Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells

Dacheng Ma¹*, Shuguang Peng¹* & Zhen Xie¹

Programmable and precise regulation of dCas9 functions in response to multiple molecular signals by using synthetic gene circuits will expand the application of the CRISPR-Cas technology. However, the application of CRISPR-Cas therapeutic circuits is still challenging due to the restrictive cargo size of existing viral delivery vehicles. Here, we construct logic AND circuits by integrating multiple split dCas9 domains, which is useful to reduce the size of synthetic circuits. In addition, we engineer sensory switches by exchanging split dCas9 domains, allowing differential regulations on one gene, or activating two different genes in response to cell-type specific microRNAs. Therefore, we provide a valuable split-dCas9 toolkit to engineer complex transcription controls, which may inspire new biomedical applications.

¹MOE Key Laboratory of Bioinformatics and Bioinformatics Division, Center for Synthetic and System Biology, Department of Automation, Tsinghua National Lab for Information Science and Technology, Tsinghua University, Beijing 100084, China. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Z.X. (email: zhenxie@tsinghua.edu.cn).
he CRISPR-associated protein 9 (Cas9) discovered from *Streptococcus pyogenes* is a multi-domain protein, which has been widely used in genome editing and transcriptional control in mammalian cells due to its superior modularity and versatility. The Cas9-DNA targeting specificity is determined by both the Cas9-associated guide RNA (gRNA) and a short protospacer adjacent motif (PAM) directly downstream of the DNA recognition site. The *S. pyogenes* Cas9 (SpCas9) protein usually consists of a recognition lobe and a nuclease lobe. The recognition lobe contains a bridge helix (residues 60–93), a REC1 (residues 94–179 and 308–713) domain and a REC2 (residues 180–207) domain, while the nuclease lobe includes a RuvC (residues 1–59, 718–769 and 909–1,098) domain, a HNH (residues 775–908) domain and a PAM-interacting (PI) fragment (residues 1,099–1,368) domain. The functional Cas9 protein can cleave the EYFP repeat region, triggering the reconstitution of inactive EYFP into the full-length active EYFP reporter gene. By using this EYFP-reconstitution reporter system, we found that all four intein-mediated split-Cas9 pairs efficiently reactivated the EYFP expression in a human embryonic kidney HEK293 cells (Supplementary Fig. 1c). The Cas9 sets split at residues 203, 468, 713 and 1,153 without intein fusion, displayed a reduced activity compared with their counterparts with appropriate intein fusions (Supplementary Fig. 1c). Next, we tested whether a similar set of split dCas9:VPR pairs can recapitulate the function of the full-length dCas9:VPR in HEK293 cells by a transient transfection experiment. Three of split pairs without intein fusions showed a reduced activation function compared with their counterparts with intein fusions (Fig. 1a). In contrast, the dCas9 protein directly split at position 1,135 was almost as active as the intact dCas9 protein (Fig. 1a). Three of the split dCas9:VPR pairs fused to intein fragments activated the reporter gene as efficiently as the full-length dCas9:VPR, while the dCas9:VPR split at residue 713 was not as efficient (Fig. 1a). These results suggest that the VPR fusion and the choice of split site might affect reconstitution of split dCas9 fragments and influence the protein activity. We further demonstrated the functional reassembly of dCas9 constituents split at either residue 713 or residue 1,153 when fused to different transcription regulatory domains, such as Krab, SunTag and VP64 (Supplementary Fig. 2).

