

Casilio: a versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling

Cell Research (2016) 26:254-257. doi:10.1038/cr.2016.3; published online 15 January 2016

Dear Editor,

The CRISPR-Cas9 system has recently been widely adopted in genome editing due to its simplicity [1-3]. Nuclease-deficient mutant dCas9 protein can be fused to effector domains and the fusion proteins can be guided by sgRNAs to genomic sites to regulate gene expression or label chromosomes [4-10]. However, only one type of effector is applied in most experiments due to the exclusive sgRNA:Cas9 pairing. Moreover, multimerization by directly fusing multiple copies of effectors with dCas9 protein to achieve sufficient effector activity is technically challenging. RNA aptamer approaches utilizing viral RNA sequences such as MS2 and PP7 have been combined with the CRISPR-Cas9 system to provide tools with improved multiplexing and multimerization functionalities [11, 12]. However, there are a limited number of well-characterized RNA aptamers. Furthermore, incorporation of three or more copies of these structured aptamers onto sgRNA reduces sgRNA expression, thus limiting the number of effector proteins that can be recruited. We established here the Casilio system by combining CRISPR-Cas9 and Pumilio RNA-binding protein. Pumilio and FBF proteins share a conserved Pumilio/FBF (PUF) RNA-binding domain which is programmable to bind a specific 8-mer RNA sequence (PUF-binding site, PBS) [13, 14] (Figure 1A). The Casilio system consists of the dCas9 protein, an sgRNA appended with one or more PUF-binding site(s) (sgRNA-PBS), and an effector fused with a PUF domain (PUF fusion) (Figure 1A). The sgRNA-PBS specifies both DNA binding via its spacer sequence and effector recruitment via PBS.

