A CRISPR/Cas9-Based System for Reprogramming Cell Lineage Specification

Syandan Chakraborty,1,5 HaYeun Ji,1,5 Ami M. Kabadi,1 Charles A. Gersbach,1,2,3 Nicolas Christoforou,1,4 and Kam W. Leong1,5,*

1Department of Biomedical Engineering, Duke University, Durham, NC 27708, USA
2Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708, USA
3Department of Orthopaedic Surgery, Duke University Medical Center, Durham, NC 27708, USA
4Department of Biomedical Engineering, Khalifa University, Abu Dhabi 127788, UAE
5Department of Biomedical Engineering, Columbia University, New York, NY 10027, USA

*Correspondence: kam.leong@duke.edu
http://dx.doi.org/10.1016/j.stemcr.2014.09.013

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

SUMMARY

Gene activation by the CRISPR/Cas9 system has the potential to enable new approaches to science and medicine, but the technology must be enhanced to robustly control cell behavior. We show that the fusion of two transactivation domains to Cas9 dramatically enhances gene activation to a level that is necessary to reprogram cell phenotype. Targeted activation of the endogenous Myod1 gene locus with this system led to stable and sustained reprogramming of mouse embryonic fibroblasts into skeletal myocytes. The levels of myogenic marker expression obtained by the activation of endogenous Myod1 gene were comparable to that achieved by overexpression of lentivirally delivered MYOD1 transcription factor.

INTRODUCTION

The type II clustered regularly interspersed palindromic repeat (CRISPR) systems and the associated Cas9 nucleases have evolved in archaea and bacteria for sequence-specific recognition of DNA targets via a single-stranded RNA intermediate (Jinek et al., 2012). In an engineered version of the CRISPR system, the Streptococcus pyogenes Cas9 nuclease is directed by guide RNAs (gRNAs) targeting 20 bp sequences adjacent to a 5’-NRG-3’ sequence motif, and the resultant cleavage has been used to edit the genome in several species (Cho et al., 2013; Cong et al., 2013; Hwang et al., 2013; Jinek et al., 2013; Mali et al., 2013b). A mutated nuclease-inactive Cas9 (dCas9) regulates gene expression by physically blocking transcription or through fusion to a transactivator (VP64, Ω subunit of RNA polymerase) or repressor domain (KRAB, SID) (Bikard et al., 2013; Cheng et al., 2013; Farzadfard et al., 2013; Gilbert et al., 2013; Kearns et al., 2014; Konermann et al., 2013; Maeder et al., 2013; Mali et al., 2013a; Perez-Pinera et al., 2013; Qi et al., 2013). While transgene overexpression has been used to achieve cellular reprogramming (Davis et al., 1987; Takahashi and Yamanaka, 2006), reprogramming via direct activation of an endogenous gene has only been recently demonstrated through the use of transcription activator-like effectors (TALEs) (Gao et al., 2013). However, difficulty in designing and codelivering multiple TALE expression constructs precludes simple screening and multiplexed gene activation that is straightforward with the dCas9-VP64 system. In this study, we used dCas9-based transactivators combined with an efficient lentivirus-based gene delivery system to induce cellular reprogramming.

RESULTS AND DISCUSSION

We tested the efficacy of a VP64dCas9-BFPVP64 fusion protein to activate expression of the endogenous Myod1 gene locus for a sufficient duration and magnitude to ultimately induce the reprogramming of mouse embryonic fibroblasts (MEFs) to skeletal myocytes (SKMs) (Figure 1A). Although we and others have previously used the single C-terminal fusion of VP64 to dCas9 to activate gene expression (Cheng et al., 2013; Farzadfard et al., 2013; Gilbert et al., 2013; Kabadi et al., 2014; Maeder et al., 2013; Perez-Pinera et al., 2013), preliminary studies indicated that this approach did not lead to levels of expression sufficient for cell reprogramming. Therefore, we tested whether two VP64 domains flanking dCas9 (VP64dCas9-BFPVP64) would yield higher Myod1 gene expression levels (Figures 1B and 1C and Figure S1A available online). dCas9 was also fused to blue fluorescent protein (BFP) to monitor expression (Figures 1B, 1C, 1S1A, and 1S1B). Lentiviral VP64dCas9-BFPVP64 was placed under the transcriptional control of a doxycycline-inducible promoter (Figures 1C and 1S1B). To avoid steric hindrance that may prevent transcriptional complex recruitment, we included flexible glycine-serine linkers adjacent to the VP64 domains (Figure S1A). We also added a third nuclear localization signal (NLS), which improved the nuclear localization of VP64dCas9-BFPVP64 by ~10-fold (Figures 1C–1F). Initially, we transfected C3H10T1/2 cells with the VP64dCas9-BFPVP64 plasmid. However, transfection was inefficient as transgene expression was detectable only in a few cells, presumably due to the large size of the plasmid (13.5 kb). We then used a lentiviral gene delivery system, allowing stable transduction of more cells as
Figure 1. Engineering an RNA-Guided, Nuclease-Inactive VP64-dCas9-BFP—VdC9BV—Fusion Protein to Enable Robust Transactivation of the Endogenous Myod1 Gene

