REVIEW

Genome-editing tools for stem cell biology

EA Vasileva¹, OU Shuvalov¹, AV Garabadgiu², G Melino²,³ and NA Barlev*¹,²

Human pluripotent stem cells provide a versatile platform for regenerative studies, drug testing and disease modeling. That the expression of only four transcription factors, Oct4, Klf4, Sox2 and c-Myc (OKSM), is sufficient for generation of induced pluripotent stem cells (iPSCs) from differentiated somatic cells has revolutionized the field and also highlighted the importance of OKSM as targets for genome editing. A number of novel genome-editing systems have been developed recently. In this review, we focus on successful applications of several such systems for generation of iPSCs. In particular, we discuss genome-editing systems based on zinc-finger fusion proteins (ZFs), transcription activator-like effectors (TALEs) and an RNA-guided DNA-specific nuclease, Cas9, derived from the bacterial defense system against viruses that utilizes clustered regularly interspaced short palindromic repeats (CRISPR).

Cell Death and Disease (2015) 6, e1831; doi:10.1038/cddis.2015.167; published online 23 July 2015

Facts

- Genome editing systems based on zinc-finger fusion proteins (ZFs), transcription activator-like effectors (TALEs) and an RNA-guided DNA-specific nuclease (Cas9) can be successfully used for generation of induced pluripotent stem cells.
- ZF-TFs and TALENs fused with different transcriptional domains can modulate expression of master genes of pluripotency, such as Oct4, Sox2, Klf4 and c-Myc.
- The CRISPR/Cas9 fusion with the histone acetyltransferase domain of p300 can reactivate on its own the epigenetically silenced locus of Oct4, which makes this system a very attractive tool for generation of iPSCs.
- As generation of iPSCs requires p53 inactivation, which, in turn, provokes tumorigenesis, it will be interesting to see whether temporal Cas9-mediated inhibition of p53 downstream targets, but not p53 itself, is sufficient to trigger dedifferentiation without affecting the quality control.

Introduction

The emergence of genome-editing technologies over the past several years has flourished the investigation of human cellular disease models. Recent achievements in generation of pluripotent stem cells (PSCs) from patients and specific differentiation of these cells into various somatic cell types greatly facilitated the studies on pathophysiology of socially important diseases (Figure 1). PSCs include embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). PSCs are able to proliferate indefinitely, to self-renew and to develop into more differentiated cell lineages offering the opportunity for human disease modeling. Pluripotency is characterized by specific configuration of chromatin and epigenetic modifications. Forced expression of four transcription master-regulators, Oct3/4, Sox2, Klf4 and c-Myc (OSKM), were able to overcome the epigenetic traits of differentiated cells and to revert them into the naive pluripotent state. Both activity and expression of these transcription factors (TFs) are repressed in normal somatic cells and hence their re-activation is instrumental for the re-programming of somatic cells into the iPSCs. The originally described direct delivery of the corresponding cDNAs into somatic cells cannot be utilized for the purpose of gene therapy, because of the possibility of DNA recombination. Another approach is via pharmacological enhancement of the downstream targets of OSKM. An alternative approach of OSKM re-activation in differentiated cells can be achieved through specific targeting of transcription activators by means of genome editing.

Open Questions

- One downside of the iPSCs generation process is its low efficacy. In this respect, what will happen when the precision of genome editing systems is combined with the power of small molecule inhibitors that reverse the epigenetic state of differentiated cells?

Abbreviations: CR, conservative region; CRISPR, clustered regularly interspaced short palindromic repeat; dCas9, dead Cas9 nuclease; CR, conservative region; CRISPR, clustered regularly interspaced short palindromic repeat; dCas9, dead Cas9 nuclease; DE, distal enhancer; ESC, embryonic stem cell; FD, functional domain; gRNA, guide RNA; iPSC, induced pluripotent stem cell; LITE, light-inducible transcriptional effector; OKSM, Oct4, Klf4, Sox2 and c-Myc; PAM, protospacer adjacent motif; PE, proximal enhancer; PSC, pluripotent stem cell; RVD, repeat variable di-residue; sgRNA, single-guide RNA; SL1-3, stem loop 1-3; TALE, transcription activator-like effector; TALEN, transcription activator-like effector nuclease; TF, transcription factor; ZF, zinc-finger protein; ZFN, zinc-finger nuclease

Received 26.3.15; revised 15.5.15; accepted 18.5.15; Edited by G Kroemer
As of today, several efficient systems of genome manipulation have been described based on various classes of DNA-binding chimeric proteins such as zinc-finger proteins (ZFs), transcription activator-like effectors (TALEs) and the guide RNA (gRNA)-driven Cas9d mutant (CRISPR) system. In this review, we discuss the exploitation of various genome editing techniques for successful and robust generation of iPSCs from human somatic cells.

