



# CRISPR-Cas12a System With Synergistic Phage Recombination Proteins for Multiplex Precision Editing in Human Cells

Chengkun Wang<sup>1,2†</sup>, Qiong Xia<sup>1,2†</sup>, Qianhe Zhang<sup>1,2†</sup>, Yuanhao Qu<sup>1,2†</sup>, Stephen Su<sup>1,2</sup>, Jason K. W. Cheng<sup>1,2</sup>, Nicholas W. Hughes<sup>1,2</sup> and Le Cong<sup>1,2,3\*</sup>

<sup>1</sup>Department of Pathology, Stanford University School of Medicine, Stanford, CA, United States, <sup>2</sup>Department of Genetics, Stanford University School of Medicine, Stanford, CA, United States, <sup>3</sup>Wu Tsai Neuroscience Institute, Stanford University, Stanford, CA, United States

## OPEN ACCESS

### Edited by:

Dali Li,  
East China Normal University, China

### Reviewed by:

Ivanka Markovic,  
University of Belgrade, Serbia  
Huiming Lu,  
University of Texas Southwestern  
Medical Center, United States

### \*Correspondence:

Le Cong  
congle@stanford.edu

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Cellular Biochemistry,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

Received: 02 June 2021

Accepted: 05 November 2021

Published: 14 June 2022

### Citation:

Wang C, Xia Q, Zhang Q, Qu Y, Su S,  
Cheng JKW, Hughes NW and Cong L  
(2022) CRISPR-Cas12a System With  
Synergistic Phage Recombination  
Proteins for Multiplex Precision Editing  
in Human Cells.  
Front. Cell Dev. Biol. 9:719705.  
doi: 10.3389/fcell.2021.719705

The development of CRISPR-based gene-editing technologies has brought an unprecedented revolution in the field of genome engineering. Cas12a, a member of the Class 2 Type V CRISPR-associated endonuclease family distinct from Cas9, has been repurposed and developed into versatile gene-editing tools with distinct PAM recognition sites and multiplexed gene targeting capability. However, with current CRISPR/Cas12a technologies, it remains a challenge to perform efficient and precise genome editing of long sequences in mammalian cells. To address this limitation, we utilized phage recombination enzymes and developed an efficient CRISPR/Cas12a tool for multiplexed precision editing in mammalian cells. Through protein engineering, we were able to recruit phage recombination proteins to Cas12a to enhance its homology-directed repair efficiencies. Our phage-recombination-assisted Cas12a system achieved up to 3-fold improvements for kilobase-scale knock-ins in human cells without compromising the specificity of the enzyme. The performance of this system compares favorably against Cas9 references, the commonly used enzyme for gene-editing tasks, with improved specificity. Additionally, we demonstrated multi-target editing with similar improved activities thanks to the RNA-processing activity of the Cas12a system. This compact, multi-target editing tool has the potential to assist in understanding multi-gene interactions. In particular, it paves the way for a gene therapy method for human diseases that complements existing tools and is suitable for polygenic disorders and diseases requiring long-sequence corrections.

**Keywords:** CRISPR, genome editing, Cas12a, RecT, single-stranded annealing protein

## INTRODUCTION

The powerful CRISPR technologies have been widely applied in various biological systems and are greatly changing the field of genome engineering (Cong et al., 2013; Jinek et al., 2013; Hsu et al., 2014; Komor et al., 2017; Knott and Doudna, 2018). Cas12a(Cpf1), a member of the Class 2 Type V CRISPR-associated endonuclease family, differs from Cas9 in several ways (Zetsche et al., 2015): it has a T-rich PAM recognition site (TTTV) while Cas9 PAM site is G-rich (NGG) (Zetsche et al., 2015; Yamano et al., 2016); its guide RNA is shorter in sequence than Cas9 (Zetsche et al., 2015); Cas12a cuts DNA distal and downstream relative to the PAM sequence and results in staggered ends

while Cas9 cuts DNA proximal and upstream and results in blunt ends; and Cas12a is reported to be more precise and induces less off-target activities than Cas9 (Zetsche et al., 2015; Kim et al., 2016; Kleinstiver et al., 2016; Yamano et al., 2016; Swarts et al., 2017). Moreover, due to its intrinsic ribonuclease activity that processes its own crRNA, Cas12a is capable of multiplex editing from a single RNA transcript (Zetsche et al., 2015; Zetsche et al., 2017).

