Nucleic Acids Research, 2024, **52**, 4739–4755
https://doi.org/10.1093/nar/gkae228
Advance access publication date: 3 April 2024
Synthetic Biology and Bioengineering

Multiplexed *in-situ* mutagenesis driven by a dCas12a-based dual-function base editor

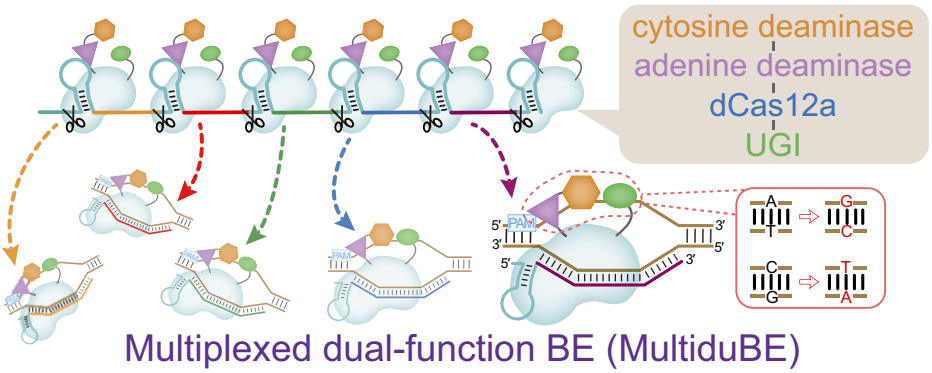
Yaokang Wu [1,2], Yang Li[1,2], Yanfeng Liu [1,2], Xiang Xiu[1,2], Jiaheng Liu[1,2], Linpei Zhang[3], Jianghua Li[1,2], Guocheng Du[2], Xueqin Lv[1,2], Jian Chen[2], Rodrigo Ledesma-Amaro [4] and Long Liu [1,2,*]

[1]Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China
[2]Science Center for Future Foods, Jiangnan University, Wuxi 214122, China
[3]Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, China
[4]Department of Bioengineering and Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, UK
[*]To whom correspondence should be addressed. Tel: +86 510 85918312; Fax: +86 510 85918309; Email: longliu@jiangnan.edu.cn

Abstract

Mutagenesis driving genetic diversity is vital for understanding and engineering biological systems. However, the lack of effective methods to generate *in-situ* mutagenesis in multiple genomic loci combinatorially limits the study of complex biological functions. Here, we design and construct MultiduBE, a dCas12a-based multiplexed dual-function base editor, in an all-in-one plasmid for performing combinatorial *in-situ* mutagenesis. Two synthetic effectors, duBE-1a and duBE-2b, are created by amalgamating the functionalities of cytosine deaminase (from hAPOBEC3A or hAID*Δ), adenine deaminase (from TadA9), and crRNA array processing (from dCas12a). Furthermore, introducing the synthetic separator Sp4 minimizes interference in the crRNA array, thereby facilitating multiplexed *in-situ* mutagenesis in both *Escherichia coli* and *Bacillus subtilis*. Guided by the corresponding crRNA arrays, MultiduBE is successfully employed for cell physiology reprogramming and metabolic regulation. A novel mutation conferring streptomycin resistance has been identified in *B. subtilis* and incorporated into the mutant strains with multiple antibiotic resistance. Moreover, surfactin and riboflavin titers of the combinatorially mutant strains improved by 42% and 15-fold, respectively, compared with the control strains with single gene mutation. Overall, MultiduBE provides a convenient and efficient way to perform multiplexed *in-situ* mutagenesis.

Graphical abstract



Introduction

Natural genetic diversity, even just single nucleotide variation (SNV), provides unique biological functionalities (1). Creating intentional genetic diversity via mutagenesis is a key way for better understanding and engineering biological systems, which can be followed by the identification of the mutants of interest by artificial selection or screening of the genetic variants (2). Traditional methods that combine *in vitro* diversification and *in vivo* selection can be time-consuming and tedious. To overcome these limitations, various *in vivo* mutagenesis systems, such as phage-assisted continuous evolution (PACE) (3), OrthoRep (powered by the orthogonal DNA polymerase) (4,5), and MutaT7 (driven by T7 RNA polymerase fused with a cytidine deaminase) (6), have been developed. However, these methods require the insertion of a copy of the target gene into a phage or plasmid, limiting their ability to induce mutations in the original genomic loci. As a consequence, the genetic and phenotypic diversity generated often differs from that naturally occurring (7).

Received: July 21, 2023. **Revised:** March 13, 2024. **Editorial Decision:** March 15, 2024. **Accepted:** March 19, 2024
© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

Traditional chemical and physical mutagenesis methods, such as ethyl methane sulfonate (EMS) or UV mutagenesis, accelerate the microbial evolution process by inducing damage to genomic DNA. However, the mutational sites are relatively random, and their identification is limited to whole-genome sequencing. The Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) systems (e.g. Cas9 and Cas12a), which can precisely recognize the target DNA with a protospacer adjacent motif (PAM) via single guide RNA (sgRNA) or CRISPR RNA (crRNA) guidance, have revolutionized the gene editing process (8). CRISPR-derived tools, coupled with specific functional domains, enable *in-situ* genomic mutagenesis similar to traditional chemical and physical mutagenesis, providing this process with improved target specificity. For instance, the EvolvR system has been built to diversify targeted genomic loci by fusing a nick-translating and error-prone DNA polymerase to the nickase version of Cas9 (nCas9, D10A or H840A) (2). Furthermore, cytosine base editors (CBEs) or adenine base editors (ABEs), which produce cytosine to thymine (C > T) or adenine to guanine (A > G) conversion in the genomic target, can be created by fusing cytosine deaminases or adenine deaminases to the catalytically dead Cas9 (dCas9) or nCas9 (D10A) (9). By combining both cytosine and adenine base editing activities, dual-function BE has also been constructed based on nCas9 to generate more variation (10–15). In addition, CRISPR systems enable multiplexed activities by utilizing multiple gRNAs/crRNAs to target a range of genomic loci (16,17), offering great potential for mutating multiple genes in combination to reveal interdependent genetic and phenotypic variations. Cas12a (or dCas12a), known for its multiplexed advantage in producing multiple crRNAs from a CRISPR array by its own RNase activity compared with Cas9 (17–19), may be more suitable for combinatorial *in-situ* mutagenesis by incorporating cytosine deaminases and adenine deaminases. However, this dual-function BE derived from dCas12a has not been built yet, and the multiplexed advantage was underutilized for the dCas12a-based CBEs and ABEs (20–25).

In this work, we design and construct the dCas12a-based multiplexed dual-function BE (MultiduBE), and employ it in the multiplexed *in-situ* mutagenesis of various biological phenotypes, aided by specific crRNA arrays (Figure 1). Through composition and structure optimization, synthetic effectors duBE-1a and duBE-2b possessing combinatorial activities of cytosine deaminase, adenine deaminase, and crRNA array processing are obtained. Further modifications, including replacing the promoter of the synthetic effector and engineering the crRNA array with a synthetic separator, allow MultiduBE to achieve the synchronized multiplexed *in-situ* mutagenesis at five different genomic loci in both *Escherichia coli* and *Bacillus subtilis*. Using *B. subtilis* as the host, we study the capacity of MultiduBE to enable phenotypic diversification by selecting fluorescent proteins (sYFP2, mKate, and mTagBFP2) and cell morphology determining proteins (FtsZ and MreB) as the targets. Additionally, we generate mutants conferring resistance to various antibiotics (tetracycline, rifampicin, spectinomycin, and streptomycin), including one that has never been reported before. Moreover, we create strains with improved either surfactin or riboflavin synthesis, which demonstrate the role of MultiduBE in metabolic diversity and regulation. As a conclusion, MultiduBE can produce genetic and phenotype variations through multiplexed *in-situ* genome mutagenesis, showing general applicability and portability in both basic and applied research.