**Systems for genome editing and manipulation of gene expression**

**ZFs nucleases and ZF-TFs:** ZFNs (zinc-finger nucleases) genome-editing system utilizes chimeric proteins that consist of highly specific ‘zinc finger’ (ZF) DNA-binding domains fused to a nuclease domain of the restriction endonuclease FokI (Figure 2a). Each finger of the DNA binding domain, which consists of tandem Cys-His2 arrays, recognizes approximately three bp of DNA. Thus, a combination of six ZFs is sufficient to
bind 18 bp of the unique target DNA sequence providing sufficient genomic specificity (Table 1). The nuclease domain can be substituted with other functional domains for manipulating the levels of gene expression (Figure 2b).

ZFN system has been successfully applied for the modification of various genomes, including plants,7 insects8 Danio rerio,9 mice,10 rats,11 pigs,12 human cell lines13 and iPSCs.13–15

ZNF system showed promising results in gene therapy of the mutation causing sickle cell anemia in human iPSCs.16 Further, the bi-allelic correction of the point mutation (Glu342Lys) in the gene for the photosensitive protein CRY2 (TALE:CRY2) of Arabidopsis thaliana.31 The second component comprises CIB1 (interaction partner with CRY2), fused to a transcription activation domain (e.g., from a viral activator VP64), CIB1:TAD. In the absence of light TALE: CRY2 binds to the promoter region of a target gene, whereas the complex CIB1:TAD remains unbound (Figure 2e). The treatment of cells with light causes a conformational change to CRY2, facilitating the recruitment of the CIB1:TAD complex to induce transcription from the target promoter (Figure 2f). In addition to the regulation of transcriptional activity, such system can also be used for the targeting of specific epigenetic chromatin modifications to specific genomic loci.20

This approach allows studying the effect of selected chromatin modifications on the expression of specific genes. For example, the TALEs domain fused to the catalytic domain of TET1 protein (ten-eleven translocation), which oxidizes 5-methylcytosine to methylated cytosine (5 mC), was reported to cause a significant demethylation in the CpG-rich chromatin.32

**TALENs and TALE-TFs:** On the basis of the TALE protein TALENs genome-editing system has been successfully applied for genome modification in plants,18 insects,19 nematodes,20 the fish,21 amphibians,22 mice,23 rats,24 rabbits,25 cancer human cell lines, hESCs and iPSCs.26,27

The most important component of this system is a site-specific DNA-binding protein TALE isolated from a pathogenic for plant organism Xanthomonas. Another TALE-like protein derived from a pathogenic bacterium Ralstonia can also be used for specific editing.28 The DNA-binding domain represents a TALE tandem repeats of 33–35 amino acids. TALE repeats have similar sequences and differ only in the two highly variable amino acids at positions 12 and 13 (RVDs, repeat variable di-residues), which form the basis for specific-nucleotide recognition.29 Four tandem repeats Asn-Asn, Asn-Ile, His-Asp and Asn-Gly are sufficient for recognition of guanine, adenine, cytosine and thymine, and, hence, for generation of TALEs with unique properties (see Table 1). The second element of this fusion is the nuclease domain of a restriction endonuclease FokI or another functional domain (e.g., VP64, TET1, KRAB, etc), which introduces specific changes to the genome (Figures 2c and d).

Recently, a new variation of the TALE system, an optogenetic LITE system (light-inducible transcriptional effectors) has been developed.30 This LITE system consists of two components (Figures 2e and f). The first one is the DNA-binding TALE domain of Xanthomonas with the photosensitive protein CRY2 (TALE:CRY2) or Arabidopsis thaliana.31

| Table 1 A brief comparative summary of ZFPs, TALENs and CRISPR/Cas9 genome-editing systems |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| **ZFPs** | **TALEs** | **CRISPR/Cas9** |
| DNA binding | ‘zinc-finger’ domain | Transcription activator-like effectors (TALE) | crRNA:tracrRNA or sgRNA |
| Nuclease domain for genome editing | FokI | FokI | Cas9 |
| Regulation of gene expression | ZF-TFs with VP16, VP64 and 2xP64 domains for Oct4 activation; KRAB domain for Oct4 both repression and activation. | TALEs with VPA and 5azadC inhibitors for upregulation of Oct4; Oct4 enhancer targeting by TALE; VP64 for induction of Oct4 transcription. | sgRNA/dCas-VP64 targeting for induction of Oct4 transcription; dCas9 fused to VP160 and sgRNA for induction of Oct4 transcription. |
| Efficiency | ++ | ++ | +++ |
| Specifity | 18–36 bp | 30-36 bp | 23–28 bp |
| Off-target | Vary | Low | Vary |
| Cytotoxicity | Vary | Low | Vary |
| The frequency of potential sites, limitations | 1 to 100 bp. Limitation: absence of a collection of 64 zinc-fingers that would cover all possible combinations of triplets. | 1 to 1 bp. Can be designed virtually for any DNA sequence. Limitation: the necessity of thymine at the 5’-end of the target sequence. | 1 to 4–8 bp. Necessity of PAM sequence: 5’-X20 NGG-3’, 5’-X20 NAG-3’ or 5’-X20 NNNNGATT-3’. |