Multiplex engineering can be useful in understanding gene-gene interactions, combinatorial effects of different genetic mutations, and biological pathways with redundancy between individual genes (Tak et al., 2017; Li et al., 2019; McCarty et al., 2020). Along with its other beneficial properties, Cas12a holds great potential in serving as a tool to solve genome engineering problems. However, similar to Cas9, Cas12a relies on endogenous homology-directed repair (HDR) mechanisms to achieve precise gene modifications, but the HDR efficiency is often very limited due to the competing error-prone repair pathways which happen much more frequently (Pickar-Oliver and Gersbach, 2019; Anzalone et al., 2020). While many previous studies have been done to improve the efficiency of base editing and short genomic alterations, large-scale genome editing using Cas12a still remains a challenge (Li et al., 2018; Kleinstiver et al., 2019; Wang et al., 2020).

Previously we have identified that phage single-stranded annealing proteins (SSAPs) could enhance Cas9-based gene editing efficiencies in mammalian cells, and among all SSAPs screened, *E. coli* RecT exhibited the largest increase (Wang et al., 2021). We termed this high-efficiency Cas9 editing method as RecT Editor *via* Designer-Cas9-Initiated Targeting (REDIT). Here, combining the efficiency of REDIT and the multi-target editing ability of Cas12a, we developed a new Cas12a-REDIT method.

## MATERIALS AND METHODS

### Plasmids Construction

The fragment encoding the scFV-RecT or GFP was Gibson assembled into MS2-P65-HSF1-GFP (addgene, 61423) using NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs), the MS2-P65-HSF1-GFP plasmid digested with BsiWI and EcoRI. The fragment encoding Cas12a-10x GCN4 was Gibson assembled into the pX330 (42230; addgene), pX330 was digested with AgeI and EcoRI. The guide RNAs were then inserted using Golden Gate cloning (guideRNA sequence used in this study could be found in the **Supplementary Table**). All plasmids were sequence-verified (Eton and Genewiz) through Sanger sequencing and will be deposited to Addgene for open access.

### Cell Culture

Human Embryonic Kidney (HEK) 293T cells was ordered from ThermoFisher; HeLa and HepG2 were obtained from American Type Culture Collection (ATCC). All three cell lines were maintained in Dulbecco's Modified Eagle's Medium (DMEM, Life Technologies), with 10% fetal bovine serum (FBS,

BenchMark), 100 U/ml penicillin, and 100 µg/ml streptomycin (Life Technologies) at 37°C with 5% CO<sub>2</sub>.

### Transfection

HEK293T cells were seeded into 96-well plates (Corning) 12 h prior to transfection at a density of 3E4 cells/well, and 250 ng of total DNA was transfected per well using Lipofectamine 3000 (Life Technologies) following the manufacturer's instruction. HeLa and HepG2 cells were seeded into 48-well plates (Corning) 1 day prior to transfection at a density of 5E4 and 3E4 cells/well respectively, and 400 ng of total DNA was transfected per well.

### Fluorescence-Activated Cell Sorting (FACS)

72 h after transfection, cells were washed with PBS and dissociated using 50 µl TrypLE Express Enzyme (Thermo Fisher Scientific). Cell suspension was then transferred to a 96-well U-bottom plate (Thermo Fisher Scientific) and centrifuged at 300 g for 5 min. After removing the supernatant, pelleted cells were resuspended with 50 µl 4% FBS in PBS, and the mKate or GFP knock-in efficiency was analyzed on a CytoFLEX flow cytometer (Beckman Coulter; Stanford Stem Cell FACS Core) within 30 min.