Materials and methods

Chemicals and reagents

The chemicals and plasmid extraction kit were obtained from Sangon Biotech (Shanghai, China). The riboflavin, FMN, and surfactin standards were bought from Yuanye Bio-Technology (Shanghai, China). T4 DNA ligase and PrimeSTAR HS DNA polymerase were obtained from Takara Biomedical Technology (Beijing, China). Taq master mix (Dye Plus) and Phanta Max Master Mix (Dye Plus) were purchased from Vazyme Biotech (Nanjing, China). Restriction enzyme *Bsa*I used for Golden Gate Assembly was purchased from New England Biolabs (Beijing, China). The seamless cloning kit was obtained from Beyotime Biotechnology (Shanghai, China). The DNA gel purification kit was obtained from Thermo Scientific (Waltham, USA). Oligonucleotides and genes were synthesized from GENEWIZ (Suzhou, China). For common cell culture, Luria-Bertani (LB) medium (10 g/l tryptone, 5 g/l yeast extract, and 5 g/l NaCl) was used. The following antibiotics were used for selections in *E. coli*: ampicillin, 100 μg/ml; kanamycin, 50 μg/ml. The following antibiotics were used for selections in *B. subtilis*: kanamycin, 50 μg/ml; chloromycetin, 5 μg/ml; tetracycline, 20 μg/ml; rifampicin, 50 μg/ml; spectinomycin, 50 μg/ml; streptomycin, 200 μg/ml.

Plasmids and strains

Strains and plasmids used in this study are provided in Supplementary Note S1 and Supplementary Note S2, respectively. The plasmids constructed in this study can be acquired from the MolecularCloud plasmid repository (MC_0101390–MC_0101405, MC_0101449–MC_0101454). The primers used for plasmids and strains construction are shown in Supplementary Note S3. Plasmid and crRNA array sequences with the necessary annotations can be downloaded using the URL shown in Supplementary Note S2 and Supplementary Note S4. For plasmid construction, *E. coli* DH5α was used as the host, and the seamless cloning kit was used according to the manufacturer's instructions.

A previously reported method was used to transform the *B. subtilis* strains into supercompetent cells (26). The *B. subtilis* strain BSZRG (G00, *B. subtilis* 168Δ*epr*::XylR-P_{xylA}-*comKS*, Δ*trpC2*::*trpC0*, Δ*gudB*::*gudB*$^+$) (27) was used for the verification of the base editors and the generation of antibiotic-resistant strains. Strain G600 (*B. subtilis* 168Δ*epr*::XylR-P_{xylA}-*comKS*, Δ*trpC2*::*trpC0*, Δ*gudB*::*gudB*$^+$, *aprE*^{0}, *nprE*^{0}, *bpr*^{0}, *mpr*^{0}, *nprB*^{0}) (27) was used as the start host strain for riboflavin synthesis emergence. Sur-s0 (*B. subtilis* G600, *sfpm0*) was constructed aided by plasmid pWLBE-gCpf1Ng-sfpM0, and used as the start host strain for surfactin synthesis regulation. Strain G00-CmYKT (*B. subtilis* G00 *ybbU*::Cm^R-*syfp2-mkate-mtagbfp2*) was constructed by transforming the DNA fragment with corresponding homologous arms (S1-CmYKT) into strain G00, and used for fluorescent and morphology assay.

Base editing and sequencing analysis

Before introducing the synthetic separator, the crRNA array was constructed using the SOMACA method (28). After introducing the separator, a modified two-round PCR based

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

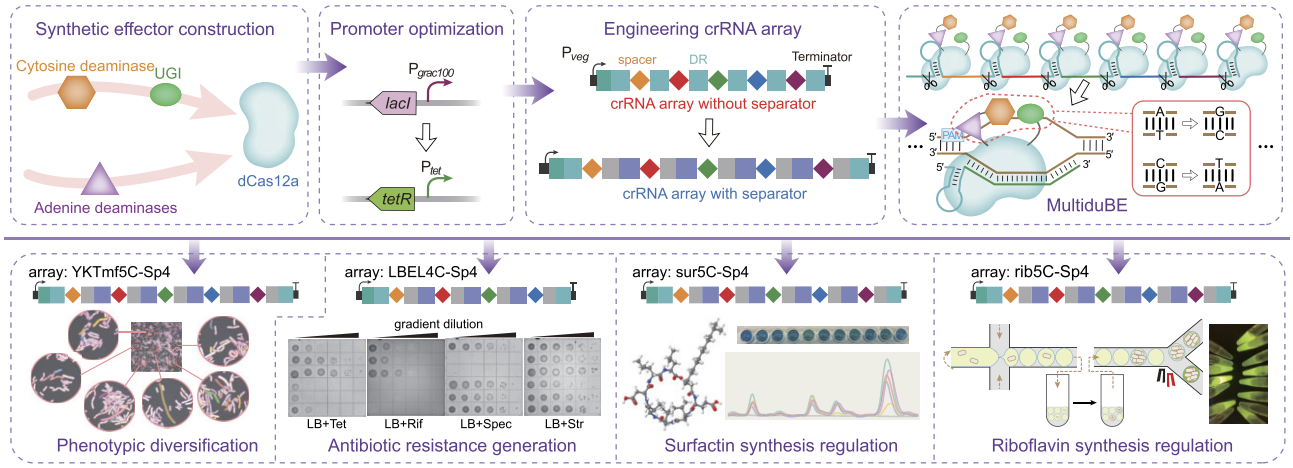


Figure 1. The workflow for the design, construction, optimization and application of the dCas12a-based multiplexed dual-function base editor (MultiduBE). MultiduBE, a versatile tool for synchronized multiplexed *in-situ* mutagenesis, was created by amalgamating the functionalities of cytosine deaminase, adenine deaminase, and dCas12a. Through the replacement of the synthetic effector's promoter and the engineering of the crRNA array with a synthetic separator, MultiduBE demonstrates its capability to achieve simultaneous in-situ mutagenesis at five distinct genomic loci in both *Escherichia coli* and *Bacillus subtilis*. MultiduBE was employed for performing multiplexed *in-situ* mutagenesis, taking generic phenotypic diversity, identification of antibiotic resistance mutation, and metabolic regulation as examples.

method was developed and applied for the assembly of the crRNA array (Supplementary Figure S1). The primers used for crRNA array construction were listed in Supplementary Note S3. The crRNA fragment in the first position of the array was produced by 1 round PCR using primers with the *Bsa*I adapter (Phanta Max 10 μl, 10 μM forward primer 5 μl, 10 μM reverse primer 5 μl, 30 cycles). The crRNA fragments except for the first position were generated using two-round PCR: In round-1, a long forward primer including 27-nt DR+, N23, and 32-nt Sp4 was designed and turned into a double strand with a general reverse primer complement to Sp (Phanta Max 10 μl, 10 μM forward primer 5 μl, 10 μM reverse primer 5 μl, 15 cycles); In round-2, the general forward primer targeting DR + with a *Bsa*I adapter and the general reverse primer targeting Sp with a *Bsa*I adapter were used, with PCR product of round-1 as the template (Phanta Max 10 μl, 10 μM forward primer 5 μl, 10 μM reverse primer 5 μl, PCR product of round-1 1 μl, 30 cycles). These PCR products containing crRNA fragments were diluted 10 times and directly used for the construction of the crRNA array by Golden Gate assembly (18). Based on this method, only the long forward primer in round-1 needed to be changed if new targets will be edited.

