LETTER

Engineering hfCas12Max for improved gene editing efficiency

Xiaoning Han,[1,2] Ying Chen,[2] Ruiheng Liu,[2] Jian-Kang Zhu,[3,4,*] and Zhiqiang Duan[2,3,*]

[1]College of Life Sciences, Shandong Normal University, Jinan 250014, China

[2]Bellagen Biotechnology Co. Ltd, Jinan 250014, China

[3]Institute of Advanced Biotechnology and School of Medicine, Southern University of Science and Technology, Shenzhen 518055, China

[4]Center for Advanced Bioindustry Technologies, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081, China

*Correspondence: zhujk@sustech.edu.cn (J.Z.); duanzq@sustech.edu.cn (Z.D.)

Received: March 18, 2024; Accepted: May 17, 2024; Published Online: May 21, 2024; https://doi.org/10.59717/j.xinn-life.2024.100067

Citation: Han X., Chen Y., Liu R., et al., (2024). Engineering hfCas12Max for improved gene editing efficiency. The Innovation Life **2(2)**: 100067.

Dear Editor,

The clustered regularly interspaced short palindromic repeats-CRISPR associated proteins (CRISPR-Cas) systems provide adaptive immunity to prokaryotes against viral pathogens.[1] These natural bacterial defense systems have been engineered to become versatile tools for genome editing in diverse fields such as gene therapy, animal and plant breeding, and genetics research.[2] Cas9 and Cas12a are highly efficient and widely used gene editing tools, in the search for other RNA-guided nucleases, many other Cas12 proteins as well as TnpBs and IscBs have been discovered from genome sequence databases.[3,4] However, these recently discovered nucleases are characterized with low editing efficiency compared to Cas9. Strategies in protein engineering have been used successfully to improve the editing efficiency, expand target site compatibility, and reduce off-target effects of RNA-guided nucleases.[5,6] For instance, protein engineering significantly improved xCas12i, yielding the hfCas12Max with reported gene editing activities surpassing Cas9 in mammalian cells.[5] The efficacy of hfCas12Max in plants has not yet been evaluated. In this study, we examined this and also tested whether the editing efficiency of hfCas12Max could be further improved. We integrated homologous mutations known to enhance the activity of Cas12i3 into the hfCas12Max scaffold, generating a hfCas12Max variant (hfCas12Max**) with robust editing efficiency in mammalian cells and plants.

The hfCas12Max comprises 1,081 amino acids, smaller than SpCas9 (1,368 aa) and comparable to the previously characterized Cas12i3 (1,045 aa). We compared the Alpha-fold2 predicted protein structures of Cas12i3 and hfCas12Max, finding that their folded structures were similar (RMSD = 2.4) overall, despite of only 29% amino acid sequence identity (Figure 1A). Guided by previous engineering of the Cas12i3 variant Cas-SF01,[6] 23 mutations conferring ≥50% increase in Cas12i3 nuclease activity were introduced at structurally homologous sites in the hfCas12Max scaffold. The results showed that 43% the mutations increased the activity of hfCas12Max, with one mutation, A825R, increasing the activity by nearly 20% and reaching an editing efficiency of 98.3% (Figure 1B). We named this quadruple mutant (N243R/E336R/A825R/D892R) as hfCas12Max**. To further validate the editing efficiency, 20 target sites on the *TTR* gene were tested using hfCas12Max**. This multi-locus assessment demonstrated an average of 28% increase in editing efficiency relative to the reported hfCas12Max (Figure 1C and D). To further evaluate hfCas12Max** on-target specificity in human cells, deep sequencing quantified indel frequencies at the *TTR* and *P2RX5* target sites, as well as computationally predicted off-target loci. This analysis showed that hfCas12Max** had a higher on-target editing efficiency while maintaining minimal off-target effects (Figure 1E).