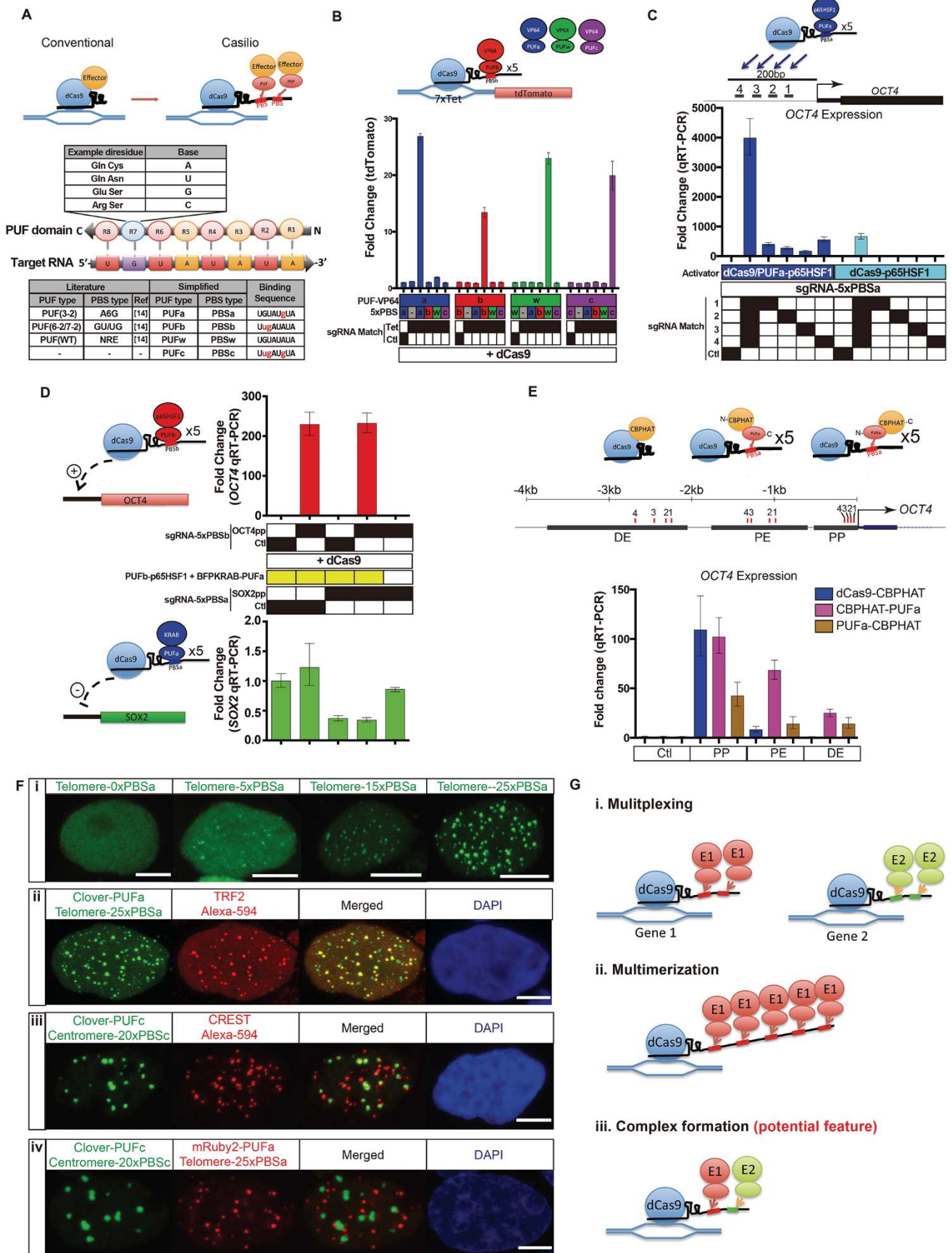
First, to test whether insertion of PBS at the 3' end of sgRNA affects its function, we generated a series of TetO-targeting (sgTetO) and control (sgCtl) sgRNAs with 0-47 copies of PBSa or PBSb, and tested their ability to direct dCas9-VP64 to activate tdTomato expression in a HEK293T/TetO::tdTomato cell line (Supplementary information, Figure S1A). All of the sgTetO-PBS but none of the sgCtl-PBS directed dCas9-VP64 to activate TetO::tdTomato, demonstrating that dCas9-VP64/sgRNA

can tolerate insertion of at least 47 copies of PBS. Next, we tested whether sgRNAs with different copy numbers of PBS can recruit PUF-VP64 to activate TetO::tdTomato reporter (Supplementary information, Figure S1B). For both PUFa-VP64/sgRNA-PBSa and PUFb-VP64/sgRNA-PBSb, the modules with 5-10 copies of PBS appended to the sgRNA activated the TetO::tdTomato most efficiently, whereas those with 15, 20, and 47 copies of PBS led to lower albeit substantial activation (Supplementary information, Figure S1B).

We then tested whether specificity of PUF-PBS pairs provides independence between Casilio modules. In addition to PUFa-VP64/sgRNA-PBSa and PUFb-VP64/sgRNA-PBSb, we constructed two additional activator modules, PUFw-VP64/sgRNA-PBSw and PUFc-VP64/sgRNA-PBSc (Figure 1A), and tested their ability to activate TetO::tdTomato expression. As shown in Figure 1B, PUF-VP64 activated TetO::tdTomato expression only when the sgRNA with the cognate PBS was provided. This demonstrates that the Casilio modules with different PUF domains can operate independently with high specificity.

A recent paper showed p65HSF1 as a potent transcriptional activation domain [15]. We constructed Casilio activator module PUFa-p65HSF1 and compared it with PUFa-VP64 for the ability to activate TetO::tdTomato. PUFa-p65HSF1 showed an increase in TetO::tdTomato activation by approximately 3-fold compared with PUFa-VP64 and even enabled activation with the presence of one copy of PBSa in the sgRNA (Supplementary information, Figure S1C). We thus switched to PUF-p65HSF1 for subsequent experiments.