(A) Illustration depicting the reprogramming of MEFs to SkMs initiated by the lentiviral transduction of MEFs with the gRNA, reverse tetracycline transactivator (M2rtTA), and a doxycycline-inducible VdC9BV.

(B) Illustration of gRNA-guided transactivation by the VdC9BV fusion protein.

(C) The lentiviral constructs utilized to activate endogenous genes.

(D) Confocal images of live HEK293T cells transfected with VdC9BV or dCas9, H2Bm Cherry fluorescent protein (nuclear-localized), and GFP (predominantly cytoplasmic) showing that inclusion of an additional NLS in VdC9BV (3 NLS) enables better nuclear localization than the 2 NLS dCas9-BFP (dC9B) parent construct. The scale bar represents 20 μm.

(E) Nucleus-to-cytoplasm (N:C) ratio of BFP fluorescence intensity quantified from confocal images of live transfected HEK293T cells. The analysis shows that the 3 NLS form of dCas9 (VdC9BV) localizes better to the nucleus than the original 2 NLS dCas9 (dC9B) form (n = 20 cells, error bars represent SEM).

(F) Lentiviral vector enables robust delivery of nuclear-localized VdC9BV. dC9B (2 NLS) transfected into C3H10T1/2 cells shows lack of nuclear localization. VdC9BV (3 NLS) enables better nuclear localization. However, few cells expressed VdC9BV when transfected into the cells. Transduction of C3H10T1/2 with the lentiviral form of VdC9BV enables higher efficiency of delivery of the construct, as evidenced by BFP fluorescence imaging of live cells and immunofluorescence staining of the FLAG epitope in fixed cells. The scale bar represents 100 μm.

(G) Lentiviral delivery of VdC9BV is more efficient than delivery by transfection. By flow cytometry analysis, ~50% of the transduced cells were found to be BFP positive compared with ~5% by transfection (two-tailed unpaired t test, p ≤ 0.0001, n = 3 biological replicates; error bars represent SEM).

(H) qRT-PCR evaluation of relative Myod1 mRNA expression in C3H10T1/2 on day 6 postinduction of VdC9BV expression in the presence of various combinations of three gRNA molecules (p ≤ 0.001, one-way ANOVA, Dunnett’s post hoc comparing all the groups to the [−] gRNA control group [M2rtTA + VdC9BV + Empty FUW], n = 3 biological replicates, error bars represent SEM).
evidenced by both BFP and immunofluorescence staining of the FLAG epitope (Figure 1F). Approximately 50% of the transduced cells were found to be BFP positive as compared with 5% by transfection (Figure 1G). To ensure efficient gRNA codelivery, we developed a lentivirus-based U6 promoter-driven gRNA delivery system (Figure 1C). In separate experiments, we observed an ~6-fold upregulation of Myod1 mRNA levels when all components were delivered by transfection (Figure S1C) compared with ~60-fold by lentiviral transduction (Figure 1H).

We initially tested this system’s capacity to induce epigenetic reprogramming toward the SkM lineage in C3H10T1/2 cells previously shown to readily undergo this transformation (Davis et al., 1987). We designed three separate gRNAs (1–3) targeted to different positions proximal to the Myod1 transcription start site and initially co-delivered all three gRNAs (Figure S1D). We found that VP64 fusion and the presence of the gRNA were essential for endogenous Myod1 gene activation (Figure S1C). Activation of the endogenous Myod1 gene in C3H10T1/2 was sufficient to initiate the SkM reprogramming process, as determined by gene expression analysis, phase contrast imaging, and immunocytochemistry (Figures 2A–2C and S1B). Myotubes were visible as early as 4.5 days after transduction. Nuclear-localized blue fluorescence was observed in the myotubes during the time VP64-dCas9-BFPVP64 was kept induced by doxycycline addition (Figure S1B). Importantly, the myotubes stained positive for the skeletal transcription factors (TFs) MYOD1 and MYOG, and the sarcomeric proteins actinin, desmin, myosin heavy chain, and titin (Figures 2C and S2A). The myotubes were also multinucleated, indicating cell fusion, one of the hallmarks of myogenic reprogramming (Figure 2C). When examining the effect of individual gRNA molecules in inducing Myod1 mRNA expression, we determined that gRNA3 alone was as potent as all three gRNAs combined (Figure 1H). We speculate that factors, including distance from transcription initiation site, low binding affinity, and binding site competition with endogenous TFs, are the possible causes of gRNA1 and gRNA2 inactivity. C3H10T1/2 cells fused in the 8 days that Myod1 was expressed under doxycycline-induced VP64-dCas9-BFPVP64 expression, as demonstrated quantitatively by a recombination-based cell fusion assay. The fusion process continued unabated even after the withdrawal of doxycycline, thereby indicating stability of this phenotypic transformation (Figure 2D).