Previous studies have indicated that enhancing the activity of Cas nucleases may also result in broadening their protospacer adjacent motif (PAM) recognition range.[5,6] To decipher the impact of the A825R substitution on PAM recognition of hfCas12Max, *E. coli* cells expressing hfCas12Max** and crRNA were transformed with a plasmid library harboring 6 bp randomized nucleotides upstream of the target. Deep sequencing revealed both hfCas12Max** and hfCas12Max preferentially recognize 5′-TTN-3′ PAMs, indicating the A825R mutation in the RuvCII domain does not alter PAM specificity (Figure 1F). Additionally, a 5′-NNN PAM recognition assay was conducted utilizing a GFP reporter system. Consistent with the bacterial results, the two nuclease versions exhibited comparably high target cleavage activities at TTN PAM sites. Different from previous work,[5] hfCas12Max showed no appreciable activity at TNN, ATN, GTN or CTN non-canonical PAM sites in this analysis (Figure 1F).

To explore applications in agriculture, the gene editing efficacy of hfCas12Max** was evaluated in plants. Rice was stably transformed via *Agrobacterium tumefaciens* and soybean hairy roots were generated using *Agrobacterium rhizogenes* to compare editing efficiencies between hfCas12Max and hfCas12Max**. Since many Cas12 family proteins possess the ability to process their own CRISPR RNA,[7] we designed a pre-crRNA array to target three endogenous rice genes, a miRNA (*OsMIR396e*), young seedling albino (*OsYSA*), and narrow leaf (*OsNAL*), to evaluate the editing efficiency. Sequence analysis of T_0 plants showed 20-95% editing efficiencies for hfCas12Max**, surpassing original hfCas12Max across the three rice genes (Figure 1G). The results suggest that unlike Cas12i3,[8] hfCas12Max is effective in processing CRISPR pre-crRNA array in rice. For the test in soybean, we used a tRNA-crRNA tandem array to target three soybean genes, JAGGED 1(*GmJag1*), fatty acid desaturase 2 (*GmFAD2-1a*), and a sugar transporter (*Gmsweet10a*). The gene editing efficiency of hfCas12Max and hfCas12Max** in soybean was lower than in rice and mammalian cells (Figure 1G), indicating that more efforts are needed to improve the gene editing efficiency in soybean. We also analyzed the genotypes and mutation types generated by hfCas12Max. Most of the editing events were biallelic and chimeric, and hfCas12Max predominantly generated relatively large deletions in plants and mammalian cells (Figure 1H), which was consistent with other type V Cas systems.[7] Together, these results suggest a potential for hfCas12Max** as an effective platform for plant gene editing and crop improvement.

In this study, structure-guided protein engineering successfully enhanced the editing efficiency of hfCas12aMax. The improved editing performance of hfCas12Max** in mammalian cells and plants renders it a versatile tool for gene therapy and plant breeding.

MATERIALS AND METHODS

Vector construction

The full-length open reading frame (ORF) of hfCas12Max was synthesized by TSINGKE Biological Technology and cloned into the pET-28a bacterial expression vector, which harbored a crRNA cassette derived from J23119. This construct enabled determination of protospacer adjacent motif (PAM) recognition for both hfCas12Max and hfCas12Max**. For editing assays in mammalian cells, a human codon-optimized hfCas12Max gene synthesized by TSINGKE was cloned into pcDNA3.3 as previously described. Vectors expressing hfCas12Max variants were generated through PCR-based site-directed mutagenesis using the pEASY Seamless Cloning and Assembly Kit (Transgen). Rice and soybean expression constructs were generated as previously described.[6]

Homology analysis of CRISPR-Cas12i

The protein structures of Cas12i3 and hfCas12Max were predicted using AlphaFold2 under default parameters. The top five predicted models were compared, with the highest ranked (Rank 0) selected for analysis. Structural alignment was performed using PyMol v2.4.0 with default settings.

