*Communication*

# Synergistic Upregulation of Target Genes by TET1 and VP64 in the dCas9–SunTag Platform

**Sumiyo Morita, Takuro Horii, Mika Kimura and Izuho Hatada \***

Laboratory of Genome Science, Biosignal Genome Resource Center, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15 Showa-machi, Maebashi 371-8512, Japan; msumiyo@gunma-u.ac.jp (S.M.); horii@gunma-u.ac.jp (T.H.); mikimura@gunma-u.ac.jp (M.K.)

\* Correspondence: hatada@gunma-u.ac.jp

check for
updates

**Abstract:** Overexpression of a gene of interest is a general approach used in both basic research and therapeutic applications. However, the conventional approach involving overexpression of exogenous genes has difficulty achieving complete genome coverage, and is also limited by the cloning capacity of viral vectors. Therefore, an alternative approach would be to drive the expression of an endogenous gene using an artificial transcriptional activator. Fusion proteins of dCas9 and a transcription activation domain, such as dCas9–VP64, are widely used for activation of endogenous genes. However, when using a single sgRNA, the activation range is low. Consequently, tiling of several sgRNAs is required for robust transcriptional activation. Here we describe the screening of factors that exhibit the best synergistic activation of gene expression with TET1 in the dCas9–SunTag format. All seven factors examined showed some synergy with TET1. Among them, VP64 gave the best results. Thus, simultaneous tethering of VP64 and TET1 to a target gene using an optimized dCas9–SunTag format synergistically activates gene expression using a single sgRNA.

**Keywords:** CRISPR/Cas9; dCas9; SunTag; Tet1; VP64

## 1. Introduction

Overexpression of a gene of interest is a general approach used in both fundamental research and therapeutic applications. Conventionally, overexpression uses an exogenous gene. In basic research, exogenous expression of a genome-wide cDNA library is the most commonly used gain-of-function approach for the systematic elucidation of gene function. In a therapeutic context, exogenous expression of several key transcription factors in differentiated cells leads to a transition in cellular state and generates induced pluripotent stem cells (iPSCs) [1]. Exogenous expression of a functional copy by a viral vector, such as an adeno-associated virus (AAV) vector, can be used in gene therapy to replace mutant genes, thereby treating human diseases caused by haploinsufficiency [2]. However, this approach is inefficient and has two crucial limitations. First, it is difficult to construct a cDNA library that covers the whole genome, in particular in a manner that encompasses isoform variance. Second, large cDNA sequences are often difficult to clone into size-limited viral expression vectors.

One alternative approach for overexpression is activation of endogenous genes by an artificial transcriptional activator consisting of a programmable DNA-binding protein and an effector domain. In this method, the programmable DNA-binding protein tethers the transcription activation domains to the promoter of the target gene [3–9]. The following are three types of DNA-binding proteins currently in wide use: zinc finger proteins [10], transcription activator-like effectors (TALEs) [11], and catalytically dead Cas9 (dCas9) [12–14]. Among these, dCas9 is the most appropriate for genome-wide gain-of-function screening due to its simplicity and versatility.

Fusions of dCas9 and a transcription activation domain (e.g., dCas9–VP64) are widely used for the activation of endogenous genes. In this approach, expression is induced by targeting the promoter region of the gene of interest. Programmable targeting of endogenous loci can be achieved by an individual single-guide RNA (sgRNA) [7–9]. However, because the activation range is low when only a single sgRNA is used, tiling of the target promoter region with several sgRNAs is often necessary to achieve robust transcriptional activation [7–9]. Consequently, it is difficult to use this format for genome-wide gain-of-function screening and gene therapy. To solve this problem, a format that recruits multiple factors to dCas9, thereby improving the activation range, was recently developed [15]. The synergistic activation mediator (SAM) system consists of dCas9–VP64, modified gRNA containing RNA aptamers, and MS2 bacteriophage coat protein fused to a transcription activation domain p65–HSF1. VP64 fused to dCas9 and MS2–p65–HSF1 recruited to dCas9 by the RNA aptamers synergistically activates transcription of the target gene.