To test the possibility that recruitment of multiple molecules of p65HSF1 by sgRNA-PBS might increase transactivation activity, we compared activation of endogenous genes *OCT4* and *SOX2* in HEK293T cells induced by the Casilio system with that induced by a direct dCas9-p65HSF1 activator using either a cocktail of 4 sgRNA-PBS per gene or individual sgRNA-PBS. Using either the sgRNA-PBS cocktail or a single sgRNA-PBS, we observed higher activation using the Casilio system



compared to the direct fusion for both *OCT4* and *SOX2* activation. We observed little to no activation by single targeting of the dCas9-p65HSF1 fusion protein to *OCT4* and *SOX2* while robust activation was observed in the corresponding Casilio experiments (Figure 1C and Supplementary information, Figure S1D). These results demonstrate the superior activity of Casilio activator over the direct fusion approach. We then sought to determine the optimal number of PBS for activation of endogenous genes and found that 5 copies of PBS produced the highest activation, resembling results from the TetO::tdTomato activation experiment (Supplementary information, Figure S1E).

To test the feasibility of the Casilio system to recruit various effectors to distinct targets, we tested repression of eGFP by a KRAB-PUFa repressor and activation of tdTomato by a PUFc-p65HSF1 activator simultaneously in a HEK293T/TetO::tdTomato/SV40::EGFP cell line (Supplementary information, Figure S1F). When both modules were applied, simultaneous activation of TetO::tdTomato and repression of SV40::EGFP were achieved (Supplementary information, Figure S1F). Next, we tested whether the expression of multiple endogenous genes can be independently regulated using this strategy. We directed the Casilio activator and repressor modules to *OCT4* and *SOX2* promoters, respectively, and observed simultaneous activation of *OCT4* and repression of *SOX2* when two sets of sgRNA-5×PBS individually targeting *OCT4* and *SOX2* were introduced into cells together with dCas9 and PUF fusions (Figure 1D).

A recent study used dCas9-tethered histone acetyltransferase (HAT) to activate enhancers [10]. We tested whether the Casilio system could recruit multiple molecules of HAT domain to increase the efficiency of epigenetic editing. We targeted the proximal promoter

(PP), proximal enhancer (PE) and distal enhancer (DE) of *OCT4* with cocktails of 4 sgRNA-PBS for each region (Figure 1E). We constructed fusions of the HAT domain from CREB-binding protein (CBP) attached to the C-terminus of dCas9 (dCas9-CBPHAT), and N-terminus (CBPHAT-PUFa) and C-terminus (PUFa-CBPHAT) of PUFa and tested their ability to activate *OCT4* expression. dCas9-CBPHAT and CBPHAT-PUFa modules have similar activity at PP. Interestingly, the Casilio modules have higher efficiency in activating *OCT4* gene via PE and DE compared to dCas9-CBPHAT, with CBPHAT-PUFa giving the highest activation (Figure 1E). In addition, CBPHAT-PUFa could mediate significant *OCT4* activation when directed by a single sgRNA-5×PBSa (Supplementary information, Figure S1G).

Another important application of dCas9 is to label genomic loci for live cell imaging [9]. We first tested whether the Casilio system can label telomeric repeats using sgTelomere (Supplementary information, Figure S2A) appended with 0, 5, 15 or 25 copies of PBSa to recruit Clover-PUFa (Figure 1F). Expression of sgTelomere-5×PBSa, -15×PBSa and -25×PBSa with dCas9 and Clover-PUFa produced fluorescent foci at telomeres (Figure 1Fi), as confirmed by their overlap with immunostaining signals of telomeric repeat-binding factor 2 (TRF2) (Figure 1Fii). Interestingly, the strength of telomere labeling increased as more copies of PBS were appended to sgTelomere-PBS (Figure 1Fi). Quantification of foci number and signal-to-noise ratio in cells expressing sgRNA with 5, 15 or 25 copies of PBSa confirmed such increase (Supplementary information, Figure S2C and S2D), suggesting that the multimerization feature of the Casilio system allows titration of labeling intensity at target loci. We further tested the ability of the Casilio system to simultaneously label centromeres and telomeres. By using sgRNA to target alpha-satellite repeats