Cytotoxity Vary Low Low

Off-target Vary Low Vary

The most important component of this system is a site-specific DNA-binding protein TALE isolated from a pathogenic for plant organism Xanthomonas. Another TALE-like protein derived from a pathogenic bacterium Ralstonia can also be used for specific editing.28 The DNA-binding domain represents a TALE tandem repeats of 33–35 amino acids. TALE repeats have similar sequences and differ only in the two highly variable amino acids at positions 12 and 13 (RVDs, repeat variable di-residues), which form the basis for specific-nucleotide recognition.29 Four tandem repeats Asn-Asn, Asn-Ile, His-Asp and Asn-Gly are sufficient for recognition of guanine, adenine, cytosine and thymine, and, hence, for generation of TALEs with unique properties (see Table 1). The second element of this fusion is the nuclease domain of a restriction endonuclease FokI or another functional domain (e.g., VP64, TET1, KRAB, etc), which introduces specific changes to the genome (Figures 2c and d).

Recently, a new variation of the TALE system, an optogenetic LITE system (light-inducible transcriptional effectors) has been developed.30 This LITE system consists of two components (Figures 2e and f). The first one is the DNA-binding TALE domain of Xanthomonas with the photosensitive protein CRY2 (TALE:CRY2) of Arabidopsis thaliana.31 The second component comprises CIB1 (interaction partner with CRY2), fused to a transcription activation domain (e.g., from a viral activator VP64), CIB1:TAD. In the absence of light TALE: CRY2 binds to the promoter region of a target gene, whereas the complex CIB1:TAD remains unbound (Figure 2e). The treatment of cells with light causes a conformational change to CRY2, facilitating the recruitment of the CIB1:TAD complex to induce transcription from the target promoter (Figure 2f). In addition to the regulation of transcriptional activity, such system can also be used for the targeting of specific epigenetic chromatin modifications to specific genomic loci.20

This approach allows studying the effect of selected chromatin modifications on the expression of specific genes. For example, the TALEs domain fused to the catalytic domain of TET1 protein (ten-eleven translocation), which oxidizes 5-methylcytosine to methylated cytosine (5 mC), was reported to cause a significant demethylation in the CpG-rich chromatin.32

CRISPR / Cas9 and dCas9-TF: The CRISPR/Cas system is a prokaryotic analog of the immune system against exogenous DNA-containing phages and plasmids.33 Although the exact mechanism of CRISPR/Cas9 action is still under investigation, it is deemed that clustered regularly interspaced short palindromic repeats (CRISPR) along with short spacer DNA fragments that derive from previous encounters with viruses are transcribed into long CRISPR RNAs (crRNAs). When combined with transactivation crRNA (tracrRNA), these crRNA:tracrRNA duplexes provide a ‘search engine’ for the Cas9 nuclease to attack specific viral DNAs34 (Figure 2g).

For the purpose of simplifying the implementation of the system in biotechnology, the crRNA:tracrRNA duplex was substituted with single-guide sgRNA24 (Figure 2h). The specificity of Cas9 homing is determined by the nuclease PAM motif and 20 nucleotides of the complementary sequence of sgRNA. PAM sequences vary among Cas orthologs: 5’-NGG-3’ PAM in Streptococcus pyogenes,35 5’-NGGNG-3’ and 5’-NAGAAW-3’ PAM in Streptococcus thermophilus,36,37 5’-NNNGATT-3’ PAM from Neisseria meningitidis.38 PAM dependence increases the specificity of CRISPR/Cas (Table 1).
The CRISPR/Cas system has successfully been applied for genome editing in plants,\textsuperscript{39} nematodes,\textsuperscript{40} insects,\textsuperscript{41} the fish,\textsuperscript{42,43} mice,\textsuperscript{44} rat,\textsuperscript{45} human cell lines,\textsuperscript{35} ESCs and iPSCs.\textsuperscript{38,46}

More recently, a new iCRISPR platform based on CRISPR/Cas and TALENs systems has been designed for quick (up to 1 months) and highly efficient production of bi-allelic knockout in hPSCs lines. First of all generation of hPSCs lines that express Cas9, the invariable component of the CRISPR/Cas system, was performed for creating such platform. For the next step the lipid-mediated transfection of small RNAs was determined as efficient for co-transfection of multiple gRNAs for multiplexed genome editing during a desirable stage of hPSC. To make the iCRISPR platform more flexible, special iCas9 hPSC lines were engineered for doxycycline-inducible expression of Cas9 through TALEN-mediated gene targeting. Thereby, this platform allows successful one-step generation of double- and triple-knockout hPSC lines as well as stage-specific inducible gene knockouts during differentiation of hPSCs.\textsuperscript{46}