In previous work, we developed a modified dCas9–SunTag format for efficient, targeted demethylation and activation of specific DNA loci using a single sgRNA [16]. SunTag is a repeated GCN4 peptide tag array that recruits multiple copies of an effector molecule fused to anti-GCN4 peptide antibody (scFv) [17]. We optimized the linker length between the tags to maximize transcriptional activation by the effector TET1, which performs the first enzymatic steps during DNA demethylation.

Here, we have described simultaneous tethering of VP64 and TET1 to a target gene using the optimized dCas9–SunTag format, resulting in synergistic activation of gene expression.

## 2. Results

### 2.1. Selection of Another Factor that Activates Gene Expression Synergistically with TET1 in the Modified dCas9–SunTag Format

In the modified dCas9–SunTag format, dCas9 is fused to a tandem array of multiple copies of a 19-amino acid (aa) GCN4 peptide tag separated by a 22-amino acid linker [16]. This format can recruit multiple copies of anti-GCN4 peptide antibody (scFv)-fused catalytic domain of TET1 (scFv-TET1) to the target promoter, leading to demethylation and activation of the gene. If anti-GCN4 peptide antibody (scFv) that is fused to another factor (scFv–Factor X) is used along with scFv–TET1, then both TET1 and Factor X can be recruited to the target with a single sgRNA (Figure 1) and should synergistically activate the target gene. Therefore, we used this format to search for other factors that synergistically activate gene expression, when used in conjunction with TET1.
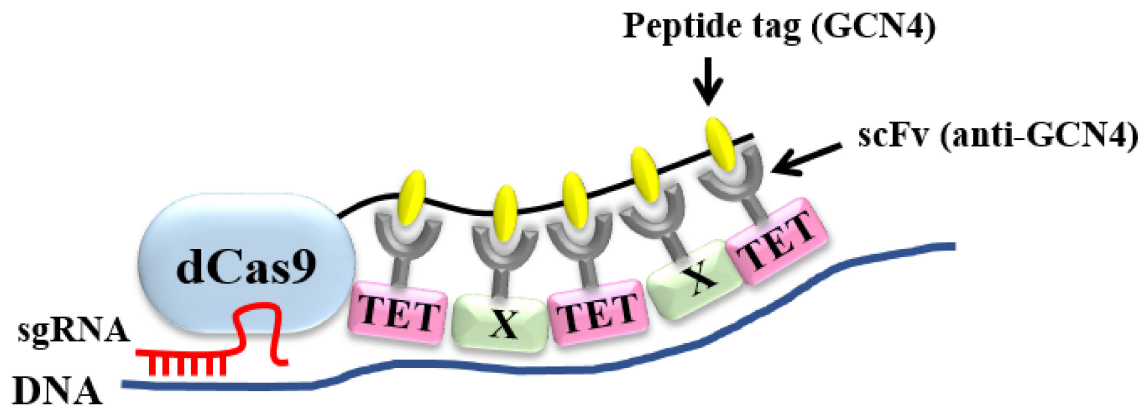
**Figure 1.** Modified dCas9–SunTag format for simultaneous recruitment of TET1 and another factor (Factor X). In the modified dCas9–SunTag format, dCas9 is fused to a tandem array of multiple copies of GCN4 peptide tag separated by a 22-amino acid linker. This format can recruit multiple copies of fusion proteins of anti-GCN4 peptide antibody (scFv) with TET1 (scFv–TET1) and Factor X (scFv–X), which should synergistically activate the target gene.