Figure 1 The Casilio system enables multiplexing and multimerization of effectors for gene regulation, epigenetic editing and chromosomal labeling. **(A)** Schematics of the Casilio system. Top: comparison of conventional dCas9 fusion and Casilio tethering of effectors. Center: the 8-repeat PUF domain and the corresponding 8-mer target RNA. Table shows PUF diresidues and the corresponding RNA base. Bottom: annotation for the 4 PUF isotypes and PBS used in this study. **(B)** The specificity of the Casilio system. Column plot shows mean fold changes (mean ± SEM; $n = 3$) of tdTomato fluorescence of cells transfected with the indicated constructs. **(C)** Multimerization of activators by the Casilio system enables robust activation of endogenous *OCT4* gene. Upper: schematic showing positions of 4 sgRNA-5×PBS used to activate *OCT4*. Lower: mean fold changes (mean ± SEM; $n = 3$) measured by qRT-PCR of *OCT4* expression induced by dCas9/PUFa-p65HSF1 or dCas9-p65HSF1 activator with the indicated cocktail of *OCT4* or control sgRNAs-5×PBSa. **(D)** The Casilio system can activate and repress different genes simultaneously. Column plots show mean fold changes of *OCT4* (upper) and *SOX2* gene expression (lower) in cells transfected with the indicated constructs. **(E)** Recruitment of the HAT domain by the Casilio system to enhancers to activate *OCT4* expression. HEK293T cells were overexpressed with different HAT fusions and cocktails of 4 sgRNA-5×PBSa targeting PP, PE or DE as indicated. qRT-PCR was performed to evaluate the fold changes of *OCT4* expression. **(F)** Simultaneous labeling of telomeres and centromeres by the Casilio system (scale bars: 5 μm). (i) Labeling of telomeres by Clover-PUFa and sgTelomere with 0, 5, 15 or 25 copies of PBSa. (ii) TRF2 immunostaining to validate labeling of telomeres by dCas9/Clover-PUFa/sgTelomere-25×PBSa. (iii) CREST immunostaining to validate labeling of centromeres by dCas9/Clover-PUFc/sgCentromere-20×PBSc. (iv) Co-labeling of centromeres and telomeres by the Casilio system. **(G)** Cartoon highlighting three major features of the Casilio system.

in centromeres (sgCentromere-20×PBSc; Supplementary information, Figure S2B), we observed specific centromere labeling as confirmed by CREST immunostaining (Figure 1Fiii). Although not all centromeres marked by CREST were detected by the Casilio system, likely due to the low number of sgCentromere-binding sites in some centromeres (Supplementary information, Figure S2B), most of the Casilio signal overlapped with CREST signals. When we co-introduced Clover-PUFc/sgCentromere-20×PBSc, Ruby-PUFa/sgTelomere-25×PBSa and dCas9 into HEK293T cells, we observed simultaneous labeling of both centromeres and telomeres in the same cells (Figure 1Fiv), demonstrating that the Casilio system can be used to independently label multiple genomic loci.