Furthermore, the CRISPR system with dead Cas9 nuclease (dCas9) protein fused with a transcription activation domain (Figures 2l and k) has been developed. Activation and repression of specific genes in hPSCs thus affecting the course of differentiation has been achieved by employing this system.\textsuperscript{47}

Regulation of Oct4 expression and pluripotency. As was mentioned earlier, the activity of five transcriptional master-regulators is critical for the maintenance of pluripotency and self-renewal of stem cells. Among those, the Oct4 (POUSF1) gene is the critical one.\textsuperscript{48–51} The Octamer-binding TF4 (Oct4) protein belongs to the family of homeodomain-containing transcription factors. Mechanistically, Oct4 not only positively affects transcription of genes required for pluripotency and self-renewal but also prevents the expression of TFs that contain a zinc-finger DNA-binding domain and the functional domain to modulate gene expression. The ZF-TF system was successfully employed to target the Oct4 gene expression. Specifically, ZFs targeting a 19-bp region between −25 and −7 bp downstream of the Oct4 promoter were fused with either the herpes simplex virus VP16 activation domain or the repression domain from the human KOX1 protein.\textsuperscript{61} Transfection with the ZF-VP16 plasmid caused moderate, but CR4.\textsuperscript{54} CR1 (proximal promoter) and the most distal conserved region CR4 are the regions critically important for regulation of Oct4 gene expression by several transcription factors, including Sp1 and RAR (Figure 3).\textsuperscript{54}

Oct4 interacts with other TFs such as Sox2 and Nanog, which are also instrumental for the maintenance of pluripotency and iPSCs reprogramming,\textsuperscript{55,66} thus forming a network of protein–protein interactions. As TFs exert their functions at least in part through the recruitment of epigenetic modifiers, it is not surprising that the promoter of Oct4 gene is regulated by DNA methylation. Dnmt3a and Dnmt3b were shown responsible for DNA methylation of the Oct4 promoter. This event is critical for triggering ESCs to differentiate.\textsuperscript{57} Oct4 promoter is methylated and hence silenced in the vast majority of somatic cells. On the contrary, this gene is expressed not only in ESCs but also in several malignancies.\textsuperscript{58} For example, reactivation of Oct4 is associated with tumor initiation in breast cancer cells\textsuperscript{59} as well as in poorly differentiated epithelial ovarian cancers.\textsuperscript{60} Exogenous delivery of specific cDNA combinations reactivates the endogenous Oct4 promoter.

It needs to be mentioned that Oct4 is required not only for the maintenance of pluripotency, but when overexpressed it triggers differentiation. Thus, Oct4 serves a gauge of the cellular state in terms of commitment to differentiation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Structure of Oct4 upstream promoter region. (a) Schematic representation of the Oct4 upstream region of the human promoters.\textsuperscript{52} CR1-4 denote Conservative Regions in the promoter of Oct4 gene (see the text). Conserved sequences are shown inside the boxes. Their locations relative to the start site are indicated below. Known transcription factors that bind these CRs are indicated. (b) Shown is the upstream region in the promoter of Oct4 gene. Specific DE and PE sites with respect to the CRs are indicated.\textsuperscript{52} Green arrow denotes the direction of Oct4 gene transcription.}
\end{figure}
reproducible activation of Oct4, whereas an overexpression of ZFs-KOX1 fusion caused a significant repression.\textsuperscript{61} Functionally, increasing or decreasing levels of Oct4 expression by more than twofold forced ES cells to differentiate into primitive endoderm and mesoderm.\textsuperscript{61}

ZF-s attached to a Kruppel-associated Box (KRAB) domain function as potent transcriptional repressors via recruitment of the histone deacetylase (NuRD) complex. The latter includes histone deacetylases (HDACs), histone methyltransferase (SETDB1) and heterochromatin protein 1 (HP1).\textsuperscript{62,63} Recently, almost complete repression of the Sox2 gene via ZFs linked to a KRAB domain has been described in breast cancer cells.\textsuperscript{64} However, in addition to their well-established role as transcriptional repressors several KRAB-containing ZF chimeras can also activate transcription.\textsuperscript{65} For example, ectopic expression of KRAB-containing ZFs strongly reactivated Oct4 expression in a panel of breast and ovarian cell lines.\textsuperscript{66} The KRAB domain is composed of two A and B boxes.

KRAB-associated protein 1 (KAB1) is one of the main co-repressors of KRAB and interacts with box A subsequently recruiting lysine methyltransferase SETDB1 to tri-methylate H3-K9 (H3K9me3).\textsuperscript{67,68} Stabilization of the repressive complex on chromatin is maintained by binding of KAP1 with HP1 through interaction with H3K9me3. How the KRAB domain interacts with transcriptional co-activators is not known yet. One possibility is that KRAB–ZFs fusion may interfere with other transcriptional repressors (e.g., DNMTs), thus mediating the ‘inhibition of inhibitors’. Irrespective of the exact mechanism, these results indicate that KRAB–ZFs can function as an activator of silenced genes in specific chromatin context.