PAM depletion assay and analysis

PAM depletion assay was conducted as previously described.[6] pET28a - hfCas12Max and negative control plasmid were transformed into *E. coli* BL21(DE3), and the positive clone was used to make competent cells by ice cold water and 10 % glycerol washing. Competent cells were transformed by

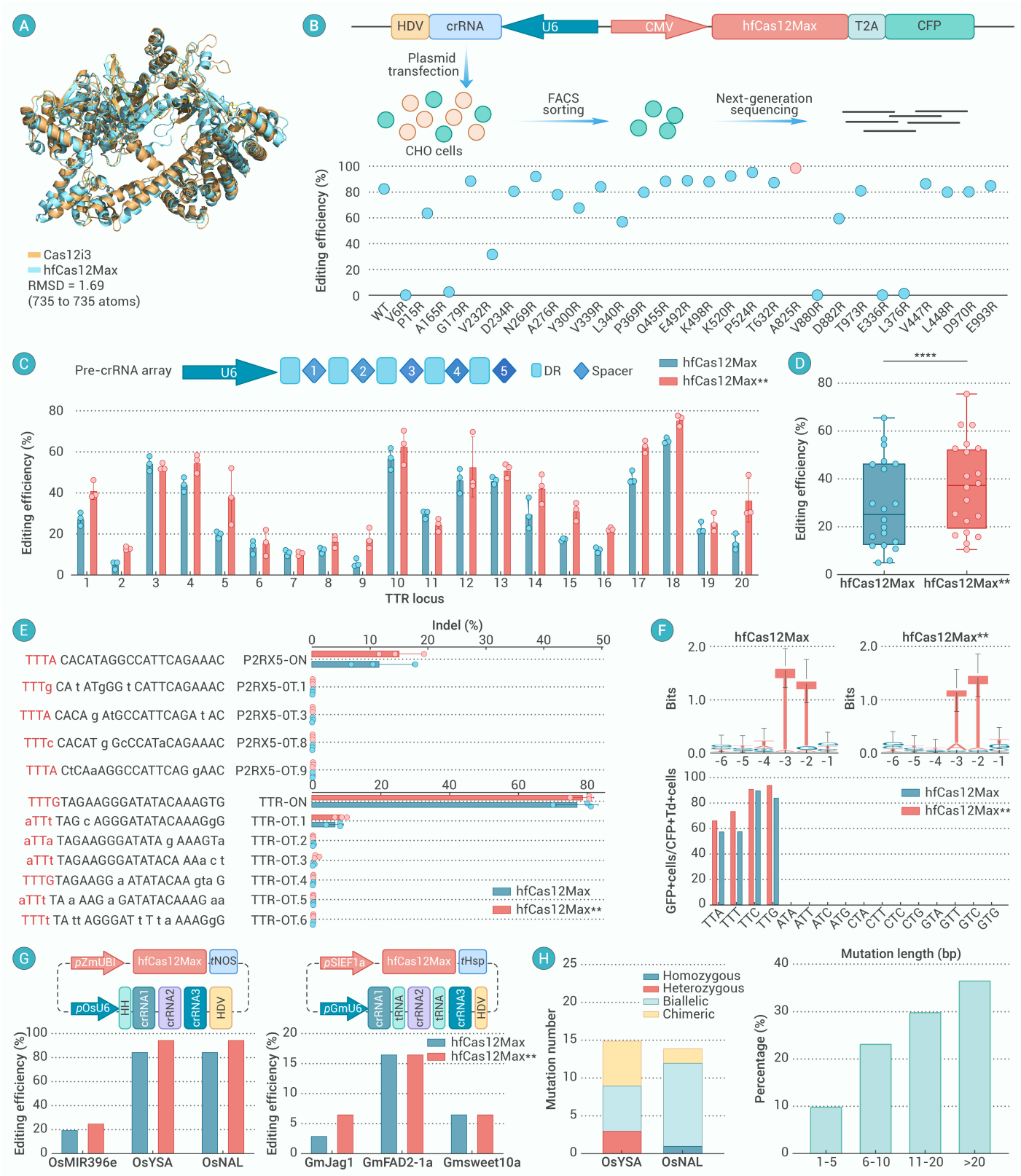

Figure 1. hfCas12Max, an engineered variant hfCas12Max, mediates high-efficiency editing in mammalian cells and plants** (A) The structural similarity between Cas12i3 (yellow) and hfCas12Max (blue) as predicted by Alpha-fold. (B) Upper panel, the experimental workflow for detecting hfCas12Max cleavage activity at endogenous loci. T2A, thosea asigna virus 2A peptides. Lower panel, the editing efficiency of hfCas12Max variants. (C) Comparison of indel efficiency of hfCas12Max and hfCas12Max** at 20 endogenous sites using a pre-crRNA array in HEK293T cells. n = 3 independent biological replicates. (D) Summary dot plots showing the activity of hfCas12Max and hfCas12Max** in HEK293T cells. Each point represents the average editing efficiency from three independent biological replicates measured at each endogenous sites. Significant difference was determined by Student's t-test. (E) Off-target efficiency of hfCas12Max** and hfCas12Max at in-silico predicted off-target sites, determined by targeted deep sequencing. PAM sequences are in red and mismatched bases are in lowercase letter. n = 3 independent biological replicates. (F) The PAM preference of hfCas12Max and hfCas12Max** was characterized using both deep sequencing data (upper panel) and the GFP reporter system (lower panel). (G) Editing efficiencies of hfCas12Max and hfCas12Max** in plants. Right panel, the rates of mutations caused by hfCas12Max and hfCas12Max** using pre-crRNA array in rice, n=20. Left panel, comparison of indel efficiency by hfCas12Max and hfCas12Max** in soybean hairy roots, n=30. The crRNA was separated by tRNA. HDV, hepatitis delta virus, HH, hammerhead ribozyme. (H) Distribution of the genotypes and mutation lengths generated by hfCas12Max at the *OsYSA*, *OsNAL*, and *OsMIR396e* target genes in rice in T_0 plantlets.

