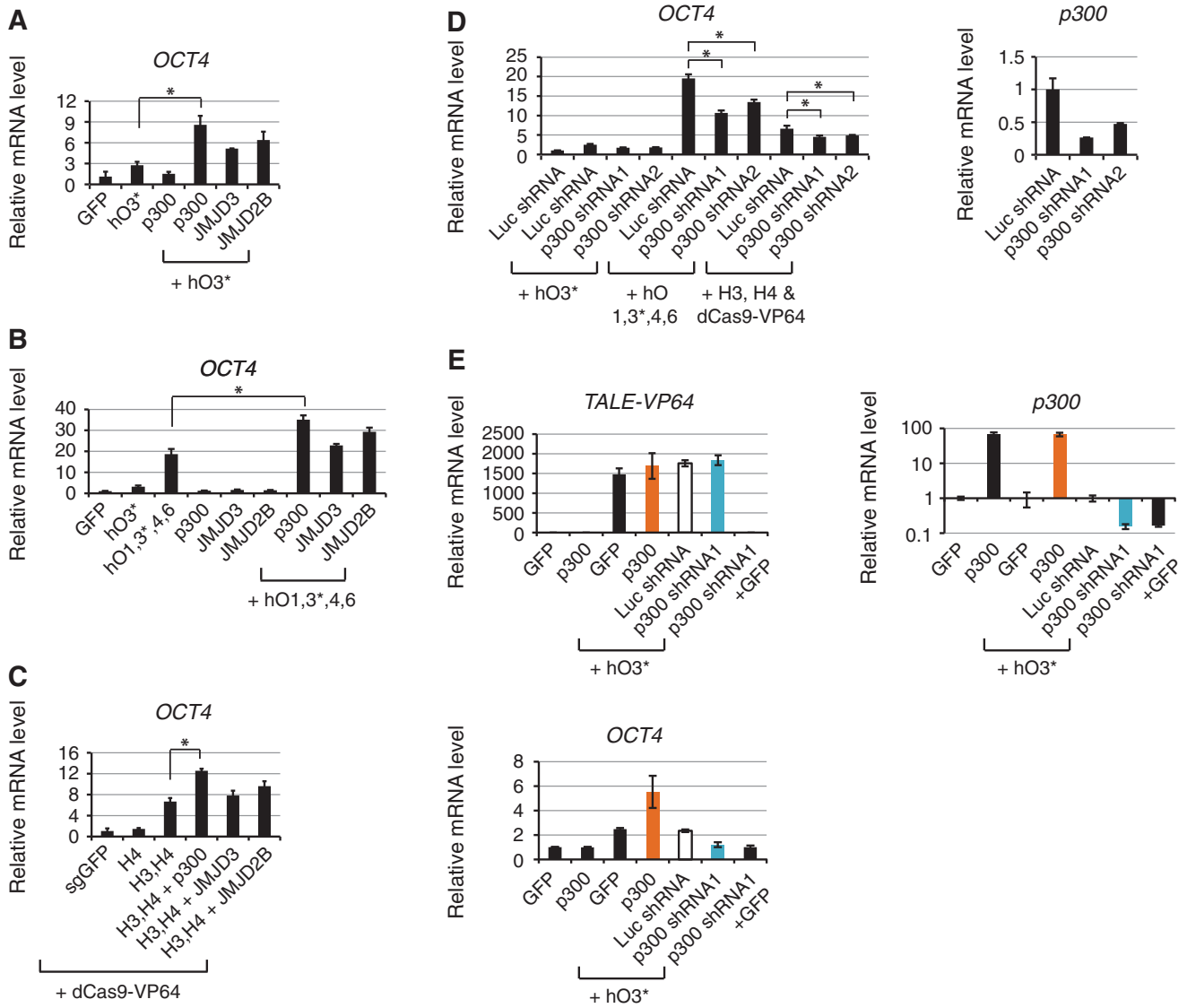

Figure 5. Enhanced activation of silenced human *OCT4* genes by p300, in combination with TALE-V64s or sgRNA/dCas9-VP64s. (A) Epigenetic modifiers p300, JMJD3 or JMJD2B were co-transfected with TALE-VP64 hO3* into HEK293T cells. Relative expression of endogenous *OCT4* gene in each sample was analyzed at 48 h after transfection. (B) p300, JMJD3 or JMJD2B were co-transfected with the combination of TALE-VP64s hO1, 3*, 4 and 6 into HEK293T cells. Expression of endogenous *OCT4* gene was analyzed using qRT-PCR at 48 h after transfection. (C) p300, JMJD3 or JMJD2B were co-transfected with sgRNAs H3 and H4 and dCas9-VP64s into HEK293T cells. Expression of endogenous *OCT4* gene was analyzed using qRT-PCR at 48 h after transfection. (D) Individual shRNAs targeting *p300* was co-transfected with the combinations of TALE-VP64 hO1, 3*, 4 and 6, or the combination of sgRNAs H3–H4 and dCas9-VP64 into HEK293T cells. Expression of endogenous *OCT4* gene (left panel) as well as the level of *p300* (right panel) were analyzed using qRT-PCR at 48 h after transfection. (E) Influence of *p300* levels on the expression of TALE-VP64s. TALE-VP64 hO3* was co-transfected with either *p300* overexpressing plasmid or *p300* shRNA plasmids into HEK293T cells. The expression of transfected TALE-VP64, level of *p300* as well as the increase of endogenous *OCT4* were examined using qRT-PCR analysis at 48 h after transfection. Data were shown as mean ± SEM ($n = 3$). *$P < 0.05$.

gene expression. Interestingly, the combined application of multiple TFs for endogenous gene activation can also overcome the potential off-target effect that might be introduced by using single activators (37). Compared to TALE-VP64s, sgRNA-guided dCas9-VP64s showed comparable activity in boosting the transcription of *Oct4* reporters, but lowered activity for inducing transcription from silenced endogenous genes. This is consistent with recent studies which used sgRNA/dCas9-TFs for activating endogenous genes in mammalian cells (18–20,35,36).