We investigated the synergy between TET1 and other factors, hereafter referred to as X. Candidate factors were transcription factors (VP64 and p65HSF1), a coactivator (p300), a putative SWI/SNF subunit (SS18), a heterochromatin relaxer (GADD45A), and pioneer factor factors (FOXA1 and PU.1). Ten genes (CARD9, KDM2B, RAB19, CNKSR1, SBNO2, SPARC, CLEC11A, HGF, TCF21, and TINAGL1) were examined for each factor. All these genes are hypermethylated in the A549 lung adenocarcinoma cell line. Expression levels were normalized against the corresponding level of actin mRNA (Figure S1) and were presented as fold changes relative to GFP-transfected experimental controls (Figure 2). When X was VP64, p65HSF1, p300, SS18, GADD45A, FOXA1, or PU.1, cells were transfected with dCas9–SunTag and scFv–TET1 (Figure 2, green bar); dCas9–SunTag and scFv–X (Figure 2, blue bar); or dCas9–SunTag, scFv–TET1 and scFv–X (Figure 2, red bar). Synergy between TET1 and X was judged as follows. The expression level in cells transfected with dCas9–SunTag and scFv–X (Figure 2, blue bar) were compared with the level in cells transfected with dCas9–SunTag and scFv–TET1 (Figure 2, green bar), and the cells that exhibited a higher expression level were selected. These expression levels were compared with those of cells transfected with dCas9–SunTag, scFv–TET1, and