ZF-TFs have been used for targeting of Oct4, Sox2, Klf4 and c-Myc genes, which are critical for the maintenance and acquisition of pluripotency. The levels of activation for these genes were comparable to the ones observed in ES cells and did not require additional active epigenetic agents.\textsuperscript{69} Over 300 promoter region-targeting ZFs fused with the p65 subunit of NF-B were designed to target 1 kb (from –800 to +200 bp) region of the promoter around transcriptional start sites for each of Oct4, Sox2, Klf4 and c-Myc. Each designed ZF-coding sequence was cloned between the N-terminal nuclear localization signal and the C-terminal NF-B p65 activation domain. To identify critical binding sites for ZF fusions to upregulate transcription, the upstream enhancer and the region downstream of the transcriptional start site were explored. Three out of the six constructs were shown by RT-PCR and western blotting to activate Oct4 more than 16-fold in HEK293 cells. The best two of the three activator of Oct4 are located in the upstream enhancer region of the gene. Also, the highest transcription activation of Sox2 and Oct4 in HEK293 was achieved with ZF-TFs containing VP64 or 2xP65 activation domains, respectively. However, the same ZFs fused to VP16 and VP64 domains failed to activate Oct4 efficiently. Apparently, the ability of ZF-TFs to activate transcription depends on the cell type, the exact functional domain fused to ZF-TFs and the chromatin context of the targeted gene.\textsuperscript{69} In this respect, it can be speculated that depending on the epigenetic state of chromatin in the target locus, different functional domains fused to ZFs may exhibit various efficacies.\textsuperscript{69}

**TALE-TFs and master regulators of pluripotency:** The TALE-TF system has successfully been employed to activate Sox2 and Klf4 genes in HEK293 cancer cell line.\textsuperscript{70} TALEs genome-editing system has also been used for reactivation of the Oct4 gene.\textsuperscript{71} Designed TALEs efficiently upregulated Oct4 transcription in ESCs, but failed to activate this gene in ESC-derived neural stem cells (NSCs) because of the repressive epigenetic state of the corresponding genomic locus. Chemical inhibition of histone deacetylases (HDAC) by VPA (valproic acid) and DNA methyltransferases by 5-azaC, respectively, greatly facilitated the effect of designed TALEs on expression of the epigenetically silenced Oct4 promoter in NSC.\textsuperscript{72} This result suggests that designed TALEs can be used for reprogramming somatic cells into iPSCs.

TALE-VP64 fusion can induce transcription of endogenous Oct4 by targeting its distal enhancer (DE). Reactivation of the endogenous Oct4 by TALE-VP64 was sufficient for epigenetic reprogramming of fibroblasts into iPSC in the absence of exogenous factors Oct4 or Nanog.\textsuperscript{73} Mechanistically, TALE-VP64 likely recruited histone acetyltransferase (HAT) p300 to acetylate histones. In this respect, both TALE-VP64 and sgRNA/dCas-VP64 chimeras were shown to interact with p300 in human and mouse cells.\textsuperscript{73}

Interestingly, TALE- and dCas9-based activators utilize different regulatory regions of the Oct4 gene. The binding region from –120 to –80 bp was the most efficient for TALE-VP64-mediated activation, while Cas9d was highly effective when targeted by sgRNA to the region from –147 to –89 bp upstream of the transcription start site (Figure 3). In line with this, a significant increase of transcriptional activation of mouse Oct4 promoters was achieved by moving the target sequences of inefficient TALE-VP64 into the –120 to –104 bp region.

Individual activators often exhibited marginal or no activity, whereas application of multiple TALE-VP64 or several sgRNA targeting the same region exhibited transcriptional synergy.\textsuperscript{73} Multiple TALE-VP64 targeting enhanced transcription of mouse Oct4 gene up to 30-fold in NIH3T cells and increased transcription of the human Oct4 up to 20-fold in HEK293T cells.\textsuperscript{73}

dCas-TFs and master regulators of pluripotency: Recently, a CRISPRi (CRISPR inference) system has been utilized for regulation of transcription.\textsuperscript{74} In this system defective dCas9 lacking the nuclease activity was used. dCas9 when co-expressed with an appropriate sgRNA disrupts transcription by interfering with the binding or elongation of the RNA polymerase complex and/or specific transcription factors.\textsuperscript{75} This system provides means for transient attenuation of gene expression without causing deep epigenetic modifications to the DNA sequence.