With further examination, we showed that the activation of endogenous *OCT4* genes induced by TALE-VP64s or sgRNA/dCas9-VP64 was transient in HEK293T cells. The *OCT4* promoter remained to be hypermethylated even when the endogenous *OCT4* transcription was significantly up-regulated by a combination of hO TALE-VP64s. Our work pinpointed a challenge in using these engineered TFs to modulate endogenous gene expression in mammalian cells, as elevated *Oct4* transcription alone might not be sufficient to reverse epigenetic repression for establishing stable expression.

In this article, we also attempted to introduce active epigenetic modifiers to facilitate the activation of silenced *Oct4* genes using TALE-VP64s or dCas9-VP64s. We found that p300, a well-known transcription co-activator that possess histone acetyltransferase activity (54,58), significantly enhanced the activation of *Oct4* genes induced by TALE-VP64s and dCas9-VP64s. This finding illustrated the potential usage of engineered TFs and epigenetic modifiers simultaneously to achieve a more efficient gene modulation.

Generation of iPSCs has been broadly demonstrated to be a slow and stochastic process. To date, it is still unclear how endogenous pluripotency genes, especially *Oct4*, are activated to reinstate the cellular pluripotency state through reprogramming. Some epigenetic regulations associated with the repression of *Oct4* genes upon differentiation have previously been revealed, such as the H3K9 methylation mediated by G9a and subsequent DNA methylation by Dnmt3a/3b proteins (59). However, shRNA silencing of G9a and genetic depletion of Dnmt3a/3b exhibited little effect on facilitating activation of silenced *Oct4* genes or enhancing cellular reprogramming (60,61), suggesting that additional mechanisms could have been involved in repressing *Oct4* genes. Thus, current platform using TALE-VP64s or sgRNA/dCas9-VP64s for direct transcriptional activation of silenced *Oct4* genes could provide a valuable tool to demystify new regulators involved in repressing *Oct4* in somatic cells.