Next, we explored whether split dCas9 fragments across different split pairs can be reconstituted into a functional dCas9. The experiment demonstrated that combinations of dCas9 IntN and IntC fragments that resulted in incomplete dCas9 proteins failed to activate TagBFp expression (Fig. 1b). In contrast, fragment combinations that provided the complete polypeptide sequence activated TagBFp expression even when the two fragments were partly redundant. Interestingly, the dCas9 pair split at residue 1,153 divided the PI domain into two fragments. We then tested whether it is possible to increase the orthogonality of the split set at residue 1,153 by introducing EQR mutations in the PI domain. As shown in Fig. 1c, the reconstitution of the split dCas9 pair at residue 1,153 with the EQR mutations only activated the mKate2 reporter gene with the NGCG PAM but not the EYFP reporter gene with the NGG PAM, while the reconstitution of the split dCas9 pair at residue 713 without mutations led to the opposite results. In addition, no cross activity was observed when either the wild-type N-terminal or C-terminal dCas9 fragment was combined with the EQR mutant C-terminal or N-terminal dCas9 constituents (Fig. 1c). These orthogonal split dCas9 pairs would potentially facilitate the construction of complex genetic circuits and logic gates.

**Results**

**Functional reconstitution of split Cas9 domains.** According to the Cas9 sequence and structural information, as well as previous reports, we selected four potential split sites where serine is at the +1 amino acid position when fused to the C-terminal Intein fragment. All four potential split sites were surface residues and located in the loop region, which may be more accessible for intein trans-splicing reaction and have less effect on the protein folding. Then, we constructed eight pairs of split Cas9 constituents that either fused to the N-terminal (IntN) and C-terminal (IntC) split inteins or not (Fig. 1a and Supplementary Fig. 1a). As shown in Supplementary Fig. 1b, we inserted a repeat sequence in the middle of enhanced yellow fluorescent protein (EYFP) reporter gene as described. The functional Cas9 protein can cleave the EYFP repeat region, triggering the reconstitution of inactive EYFP into the full-length active EYFP reporter gene. By using this EYFP-reconstitution reporter system, we found that all four intein-mediated split-Cas9 pairs efficiently reactivated the EYFP expression in a human embryonic kidney HEK293 cells (Supplementary Fig. 1c). The Cas9 sets split at residues 203, 468, 713 and 1,153 without intein fusion, displayed a reduced activity compared with their counterparts with appropriate intein fusions (Supplementary Fig. 1c). A similar set of split dCas9:VPR pairs can recapitulate the function of the full-length dCas9:VPR in HEK293 cells by a transient transfection experiment. Three of split pairs without intein fusions showed a reduced activation function compared with their counterparts with intein fusions (Fig. 1a). In contrast, the dCas9 protein directly split at position 1,135 was almost as active as the intact dCas9 protein (Fig. 1a). Three of the split dCas9:VPR pairs fused to intein fragments activated the reporter gene as efficiently as the full-length dCas9:VPR, while the dCas9:VPR split at residue 713 was not as efficient (Fig. 1a). These results suggest that the VPR fusion and the choice of split site might affect reconstitution of split dCas9 fragments and influence the protein activity. We further demonstrated the functional reassembly of dCas9 constituents split at either residue 713 or residue 1,153 when fused to different transcription regulatory domains, such as Krab, SunTag and VP64 (Supplementary Fig. 2).

Next, we explored whether split dCas9 fragments across different split pairs can be reconstituted into a functional dCas9. The experiment demonstrated that combinations of dCas9 IntN and IntC fragments that resulted in incomplete dCas9 proteins failed to activate TagBFp expression (Fig. 1b). In contrast, fragment combinations that provided the complete polypeptide sequence activated TagBFp expression even when the two fragments were partly redundant. Interestingly, the dCas9 pair split at residue 1,153 divided the PI domain into two fragments. We then tested whether it is possible to increase the orthogonality of the split set at residue 1,153 by introducing EQR mutations in the PI domain. As shown in Fig. 1c, the reconstitution of the split dCas9 pair at residue 1,153 with the EQR mutations only activated the mKate2 reporter gene with the NGCG PAM but not the EYFP reporter gene with the NGG PAM, while the reconstitution of the split dCas9 pair at residue 713 without mutations led to the opposite results. In addition, no cross activity was observed when either the wild-type N-terminal or C-terminal dCas9 fragment was combined with the EQR mutant C-terminal or N-terminal dCas9 constituents (Fig. 1c). These orthogonal split dCas9 pairs would potentially facilitate the construction of complex genetic circuits and logic gates.