Recently, a multiplexed activation of endogenous Oct4, Sox2 and ilrin genes by an inducible Tet-on CRISPR/dCas9 system has been developed for human and mouse cells.\textsuperscript{76} This system is based on the dCas9 protein fused to several copies of the viral transcription activation domain VP16. It was shown that dCas9-VP160 (10 copies of the VP16 minimal activation domain) efficiently activated endogenous genes when targeted by specific sgRNA to the region within 300 bp upstream of the transcriptional start site. The most efficient gene activation was achieved by clusters of 3-4 sgRNAs binding to the proximal promoters, suggesting a synergistic
mode of action. Simultaneous induction of at least three different endogenous genes was achieved with the CRISPR-on system in this study. A recent report shed some light on the mechanistic differences in gene regulation by TALEs and Cas9d proteins. While TALE-TFs and CRISPR systems were comparable in their ability to repress transcription of endogenous Oct4 and Nanog genes, TALE-TFs were much superior in their ability to activate transcription during the reprogramming of both MEFs and EpiSCs. Expression of Cas9d alone failed to reprogram cells into iPSCs despite modest upregulation of mRNA expression and positive effect in the luciferase reporter assay. It was likely due to an inefficient recruitment of p300 HAT to acetylate histones at the target site.

In line with this notion, a recent report described a fusion construct between dCas9 and the histone acetyltransferase (HAT) domain of p300 as a powerful transcription activator. Specifically, the dCas9-p300 HAT protein targeted by a pool of gRNAs to the PE of Oct4 gene 30-fold more potently activated transcription compared with dCas9-VP64. Moreover, this approach may be transferable to other genome-editing systems (ZF-TFs, TALE-TFs), thus making it a versatile technology for targeted gene activation.

Conclusions

Systems for genome editing and manipulation with gene expression based on DNA-binding ZFs, TALEs and CRISPR/Cas9 molecules fused to special functional or nuclease domain could be used in various areas of modern bioengineering. In particular, genome-editing systems represent a promising approach for generation of iPSCs (see Table 1 for comparison).

Importantly, genome-editing systems have a significant advantage over the existing OSKM scheme of generating iPSCs. The problem with OSKM is the induction of genomic instability and tumor formation especially by c-Myc and, to a lesser extent, Klf4. Comparative analysis of stem cells reprogrammed by expressing c-Myc revealed genomic deletions and amplifications, characteristic of oncogene-induced DNA replication stress. One of the critical effectors of c-Myc overexpression is the major mammalian tumor suppressor TP53. TP53 is the guardian of genome protecting the organism from cancer as well as infertility or aging.

Among a large number of regulated genes, p53 activates expression of the p21 gene, whose product, in turn, blocks proliferation and triggers differentiation of pluripotent cells. To circumvent this problem several approaches have been described. For example, direct inactivation of p53 significantly increased the efficacy of iPSCs generation. Alternatively, inhibition of Notch signaling whose downstream target is p21 with small molecules also facilitated iPSCs generation. However, inactivation of p53 results in genomic instability and inactivation of Notch promotes differentiation of iPSCs into neural progenitors. In this respect, genome editing of downstream targets that prevent de-differentiation, for example, temporal inactivation of p21 or PUMA, would seem an ideal way to control these unwanted biological effects. Obviously, as any experimental system, genome-editing systems have their own limitations, that is, their efficacy varies greatly depending on the chromatin accessibility of a regulatory region selected for targeting, its proximity to the promoter or enhancer of the gene of interest, accessibility to other TFs for binding and so on. However, these obstacles will be avoidable in future once the working range and preferable epigenetic makeup of chromatin is determined for each particular targeting TF fusion.

Another potential limitation of this approach is based on the fact that most of the genome-editing systems fail to reactivate on their own epigenetically silenced genetic loci, such as Oct4. However, very recent data on epigenetic reactivation by targeting of the dCas9-p300 HAT domain fusion provide optimism on this end. Furthermore, there is a wealth of data arguing that isolated activation of the Oct4 gene is not sufficient for generation of iPSCs. In this respect, it may be beneficial to combine the precision of genome-editing tools with a wider effect of pharmacological inhibitors. Future studies should test this intriguing possibility, which will then broaden the area of applications for the genome-editing tools.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements. EAV, OUS and NAB appreciate the support from the Russian Science Foundation (project No. 14-50-00006). Financial support from the Federal Agency of Scientific Organizations (Russia) is also appreciated. AVG and GM are supported by the Government grant 11.G34.31.0069 to GM.

Author Contributions

EAV, OUS and NAB conceived and wrote the manuscript. AVG and GM contributed intellectually to the section on Cas9.