### Next-Generation Sequencing (NGS) Library Preparation

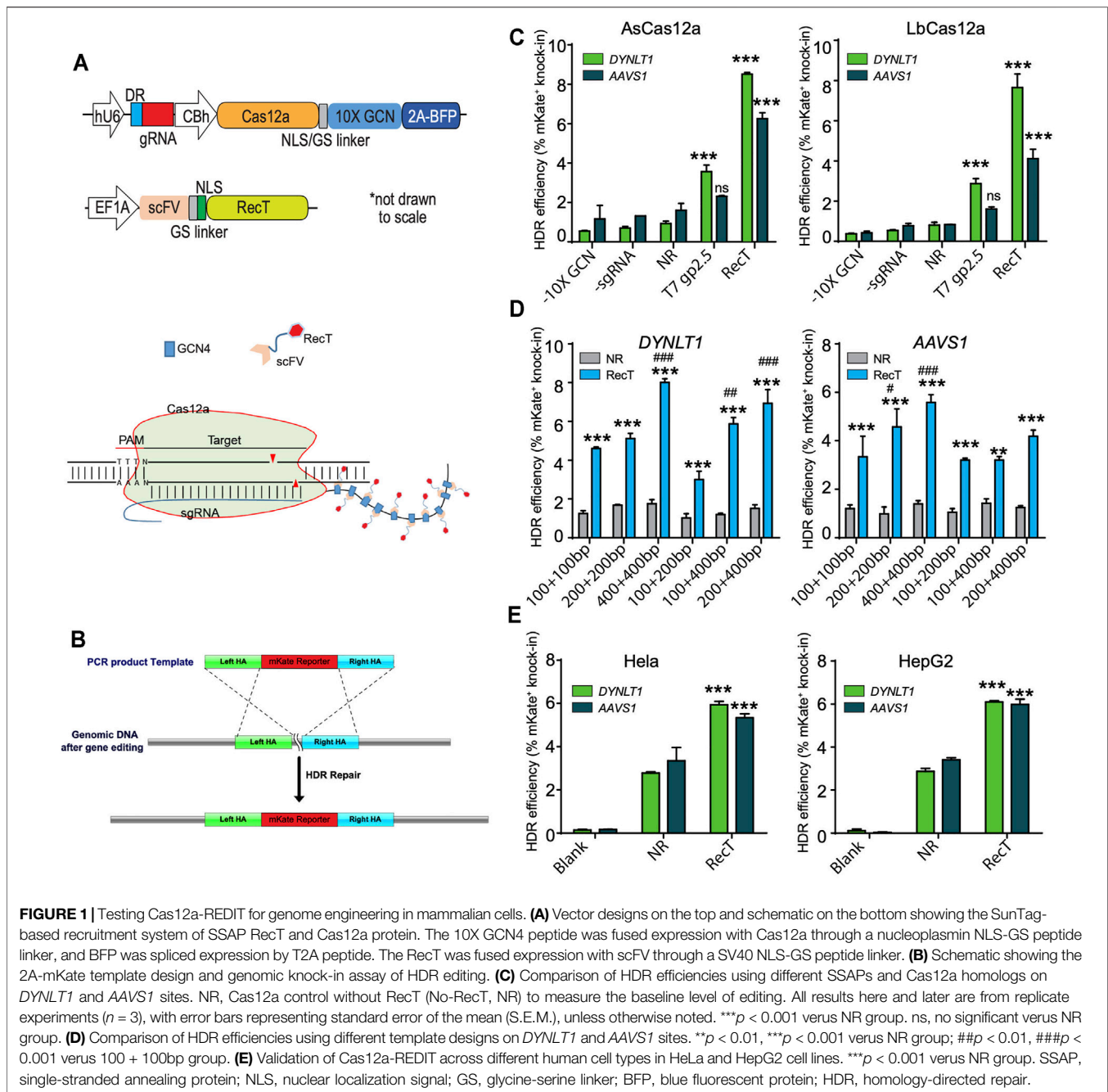
48–72 h after transfection, genomic DNA (gDNA) was extracted using QuickExtract DNA Extraction Solution (Lucigen). 200 ng total gDNA was used for NGS library preparation. Genes of interest were amplified using specific primers (**Supplementary Table**) for the first round of PCR reaction, and Illumina adapters and index barcodes were added in the second round of PCR. Final PCR products were run on a 2% E-gel (Thermo Fisher Scientific) and extracted using the Monarch DNA Gel Extraction Kit (New England BioLab). The purified product was quantified with Qubit dsDNA HS Assay Kit (Thermo Fisher) and loaded for sequencing using Illumina MiSeq paired-end PE300 kits following the manufacturer's instructions.