When the crRNA array was ligated in the all-in-one BE plasmid, the generated plasmid was verified by sequencing, transformed into *B. subtilis* strain G00 and cultured in a LB-kanamycin agar plate at 30°C. A single colony was then picked from the plate and inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media for pre-culture about 10 h, before being inoculated in other three parallel 14-ml sterile tubes containing 2 ml LB-kanamycin media with corresponding inducer (1 mM IPTG or 0.5 μM aTC). After 24 h induction, the targets on the genome of the 3 parallel cultures were amplified separately for analysis. For the resolution of the genome editing results, the Sanger sequencing method shown in Supplementary Figure S2A or the next-generation sequencing (NGS) method shown in Supplementary Figure S2B was employed. For Sanger sequencing, the target genes on the genome were amplified by using 2 μl culture from each of the 3 parallel as the template, and the BEAT software was employed to calculate the mutation rates (29). For high-throughput targeted amplicon sequencing, 4 fragments containing the 5 targets on *aprE* and *nprE* were amplified and equal mole mixed for sequencing, which was conducted for each of the 3 parallel cultures reads were used for each sam. In addition, potential off-target sites (POTs) were identified by the Cas-OFFinder algorithm (30) and subsequently amplified for NGS analysis. Library construction, sequencing, and data acquisition were performed on the Illumina NovaSeq platform by GENEWIZ (Suzhou, China), and the results were processed and handled using the Geneious software (> 500000 reads were used for each sample). The primers used for generating amplicons are shown in Supplementary Note S3. After induction, the whole genome resequencing (>500 × coverage depth), RNA-seq, and miRNA sequencing of the culture were also performed and analyzed by GENEWIZ (Suzhou, China). As shown in Supplementary Table S1, the strain G00 possesses numerous single nucleotide variations (SNVs) when compared to the reference genome of *B. subtilis* 168. These variations may have arisen due to long-term spontaneous mutations in the laboratory.

Assay of fluorescence intensity

The recombinant *E. coli* and *B. subtilis* strains with the fluorescence proteins were pre-cultured in LB medium for 10 h and further inoculated into 200 μl LB medium with 1% proportion at 96-well plates (corning 3603). The 96-well plates were subsequently cultured at 37°C with shaking at 750 rpm. The GFP fluorescence (excitation, 490 nm; emission, 530 nm), mCherry fluorescence (excitation, 580 nm; emission, 610 nm), and OD_{600} were measured using a microplate Multi-Mode Reader (BioTek, Cytation 3) directly (18). To calculate the relative fluorescence intensities, background OD of the medium (OD_{bg}) and background fluorescence of the strain without fluorescent protein expression (FP_{bg}) was excluded, and the equation (1) was applied as follows:

$$\left(\frac{FP}{OD}\right)_{corrected} = \frac{FP - FP_{bg}}{OD - OD_{bg}} \quad (1)$$

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

Flow cytometry assay and microscopy imaging

For the verification of MultiduBE promoted fluorescent and morphology diversification, the crRNA array YKTmf5C-Sp4 was assembled into pWLT-duBE-1a and pWLT-duBE-2b and then transformed into strain G00-CmYKT. Two single colonies were picked from the two agar plates and inoculated in a 14-ml sterile tube containing 2 ml LB media with kanamycin for pre-culture about 10 h. Subsequently, the pre-culture was inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media without aTC and a 14-ml sterile tube containing 2 ml LB-kanamycin media with 0.5 μM aTC, respectively. The two group cultures (without induction and with induction) were analyzed by flow cytometry and Laser Scanning Confocal Microscopy (LSCM).

Flow cytometry was performed on a BD FACSArica III with BD FACSDiva software (BD Biosciences). Specifically, 1 ml cultured cells were taken and washed three times with 0.012 M PBS (pH 7.4). Then, the cells resuspended to an OD_{600} of 0.3 were run at a rate of 0.5 μl/s, and the events were gated based on forward and side scatters to reduce false events (18)(Supplementary Figure S3). After gating, at least 10 000 events were collected for the analysis of each sample. All data were exported in FCS3 format and processed using FlowJo_V10 software. LSCM was carried out by a Nikon AX-SIM S microscope (Nikon, Tokyo, Japan) fitted with a 60× oil-immersion objective. The fluorescent proteins mTagBFP2, sYFP2, and mKate were excited by 405-, 488- and 561-nm laser lines of an argon-ion laser line of a He–Ne laser, respectively. Fluorescence emissions were detected with spectral detector sets BP 429–474 for mTagBFP2, BP 500–550 for sYFP2, and BP 570–616 for mKate. Image analysis was carried out on the NIS-Elements Viewer software.

Multiplexed *in-situ* mutagenesis and screen method

For antibiotic resistance generation, corresponding crRNAs were designed by referencing reported mutation sites in *B. subtilis* or aligning the genes of *B. subtilis* with those of other resistant mutant hosts. The MultiduBE plasmids containing crRNA arrays tetL-rpoB-Sp4, tetL-rpsE-Sp4, tetL-mtrp-Sp4, or LBEL4C-Sp4 were transformed into strain G00. The results colonies were scraped from the agar plate and inoculated in a 50-ml sterile tube containing 10 ml LB-kanamycin media for pre-culture about 10 h, before being inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media with 0.5 μM aTC. After induction, the mutation library was washed to remove aTC, and then inoculated into a 14-ml sterile tube containing 2 ml LB-kanamycin media with corresponding antibiotics and cultured at 30°C. The grown cells were streaked on antibiotics agar plate and cultured at 30°C for single colonies isolation and screening. The MultiduBE plasmid in the screened strain was eliminated by culturing the single colony in a 14-ml sterile tube containing 2 ml LB media without kanamycin at 50°C, followed by streaking on the LB agar plate without kanamycin and culturing at 50°C (Supplementary Figure S4A).

For surfactin synthesis regulation, the MultiduBE plasmids pWLT-duBE-1a and pWLT-duBE-2b containing crRNA array sur5C-Sp4 were transformed into strain G600. Two single colonies were picked from the two agar plates containing the BE plasmid and inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media for pre-culture about 10 h, before being inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media with 0.5 μM aTC. After induction, the mutation library was washed to remove aTC, and then spread on LB-kanamycin agar plates and cultured at 30°C for single colonies isolation. The surfactin overproducing strains were screened using a chromatic visible screening method (31), in which bromothymol blue (BTB) acts as a colour indicator and cetylpyridinium chloride (CPC) acts as a mediator, in 96-well plates (corning 3596). The difference in surfactin concentrations can be exactly reflected by a regular chromatic response from faint yellow-green to dark green and bright blue (Supplementary Figure S4B). The MultiduBE plasmid in the 12 screened strains Sur-s1∼s12 was also eliminated by treating at 50°C.

For riboflavin synthesis strain emergence, the MultiduBE plasmids pWLT-duBE-1a and pWLT-duBE-2b containing crRNA array rib5C-Sp4 were transformed into strain G600. Two single colonies were picked from the two agar plates containing the BE plasmid and inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media for pre-culture about 10 h, before inoculated in a 14-ml sterile tube containing 2 ml LB-kanamycin media with 0.5 μM aTC. After induction, riboflavin-producing strains in the mutation library were screened using FADS conducted in a microfluidic devices including droplet generator, droplet injector, droplet sorter, and observing chamber (32). The cells were collected by centrifugation (10 000 × *g* and 4°C for 2 min), washed with fresh LB-kanamycin medium, and then diluted with LB-kanamycin medium with 30 g/l glucose to an OD_{600} of 0.05. In this cell density, 10% of droplets (diameter was ∼30 μm) will contain a single cell and less than 0.5% will contain more than one cell except for the remaining empty droplets, owing to the Poisson distribution of cells in the droplets. The cells were collected in a 1-ml syringe, and the oil phase consisting of HFE 7500 (3M, USA) with 2% (v/v) Pico-Surf surfactant (5% in HFE 7500, Sphere Fluidics) was also put in another 1-ml syringe. The connection between the syringe and the microfluidic device was made using polyethylene tubing with an inner diameter of 0.38 mm and an outer diameter of 1.03 mm. The flow rates of the cell solution and spacing oil were controlled by driving the syringes with two pumps (PHD2000, Harvard, USA), and water-in-oil droplets were generated due to the shearing action of the oil phase. The generated droplets were incubated at 30°C with 220 rpm for 24 h in an inverted syringe. The incubated droplets were injected into the droplet sorter, and the fluorescence of the droplet was excited by white LED light with the corresponding filters (excitation, 495 nm; emission, 525 nm) in an inverted microscope (DMi-8, Leica). The emission fluorescent light was detected by a photomultiplier tube (PMT; H10721, Hamamatsu) and analyzed by Arduino DUE. The collected droplets were put in a 1.5-ml tube containing 100 μl demulsifier 1H,1H,2H,2H-perfluoro-1-octanol (Aladdin), and 200 μl LB-kanamycin medium. The aqueous phase containing the collected cells was spread on LB-kanamycin agar plates and cultured at 30°C. The isolated single colonies were inoculated into 96-well plates (corning 3603) for fluorescence determination (excitation, 490 nm; emission, 530 nm). The MultiduBE plasmid in the 14 screened strains Rib-s1∼s14 was also eliminated by treating at 50°C (Supplementary Figure S4C).