We tested whether coexpression of a single targeting gRNA molecule (gRNA3) along with VP64-dCas9-BFPVP64 would be sufficient to induce reprogramming of primary MEFs (Figure 2E). Using gene expression analysis, we readily detected transcriptional activation of the mature myotube markers MyI1, Ckm, Desmin, and ChmA1 (Figure 2F). We detected formation of multinucleated myotubes expressing skeletal TFs and striated sarcomeres indicating a high degree of cytoskeletal organization and maturity (Figures 2G and S2B). Within 3 weeks of cell transduction, we observed myotubes exhibiting spontaneous intermittent twitching (Movie S1).

To test whether transient expression of VP64-dCas9-BFPVP64 is sufficient to induce sustained expression of Myod1, we induced transgene expression between days 2 and 10 posttransduction. We detected an initial robust upregulation of VP64-dCas9-BFPVP64 expression followed by a gradual downregulation suggesting transgene silencing in the reprogramming cells (Figure 3A). Doxycycline removal coincided with rapid VP64-dCas9-BFPVP64 decline to levels measured prior to initial induction. VP64-dCas9-BFPVP64 expression also coincided with an increase in Myod1 expression. Importantly, Myod1 mRNA levels remained elevated for the duration of the 18-day experiment even after doxycycline removal (day 10), suggesting that maintenance of endogenous activation of Myod1 is stable and independent of VP64-dCas9-BFPVP64 activity. dCas9 with VP64 fused to its C terminus (dCas9-BFPVP64) failed to activate the Myod1 locus to levels sufficient to initiate cellular reprogramming even in the presence of three gRNAs (Figures 3B, 3C, and S3A). However, in the presence of three gRNAs, the mean fold-activation level by dCas9-BFPVP64 was higher than with a single gRNA (Figure 3C), consistent with recent reports of synergistic effects of multiple gRNAs with dCas9 and a single VP64 fusion (Cheng et al., 2013; Maeder et al., 2013; Mali et al., 2013a; Perez-Pinera et al., 2013). The N-terminal-only VP64 fusion protein (VP64-dCas9-BFP) also failed to match the ability of VP64-dCas9-BFPVP64 to activate sufficient endogenous Myod1 for reprogramming in the presence of a single gRNA (Figures 3B, 3D, and S3A). Interestingly, VP64-dCas9-BFPVP64 was able to activate the endogenous human MYOD1 locus significantly in the presence of a single gRNA, whereas the N-terminal-only VP64 fusion protein (VP64-dCas9-BFP) and the C-terminal-only VP64 fusion protein (dCas9-BFPVP64) failed to do so (Figure 3E). These observations suggest that the inclusion of both VP64 activator domains fused at the two termini of dCas9-BFP significantly improves the activation capacity of this targeting platform. However, the endogenous human MYOD1 transactivation was an order of magnitude lower than that observed in the murine system. Transactivation of higher levels of MYOD1 mRNA may require extensive optimization of the gRNA target sequences in the future.

To test whether the direction of the genomic strand targeted by the gRNA molecule was a factor in its capacity to recruit dCas9 and activate expression of the endogenous locus, we designed a gRNA molecule (gRNA4) targeting the minus strand of the same region for which gRNA3 was
designed (PAM separated by 14 bases) (Figure S1D). Interestingly, gRNA4 was also able to activate Myod1 expression and initiate reprogramming, indicating the system is insensitive to the target strand (Figures 3F and S3B).