More recently, the potential of using TALE-VP64s targeting *Oct4* to promote epigenetic reprogramming and enhance iPSC generation has been examined in Gao *et al.*'s study (32). Indeed, TALE-VP64s targeting mouse *Oct4* enhancer facilitated the activation of silenced *Oct4* genes and enhanced the reprogramming efficiency from mouse fibroblasts into iPSCs; however, activation of *Oct4* by TALE-VP64s alone could not induce successful reprogramming (32). Our data provided an explanation for this observation as TALE-VP64s alone might not be able to overcome epigenetic repression and establish sustainable expression of *Oct4* gene in fibroblasts. Further investigation into the mechanism of *Oct4* silencing and reactivation is needed to achieve more efficient activation of *Oct4* gene as well as iPSC generation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Huck-Hui Ng and Jianwei Ren for critical comments on the manuscript.

FUNDING

Research Grants Council of Hong Kong [CUHK 464411 to B.F; CUHK 478812 to B.F.; HKUST T13-607/12R to Y.I.]; National Natural Science Foundation of China [NSFC 31171433 to B.F. (in part)]; Shenzhen Basic Research Funding [JC201104220293A to B.F. (in part)]; National Science Foundation, USA [DBI0844749 to J.G. (in part)]. Funding for open access charges: Research Grants Council of Hong Kong.

Conflict of interest statement. None declared.

REFERENCES

1. Blancafort,P., Magnenat,L. and Barbas,C.F. 3rd (2003) Scanning the human genome with combinatorial transcription factor libraries. *Nat. Biotechnol.*, **21**, 269–274.
2. Sadowski,I., Ma,J., Triezenberg,S. and Ptashne,M. (1988) GAL4-VP16 is an unusually potent transcriptional activator. *Nature*, **335**, 563–564.
3. Yao,F., Svensjo,T., Winkler,T., Lu,M., Eriksson,C. and Eriksson,E. (1998) Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Hum. Gene Ther.*, **9**, 1939–1950.
4. Maeder,M.L., Thibodeau-Beganny,S., Osiak,A., Wright,D.A., Anthony,R.M., Eichtinger,M., Jiang,T., Foley,J.E., Winfrey,R.J., Townsend,J.A. *et al.* (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell*, **31**, 294–301.
5. 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. Cong,L., Zhou,R., Kuo,Y.C., Cunniff,M. and Zhang,F. (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.*, **3**, 968.
8. Christian,M.L., Demorest,Z.L., Starker,C.G., Osborn,M.J., Nyquist,M.D., Zhang,Y., Carlson,D.F., Bradley,P., Bogdanove,A.J. and Voytas,D.F. (2012) Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One*, **7**, e45383.
9. Witzgall,R., O'Leary,E., Leaf,A., Onaldi,D. and Bonventre,J.V. (1994) The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl Acad. Sci. USA*, **91**, 4514–4518.
10. 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.
11. Li,Y., Moore,R., Guinn,M. and Bleris,L. (2012) Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Sci. Rep.*, **2**, 897.
12. Bhaya,D., Davison,M. and Barrangou,R. (2011) CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet.*, **45**, 273–297.
13. Wiedenheft,B., Sternberg,S.H. and Doudna,J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331–338.
14. 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.
15. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
16. 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.

17. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
18. 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**, 422–451.
19. 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.
20. Bikard,D., Jiang,W., Samai,P., Hochschild,A., Zhang,F. and Marraffini,L.A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.*, **141**, 7429–7437.
21. Nichols,J., Zevnik,B., Anastassiadis,K., Niwa,H., Klewe-Nebenius,D., Chambers,I., Scholer,H. and Smith,A. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*, **95**, 379–391.
22. Niwa,H., Miyazaki,J. and Smith,A.G. (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.*, **24**, 372–376.
23. Takahashi,K. and Yamanaka,S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**, 663–676.
24. 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.
25. Kim,J.B., Sebastiano,V., Wu,G., Arauzo-Bravo,M.J., Sasse,P., Gentile,L., Ko,K., Ruau,D., Ehrich,M., van den Boom,D. *et al.* (2009) Oct4-induced pluripotency in adult neural stem cells. *Cell*, **136**, 411–419.
26. Yeom,Y.I., Fuhrmann,G., Ovitt,C.E., Brehm,A., Ohbo,K., Gross,M., Hubner,K. and Scholer,H.R. (1996) Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development*, **122**, 881–894.
27. Nordhoff,V., Hubner,K., Bauer,A., Orlova,I., Malapetsa,A. and Scholer,H.R. (2001) Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mamm. Genome Official J. Int. Mamm. Genome Soc.*, **12**, 309–317.
28. Freberg,C.T., Dahl,J.A., Timoskainen,S. and Collas,P. (2007) Epigenetic reprogramming of OCT4 and NANOG regulatory regions by embryonal carcinoma cell extract. *Mol. Biol. Cell*, **18**, 1543–1553.
29. Bao,S., Tang,F., Li,X., Hayashi,K., Gillich,A., Lao,K. and Surani,M.A. (2009) Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature*, **461**, 1292–1295.
30. Polo,J.M., Anderssen,E., Walsh,R.M., Schwarz,B.A., Nefzger,C.M., Lim,S.M., Borkent,M., Apostolou,E., Alaei,S., Cloutier,J. *et al.* (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell*, **151**, 1617–1632.
31. Bultmann,S., Morbitzer,R., Schmidt,C.S., Thanisch,K., Spada,F., Elsaesser,J., Lahaye,T. and Leonhardt,H. (2012) Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res.*, **40**, 5368–5377.
32. Gao,X., Yang,J., Tsang,J.C.H., Ooi,J., Wu,D. and Liu,P. (2013) Reprogramming to Pluripotency Using Designer TALE Transcription Factors Targeting Enhancers. *Stem. Cell Reports*, **1**, 183–197.
33. Maeder,M.L., Linder,S.J., Reyon,D., Angstman,J.F., Fu,Y., Sander,J.D. and Joung,J.K. (2013) Robust, synergistic regulation of human gene expression using TALE activators. *Nat. Methods*, **10**, 243–245.
34. Perez-Pinera,P., Ousterout,D.G., Brunger,J.M., Farin,A.M., Glass,K.A., Guilak,F., Crawford,G.E., Hartemink,A.J. and Gersbach,C.A. (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat. Methods*, **10**, 239–242.
35. 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.
36. Maeder,M.L., Linder,S.J., Cascio,V.M., Fu,Y., Ho,Q.H. and Joung,J.K. (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods*, **10**, 977–979.
37. 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.
38. Cheng,A.W., Wang,H., Yang,H., Shi,L., Katz,Y., Theunissen,T.W., Rangarajan,S., Shivalila,C.S., Dadon,D.B. and Jaenisch,R. (2013) Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.*, **23**, 1163–1171.
39. Kwan,K.M., Fujimoto,E., Grabher,C., Mangum,B.D., Hardy,M.E., Campbell,D.S., Parant,J.M., Yost,H.J., Kanki,J.P. and Chien,C.B. (2007) The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.*, **236**, 3088–3099.
40. Qi,D. and Scholthof,K.B. (2008) A one-step PCR-based method for rapid and efficient site-directed fragment deletion, insertion, and substitution mutagenesis. *J. Virol. Methods*, **149**, 85–90.
41. Hockemeyer,D., Soldner,F., Cook,E.G., Gao,Q., Mitalipova,M. and Jaenisch,R. (2008) A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem. Cell*, **3**, 346–353.
42. 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.
43. Lei,Y., Guo,X., Liu,Y., Cao,Y., Deng,Y., Chen,X., Cheng,C.H., Dawid,I.B., Chen,Y. and Zhao,H. (2012) Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc. Natl Acad. Sci. USA*, **109**, 17484–17489.
44. 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.
45. Feng,B., Jiang,J., Kraus,P., Ng,J.H., Heng,J.C., Chan,Y.S., Yaw,L.P., Zhang,W., Loh,Y.H., Han,J. *et al.* (2009) Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.*, **11**, 197–203.
46. Chia,N.Y., Chan,Y.S., Feng,B., Lu,X., Orlov,Y.L., Moreau,D., Kumar,P., Yang,L., Jiang,J., Lau,M.S. *et al.* (2010) A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature*, **468**, 316–320.
47. Chen,X., Xu,H., Yuan,P., Fang,F., Huss,M., Vega,V.B., Wong,E., Orlov,Y.L., Zhang,W., Jiang,J. *et al.* (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, **133**, 1106–1117.
48. Wiznerowicz,M. and Trono,D. (2003) Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference. *J. Virol.*, **77**, 8957–8961.
49. Deng,D., Yin,P., Yan,C., Pan,X., Gong,X., Qi,S., Xie,T., Mahfouz,M., Zhu,J.K., Yan,N. *et al.* (2012) Recognition of methylated DNA by TAL effectors. *Cell Res.*, **22**, 1502–1504.
50. Valton,J., Dupuy,A., Daboussi,F., Thomas,S., Marechal,A., Macmaster,R., Melliand,K., Juillerat,A. and Duchateau,P. (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem.*, **287**, 38427–38432.
51. Buganim,Y., Faddah,D.A., Cheng,A.W., Itskovich,E., Markoulaki,S., Ganz,K., Klemm,S.L., van Oudenaarden,A. and Jaenisch,R. (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell*, **150**, 1209–1222.
52. Swigut,T. and Wysocka,J. (2007) H3K27 demethylases, at long last. *Cell*, **131**, 29–32.
53. Berry,W.L. and Janknecht,R. (2013) KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. *Cancer Res.*, **73**, 2936–2942.