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

Shake flasks culture and analytical methods

The medium used for shake flask culture of riboflavin or surfactin contains 80 g/l glucose, 6 g/l urea, 12 g/l yeast extract, 6 g/l tryptone, 12.5 g/l $K_2HPO_4{\cdot}3H_2O$, 2.5 g/l KH_2PO_4, and 3 g/l $MgSO_4{\cdot}7H_2O$. To prepare the seed cultures, single colonies were picked into 10 ml LB medium in 250-ml shake flasks from the LB plates with the engineered *B. subtilis* strains. After being cultured at 37°C and 220 rpm for 10 h, the seed cultures were further inoculated into a 50 ml fermentation medium with 5% proportion at a 250-ml baffled flask and grown at 37°C and 220 rpm for 60 h (Supplementary Figure S4D), and three replicates were set for each strain. To measure cell densities (absorbance at 600 nm, OD_{600}), glucose, and product (riboflavin or surfactin), 1 ml cell suspension was sampled every 12 h.

Glucose in the fermentation broth was analyzed by a glucose-lactate analyzer (M100, Shenzhen Sieman Technology Co., Ltd, Shenzhen, China). For surfactin determination, 200 μl of the fermentation culture was added into 800 μl 0.1 M phosphate buffer solution (PBS, pH 8.0). After centrifugation at 10000 × *g* for 2 min, the supernatant was filtered through a 0.2 μm membrane. A 10 μl aliquot was injected into an HPLC system (Agilent 1260, Santa Clara, CA, USA) equipped with a variable wavelength detector (VWD, 205 nm) and a C18 column (4.6 × 250 mm; 5 μm; Waters, Ireland)(31). The mobile phases were 10% water and 90% methanol containing 0.1% trifluoroacetic acid (TFA). The total flow rate of the mobile phase was kept at 1.0 ml/min and the column temperature was kept at 40°C. The total peak area of 4 surfactin isoforms was calculated for quantification of the total concentration of surfactin according to the concentration standard curve. For analyzing the expression of the *srfA* gene cluster, the plasmid pWLBE-gCpf1Ng-srfA-NmC was constructed to integrate mCherry into the genomic *srfA* gene cluster loci of specific strains. For analyzing the expression of *spoIVB*, *spoIVB* with its promoter from Sur-s0 or Sur-s1 was fused with mCherry and generated plasmid pHT-SpoIVBs0-NmC or pHT-SpoIVBs1-NmC. Similarly, *remA* with its promoter from Sur-s0 or Sur-s1 was fused with sfGFP and generated plasmids pHT-RemAs0-NmC and pHT-RemAs1-NmC.

For riboflavin measurement, 200 μl of the fermentation culture was added into 800 μl of 0.05 M NaOH. After centrifugation at 10 000 × *g* for 2 min, the supernatant was collected and diluted with 0.1 M acetate–sodium acetate buffer (pH 4.42). Subsequently, absorption at 444 nm was measured (33). The concentration of riboflavin was calculated using the following validated standard equation: $Y = (OD_{444}\text{-}0.0203) \times DF/0.0163$ ($R^2 = 0.9997$; OD_{444}, the absorbance value at 444 nm; *Y*, the concentration of riboflavin (mg/l); DF, dilution fold; OD_{444} was controlled within the range of 0.1–0.8 by dilution).

The *B. subtilis* cells sampled at 24 h were used for the determining activity of flavokinase encoded by *ribC*. The cells were collected by centrifugation (10 000 × *g* and 4°C for 2 min), and the pellet was washed with buffer A (0.1 M PBS (pH 7.5), 1 mM dithiothreitol, and 0.1 mM EDTA) (34). The cells were resuspended in buffer A and disrupted with ultrasonic oscillation (VCX750, Sonics, CT, USA). After centrifugation at 10 000 × *g* and 4°C for 10 min, an aliquot of the supernatant was directly used in the flavokinase assay, and the protein concentration of the supernatant was determined by the Bradford method. Flavokinase activity was measured in a final volume of 1 ml of 0.1 M PBS (pH 7.5) containing 50 mM riboflavin, 3 mM ATP, 15 mM $MgCl_2$, and 10 mM Na_2SO_3 (34). The mixture was preincubated at 37°C for 5 min, and the reaction was started by the addition of the enzyme. The reaction was performed at 37°C for 20 min and stopped by adding 25 uL of 1.5 M trichloroacetic acid. After centrifugation at 10 000 × *g* for 2 min, the supernatant was filtered through a 0.2 μm membrane. A 10 μl aliquot was injected into a high-performance liquid chromatography (HPLC) system (Agilent 1260, Santa Clara, CA, USA) equipped with a fluorescence detector (FLD; excitation, 450 nm; emission, 520 nm) and a C18 column (4.6 × 250 mm; 5 μm; Waters, Ireland). The mobile phases were 35% methanol and 65% ammonium acetate (5 mM). The total flow rate of the mobile phase was kept at 1.0 ml/min and the column temperature was kept at 25°C. One unit of flavokinase activity was defined as the amount of enzyme that produced 1nmol FMN from riboflavin and ATP per min under the assay conditions. The specific activities of different strains were compared by using the flavokinase activity and protein concentration measured above. For analyzing the expression of the *pur* operon, the plasmid pWLBE-gCpf1Ng-pur-NmC was constructed to integrate mCherry into the genomic *pur* operon loci of specific strains. Similarly, the plasmid pWLBE-gCpf1Ng-rib-NmC was constructed for analyzing the expression of the *rib* operon of specific strains, and the plasmid pWLBE-gCpf1Ng-zwf-NmC was built to fuse mCherry to the C-terminus of ZWF on the genome of specific strains.

Statistical analysis

All experiments were independently carried out at least three times unless otherwise specified, and the results were expressed as the mean value. Two-sided *t*-test in Excel (Microsoft 365) was used to perform the statistical evaluation (*P*-value), and $P > 0.05$, $P < 0.05$ and $P < 0.01$ were presented by no significance (n.s.), * and **, respectively. The graphs of results were plotted using the OriginPro software.

Results

Coupling cytosine or adenine deaminases with dCas12a

In previous studies, *Lachnospiraceae* bacterium dCas12a (dLbCas12a) was commonly used for CBE or ABE construction (20–24). Here, *Francisella novicida* dCas12a (dFnCas12a, D917A), which possesses a lower target requirement (PAM = TTV, V = A, G, C) than that of dLbCas12a PAM = TTTV), was chosen for the building of BE. Although FnCas12a has been reported to have lower activity compared with LbCas12a and AsCas12a (*Acidaminococcus* sp.) in human cells (35), it has been widely used in bacteria (36) and plants (37). To assess the compatibility between different cytosine or adenine deaminases and dFnCas12a in *B. subtilis*, we developed a suite of all-in-one plasmids. These plasmids were derived from a custom-made plasmid backbone pWLBE-dCas12a-N (Supplementary Figure S5A), in which dCas12a expression is regulated by the isopropyl β-D-1-thiogalactopyranoside (IPTG) induced promoter $P_{grac100}$, and a crRNA array implantation region is placed downstream of a constitutive promoter P_{veg} that with a precise transcriptional start site (commonly used for *B. subtilis* and also exhibits activity in *E. coli*) (27). The crRNA array implantation region contains a pair of head-to-head *Bsa*I cleavage sites flanked by two directed repeats (DRs), which allows for the rapid assem-

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

bly of the required crRNA array (28) (Supplementary Figure S5B). The crRNA array consists of an alternant DR and a spacer, in which the DR is specifically recognized by Cas12a and the spacer (also known as guide region) serves to guide Cas12a to the complementary target DNA strand. Additionally, a temperature-sensitive replication origin pE194ts is employed to enable the elimination of the plasmid after editing the genome. Subsequently, we introduced cytosine deaminase or adenine deaminase into the plasmid backbone for genome editing validation (Supplementary Figure S5C). Moreover, we also co-introduced a uracil glycosylase inhibitor (UGI), which can prevent the excision of the produced uracil (U) base by the DNA repair machinery(9), with cytosine deaminase to investigate whether it could enhance the efficiency of genome editing (Figure 2A). After transforming the plasmid containing the corresponding deaminases and crRNA array into *B. subtilis*, IPTG was added for induction, and the editing outcomes were analyzed through Sanger sequencing and BEAT software (29) (Supplementary Figure S2A).