We also evaluated the effect of BFP fusion on the activity of Vp64-dCas9-BFP-Vp64. Following transduction at equivalent multiplicity of infection (MOIs), the construct without BFP showed statistically higher levels of dCas9 expression, although it was less efficacious than the BFP-fused form in inducing expression of Myod1 (Figures 3B and 3G). However, both forms led to reprogramming (Figure S3C). The omission of BFP from the C-terminal-only VP64 fusion (VP64dCas9-BFPVP64) in the presence of three gRNAs targeting the Myod1 locus, results in upregulation of major SkM genes in C3H10T1/2 cells detected by RT-PCR on day 18 posttransduction.

Reprogramming of C3H10T1/2 cells by VdC9BV is characterized by significantly higher levels of Myod1 and Myog mRNA detected by qRT-PCR on day 18 posttransduction (two-way ANOVA, Bonferroni post tests, p ≤ 0.001 for both Myod1 and Myog, n = 3 biological replicates, error bars represent SEM). Fold change in expression is relative to the (−) gRNA control group.

Reprogrammed myotubes were absent in the control group (phase contrast [PC] imaging and ICC). The scale bar represents 100 μm.

Cell fusion assay. Cell fusion in the reprogramming cells was quantified by detecting the luciferase released after mixing two lentivirally transduced cell populations: one expressing Cre recombinase (LV-Cre) and a second containing a Cre recombinase responsive Floxed-Stuffer Luciferase cassette (LV-Floxed Luc). Fold change in relative light units (RLUs) was calculated relative to a Luc-only group (VdC9BV, M2rtTA, and LV-Floxed Luc). The RLU fold change of the (+) gRNA3 group is significantly higher on day 22 than on day 14 (two-way repeated measures [RM] ANOVA, Bonferroni post hoc test, p ≤ 0.001, n = 3 biological replicates, error bars = SEM). Overall, the (+) gRNA3 group has significantly higher luciferase activity than other groups (two-way RM ANOVA, p ≤ 0.0001).

Timeline denoting the main steps of the MEF reprogramming process.

Gene expression analysis (RT-PCR) performed on reprogrammed and control MEFs.

Imaging of the reprogrammed MEFs shows the presence of multinucleated myotubes containing SkM markers. No myogenesis was observed in the (−) control group. Confocal imaging shows cross-striated organization of ACTN1 and F-actin.

(F) Gene expression analysis (RT-PCR) performed on reprogrammed and control MEFs.

(G) Imaging of the reprogrammed MEFs shows the presence of multinucleated myotubes containing SkM markers. No myogenesis was observed in the (−) control group. Confocal imaging shows cross-striated organization of ACTN1 and F-actin.
protein (dCas9 VP64) failed to impart ability to activate the endogenous Myod1 locus (Figure S3D). We speculate that the positive effect of BFP on Myod1 expression may be due to increased spacing between the two VP64 domains or increased flexibility of the domains.

We also compared the VP64 dCas9-BFP VP64-mediated activation of endogenous Myod1 gene to transgenic MYOD1 overexpression for its ability to reprogram cells. C3H10T1/2 cells were transduced with VP64 dCas9-BFP VP64 and M2rtTA. The resultant BFP-positive cells were sorted for BFP expression and the BFP+ cells were transduced with either gRNA or doxycycline-inducible transgenic human MYOD1 at equivalent MOIs (MOI = 10) (Figures 4A and 4B). During the induction phase of the reprogramming...
process, endogenous Myod1 expression over the entire duration was significantly higher in the VP64dCas9-BFPVP64/gRNA-mediated activation group than the transgenic MYOD1 overexpression group. Expression kinetics of Myog and Desmin were similar in both the groups (Figure 4C).
The effects of limiting the activation of the \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64}/gRNA system (by varying the duration of doxycycline exposure) on myogenic gene expression were compared with the MYOD1 overexpression system. Results showed that even a single day of doxycycline exposure (from day 2 to day 3 posttransduction) was adequate to activate the downstream myogenic genes (Myog and Desmin) to similar levels in both the groups and similar to levels achieved after 8 days of induction (Figure 4D). Almost all the markers on day 18 posttransduction after a maximum 8 days of induction had similar expression levels in both the groups (Figures 4D and S4A). However, both Myod1 expression and the percentage of MYOD1 TF\textsuperscript{+} nuclei were approximately 3-fold higher in the Myod1 transactivation group than the MYOD1 overexpression group (Figures 4D, 4E, and S4B). It indicates that the higher expression of Myod1 on quantitative RT-PCR (qRT-PCR) (qRTPCR) was probably a result of more MYOD1 TF\textsuperscript{+} cells rather than higher expression levels in individual cells. Myog expression in C3H10T1/2 followed a pattern similar to Myod1 (Figures 4D and 4E). Similar differential Myod1 expression was also observed in the reprogrammed MEFs. However, expression of all the other myogenic markers and the percentage of MYOG\textsuperscript{+} nuclei were similar in both groups (Figures S4C and S4D). It may be speculated that \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64}/gRNA action renders the endogenous Myod1 locus more receptive to MYOD1 TF-positive feedback response (Zingg et al., 1994). However, other cellular factors needed in conjunction with MYOD1 TF to initiate skeletal reprogramming may not be present in adequate quantities, thereby explaining the lack of a significant difference for other downstream late myogenic genes.