54. Ogryzko,V.V., Schiltz,R.L., Russanova,V., Howard,B.H. and Nakatani,Y. (1996) The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*, **87**, 953–959.
55. Hagmann,M., Georgiev,O. and Schaffner,W. (1997) The VP16 paradox: herpes simplex virus VP16 contains a long-range activation domain but within the natural multiprotein complex activates only from promoter-proximal positions. *J. Virol.*, **71**, 5952–5962.
56. Tremblay,J.P., Chapdelaine,P., Coulombe,Z. and Rousseau,J. (2012) Transcription activator-like effector proteins induce the expression of the frataxin gene. *Hum. Gene Ther.*, **23**, 883–890.
57. Hori,R.T., Xu,S., Hu,X. and Pyo,S. (2004) TFIIB-facilitated recruitment of preinitiation complexes by a TAF-independent mechanism. *Nucleic Acids Res.*, **32**, 3856–3863.
58. Chan,H.M. and La Thangue,N.B. (2001) p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.*, **114**, 2363–2373.
59. Feldman,N., Gerson,A., Fang,J., Li,E., Zhang,Y., Shinkai,Y., Cedar,H. and Bergman,Y. (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.*, **8**, 188–194.
60. Onder,T.T., Kara,N., Cherry,A., Sinha,A.U., Zhu,N., Bernt,K.M., Cahan,P., Marcarci,B.O., Unternaehrer,J., Gupta,P.B. *et al.* (2012) Chromatin-modifying enzymes as modulators of reprogramming. *Nature*, **483**, 598–602.
61. Pawlak,M. and Jaenisch,R. (2011) De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev.*, **25**, 1035–1040.