Cytosine deaminases, including one from the family with standard editing window (rAPOBEC1) and four from the families with wider editing windows (hAPOBEC3A, PmCDA1, hAID*Δ and tCDA1EQ) (9,38), were codon optimized for *B. subtilis* and fused to the N-terminus of dCas12a, while UGI was fused to the C-terminus (Figure 2A). In the R-loop region generated by dCas12a, the crRNA forms a double-stranded structure with the target DNA strand through complementary base pairing, while the non-targeted single strand protospacer region can undergo C > T conversion aided by the fused cytosine deaminase (Figure 2A). To test the efficiency, the nonessential gene *aprE* was selected and a crRNA targeting this gene was designed. As shown in Figure 2B, we observed C > T conversion by fusing hAPOBEC3A, PmCDA1 or hAID*Δ to dCas12a, and that the incorporation of UGI enhanced editing efficiency. Among them, pWLBE-hAPOBEC3A-dCas12a-U (U means with UGI) possess the largest editing windows (at protospacer positions C8, C14, C15), and pWLBE-hAID*Δ-dCas12a-U exhibited the strongest editing efficiency (54% at position C8). To test whether dCas12a fused with cytosine deaminase and UGI retains RNase activity, enabling the maturation of the crRNA array and the recognition of multiple targets, therefore allowing multiplexed editing, 5 targets were selected on the genes *aprE* and *nprE*, and a corresponding five-member crRNA array an5C was constructed (Figure 2C and Supplementary Figure S6A). We transformed plasmids containing various cytosine deaminases and the crRNA array an5C into *B. subtilis* to explore the editing outcome. Among these, pWLBE-PmCDA1-dCas12a-U was limited to inducing C > T conversion in only 1 target on the genome, whereas pWLBE-hAPOBEC3A-dCas12a-U and pWLBE-hAID*Δ-dCas12a-U can act simultaneously in 4 and 5 targets, respectively (Figure 2D). Other six functional CDA-like (CDAL) deaminases CDAN1∼6, which have significant sequence divergence from classical AID/APOBEC family (39), were also codon optimized for *B. subtilis* and fused to the N-terminus of dCas12a-UGI to perform multiplexed editing (Supplementary Figure S6B). Among them, pWLBE-CDAN6-dCas12a-U with LjCDAL2_1 can perform C > T conversion in target 1 and target 2 simultaneously (Supplementary Figure S6C). To construct multiplexed ABEs, several evolved variants of adenine deaminase TadA (TadA7.10, TadA8.17, TadA8.20, TadA8e, and Tad9) (40), which can generate A > G conversion on the non-targeted single strand protospacer region (Supplementary Figure S6D), were fused to the N-terminus of dCas12a, respectively (Figure 2E). Among them, pWLBE-TadA8.20-dCas12a-N (N means without UGI), pWLBE-TadA8e-dCas12a-N, and pWLBE-TadA9-dCas12a-N can achieve A > G conversion in 2, 3 and 4 targets, respectively (Figure 2F). Furthermore, the recently reported TadA variants (T_{AD}AC-3.1 and TadA-dual) with both adenine and cytosine deaminase activities (41,42) were also incorporated into the dCas12a-based BE (Supplementary Figure S6E). However, C > T or A > G conversion was not observed in any of the five targets (Supplementary Figure S6F).

Construction and optimization of the dCas12a-based MultiduBE system

To construct the dual BE system MuitiduBE, the above screened three cytosine deaminases (hAPOBEC3A, hAID*Δ, and LjCDAL2_1) and one adenine deaminase (TadA9) were incorporated into the plasmid backbone pWLBE-dCas12a-U. Guided by the five-member crRNA array an5C, the editing outcomes were analyzed. As shown in Figure 3A, pWLBE-duBE-1a and pWLBE-duBE-2b can perform dual base editing (C > T and A > G) in 4 targets simultaneously, and the cytosine deaminases in duBE-1a and duBE-2b keep their sites preferences. Although pWLBE-duBE-3b and pWLBE-duBE-3c can also edit 4 targets, only A > G conversions were found. The reachable editing targets of other synthetic effectors were less than 4. Then, we replaced the IPTG-induced promoter $P_{grac100}$ in pWLBE serial plasmids with an anhydrotetracycline (aTC)-induced promoter P_{tet}, which has low leakage and high dynamic range in both *E. coli* and *B. subtilis* (Supplementary Figure S7). As shown in Figure 3B, the reachable editing targets were improved from 4 to 5 for pWLT-TadA9-dCas12a-N and pWLT-duBE-1a. Moreover, the total conversion rate for each target of pWLT-hAID*Δ-dCas12a-U and pWLT-TadA9-dCas12a-N were both over 6%. We also analyzed the expression profiles of duBE controlled by $P_{grac100}$ and P_{tet}. As shown in Supplementary Figure S8, an increase in expression level was observed after replacing $P_{grac100}$ with P_{tet}, and both of the systems will be turned off after removing the inducer.

In natural crRNA arrays, there is a 16- to 18-nt short separator in front of the DR (Supplementary Figure S9A). The separator is nonessential for Cas12a targeting, and therefore we have omitted it, only including DR and spacer in the crRNA array, as it has been done in previous studies (Supplementary Figure S9B) (43). However, as shown in Figure 3B, the editing activities of MultiduBE on target 5 were very low (2% for pWLT-duBE-1a and 0% for pWLT-duBE-2b). This may result from interference between adjacent crRNAs after removing the separator, as suggested in another study on CRISPR activation (CRISPRa) (44). The addition of a very short 4-nt synthetic separator to isolate interference in the crRNA array has been shown to enhance the multiplex activity of dCas12a-based CRISPRa in human cells (Supplementary Figure S9C) (44). Therefore, we hypothesized that introducing a synthetic separator to reduce interference between adjacent crRNAs could increase the multiplex mutation rate mediated by the dCas12a-based duBE in bacterial cells (Supplementary Figure S9D). For matching the two-round PCR based crRNA array construction method (Supplementary Figure S1), five synthetic separators (Sp1∼Sp5), with Tm values ranging from