**Construction of three-input logic AND circuit.** On the basis of the above results, we started to engineer logic AND circuit by using the dCas9 constituents split at residue 1,153 and the SunTag repetitive peptide scaffold that contains ten ScFv binding motifs (Fig. 2a). The ScFv along with a small solubility tag GB1 and VP64 fragments were respectively fused to FK506 binding protein (FKBP) and FKBP rapamycin binding (FRB) domains with a T2089I mutation derived from the mammalian target of rapamycin (mTOR)21,25. Therefore, the resulted fusion proteins,
IntN and CAG EYFP binding sites with NGG PAM sequences. (dCas9 domains for transcription activation (lower left). The mKate2 was used as an internal control. The TRE promoter contains seven repeats of gRNAa (mutations in the PAM recognition domain. The ModTRE2 promoter contains seven repeats of gRNAb binding sites with NGCG PAM sequences.

Two-input and one-output sensory switch. We recently developed a TALER sensory switch controlled by two different shRNAs and microRNAs (ref. 25). To test the domain exchange of dCas9 constituents, we respectively fused IntC:dCas9C:VPR and IntC:dCas9C:Krab to TALER14 and TALER9, which reconstituted with a constitutive dCas9N:ntn to activate or repress the expression of the EYFP reporter gene by competitively binding to the TRE promoter. The VPR activation domain was chosen because the activation efficiency is greater than both VP64 and Suntag activation domains in our experimental setup (Supplementary Fig. 2b). As shown in Supplementary Fig. 3b, and Suntag activation domains in our experimental setup chosen because the activation efficiency is greater than both VP64 and Suntag activation domains in our experimental setup.

Figure 1 | Characterization of split dCas9 domains. (a) Diagram of dCas9 domain organization and split sites (top left), and reconstitution of split dCas9 domains for transcription activation (lower left). The mKate2 was used as an internal control. The TRE promoter contains seven repeats of gRNAa binding sites with N GG PAM sequences. (b) Orthogonality of split dCas9 domains. (c) Improved orthogonality of split dCas9 domains by introducing mutations in the PAM recognition domain. The ModTRE2 promoter contains seven repeats of gRNA binding sites with NGCG PAM sequences. (a,c) Each bar shows mean fold change (mean ± s.e.m.; n = 3) of indicated fluorescence measured by using flow cytometry 48 h after transfection.
The dCas9 constituents are split at residue 1153. (using two split dCas9 fragments (dCas9N:IntN and IntC:dCas9C:Suntag) and the rapalog that promote the dimerization of ScFv:GB1:FKBP and FRB*:VP64. The dCas9 constituents are split at residue 713 and 1153. (a,b) Data are shown as the mean fold change (mean ± s.e.m.; n = 3) of TagBFP fluorescence measured by using flow cytometer 48 h after transfection into HEK293 cells.

**Two-input and two-output sensory switch.** To test whether the sensory switch can be used to activate two different output genes in response to two different shRNAs, we replaced the IntC:dCas9C:Krab with the orthogonal activator IntC:dCas9C-VRER:VPR, increasing the ON/OFF ratio to 10-fold in response to miR21/miR18a input combination (Fig. 3d). In addition, the sensory switch responded to the miR21 input in a dosage dependent manner (Supplementary Fig. 4).