electroporation with 200 ng library plasmids harboring 6 bp randomized nucleotides positioned upstream of the 5' target sequence. After electroporation, the BL21 cells were plated onto selection media containing 50 μg/mL kanamycin, 50 μg/mL chloramphenicol, and 0.1 mM IPTG for 24 h. Subsequently, PAM depletion plasmids were isolated from the cultured cells using a QIAprep Spin Miniprep Kit (Qiagen). After Amplicon sequencing of the targeted plasmid, the reads containing 10 nt 5' sequence of the 6 bp PAM region were extracted and analyzed. The frequency of each PAM was calculated and normalized to the total reads for each sample. Enriched PAMs which were considered significantly > 2.5-fold from the negative control (experiments setup without functional hfCas12Max complex) were utilized for generating PAM sequence logos using WebLogo.

Off target analysis

Potential off-target sites (OTs) for the crRNA targeting *TTR* and *P2RX5* were predicted using the online tool CAS-OFFinder (http://www.rgenome.net/cas-offinder/: http://www.rgenome.net/cas-offinder/).[9] The predicted OTs were subsequently amplified by PCR and sequenced using next-generation sequencing (NGS) and the data was analyzed with the Hi-TOM platform.[10]

Cell culture, transfection and flow cytometry analysis

HEK293T (CCL-61, ATCC)) and CHO cells were cultured in DMEM (Gibco) with 10% (v/v) FBS (Gemini) and 1% (v/v) penicillin streptomycin (Gibco) in an atmosphere of 5.0 % CO2 at 37 °C. 1.0×10^5 cells were seeded onto each well of 24-well plates, cultured to ~70% confluency. 600 ng pCAG-T2A-EGFP plasmid expressing hfCas12Max or variant with guide RNAs which targeted the endogenous *FUT8* gene was transfected into the cells using Lipofectamine 3000 (Invitrogen) according to the manufacturer's recommended protocol. After 48 hours, transfected cells were digested by Trypsin-EDTA (0.05%) (Gibco), pelleted at 135x g for 3 mins. 80,000 GFP+ cells were sorted via BD FACSMelodyTM Cell Sorter for subsequent analysis. For NNN PAM screening, 200,000 293T cells per well were cultured in 6-well plates to 70-80% confluence over 18 hours, followed by transfection with 3 μg total plasmid (1.5 μg hfCas12Max plasmid + 1.5 μg reporter plasmid). 80,000 cells were FACS sorted for quantification of CFP, mCherry and EGFP fluorescence.