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

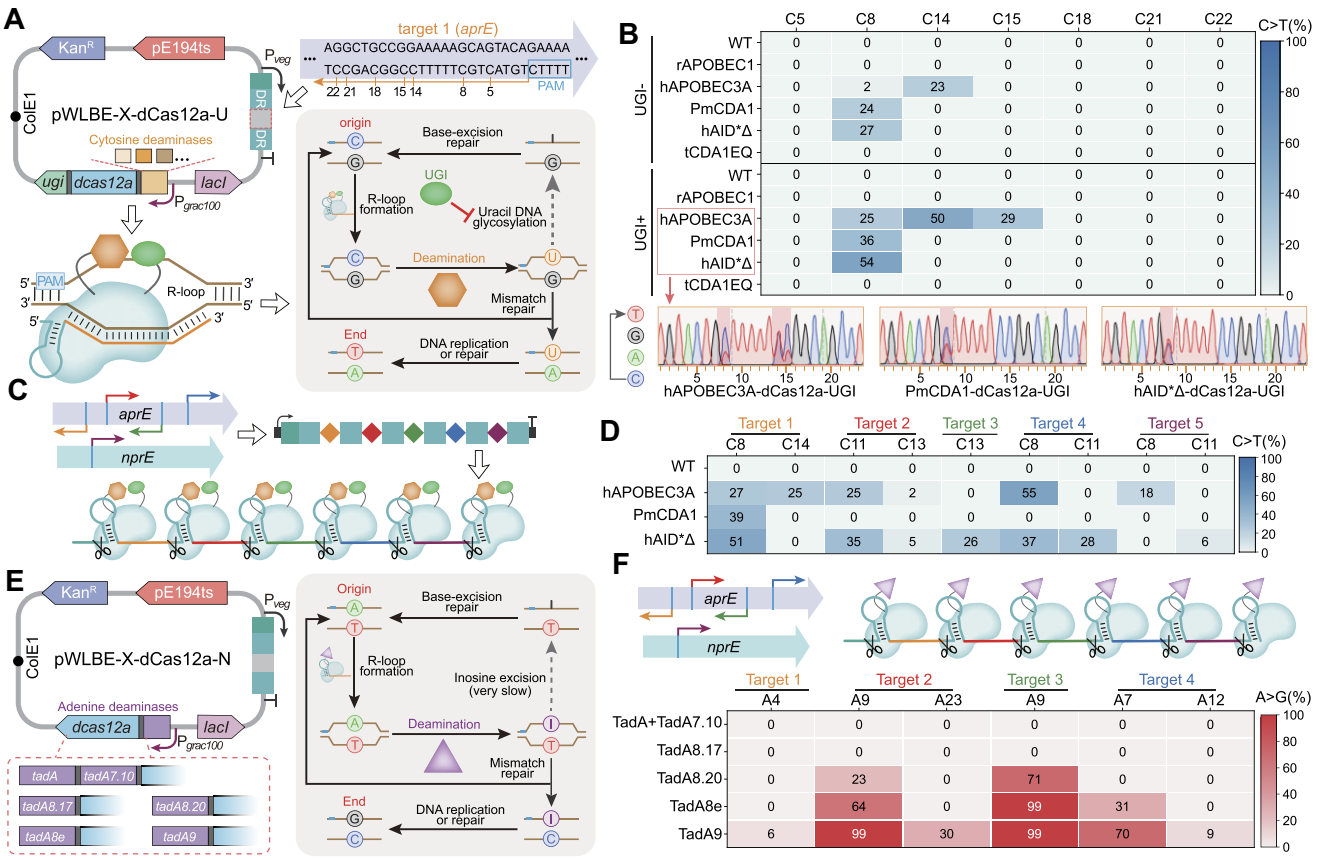


Figure 2. Integration of cytosine or adenine deaminases with dCas12a. (**A**) Schematic diagram for verifying compatibility between various cytosine deaminases and dCas12a. (**B**) Editing outcomes of the dCas12a-based cytosine base editors (CBEs) on target gene *aprE*. (**C**) A five-member crRNA array an5C target five sites on genes *aprE* and *nprE* was constructed to guide dCas12a-based cytosine base editors (CBEs) for multiplexed editing. (**D**) Editing outcomes of the dCas12a-based CBEs on target genes *aprE* and *nprE* guided by an5C. (**E**) Schematic diagram for verifying compatibility between evolved variants of adenine deaminase TadA and dCas12a. (**F**) Editing outcomes of the dCas12a-based adenine base editors (ABEs) in target genes *aprE* and *nprE* guided by an5C. Data are presented as mean values from three independent biological replicates ($n = 3$).

55 to 60°C, were generated by modifying the natural FnCas12a separator (FnSp) or assembling separators from different hosts (Figure 3C and Supplementary Figure S10A). In addition, the DR region in the array has also been appropriately extended from 19-nt to 27-nt at the 5′-terminus (Figure 3C and Supplementary Figure S10B) to improve the Tm value to 55°C, which does not affect the formation of the crRNA stem-loop.

For pWLT-duBE-1a, adding Sp1, Sp2, Sp3 and Sp4 into the array improved the conversion rates at position C8 in target 5 from 2% to 43, 24, 38 and 43%, respectively (Figure 3D). As to pWLT-duBE-2b, adding Sp1, Sp3, Sp4 and Sp5 into the array improved the conversion rates at position C11 in target 5 from 0% to 9, 5, 7 and 6%, respectively (Figure 3D). After comprehensively comparing the improvement of editing efficiencies in other targets, Sp4 was chosen as the insulator for the crRNA array of MultiduBE. Subsequently, a crRNA array an5cN-Sp4 containing spacers of varying lengths (14- to 26-nt) was also constructed (Figure 3E). We found that a crRNA larger than 17-nt was enough to guide the MultiduBE, although the editing efficiency guided by an5cN-Sp4 was lower than that guided by an5c-Sp4 (Figure 3D). As shown in Supplementary Figure S10C, we also built crRNA arrays an5C-Sp4t (containing Sp4 and the original 19-nt DR), an5C-SpN (containing natural FnSp and the original 19-nt DR), and an5C-SpQ (containing the short 4-nt Sp and the original 19-nt DR). Both duBE-1a and duBE-2b guided by these arrays could perform editing in the five targets, but the editing efficiencies are all lower than those guided by an5C-Sp4 (Supplementary Figure S10D).

We then conducted next generation sequencing (NGS) analysis of the mature RNA generated by an5C and an5C-Sp4, and found that the introduction of Sp4 indeed promoted the processing of the crRNA array, resulting in the mature crRNAs containing only DR and spacer (Supplementary Figure S10E). This effect was particularly noticeable for the last crRNA, which exhibited poor maturation in cn5C. As shown in Supplementary Figure S11A, by changing the order of crRNAs in an5C and an5C-Sp4, crRNA arrays an5C-R1, an5C-R2, an5C-R3, an5C-R1-Sp4, an5C-R2-Sp4, an5C-R3-Sp4 have been constructed. In addition, arrays with individual crRNA were also built. As expected, the highest base conversion rates were observed in all the targets guided by the individual crRNA. In each of the crRNA arrays, placing the crRNA in the first position could maintain its maximal activity even without Sp4, while placing it in a relatively later position required the assistance of Sp4 to maintain higher editing efficiency (Supplementary Figure S11B).

Whole genome resequencing was also performed after treating BSZRG (G00, *B. subtilis* 168Δ*epr*::XylR-P$_{xylA}$-*comKS*, Δ*trpC2*::*trpC0*, Δ*gudB*::*gudB*$^{+}$) strain (27) with MultiduBE. As shown in Supplementary Table S1, only one low-