### High-Throughput Sequencing Data Analysis

The sequencing data was analyzed using CRISPResso2 as previously described (Clement et al., 2019; Wang et al., 2021). Briefly, the sequencing reads were demultiplexed, trimmed and merged, then aligned to reference or expected HDR amplicons. Only reads perfectly matched to the expected amplicon were considered for HDR quantification. The computation work was supported by the SCG cluster hosted by the Genetics Bioinformatics Service Center (GBSC) at the Department of Genetics of Stanford. All sequencing data have been deposited by the authors into NCBI SRA archive, and could be identified by the accession ID PRJNA735023.

### Statistical Analysis

Unless otherwise stated, all statistical analysis and comparison were performed using *t*-test, with 1% false-discovery-rate (FDR) using a two-stage step-up method of Benjamini, Krieger and



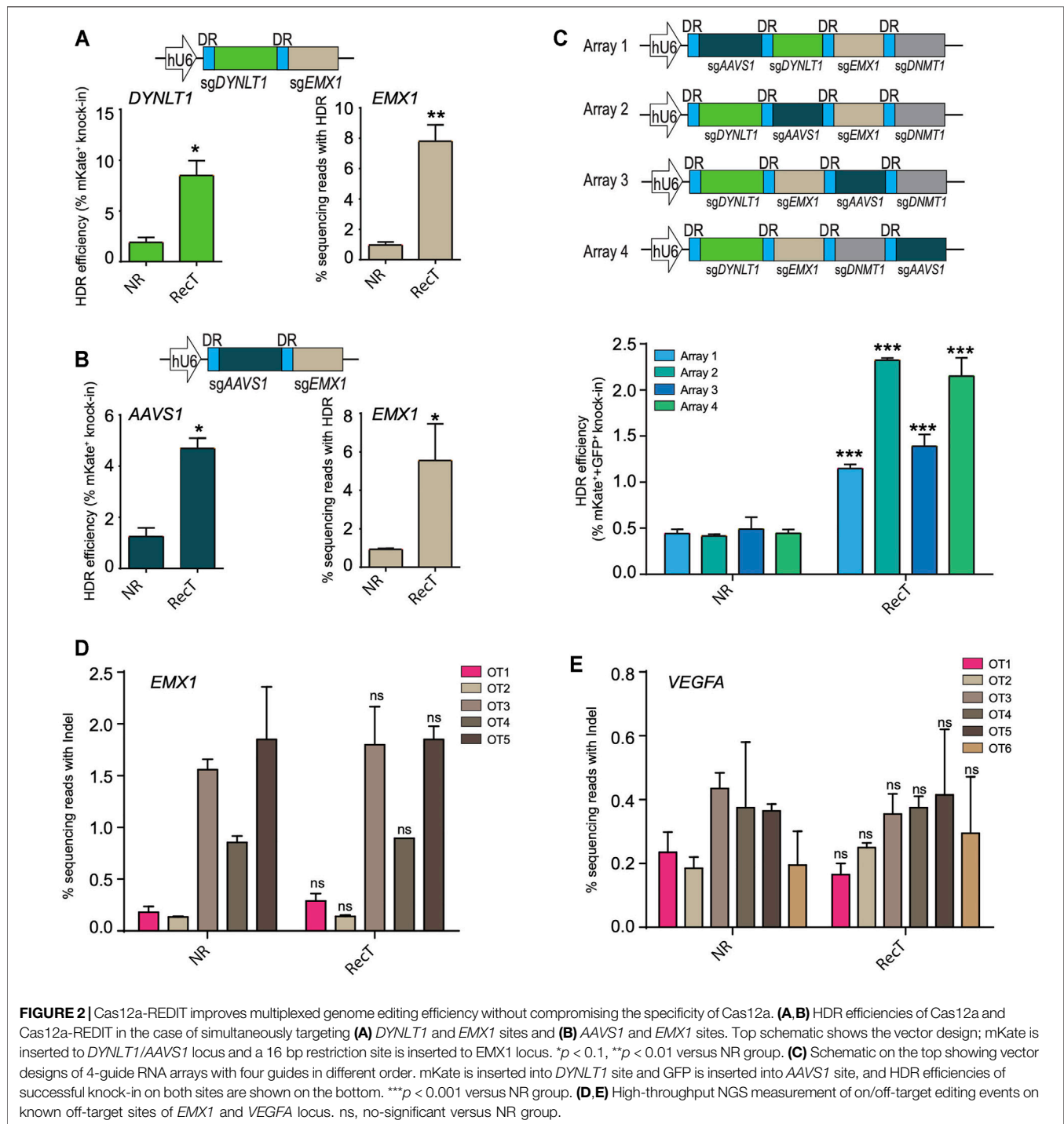
Yekutieli. All experiments were performed in triplicates unless otherwise noted to ensure sufficient statistical power in the analysis.