scFv–X (Figure 2, red bar). The results are summarized in Figure 3. The SunTag system with only TET1 yielded 1.5- to 78-fold upregulation (average: 20-fold; Figure 2, green bar). On the other hand, the SunTag system with TET1 and VP64 yielded 3.5- to 1139-fold upregulation (average: 212-fold; Figure 2, red bar). In eight of ten genes examined, synergy between TET1 and VP64 was observed at a rate of 1.8- to 21-fold (average: 6.5-fold; Figures 2 and 3). Transcription factor p65HSF1 exhibited synergy with TET1 in five of ten genes examined. Other factors exhibited synergy with TET1 in no more than three genes. Thus, relative to other factors, transcription factors yielded a stronger synergistic effect.
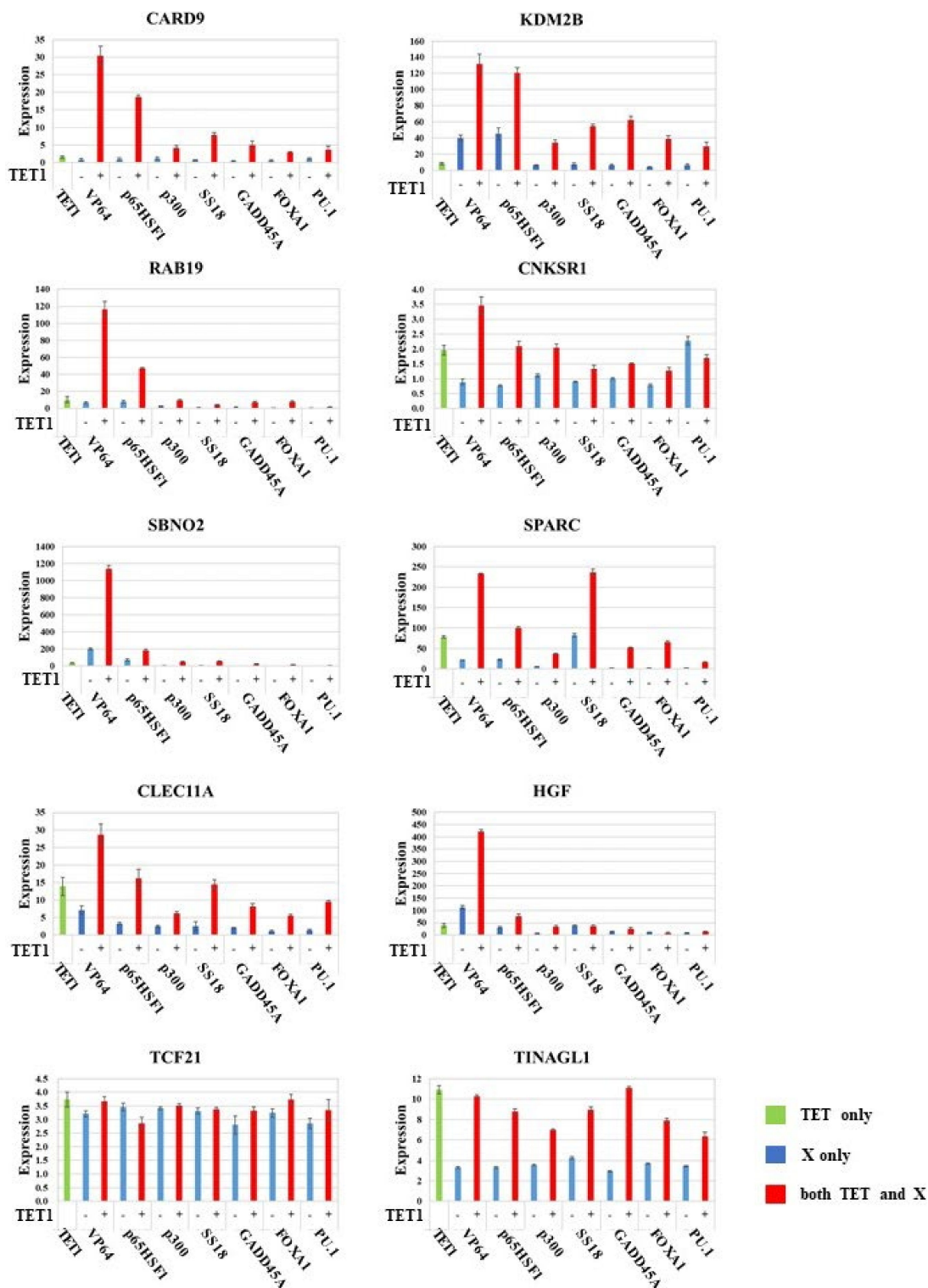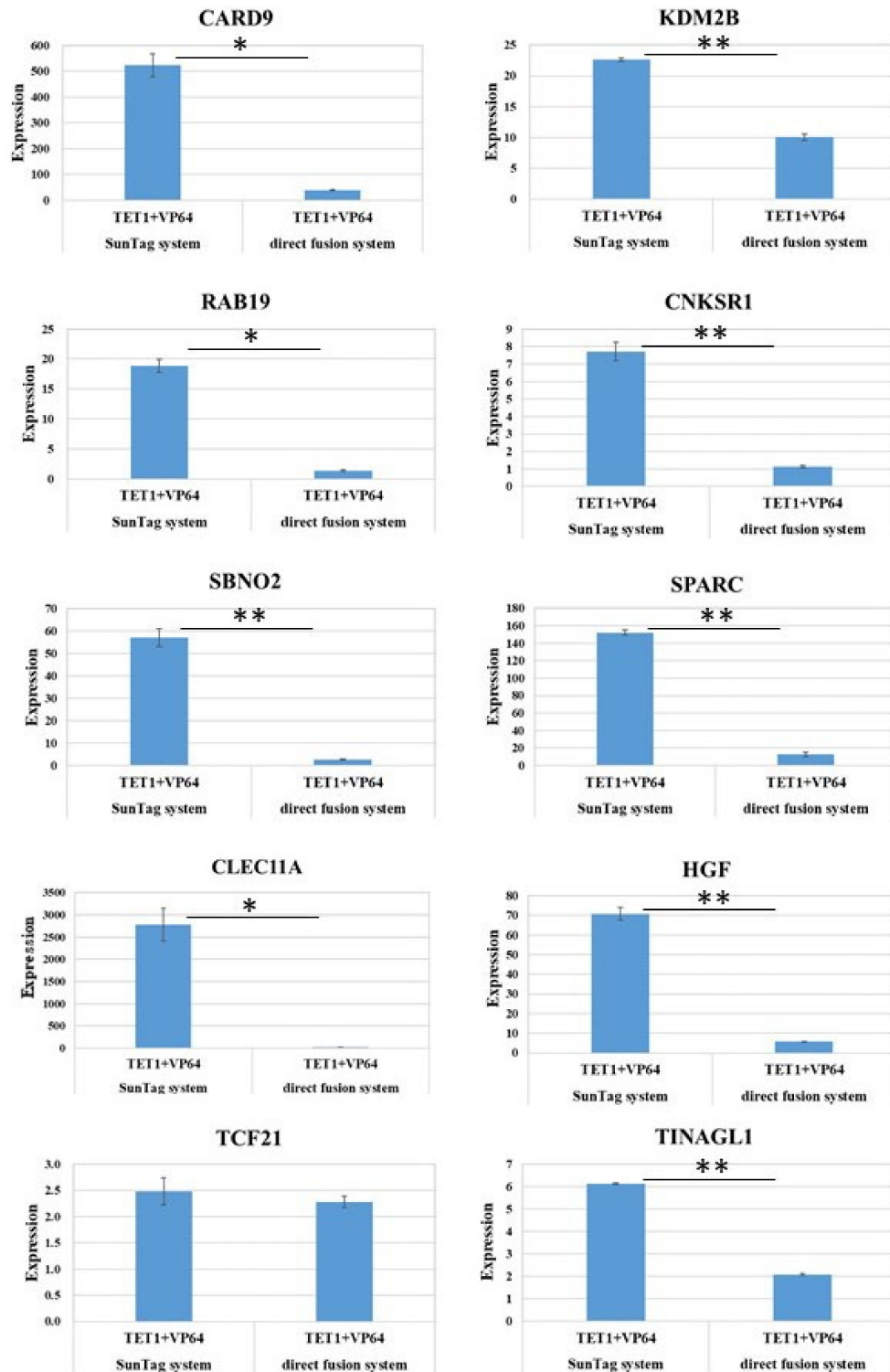
**Figure 2.** Expression levels in A549 cells transfected with dCas9–SunTag and scFv–TET1 (green bar); dCas9–SunTag and scFv–X (blue bar); and dCas9–SunTag, scFv–TET1, and scFv–X (red bar). Expression levels, determined by RT-PCR, were normalized against the corresponding level of actin mRNA and are presented as fold changes relative to GFP (green fluorescent protein)-transfected experimental controls. Results obtained using sgRNAs targeting *CARD9*, *KDM2B*, *RAB19*, *CNKSR1*, *SBNO2*, *SPARC*, *CLEC11A*, *HGF*, *TCF21*, and *TINAGL1* are shown.