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

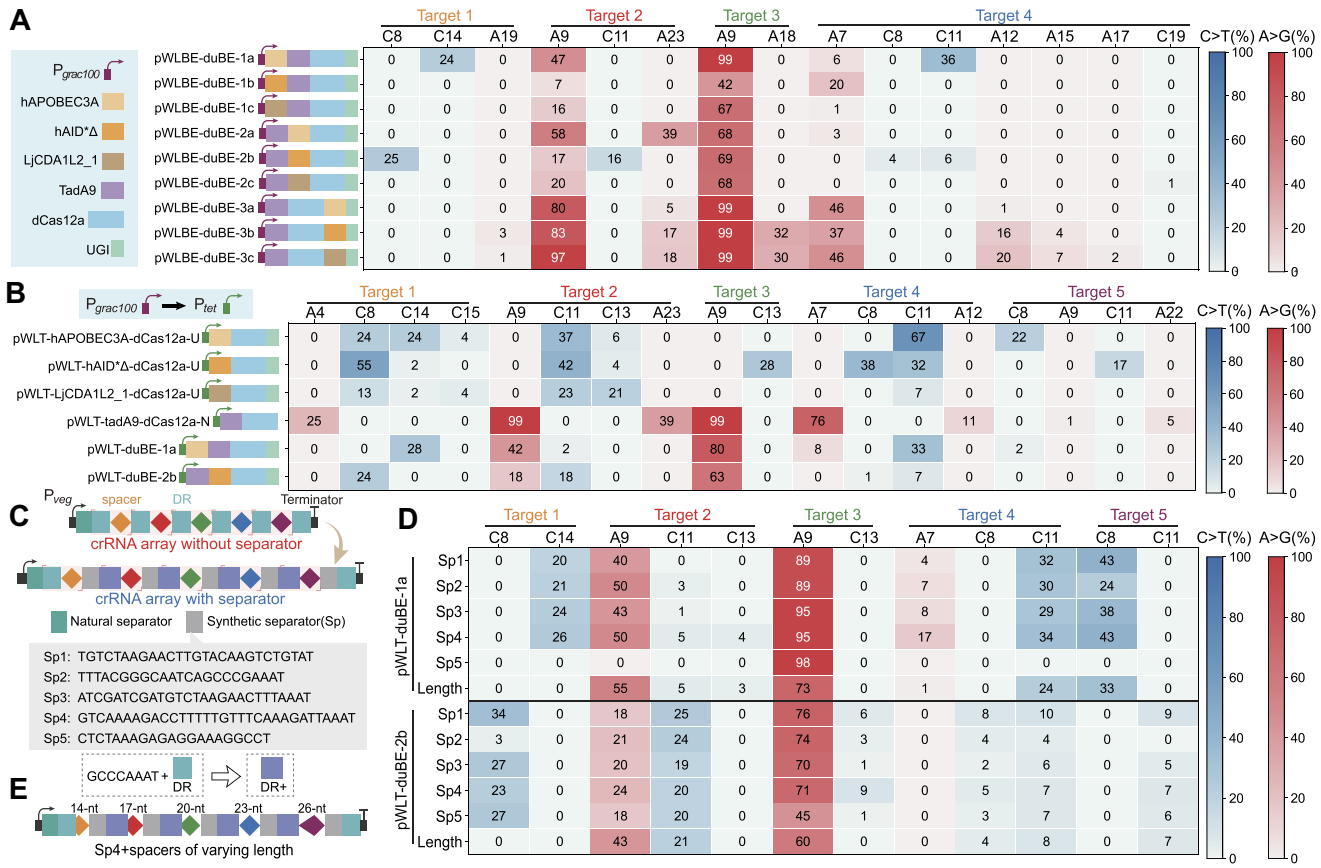


Figure 3. Construction and optimization of the dCas12a-based MultiduBE. (**A**) Composition and structure optimization of the dCas12a-based MultiduBE. (**B**) Replacing promoter of the dCas12a-based synthetic effector in pWLBE serial plasmids generating the pWLT serial plasmids. (**C**) Engineering the crRNA array for improving the editing efficiency. (**D**) Editing outcomes on target genes *aprE* and *nprE* guided by the modified five-member crRNA arrays. (**E**) A crRNA array an5cN-Sp4 containing spacers of varying lengths (14- to 26-nt) was designed and assembled into pWLT-duBE-1a and pWLT-duBE-2b from multiplexed editing. Data are presented as mean values from three independent biological replicates ($n = 3$).

frequency C > T or A > G conversion was found out of the target genes for pWLT-duBE-1a-an5C, pWLT-duBE-1a-an5C-Sp4, and pWLT-duBE-2b-an5C-Sp4, and the mutation found in variational genomic position indicates that the crRNA-independent off-target occurred with the overexpression of cytosine and adenine deaminases. Moreover, the higher mutation frequencies in *aprE* and *nprE* still indicate that MultiduBE has a high specific targeting activity. Editing in potential off-target sites (POTs) predicted by the Cas-OFFinder algorithm (30) was also analyzed by tNGS (>500 000 reads were used), and no base conversion was observed (Supplementary Table S2). As shown in Supplementary Figure S12, transcriptome-wide RNA-seq was also employed, and off-target RNA editing, which was reported in human cells before with the nCas9-based BEs (45), was not observed here.

Multiplexed *in-situ* mutagenesis mediated by MultiduBE

To evaluate the multiplexed *in-situ* mutagenesis ability of MultiduBE, we further analyzed the detailed editing outcomes of pWLT-duBE-1a and pWLT-duBE-2b using high-throughput targeted amplicon sequencing (Supplementary Figure S2B). More than 500000 reads were used for the analysis of mutagenesis in each target, and the top 10 alleles guided by an5C or an5C-Sp4 were compared (Supplementary Table S3). As shown in Figure 4, the composition of mutations produced by pWLT-duBE-1a and pWLT-duBE-2b were different. The introduction of Sp4 slightly decreased the editing efficiency in target 1 for both duBE-1a and duBE-2b, which may be a result of the increased competition with other mature crRNAs that reduced the relative proportion of dCas12a bound to the first crRNA. For target 5, adding Sp4 could significantly improve the editing efficiencies for both MultiduBEs, which was consistent with the previous results analyzed by Sanger sequencing (Figure 3D). For example, the C8T (C > T conversion at position 8 of the protospacer) conversion in target 5 produced by pWLBE-duBE-1a was improved from 11.05 to 40.41%, and the C8T + C11T conversion in target 5 produced by pWLBE-duBE-2b was improved from 1.50 to 5.28%.

As shown in Figure 4A, when pWLT-duBE-1a was used for multiplexed editing, the conversion rate of A > G in target 1 was very low (0.87% for an5C and 0.86% for an5C-Sp4). In target 1, 8 out of the 9 mutation alleles only possess C > T conversion, and no combinatorial C > T and A > G mutation was found. In target 2 to target 5, the combinatorial C > T and A > G mutations were observed. In addition, the bystander editing out-of-protospacer (BEOP) was also found in targets 1, 3 and 5 with a lower frequency (less than 2%). For example, a C > T conversion occurred in the complementary DNA strand at position 40 of the protospacer (C40T*) in target 1 (0.31% for an5C and 0.32% an5C-Sp4). As shown in Figure 4B, when pWLT-duBE-2b was used for multiplexed