## RESULTS AND DISCUSSION

We fused RecT to Cas12a *via* the previously reported SunTag-based recruitment system (Tanenbaum et al., 2014; Huang et al., 2017). In our design, Cas12a linking a tandem repeat of 10 copies

of GCN4 peptides were cloned into a mammalian expression vector. The RecT protein (GFP protein in control) linking a single chain variable fragment (ScFv) of the anti-GCN4 antibody, was cloned into another vector. Upon the co-expression of two plasmids, ScFv antibodies would bind to their cognate peptide pair GCN4, recruiting RecT to the Cas12a protein (Figure 1A).

We tested our system *via* a genomic knock-in assay where a repair template of a T2A-mKate cassette (~1 kb in length) was provided. The template bears homology arms (HA) on both sides that match the last exon (in-frame before the stop codon) of the



target genes (**Figure 1B**). HDR efficiency was then quantified by fluorescence-activated cell sorting (FACS) where the percentage of mKate-positive cells represents successful knock-in.

Many types of Cas12a orthologs have been found across different bacteria; among those, *Acidaminococcus sp.* Cas12a (AsCas12a) and *L. bacterium* Cas12a (LbCas12a) are the two

that are most efficient and commonly used (Zetsche et al., 2015). Therefore, we tested our design using these two orthologs in the HEK293T cell line. As expected, across two different gene targets (*DYNLT1* and *AASV1*), there were significant increases of HDR efficiency when RecT was recruited compared to the control in both of the Cas12a variants. Moreover, AsCas12a performed



slightly better compared to LbCas12a. Hence, we further optimized our design using AsCas12a in the rest of study (Figure 1C).

As we have previously shown that the template HA length is a deciding factor of the HDR efficiency (Wang et al., 2021), we investigated such influence in Cas12a-REDIT by testing different HA designs. We found that symmetric HA performed better than asymmetric HA in general, and while our design could work with as short as ~100 bp template HA on each side, the efficiency increased as the HA length increased. A template that had ~400 bp HA on both sides performed the best among all six designs tested (Figure 1D).

Next, we examined our designs in two more different human cell lines, Hela (cervix-derived) and HepG2 (liver-derived), in order to evaluate the generalizability of our system. As expected, there were notable improvements of nearly 2-fold in the HDR efficiency (Figure 1E).

To investigate whether the improvement persisted in the case of multi-target editing, two genes *DYNLT1* and *EMX1* were targeted at the same time where mKate was inserted into *DYNLT1* and a 16bp insertion was made into *EMX1*. For the latter, next-generation sequencing was performed to assay the editing efficiency. There was a nearly four-fold increase of editing efficiency for both long and short knock-in editing when using the Cas12a REDIT method (Figure 2A). This enhancement of editing efficiencies in multiple-target settings were further validated by changing *DYNLT1* locus to *AAVS1* (Figure 2B).

Multiple genes can be edited by expressing a single CRISPR-Cas12a crRNA array together with Cas12a protein, taking advantage of its dual DNase and RNase activity (Zetsche et al., 2017). To further understand the potential of multiple-target editing using this system, we then targeted four genes (*AAVS1*, *DYNLT1*, *EMX1*, *DNMT1*) using a crRNA array and inserted mKate into *DYNLT1* locus and GFP into *AAVS1* locus at the same time. Focusing on fluorescent tag knock-in editing, we tested four different array designs with four guides in distinct orders, and achieved successful simultaneous HDR editing of both mKate and GFP with an approximate 2–4.5-fold increase in efficiency using Cas12a-REDIT (Figure 2C). While we maintained the guide sequence across the designs, we observed different editing efficiencies (Figure 2C). This is likely due to the varying processing efficiencies of crRNA array by Cas12a protein. This processing could be influenced by multiple factors. On one hand, the positioning of the crRNA within an array could influence the processing efficiencies, thus leading to variable expression levels of individual guide depending on the position, explaining the results we observed (Campa et al., 2019). On the other hand, recent work (Magnusson et al., 2021) also reported that the guide sequence could also influence the processing efficiencies of Cas12a crRNA arrays, which could also contribute to the differences in editing outcomes when we change the position of crRNAs within an array.