**Discussion**

In this study, we demonstrated that split dCas9 domains can be reconstituted for transcriptional regulations in cultured human cells (Fig. 1), allowing modular and efficient construction of three-input logic AND circuits (Fig. 2). By using the split dCas9 and SunTag system, it is possible to easily increase the number of inputs up to seven, including three split dCas9 domains, two SunTag fragments, the rapalog and the gRNA. In addition, we developed an orthogonal split dCas9 pair by introducing mutations in the PI domains, which recognized the NGCG PAM sequences instead of the NGG PAM sequences (Fig. 1c). These orthogonal split dCas9 pairs would be a useful toolkit to construct complex and layered logic gates with multiple inputs. In addition, a similar strategy can be applied to engineering split Cas9 pairs with nuclease or nickase activity.

Recently, several strategies have been developed to control the Cas9/dCas9 activity by using small molecules and light signals. Connecting tissue and cellular specific inputs to regulate Cas9/dCas9 activity can facilitate the application of the CRISPR-Cas system in basic and translational biomedical researches. In this study, we developed sensory switches by exchanging split dCas9 domains, allowing differential regulations...
on one gene, or activating two different genes in response to cell-type specific microRNAs (Figs 3 and 4). We anticipate that combining our sensory switch with other tissue and cellular inputs will inspire new approaches for more complex regulations on the Cas9/dCas9 function.

It has been shown that split Cas9 system can be delivered in vivo by using recombinant adenovirus-associate viruses (rAAV)\textsuperscript{11,12}. Our circuit design principles provided an useful method to reduce the size of synthetic circuits by integrating and swapping split Cas9/dCas9 domains fused with different functional domains. It will be interesting to adapt our split Cas9/dCas9 system to rAAV delivery system, and test the Cas9/dCas9 activity to edit and regulate endogenous genes in vivo. Such a CRISPR-Cas9 system will be particularly appealing in biomedical applications in which viral delivery vehicles with a restrictive cargo size are preferred.

Methods

Reagents and enzymes. Restriction endonuclease, polynucleotide kinase (PNK), T4 DNA ligase, Quick DNA ligase and Q5 High-Fidelity DNA Polymerase were
Figure 4 | Two-input and two-output sensory switch. (a) Schematic representation of a two-input and two-output sensory switch by swapping split dCas9 domains (VRER) that recognize two different PAM sequences. The dCas9 constituents are split at residue 713. The light blue rectangle represents the mutant dCas9 domain (VRER) that can recognize the NGCG PAM sequences but not the NGG PAM sequences. The ModTRE2 promoter contains seven repeats of gRNA binding sites with the NGCG PAM sequences upstream of a miniCMV core. (b) Setting states of the sensory switch by artificial shRNA-FF5 and shRNA-FF4. Representative scatter plots and microscopic images are shown in the right panel. Each red scale bar in images represents 50 μm. (c) The response of the sensory switch to varying amount of shRNA-FF4. The solid line was plotted by qplot function in R package. (b,c) Each bar or data point shows mean fold changes (mean ± s.e.m.; n = 3) of EYFP or mKate2 measured by using flow cytometer 48 h after transfection.

Flow cytometry. Cells were trypsinized 48 h after transfection and centrifuged at 300g for 7 min at 4°C. The supernatant was removed, and the cells were resuspended in 1 × PBS that did not contain calcium or magnesium. Fortessa flow analyser (BD Biosciences) was used for fluorescence-activated cell sorting analysis with the following settings. TagBFP was measured using a 405 nm laser and a 450/50 filter with a photomultiplier tube (PMT) set at 275 V. EYFP was measured with a 488 nm laser and a 530/30 filter using a PMT set at 240 V. mKate2 was measured with a 561 nm laser and a 670/30 filter using a PMT set at 450 V. For each sample, ~1 × 10^4 to ~5 × 10^4 cell events were collected.

Fluorescence microscopy. The cells were grown on class bottom cell culture dish (p 20 nm) in complete media for transfection. Approximately 48 h after transfection, fluorescent images of cultured cells were captured by Zeiss LSM780. The filter sets (Chroma) are as follows. TagBFP fluorescence was measured with an excitation at 405 nm and an emission in the range of 410–507 nm. EYFP fluorescence was measured with an excitation at 543 nm and an emission in the range of 519–568 nm. mKate2 fluorescence was measured with an excitation at 360 nm and an emission in the range of 571–651 nm. Image acquisition and post-acquisition analysis were performed using ZEN 2011.