1. Thomson JA. Embryonic stem cell lines derived from human blastocysts. Science 1998; 282: 1145–1147.
2. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006; 126: 663–676.
3. Wei X, Chen Y, Xu Y, Zhan Y, Zhang R, Wang M et al. Small molecule compound induces chromatin de-condensation and facilitates induced pluripotent stem cell generation. J Mol Cell Biol 2014; 6: 409–420.
4. Ma J. Transcriptional activators and activation mechanisms. Protein Cell 2011; 2: 879–888.
5. Huangfu D, Meahr R, Guo W, Eijkelenboom A, Sinitow M, Chen AE et al. Induction of pluripotent stem-cell line profiles by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 2008; 26: 795–797.
6. Chen J, Gao Y, Huang H, Xu K, Chen X, Jiang Y et al. The combination of Tert with Oct4 generates high-quality mouse induced pluripotent stem cells (iPSCs). Stem Cell 2014; 1: 1–16.
7. Lloyd A, Plaisier CL, Carril D, Drews GN. Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proc Natl Acad Sci USA 2005; 102: 2232–2237.
8. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics 2006; 172: 2391–2403.
9. Doyon Y, McCanmon JM, Miller JC, Farij F, Ngo C, Kalbian GE et al. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 2008; 26: 760–766.
10. Carbery ID, JI D, Harrington A, Brown V, Weinstein EJ, Liaw L et al. Targeted genome modification in mice using zinc-finger nucleases. Genetics 2010; 186: 451–459.
11. Mashimo T, Takizawa A, Voigt B, Yoshimi K, Hira H, Kuraomoto T et al. Generation of knockout rats with X-linked severe combined immunodeficiency (X SCID) using zinc-finger nucleases. PLoS One 2010; 5: e8870.
12. Hauschild J, Petersen B, Santiago Y, Queisser A, Carnwath JW, Lucas-hahn A et al. Correction for Hauschild etc., Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. Proc Natl Acad Sci USA 2011; 108: 15010–15015.
13. Umov FD, Miller JC, Lee YL, Beaulieu GM, Rock JM, Augustus S et al. High efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 2005; 435: 646–651.
14. Umov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 2010; 11: 636–646.
15. Soltner F, Lagaronne J, Cheng AW, Hockenberry D, Gao Q, Alagappan R et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. Cell 2011; 146: 318–331.

16. Sebastiano V, Maeder ML, Angstman JF, Haddad B, Khyter C, Yeo DT et al. NIH Public Access 2012; 29: 1717–1728.

17. Yusa K, Rashid ST, Strick-marchand H, Varela I, Liu P, Darce S et al. Targeted gene correction of α-antitrypsin deficiency in induced pluripotent stem cells. Nature 2012; 478: 391–394.

18. Christian M, Qi Y, Zhang Y, Voytas DF. Targeted mutagenesis of Arabidopsis thaliana using engineered TAL effector nucleases. G3 (Bethesda) 2013; 3: 1697–1705.

19. Kato-sayanagi T, Ammendola A, Sarmay M, Weiss SC, Savers C, Paro R. An efficient strategy for TALEN-mediated genome engineering in Drosophila. Nucleic Acid Res 2013; 41: e163.

20. Cheng Z, Yi P, Wang X, Chai Y, Feng G, Yang Y et al. Conditional targeted genome editing using somatically expressed TALENs in C elegans. Nat Biotechnol 2013; 31: 934–937.

21. Cade L, Reynor D, Hwang WY, Tsal SG, Patel S, Khyter C et al. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. Nucleic Acid Res 2012; 40: 8011–8010.

22. Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X et al. Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Nat Acad Sci USA 2012; 109: 17464–17469.

23. Qiu Z, Liu M, Chen Z, Shao Y, Pan H, Wei G et al. High-efficiency and heritable gene targeting in mice by transcription activator-like effector nucleases. Nucleic Acid Res 2013; 41: e120.

24. Tesson L, Usal C, Ménoret S, Leung E, Niles BJ, Remy S et al. Generation of RAG 1 and 2-deficient rabbits by embryo microinjection of TALENs. Cell 2011; 146: 1059–1062.

25. Hockenberry D, Hwang W, Kian S, Lai G, Csos J, Cassady JP et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 2011; 29: 731–734.

26. Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res 2013; 41: e63.

27. Li L, Alet F, Plate I, Ali Z, Platek M, Acuida M et al. Characterization and DNA-binding specificities of Ratostria TAL-like effectors. Mol Plant 2013; 6: 1318–1330.

28. Moscow MJ, Bogdanove AJ. A simple ciprofloxin DNA gel recognition by TAL effectors. Science 2009; 326: 1501.

29. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L et al. Optical control of mammalian endogenous transcription and epigenetic states. Nature 2013; 500: 472–476.

30. Liu H, Yi X, Li K, Kleint J, Yang H, Lissner D et al. Photocleaved CR2Y interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. Science 2008; 322: 1051–1053.

31. Mendenhall EM, Williamson KE, Reynor D, Zou JY, Ram O, Joung JK et al. Locus-specific editing of histone modifications at endogenous enhancers. Nat Biotechnol 2013; 31: 1133–1136.

32. Mackorov KS, Grishin NV, Shabalina SA, Wolf YI, Koonin E V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogy with eukaryotic RNAsi, and hypothetical mechanisms of action. Biochim Biophys Acta 2006; 11: 777–785.