This study shows that the dual fusion of the VP64 transactivation domain to both the N and C terminus of dCas9 enables a high level of endogenous Myod1 activation for the direct conversion of primary murine fibroblasts into SkMs. This potentiation of the transactivation process can be explained by an increased probability of the TFs homing onto two VP64 domains compared with just one VP64 domain. As a result, transcription can be initiated more frequently in the presence of \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64}. Moreover, a synergistic transactivation effect resulting from favorable interactions of the TF complexes assembled on both the terminus of \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64} can also help in explaining its potency. Improved nuclear localization of dCas9 afforded by increasing the number of NLS sequences, efficient lentiviral transgene delivery, inclusion of a BFP spacer sequence to decrease steric hindrance, and the identification of an efficient gRNA together complemented the effect of an additional VP64 domain in \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64} to efficiently transactivate the endogenous Myod1 locus. The results also demonstrate that CRISPR/Cas9-based transactivation performs comparably to the traditional MYOD1 overexpression-based skeletal reprogramming in upregulating some major myogenic genes. It augurs well for the potential use of this tool in reprogramming protocols that require complex and multiple TF activation.

In conclusion, we expect that this \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64} platform, which is significantly more potent than previous versions, will encourage the adoption of CRISPR/Cas9-based TF technology to achieve multiplexed gene activation for reprogramming as well as nonreprogramming applications in basic science, biotechnology, and medicine.

**EXPERIMENTAL PROCEDURES**

**gRNA Design**

The gRNAs were designed by utilizing UCSC genome browser tracks (Feng Zhang Lab, MIT). The mouse Myod1 locus-specific gRNAs 1–3 were selected to represent specific locations at a decreasing distance from the designated Myod1 transcription initiation site, respectively. gRNA4 was selected to target the genome in the immediate vicinity of gRNA3, albeit on the opposite strand (Figure S1D). The human gRNA was picked to target the genome within 200 bases upstream of the MYOD1 transcription initiation site.

**\textsuperscript{V64}dCas9-BFP\textsuperscript{VP64} Plasmid Construction**

pdCas9::BFP-humanized plasmid (Addgene plasmid 44247) was used as the source of the dCas9-BFP fusion cassette. gBlock DNA fragments (Integrated DNA Technologies) were utilized to generate the \textsuperscript{V64}dCas9-BFP\textsuperscript{VP64} construct from dCas9-BFP and subsequently cloned into a lentiviral vector.

**Cell Culture, Transfection, and Viral Transduction**

All of the cell types used in this study were initially cultured with a high serum media (10% fetal bovine serum in Dulbecco’s modified Eagle’s medium-high glucose [DMEM-HG]). The cells were seeded at a density of 25,000 per well of a 12-well plate for transfection or transduction with components of the reprogramming system. Liposome-based transfection and lentiviral transduction were done while maintaining a constant dosage among the tests and controls. Low serum medium (2% horse serum in DMEM-HG) was used to induce skeletal differentiation.

**Cell Staining and qRT-PCR**

Four percent paraformaldehyde fixed cells were permeabilized with 0.2% Triton X-100 and subsequently stained with primary antibodies and fluorescent-labeled secondary antibodies. Relative quantification of gene expression (delta-delta Ct method) was performed on the RT-PCR data obtained by using SYBR green chemistry.

**Statistical Analysis**

Statistical analysis was done by GraphPad Prism 5.0 software. The convention followed to denote significance: *p \leq 0.05, **p \leq 0.01, and ***p \leq 0.001.
SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2014.09.013.

ACKNOWLEDGMENTS

Support from NIH (K.W.L.: EB015300, AI096305, UH2TR000505; C.A.G.: DP2OD008856, R01DA036865), National Science Foundation (CBET-1151035), American Heart Association (10SDG3060033), and the Muscular Dystrophy Association (22347, 27152, and 14883, respectively). C.A.G. is an inventor on patent applications related to genome engineering and a scientific advisor to Editas Medicine.

Received: March 27, 2014
Revised: September 16, 2014
Accepted: September 17, 2014
Published: October 23, 2014

REFERENCES