After validating the multi-target editing of the Cas12a-REDIT method, we additionally assayed the off-target effect of our design by profiling the editing events on potential off-target sites (OTS). We examined two genes, *EMX1* and *VEGFA*, and OTS were predicted using <http://crispor.org>. Based on the score ranking,

five *EMX1* OTS and six *VEGFA* OTS were selected (sequences listed in **Supplementary Table**), and editing efficiencies on these OTS were determined by next-generation sequencing. For both *EMX1* and *VEGFA*, our system exhibited no markable increase in the off-target editing (Figures 2D,E).

In summary, by employing the SSAP activity of phage RecT, we engineered a Cas12a-based system that can increase the efficiency of kilo-base scale HDR editing without compromising its specificity, and such improvements could also be achieved in multiplex genome editing.

In the longer term, additional method development could enhance the utility of this multiplex genome knock-in system. Specifically, many emerging areas of gene-editing research will have significant synergy with our work: protein engineering to define minimal-functional Cas12a-RecT editor including to optimize its efficiency; more powerful delivery systems including for the repair templates, which is very challenging for *in vivo* application of this multiplex editor. These additional advancements will further enhance our approach and help to achieve high-efficiency, multi-target long-sequence editing in human cells.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI BioProject PRJNA735023.

## AUTHOR CONTRIBUTIONS

LC conceived of the project. CW, QX, and QZ performed genome engineering experiments. QX, CW, JC, and LC performed the off-target sequencing. CW and QX performed deep amplicon sequencing experiments. JC, CW, and LC performed computational analysis of sequencing data. LC supervised the research and experimental design. QZ, CW, and LC wrote the manuscript with input from all authors.

## FUNDING

This work was supported by the National Institutes of Health (1R35HG011316 to LC); a Donald and Delia Baxter Foundation Faculty Scholar award (LC). The computational work is supported by NIH grant 1S10OD023452 to Stanford Genomics Cluster (SCG). We thank the entire Le Cong and Michael Cleary laboratory for their support, and for comments on the manuscript. We thank Feng Pan, Xin Yang, Li Zhu for generous help.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.719705/full#supplementary-material>