33. Jin et al. Chylinski K, Fontana I, Hauer M, Douadjia J, Charpentier E. A programmable dual-center zinc finger nuclease for genome engineering in human cells. Nat Biotechnol 2013; 31: 1199–1204.

34. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 2011; 29: 149–154.

35. Baltz H, Buxel C. Worm et al. Multiplexed activation of CRISPR/Cas9 genes to target 100 simultaneously. Cell 2014; 157: 841–851.

36. Hsu P D, Lander E S, Zhang F. Cas9, a single-guide RNA-directedCas9-mediated interference in Streptococcus thermophilus. RNA Biol 2013; 10: 841–851.

37. Hou Z, Zhang Y, Popescu NE, Howden SE, Chu L-F, Sontheimer EJ et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Nat Acad Sci USA 2013; 110: 15644–5.

38. Jiang W, Zhu H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9 sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acid Res 2013; 41: e168.

39. Liu P, Long L, Xiong K, Yu B, Chang N, Xiong-JW et al. Heritable/conditional gene editing in C elegans using a Cas9/CsCas9 feeding system. Cell Res 2013; 4: 220–228.

40. Bassett AR, Tibbitt C, Ponting CP, Liu JL. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep 2013; 4: 220–228.

41. Hwang WJ, Fu Y, Reynor D, Maeder ML, Tsal SG, Sander JD et al. Efficient genome editing in zebrafish using a CRISPR/Cas9 system. Nat Biotechnol 2013; 31: 681–687.

42. Ota S, Hisano Y, Ikawa Y, Kawaihara A. Multiple genome modifications by the CRISPR/Cas9 system in zebrafish. Gene Cell 2014; 1–10.

43. Shen B, Zhang J, Wu H, Wang M, Ji K, Liu Z et al. Generation of gene-modified mice via Cas9/RNA-guided gene targeting. Cell Res 2013; 23: 725–733.

44. Li D, Ou Z, Shao Y, Chen Y, Guan Y, Liu M et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol 2013; 31: 681–687.

45. Gonzalez F, Zhu Z, Shi Z-D, Lelli K, Verma N, Li QV et al. An CRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell 2014; 1–12.
77. Gao X, Tsang JCH, Gaba F, Wu D, Lu L, Liu P. Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. Nucleic Acid Res 2014; 14: e155.

78. Hilton IB, D’Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol 2015; 33: 510–517.

79. Gorse A, Li Z, Fung H-L, Yeung J, Agarwal S, Antosiewicz-Bourget J et al. Somatic coding mutations in human induced pluripotent stem cells. Nature 2011; 471: 63–67.

80. Pasi CE, Dereki-Öz A, Negrini S, Friedli M, Fragola G, Lombardo A et al. Genomic instability in induced stem cells. Cell Death Differ 2011; 18: 745–753.

81. Phesse TJ, Myant KB, Cole AM, Ridgway RA, Pearson H, Muncan V et al. Endogenous c-Myc is essential for p53-induced apoptosis in response to DNA damage in vivo. Cell Death Differ 2014; 21: 956–966.

82. Ruffini A, Tucci P, Celardo I, Melino G. Senescence and aging: the critical roles of p53. Oncogene 2013; 32: 5129–5143.

83. Anttonov AV, Krestyaninova M, Knight RA, Rodchenkov I, Melino G, Barlev NA. PPISURV: a novel bioinformatics tool for uncovering the hidden role of specific genes in cancer survival outcome. Oncogene 2014; 33: 1621–1628.

84. Marouco D, Garabadgiu AV, Melino G, Barlev NA. Lysine-specific modifications of p53: a matter of life and death? Oncotarget 2013; 4: 1556–1571.

85. Lezina L, Aksenova V, Ivanova T, Purmessur N, Antonov a V, Tentler D et al. KMTase Set7/9 is a critical regulator of E2F1 activity upon genotoxic stress. Cell Death Differ 2014; 21: 1899–1909.

86. Marión RM, Strati K, Li H, Murga M, Blanco R, Ortega S et al. A p53-mediated DNA damage response limits reprogramming to ensure iPSC cell genomic integrity. Nature 2009; 460: 1149–1153.

87. Ichida JK, Toc J, TCW J, Williams L a, Carter AC, Shi Y et al. Notch inhibition allows oncogene-independent generation of iPS cells. Nat Chem Biol 2014; 10: 632–639.

88. Chen C-Y, Liao W, Lou Y-L, Li O, Hu B, Wang Y et al. Inhibition of Notch signaling facilitates the differentiation of human-induced pluripotent stem cells into neural stem cells. Mol Cell Biochem 2014; 395: 291–298.

89. Li Y, Feng H, Gu H, Lewis DW, Yuan Y, Zhang L et al. The p53-PUMA axis suppresses iPSC generation. Nat Commun 2013; 4: 2174.