Agrobacterium rhizogenes-mediated transformation.

Soybean seeds were planted into a pot with sterile wet nutrient soil and vermiculite mixed 2:1. Plants were grown for 7~14 days with 12 h light/12 h dark illumination, at 25°C and approximately 70% relative humidity. One cm of the hypocotyl of soybean seedlings was cut diagonally with a sterile scalpel. The slant cut of hypocotyl was applied with *Agrobacterium* K599 containing the expression vector, directly planted into pot with sterile vermiculite and watered with 10 ml of K599 *Agrobacterium* solution, subsequently maintained at high humidity for 16 days to promote hairy root growth.

Genome editing in rice

Rice stable transformation was carried out as previously described.[6] The hfCas12Max or hfCas12Max** expression constructs were introduced into *Agrobacterium tumefaciens* strain EHA105, and then transformed into calli induced from the elite rice variety WZJ006 (*Oryza sativa* L. ssp. *japonica*). The genomic DNA was extracted from the regenerated plantlets, and the target sequence was amplified by PCR for sequencing.

Targeted deep sequencing

Genomic DNA of GFP-positive CHO or HEK293T cells was extracted using the Quick Extract DNA Extraction Solution (QE09050, Epicentre). Targeted sequencing primers were synthesized, and the specific region was amplified using Phanta Max Super-Fidelity DNA Polymerase (Vazyme Biotech, Nanjing, China) by nested PCR. The resulting PCR products (220~230 bp) were purified and sequenced by next-generation sequencing (NGS) and analyzed with the Hi-TOM platform.[10]

REFERENCES

1. Barrangou, R., Fremaux, C., Deveau, H., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science **315**: 1709–1712. DOI: 10.1126/science.1138140.
2. Wang, J.Y., and Doudna, J.A. (2023). CRISPR technology: A decade of genome editing is only the beginning. Science **379**: eadd8643. DOI: 10.1126/science.add8643.
3. Karvelis, T., Druteika, G., Bigelyte, G., et al. (2021). Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. Nature **599**: 692–696. DOI: 10.1038/s41586-021-04058-1.
4. Altae-Tran, H., Kannan, S., Demircioglu, F.E., et al. (2021). The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. Science **374**: 57–65. DOI: 10.1126/science.abj6856.
5. Zhang, H., Kong, X., Xue, M., et al. (2023). An engineered xCas12i with high activity, high specificity, and broad PAM range. Protein Cell **14**: 538–543. DOI: 10.1093/procel/pwac052.
6. Duan, Z., Liang, Y., Sun, J., et al. (2024). An engineered Cas12i nuclease that is an efficient genome editing tool in animals and plants. The Innovation **5**: 100564. DOI: 10.1016/j.xinn.2024.100564.
7. Fonfara, I., Richter, H., Bratovič, M., et al. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature **532**: 517–521. DOI: 10.1038/nature17945.
8. Lv, P., Su, F., Chen, F., et al. (2024). Genome editing in rice using CRISPR/Cas12i3. Plant Biotechnol. J. **22**: 379–385. DOI: 10.1111/pbi.14192.
9. Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics **30**: 1473–1475. DOI: 10.1093/bioinformatics/btu048.
10. Liu, Q., Wang, C., Jiao, X.Z., et al. (2019). Hi-TOM: A platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. Sci. China Life Sci. **62**: 1–7. DOI: 10.1007/s11427-018-9402-9.

FUNDING AND ACKNOWLEDGMENTS

This work was supported by Shandong BellaGen Biotechnology Co., Ltd. The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

J-K.Z., X.H. and Z.D. supervised the project, designed the research, and participated in data analysis. X.H., Y.C., R.L., performed the experiments and analyzed the data. Z.D and J-K.Z. wrote the manuscript.

DECLARATION OF INTERESTS

J-K.Z. is an Editorial Board member of The Innovation Life. He was blinded from reviewing or making final decisions on the manuscript. The article was subject to the journal's standard procedures, with peer review handled independently of this Editorial Board member and their research groups. The other authors declare that they have no conflicts of interest.

DATA AND CODE AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.