## REFERENCES

- Anzalone, A. V., Koblan, L. W., and Liu, D. R. (2020). Genome Editing with CRISPR-Cas Nucleases, Base Editors, Transposases and Prime Editors. *Nat. Biotechnol.* 38, 824–844. doi:10.1038/s41587-020-0561-9
- Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D., and Platt, R. J. (2019). Multiplexed Genome Engineering by Cas12a and CRISPR Arrays Encoded on Single Transcripts. *Nat. Methods* 16, 887–893. doi:10.1038/s41592-019-0508-6
- Clement, K., Rees, H., Canver, M. C., Gehrke, J. M., Farouni, R., Hsu, J. Y., et al. (2019). CRISPResso2 Provides Accurate and Rapid Genome Editing Sequence Analysis. *Nat. Biotechnol.* 37 (3), 224–226. doi:10.1038/s41587-019-0032-3
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339, 819–823. doi:10.1126/science.1231143
- Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157, 1262–1278. doi:10.1016/j.cell.2014.05.010
- Huang, Y.-H., Su, J., Lei, Y., Brunetti, L., Gundry, M. C., Zhang, X., et al. (2017). DNA Epigenome Editing Using CRISPR-Cas SunTag-Directed DNMT3A. *Genome Biol.* 18, 176. doi:10.1186/s13059-017-1306-z
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-Programmed Genome Editing in Human Cells. *eLife* 2, e00471. doi:10.7554/eLife.00471
- Kim, D., Kim, J., Hur, J. K., Been, K. W., Yoon, S.-h., and Kim, J.-S. (2016). Genome-Wide Analysis Reveals Specificities of Cpf1 Endonucleases in Human Cells. *Nat. Biotechnol.* 34 (8), 863–868. doi:10.1038/nbt.3609
- Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, Y. E., Hsu, J. Y., Clement, K., et al. (2019). Engineered CRISPR-Cas12a Variants with Increased Activities and Improved Targeting Ranges for Gene, Epigenetic and Base Editing. *Nat. Biotechnol.* 37, 276–282. doi:10.1038/s41587-018-0011-0
- Kleinstiver, B. P., Tsai, S. Q., Prew, M. S., Nguyen, N. T., Welch, M. M., Lopez, J. M., et al. (2016). Genome-Wide Specificities of CRISPR-Cas Cpf1 Nucleases in Human Cells. *Nat. Biotechnol.* 34 (8), 869–874. doi:10.1038/nbt.3620
- Knott, G. J., and Doudna, J. A. (2018). CRISPR-Cas Guides the Future of Genetic Engineering. *Science* 361, 866–869. doi:10.1126/science.aat5011
- Komor, A. C., Badran, A. H., and Liu, D. R. (2017). CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* 168, 20–36. doi:10.1016/j.cell.2016.10.044
- Li, T., Zhu, L., Xiao, B., Gong, Z., Liao, Q., and Guo, J. (2019). CRISPR-Cpf1-Mediated Genome Editing and Gene Regulation in Human Cells. *Biotechnol. Adv.* 37 (1), 21–27. doi:10.1016/j.biotechadv.2018.10.013
- Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., et al. (2018). Base Editing with a Cpf1-Cytidine Deaminase Fusion. *Nat. Biotechnol.* 36, 324–327. doi:10.1038/nbt.4102
- Magnusson, J. P., Rios, A. R., Wu, L., and Qi, L. S. (2021). Enhanced Cas12a Multi-Gene Regulation Using a CRISPR Array Separator. *Elife* 9 (10), e66406. doi:10.7554/elife.66406
- McCarty, N. S., Graham, A. E., Studená, L., and Ledesma-Amaro, R. (2020). Multiplexed CRISPR Technologies for Gene Editing and Transcriptional Regulation. *Nat. Commun.* 11, 1281. doi:10.1038/s41467-020-15053-x
- Pickar-Oliver, A., and Gersbach, C. A. (2019). The Next Generation of CRISPR-Cas Technologies and Applications. *Nat. Rev. Mol. Cell Biol.* 20, 490–507. doi:10.1038/s41580-019-0131-5
- Swarts, D. C., van der Oost, J., and Jinek, M. (2017). Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Mol. Cell* 66 (2), 221–233. doi:10.1016/j.molcel.2017.03.016
- Tak, Y. E., Kleinstiver, B. P., Nuñez, J. K., Hsu, J. Y., Horng, J. E., Gong, J., et al. (2017). Inducible and Multiplex Gene Regulation Using CRISPR-Cpf1-Based Transcription Factors. *Nat. Methods* 14, 1163–1166. doi:10.1038/nmeth.4483
- Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., and Vale, R. D. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell* 159, 635–646. doi:10.1016/j.cell.2014.09.039
- Wang, C., Cheng, J. K. W., Zhang, Q., Hughes, N. W., Xia, Q., Winslow, M. M., et al. (2021). Microbial Single-Strand Annealing Proteins Enable CRISPR Gene-Editing Tools with Improved Knock-In Efficiencies and Reduced Off-Target Effects. *Nucleic Acids Res.* 49, e36. doi:10.1093/nar/gkaa1264
- Wang, X., Ding, C., Yu, W., Wang, Y., He, S., Yang, B., et al. (2020). Cas12a Base Editors Induce Efficient and Specific Editing with Low DNA Damage Response. *Cel. Rep.* 31 (9), 107723. doi:10.1016/j.celrep.2020.107723
- Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., et al. (2016). Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell* 165, 949–962. doi:10.1016/j.cell.2016.04.003
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* 163, 759–771. doi:10.1016/j.cell.2015.09.038
- Zetsche, B., Heidenreich, M., Mohanraj, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex Gene Editing by CRISPR-Cpf1 Using a Single crRNA Array. *Nat. Biotechnol.* 35 (1), 31–34. doi:10.1038/nbt.3737

**Conflict of Interest:** Stanford University has filed patent applications based on this work. The authors will make the reagents openly available to the academic community through Addgene.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Xia, Zhang, Qu, Su, Cheng, Hughes and Cong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.