Data availability. The authors declare that all data supporting the findings of this study are available within the article and its Supplementary Information files or are available from the corresponding author upon request.

Plasmid DNA constructs. The methods used for plasmid DNA construction are summarized in Supplementary Information. The oligonucleotide primer sequences (purchased from Genewiz Inc.) are listed in Supplementary Table 2. The protein and DNA sequences of plasmid DNA constructs are listed in Supplementary Data 1. Diagrams of the TRE and derivative promoters (ModTRE1 and ModTRE2) are shown in Supplementary Fig. 7.

Cell culture and transfection. HEK293 (293-H) cell line was purchased from Life Technologies. Hela S3 cell line was purchased from ATCC. HEK293 or Hela S3 cells were cultured in high-glucose DMEM complete media (Dulbecco’s modified Eagle’s medium (DMEM), 4.5 g l^-1 glucose, 0.045 units ml^-1 of penicillin, 0.045 g ml^-1 streptomycin, and 10% FBS (Life Technologies)) at 37°C, 100% humidity, and 5% CO2.

One day before transfection, ~1.2 × 10^5 HEK293 cells in 0.5 ml of high-glucose DMEM complete media were seeded into each well of 24-well plastic plates (Falcon). Shortly before transfection, the medium was replaced with fresh DMEM complete media. The transfection experiments were performed by using either Lipofectamine LTX (Life Technologies) or Attractene transfection reagent (Qagen) by following the manufacturer’s protocol. The amount of plasmid DNAs used in transfection experiments is listed in Supplementary Table 3. Cells were cultured for 2 days before flow cytometry analysis.

Flow cytometry. Cells were trypsinized 48 h after transfection and centrifuged at 300g for 7 min at 4°C. The supernatant was removed, and the cells were resuspended in 1 × PBS that did not contain calcium or magnesium. Fortessa flow analyser (BD Biosciences) was used for fluorescence-activated cell sorting analysis with the following settings. TagBFP was measured using a 405 nm laser and a 450/50 filter with a photomultiplier tube (PMT) set at 275 V. EYFP was measured with a 488 nm laser and a 530/30 filter using a PMT set at 240 V. mKate2 was measured with a 561 nm laser and a 670/30 filter using a PMT set at 450 V. For each sample, ~1 × 10^4 to ~5 × 10^4 cell events were collected.

Fluorescence microscopy. The cells were grown on class bottom cell culture dish (p 20 nm) in complete media for transfection. Approximately 48 h after transfection, fluorescent images of cultured cells were captured by Zeiss LSM780. The filter sets (Chroma) are as follows. TagBFP fluorescence was measured with an excitation at 405 nm and an emission in the range of 410–507 nm. EYFP fluorescence was measured with an excitation at 543 nm and an emission in the range of 519–568 nm. mKate2 fluorescence was measured with an excitation at 360 nm and an emission in the range of 571–651 nm. Image acquisition and post-acquisition analysis were performed using ZEN 2011.

Data availability. The authors declare that all data supporting the findings of this study are available within the article and its Supplementary Information files or are available from the corresponding author upon request.
References