## 2.2. Comparison of Synergistic Effects between Modified SunTag and dCas9 Direct Fusion

Direct fusions of dCas9 and a transcription activation domain, such as dCas9–VP64, are widely used for the activation of endogenous genes [7–9]. However, the activation range is low when an individual single-guide RNA (sgRNA) is used. Activation by dCas9–VP64 is improved when it is used with dCas9-TET1 [18]. We compared the activation level of this direct fusion system with that of our SunTag system. Using the same single sgRNA for each gene, we compared the expression levels in cells transfected with dCas9–SunTag, scFv–TET1, and scFv–VP64, with those in cells transfected with dCas9–TET1 and dCas9–VP64. For nine of ten genes examined, the SunTag system yielded significantly better activation (Figure 4).

| | CARD9 | KDM2B | RAB19 | CNKSR1 | SBNO2 | SPARC | CLEC11A | HGF | TCF21 | TINAGL1 |
|---|---|---|---|---|---|---|---|---|---|---|
| VP64 | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | |
| p65HSF1 | ■ | ■ | ■ | | ■ | ■ | | | | |
| p300 | ■ | ■ | | | | | | | | |
| SS18 | ■ | ■ | | | | ■ | | | | |
| GADD45A | | ■ | | | | | | | | |
| FOXA1 | ■ | ■ | | | | | | | | |
| PU.1 | | ■ | | | | | | | | |

**Figure 3.** Synergy of TET1 and other factors in the modified SunTag format. Red cells indicate significant synergy between TET1 and another factor ($p < 0.05$). Each row indicates a factor (VP64, p65HSF1, p300, SS18, GADD45A, FOXA1, or PU.1), and each column indicates a gene (*CARD9*, *KDM2B*, *RAB19*, *CNKSR1*, *SBNO2*, *SPARC*, *CLEC11A*, *HGF*, *TCF21*, or *TINAGL1*).