In conclusion, here we show that the Casilio system can achieve multiplexing and multimerization of proteins such as transcriptional regulators, epigenetic modifiers and fluorescent proteins at defined genomic loci (Figure 1G). The main advantages of the Casilio system include: (I) Multiplexing. Different Casilio modules can be simultaneously delivered into a cell and each can operate at their defined target sites with independent function. Since PUF domains can be easily programmed to recognize any 8-mer RNA motifs, this greatly expands the potential number of independent Casilio modules. (II) Multimerization. Simplicity of the linear 8-mer PBS motif allows extensive multimerization of PUF fusions on sgRNA-PBS without hindering sgRNA transcription or dCas9/sgRNA DNA-binding activity. This feature allows multiple molecules of PUF fusions to be assembled on the sgRNA, achieving localized concentration of effectors or protein tags. This is particularly beneficial for fluorescent imaging and transcriptional regulation. This feature may facilitate labeling of non-repeat sequences where tiling of > 30 sgRNAs were required [9]. (III) Complex formation. With further development and optimization, the sgRNA-PBS has the potential to act as an RNA scaffold for PUF-directed assembly of stoichiometrically defined protein complexes. Specifically, different numbers of PBS with varying specificities can be appended to the sgRNA to enable the assembly of multi-protein complex with defined stoichiometry and defined ordering along the sgRNA-PBS (Figure 1G). Considering these features, we believe that the Casilio system will become a powerful tool for studying gene function and chromosome structure.

Acknowledgments

We thank Mary Ann Handel and Dave Mellert for their critical

comments on the manuscript. We thank the scientific services at the Jackson Laboratory for technical supports, specifically, Mark Lessard, Stephen Douthwright, Will Schott, Ted Duffy, and Rick Maser. Yasuhiro Fujiwara was supported by NIH (HD33816). HW was supported by the National Natural Science Foundation of China (31471215), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA01010409), and the Hi-Tech Research and Development Program of China (863 project; 2015AA020307).

Albert W Cheng^{1,2,3,*}, Nathaniel Jillette^{1,2,*},
Phoebe Lee¹, Dylan Plaskon¹, Yasuhiro Fujiwara¹,
Wenbo Wang¹, Aziz Taghbalout², Haoyi Wang^{1,4}

¹The Jackson Laboratory, Bar Harbor, ME 04609, USA; ²The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA; ³Department of Genetics and Genome Sciences, University of Connecticut School of Medicine, Farmington, CT 06030, USA; ⁴State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

*These two authors contributed equally to the work.

Correspondence: Haoyi Wang^a, Albert W Cheng^b

^aE-mail: wanghaoyi@ioz.ac.cn

^bE-mail: albert.cheng@jax.org

References

- 1 Cong L, Ran FA, Cox D, *et al.* *Science* 2013; **339**:819-823.
- 2 Jinek M, Chylinski K, Fonfara I, *et al.* *Science* 2012; **337**:816-821.
- 3 Mali P, Yang L, Esvelt KM, *et al.* *Science* 2013; **339**:823-826.
- 4 Gilbert LA, Larson MH, Morsut L, *et al.* *Cell* 2013; **154**:442-451.
- 5 Maeder ML, Linder SJ, Cascio VM, *et al.* *Nat Methods* 2013; **10**:977-979.
- 6 Perez-Pinera P, Kocak DD, Vockley CM, *et al.* *Nat Methods* 2013; **10**:973-976.
- 7 Qi LS, Larson MH, Gilbert LA, *et al.* *Cell* 2013; **152**:1173-1183.
- 8 Cheng AW, Wang H, Yang H, *et al.* *Cell Res* 2013; **23**:1163-1171.
- 9 Chen B, Gilbert LA, Cimini BA, *et al.* *Cell* 2013; **155**:1479-1491.
- 10 Hilton IB, D'Ippolito AM, Vockley CM, *et al.* *Nat Biotechnol* 2015.
- 11 Zalatan JG, Lee ME, Almeida R, *et al.* *Cell* 2015; **160**:339-350.
- 12 Shechner DM, Hacisuleyman E, Younger ST, *et al.* *Nat Methods* 2015; **12**:664-670.
- 13 Chen Y, Varani G. *FEBS J* 2013; **280**:3734-3754.
- 14 Wang Y, Cheong CG, Hall TMT, *et al.* *Nat Methods* 2009; **6**:825-830.
- 15 Konermann S, Brigham MD, Trevino AE, *et al.* *Nature* 2014; **517**:583-588.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>