1. Gilbert, L. A. et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell 154, 442–451 (2012).
2. Cong, L. et al. Multiplex genome engineering using CRISPR-Cas systems. Science 339, 819–823 (2013).
3. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183 (2013).
4. Farzadfard, F., Perli, S. D. & Lu, T. K. Tunable and multifunctional eukaryotic transcription factors based on CRISPR-Cas. ACS Synth. Biol. 2, 604–613 (2013).
5. Jinek, M. et al. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–822 (2012).
6. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA 109, E2579–E2586 (2012).
7. Nishimasu, H. et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156, 935–949 (2014).
8. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature 513, 569–573 (2014).
9. Kleinsteve, B. P. et al. Engineered CRISPR-Cas9 nuclease with altered PAM specificities. Nature 523, 481–485 (2015).
10. Wright, A. V. et al. Rational design of a split-Cas9 enzyme complex. Proc. Natl. Acad. Sci. USA 112, 2984–2989 (2015).
11. Truong, D.-J. J. et al. Development of an intein-mediated split-Cas9 system for gene therapy. Nucleic Acids Res. 43, 6450–6458 (2015).
12. Fine, E. J. et al. Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes. Sci. Rep. 5, 10777 (2015).
13. Zetsche, B., Volz, S. E. & Zhang, F. A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat. Biotechnol. 33, 9–11 (2015).
14. Nihongaki, Y., Kawano, F., Nakajima, T. & Sato, M. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat. Biotechnol. 9, 1–8 (2015).
15. Nguyen, D. P. et al. Ligand-binding domains of nuclear receptors facilitate tight control of split CRISPR activity. Nat. Commun. 7, 12009 (2016).
16. Nishimasa, H. et al. Crystal structure of Staphylococcus aureus Cas9. Cell 162, 1113–1126 (2015).
17. Kiani, S. et al. CRISPR transcriptional repression devices and layered circuits in mammalian cells. Nat. Methods 11, 723–726 (2014).
18. Nissim, L., Perli, S. D., Fridkin, A., Perez-pineria, P. & Lu, T. K. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas9 toolkit in human cells. Mol. Cell 54, 698–710 (2014).
19. Kiani, S. et al. Cas9 gRNA engineering for genome editing, activation and repression. Nat. Methods 12, 1051–1054 (2015).
20. Liu, Y. et al. Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells. Nat. Commun. 5, 1–7 (2014).
21. 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 159, 635–646 (2014).
22. Zalatan, J. G. et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 160, 339–350 (2015).
23. Mashiko, D. et al. Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. Sci. Rep. 3, 3355 (2013).
24. Rivera, V. M. et al. A humanized system for pharmacologic control of gene expression. Nat. Med. 2, 1028–1032 (1996).
25. Li, Y. et al. Modular construction of mammalian gene circuits using TALE transcriptional repressors. Nat. Chem. Biol. 11, 1–28 (2015).
26. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-input RNA-based logic circuit for identification of specific cancer cells. Science 333, 1307–1311 (2011).
27. Davis, K. M., Pattanayak, V., Thompson, D. B., Zuris, J. A. & Liu, D. R. Small molecule–triggered Cas9 protein with improved genome-editing specificity. Nat. Chem. Biol. 11, 316–318 (2015).
28. Oakes, B. L. et al. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. Nat. Biotechnol. 34, 646–651 (2016).
29. Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat. Chem. Biol. 11, 198–200 (2015).
30. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational regulation of Cas9 during G1 enhances homology-directed repair. Cell Rep. 14, 1555–1566 (2016).

Acknowledgements

We thank members of Xie Lab for helpful discussions. We thank Lei Stanley Qi for providing the Cas9-Suntag and ScFv-GBI constructs as gifts. The research is supported by the National Key Basic Research Program of China (2014CB745200) and Basic Research Program of Tsinghua National Lab for Information Science and Technology (Z.X.).

Author contributions

Z.X., D.M. and S.P. conceived of the ideas implemented in this work. D.M. and S.P. performed experiments. Z.X., D.M. and S.P. analysed the data. Z.X. supervised the project. Z.X., D.M. and S.P. wrote the paper.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: Z.X., D.M. and S.P. have filed a patent application to State Intellectual Property Office of China based on the findings in this work (patent number: 201610341363.0). The remaining authors declare no competing financial interests.

Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Ma, D. et al. Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. Nat. Commun. 7, 13056 doi: 10.1038/ncomms13056 (2016).