**Figure 4.** Comparison of synergy in modified SunTag and dCas9 direct fusion. Expression levels are shown for cells transfected with dCas9–SunTag, scFv–TET1, and scFv–VP64 (left); or dCas9–TET1 and dCas9–VP64 (right). Expression levels, as determined by RT-PCR, are normalized against the level of actin and presented as fold changes relative to GFP-transfected experimental controls. Results using sgRNAs for *CARD9*, *KDM2B*, *RAB19*, *CNKSR1*, *SBNO2*, *SPARC*, *CLEC11A*, *HGF*, *TCF21*, and *TINAGL1* are shown. * $p < 0.05$, ** $p < 0.01$.

## 3. Discussion

In the modified dCas9–SunTag format, multiple factors can be recruited to the target gene to synergistically activate transcription (Figure 1). We sought to identify the factor that exhibited the best synergy with TET1 in the dCas9–SunTag format. All seven of the factors that we tested showed some synergy with TET1 in at least one of ten genes examined. Among them, VP64 gave the best results, exhibiting a synergistic effect with TET1 in eight of ten genes examined (Figures 2 and 3). Although it remains unclear why VP64 gave the best synergy, it may be due to the fact that VP64 had the lowest molecular weight of the seven factors examined, and consequently may cause the least steric hindrance. Consistent with this, SunTag with a short linker between the tags yields poorer activation as a result of steric hindrance in our previous work [16]. A linker of 5 amino acids, 22 amino acids, and 43 amino acids were compared in the SunTag system with TET1. The 22-amino acid linker gives the best result and the 5-amino acid linker is the worst [16]. Thus, different linker lengths could change the activity of the factors. Judging from the fact that the molecular weight of TET1 is greater than all the factor X including VP64, which is the smallest among the factor X, the 22-amino acid linker would also give the best result in the SunTag system with TET1 and VP64.

Activated expression levels vary among genes. The expression level of each gene could be affected by several factors such as, structure of the promoter and efficiency of gRNAs. The expression levels observed in the SunTag system with TET1 and VP64 showed mild correlation with CpG observed/expected (O/E) ratio around sgRNAs (R = 0.57). This suggests that CpG-rich promoters tend to be more activated.

Our SunTag system yielded better activation than the direct fusion system (dCas9–TET1 and dCas9–VP64) (Figure 4) using the same single sgRNA. The activation range of the direct fusion system is low when using a single sgRNA; consequently, tiling with multiple sgRNAs is usually necessary to achieve robust transcriptional activation [18]. Thus, one of the merits of our system is that only a single sgRNA is required for robust activation. This makes the system simple and versatile, and thus very useful for genome-wide screening or therapeutic applications. Another merit of our system is the shorter length of the constructs, due to the relatively small size of the SunTag. Constructs for the fusion system must be at least the sum of the lengths of the dCas9 and TET1 genes. Unfortunately, in this regard, dCas9 (4.2 kb) and catalytic domain of TET1 (2.2 kb) are long DNAs, and this limits their use in gene therapy, which usually uses viral vectors with limited packaging capacity (4.7 kb for AAV). Even using shorter dCas9 from *Staphylococcus aureus* (dSaCas9, 3.2 kb), dSaCas9-TET1 (totally 5.4 kb) exceeds the size limit of the AAV vector (4.7 kb). On the other hand, the length of a construct for the SunTag system is the sum of the lengths of dCas9 and SunTag (0.5 kb), the latter of which is quite short.

Overexpression of endogenous genes is becoming increasingly important for therapeutic applications. For example, overexpression of a functional endogenous copy has the potential to rescue human diseases caused by haploinsufficiency [19]. Alternatively, overexpression of a protein similar to one encoded by a mutant gene could treat human diseases caused by recessive mutations [20]. For example, congenital muscular dystrophy type 1A (MDC1A) is an autosomal recessive disorder caused by mutations in *LAMA2* that cause production of nonfunctional laminin-$\alpha$2. Viral overexpression of *Lama1*, which encodes a similar protein, in a mouse model of MDC1A improved disease symptoms and slowed progression [20]. Thus, the development of a system that yields robust activation and yet is small enough to be cloned into a virus vector, is an important priority for therapeutic applications. Therefore, we anticipate that our modified SunTag system will facilitate advances in gene therapy.

## 4. Materials and Methods

### 4.1. Construction of sgRNAs

For each target gene, a unique 20 bp sequence was selected around the transcription start site. Selected gRNAs were cloned under control of the human U6 promoter (gRNA_Cloning Vector BbsI, Addgene 128433). The target sequences are described in the Table S1.

### 4.2. Cell Culture

A549 (RIKEN BRC, Tsukuba, Japan) cells were cultured at 37 °C under 5% $CO_2$ in minimum essential medium (MEM) (M4655-500ML, Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS) and non-essential amino acids. A549 cells were transfected with Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). Blasticidin-S (Invitrogen, Carlsbad, CA, USA) was added to the culture medium at a final concentration of 2 µg/mL 48 h after transfection. After 3 days of selection, cells were harvested and subjected to analysis. The molar ratio of the dCas9–SunTag vector, scFv–TET1 vector and scFv–X vector to sgRNA vector in the transfection was 1:1:1:1.5, respectively. In case of the direct fusion system, the molar ratio of the dCas9–TET1 vector and dCas9–VP64 vector to sgRNA vector in the transfection was 1:1:1.5, respectively. All constructs except for the sgRNA vector were expressed under the control of the CAG promoter.

### 4.3. Quantitative RT-PCR Analysis

Total RNA was isolated from cells using the AllPrep DNA/RNA micro kit (Qiagen, Hilden, Germany). Gene expression was measured with a LightCycler 98 (Roche, Roche, Basel, Switzerland) using TB Green Premix Ex Taq II (Takara, Kusatsu, Japan). Expression levels were normalized against the corresponding level of actin mRNA and were presented as fold changes relative to GFP-transfected experimental controls. Primer sequences are described in the Table S1.

**Supplementary Materials:** Supplementary Materials can be found at http://www.mdpi.com/1422-0067/21/5/1574/s1. Figure S1: Expression levels (not presented as fold changes to GFP-transfected experimental controls) in A549 cells transfected with dCas9–SunTag and scFv–TET1 (green bar); dCas9–SunTag and scFv–X (blue bar); and dCas9–SunTag, scFv–TET1, and scFv–X (red bar); Table S1: Target sequences and primer sequences used for RT-PCR.

**Author Contributions:** Conceptualization, S.M. and I.H.; methodology, S.M., M.K., and T.H.; investigation, S.M. and I.H.; data curation, S.M., M.K., T.H., and I.H.; writing—original draft preparation, S.M. and I.H.; review and editing, I.H.; supervision, I.H.; funding acquisition, I.H. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest.

### Abbreviations

| | |
|---|---|
| AAV | adeno-associated virus |
| TALEs | transcription activator-like effectors |
| dCas9 | catalytically dead Cas9 |
| sgRNAs | single-guide RNAs |
| SAM | synergistic activation mediator |

### References

1. Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676. [CrossRef] [PubMed]
2. Kublaoui, B.M.; Holder, J.L., Jr.; Tolson, K.P.; Gemelli, T.; Zinn, A.R. SIM1 overexpression partially rescues agouti yellow and diet-induced obesity by normalizing food intake. *Endocrinology* **2006**, *147*, 4542–4549. [CrossRef] [PubMed]
3. Beerli, R.R.; Segal, D.J.; Dreier, B.; Barbas, C.F., 3rd. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14628–14633. [CrossRef] [PubMed]

4.  Zhang, F.; Cong, L.; Lodato, S.; Kosuri, S.; Church, G.M.; Arlotta, P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **2011**, *29*, 149–153. [CrossRef] [PubMed]

5.  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. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **2013**, *154*, 442–451. [CrossRef] [PubMed]

6.  Konermann, S.; Brigham, M.D.; Trevino, A.; Hsu, P.D.; Heidenreich, M.; Cong, L.; Platt, R.J.; Scott, D.A.; Church, G.M.; Zhang, F. Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **2013**, *500*, 472–476. [CrossRef] [PubMed]

7.  Maeder, M.L.; Linder, S.J.; Cascio, V.M.; Fu, Y.; Ho, Q.H.; Joung, J.K. CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **2013**, *10*, 977–979. [CrossRef] [PubMed]

8.  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. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **2013**, *10*, 973–976. [CrossRef] [PubMed]

9.  Mali, P.; Aach, J.; Stranges, P.B.; Esvelt, K.M.; Moosburner, M.; Kosuri, S.; Yang, L.; Church, G.M. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **2013**, *31*, 833–838. [CrossRef] [PubMed]

10. Porteus, M.H.; Carroll, D. Gene targeting using zinc finger nucleases. *Nat. Biotechnol.* **2005**, *23*, 967–973. [CrossRef] [PubMed]

11. Miller, J.C.; Tan, S.; Qiao, G.; Barlow, K.A.; Wang, J.; Xia, D.F.; Meng, X.; Paschon, D.E.; Leung, E.; Hinkley, S.J.; et al. A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **2011**, *29*, 143–148. [CrossRef] [PubMed]

12. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821. [CrossRef] [PubMed]

13. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-guided human genome engineering via Cas9. *Science* **2013**, *339*, 823–826. [CrossRef] [PubMed]

14. 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. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819–823. [CrossRef] [PubMed]

15. Konermann, S.; Brigham, M.D.; Trevino, A.E.; Joung, J.; Abudayyeh, O.O.; Barcena, C.; Hsu, P.D.; Habib, N.; Gootenberg, J.S.; Nishimasu, H.; et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **2015**, *517*, 583–588. [CrossRef] [PubMed]

16. Morita, S.; Noguchi, H.; Horii, T.; Nakabayashi, K.; Kimura, M.; Okamura, K.; Sakai, A.; Nakashima, H.; Hata, K.; Nakashima, K.; et al. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nat. Biotechnol.* **2016**, *34*, 1060–1065. [CrossRef] [PubMed]

17. Tanenbaum, M.E.; Gilbert, L.A.; Qi, L.S.; Weissman, J.S.; Vale, R.D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **2014**, *159*, 635–646. [CrossRef] [PubMed]

18. Baumann, V.; Wiesbeck, M.; Breunig, C.T.; Braun, J.M.; Koferle, A.; Ninkovic, J.; Gotz, M.; Stricker, S.H. Targeted removal of epigenetic barriers during transcriptional reprogramming. *Nat. Commun.* **2019**, *10*, 2119. [CrossRef] [PubMed]

19. Matharu, N.; Rattanasopha, S.; Tamura, S.; Maliskova, L.; Wang, Y.; Bernard, A.; Hardin, A.; Eckalbar, W.L.; Vaisse, C.; Ahituv, N. CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **2019**, *363*. [CrossRef] [PubMed]

20. Kemaladewi, D.U.; Bassi, P.S.; Erwood, S.; Al-Basha, D.; Gawlik, K.I.; Lindsay, K.; Hyatt, E.; Kember, R.; Place, K.M.; Marks, R.M.; et al. A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. *Nature* **2019**, *572*, 125–130. [CrossRef] [PubMed]