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

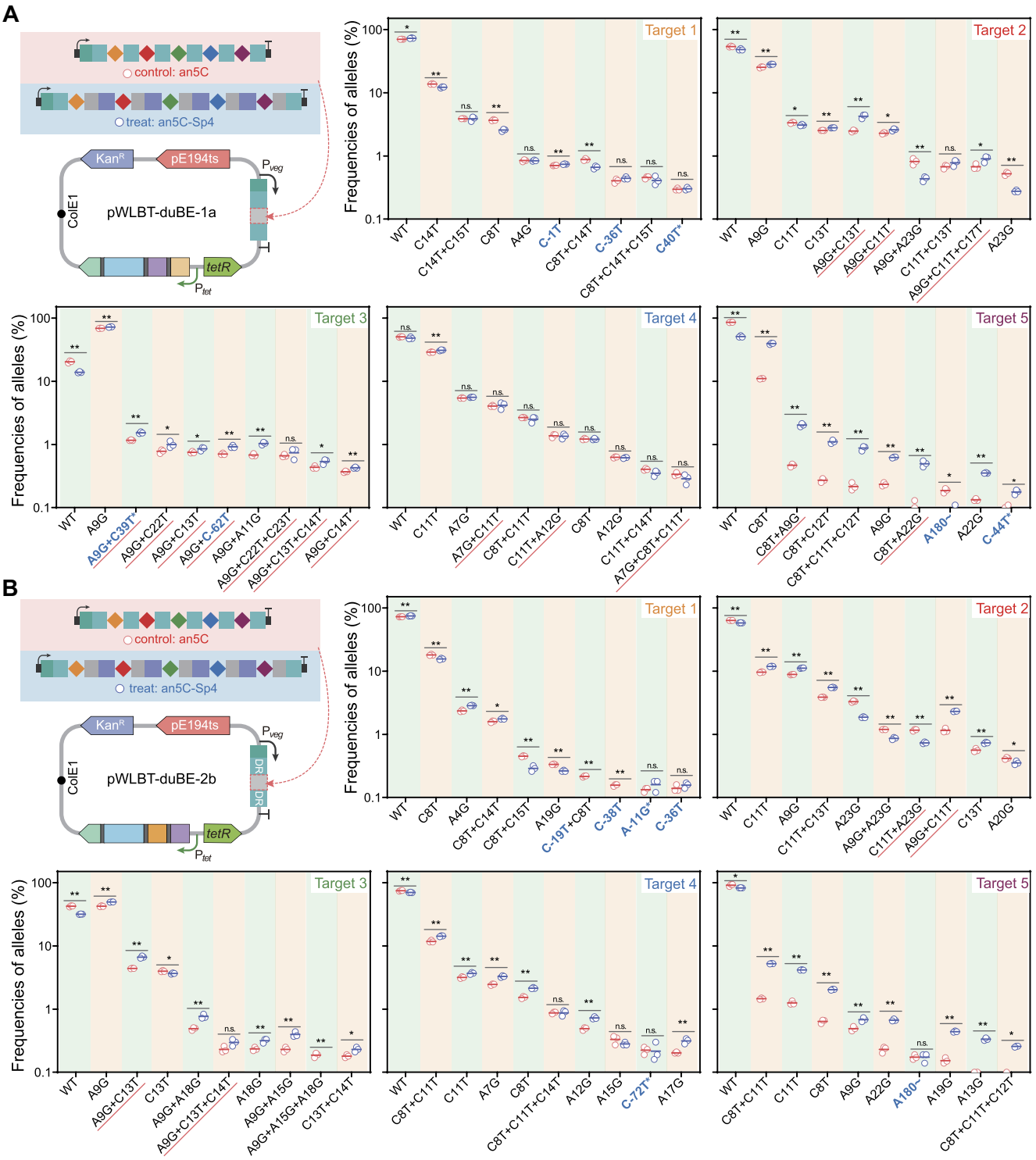


Figure 4. Detailed editing outcomes of the MultiduBE-promoted multiplexed *in-situ* mutagenesis. (**A**) High-throughput targeted amplicon sequencing (tNGS) analysis for pWLT-duBE-1a guided by an5C or an5C-Sp4. (**B**) The tNGS analysis for pWLT-duBE-2b guided by an5C or an5C-Sp4. The top 10 alleles produced from an5C and an5C-Sp4 in each target were selected and compared. The asterisk means C > T or A > G conversion occurred in the complementary DNA strand of protospacer. The alleles marked in blue mean the bystander editing out-of-protospacer (BEOP), and those with red underline mean combinatorial C > T and A > G mutants. The short lines represent mean values from three independent biological replicates ($n = 3$), and the circles represent individual data points. More than 500 000 reads were used for each sample. See Supplementary Table S3 for details including the reads of NGS.

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

editing, both the conversion rate and amount of A > G in target 1 were higher than that of pWLT-duBE-1a, but the combinatorial C > T and A > G mutation was also not found in this target. Furthermore, the combinatorial C > T and A > G mutations were only found in target 2 and target 3, and a low frequency (<0.5%) BEOP was found in target 1, target 4, and target 5.

In addition to the population-level analysis described above, mutations in the targeting genes *aprE* and *nprE* treated by two MultiduBEs (duBE-1a and duBE-2b) with an5C-Sp4 were also analyzed at the single-colony level (Supplementary Figure S13A). Among the 8 randomly chosen single colonies, 7 different mutant strains and 1 wild type strain were found both for pWLT-duBE-1a-an5C-Sp4 and pWLT-duBE-2b-an5C-Sp4 (Supplementary Figure S13B, C). In addition, duBE-1a generated a strain with combinatorial mutations in all 5 targets (Supplementary Figure S13B). For duBE-2b, although mutations at each target were detected, in the same colony, mutations occurred at most in only 3 targets. This aligns with the theory that the probability of combinatorial mutations occurring in the same cell becomes exceedingly low after the superposition of probabilities. These results demonstrate the efficiency and diversity of MultiduBE for multiplexed *in-situ* mutagenesis. To prove the universality of MultiduBEs we also tested it in *E. coli*, and found that pWLT-duBE-1a enables multi-target dual-function editing (C > T and A > G), and the simultaneous mutagenesis in 5 targets can be achieved after introducing Sp4 into the crRNA array xy5C (Supplementary Figure S14).

Multiplexed *in-situ* mutagenesis for cell physiology reprogramming

As demonstrated in Figure 5A, the generic application of the MultiduBE in phenotypic diversification was verified by constructing a five-member crRNA array YKTmf5C-Sp4 targeting three fluorescent proteins (sYFP2, mKate and mTagBFP2) and two cell morphology determining proteins (FtsZ and MreB) (46). The crRNAs were designed to target the 5′-untranslated region (UTR) or N-terminal coding sequence (NCS), resulting in an expression change for these genes (Figure 5B). As depicted in Supplementary Figure S15A, two-dimensional representations of flow cytometry revealed a more heterogeneous population, particularly in the expression of mKate and mTagBFP2, after treatment with MultiduBE. Additionally, confocal images were acquired to visualize the phenotypic changes in the cells (Supplementary Figure S15B). As anticipated, MultiduBE treatment produced various combinations of colors and shapes without selective pressure (Figure 5C).

Alterations in a single nucleotide of the bacteria genome can have an extensive impact on cell phenotype (1). For instance, point mutations in specific genes are known to be linked to antibiotic resistance (47). Hence, we opted to validate the multiplexed *in-situ* mutagenesis activity of MultiduBE by targeting endogenous genes related to antibiotic resistance. Previous studies in *B. subtilis* have demonstrated that rifampicin, spectinomycin, streptomycin, and tetracycline resistance can be produced through mutations of the *rpoB* gene (encoding RNA polymerase beta subunit, H482R/Y) (48), *rpsE* gene (encoding ribosomal protein S5, K23E) (1), *mthA* gene (encoding methylthioadenine/S-adenosylhomocysteine nucleosidase, nonsense mutation) (49), and the promoter region of *tetL* gene (encoding tetracycline efflux protein leader peptide) (50), respectively. In addition, investigations in other hosts also indicated potential mutations for conferring spectinomycin and streptomycin resistance in *B. subtilis* (Supplementary Figure S16A). The ribosomal protein S5 (encoded by *rpsE*) from *Neisseria gonorrhoeae* (51,52) and *Streptomyces filamentosus* (53) were aligned with RpsE of *B. subtilis* to assign potential targets for conferring spectinomycin resistance (V21, A22, K26, and R30), and ribosomal protein S12 (encoded by *rpsL*) from *Borrelia burgdorferi* (54), *Mycobacterium tuberculosis* (55), and *Streptomyces coelicolor* (56) were aligned with RpsL of *B. subtilis* to assign potential targets for conferring streptomycin resistance (K55 and K101). We then designed the corresponding crRNAs (Figure 5D).

We first constructed a two-member crRNA array targeting P_{tetL} and *rpoB* to investigate whether MultiduBE could generate resistances to tetracycline and rifampicin. Furthermore, we selected 4 different crRNAs targeting the potential mutation sites in *rpsE* and combined them with the crRNA targeting P_{tetL} to form a minipool library containing 4 two-member crRNA arrays, which was used to generate resistances to tetracycline and spectinomycin. We also designed 2 crRNAs in *rpsL* and 1 crRNA in *mthA* and combined them with the crRNA targeting P_{tetL} to form a minipool library including 3 two-member crRNA arrays, which was used to generate resistances to tetracycline and streptomycin. As the mutagenesis spectrums of pWLT-duBE-1a and pWLT-duBE-2b were different, we simultaneously ligated the above two-member crRNA arrays into a mixture of pWLT-duBE-1a and pWLT-duBE-2b. The generated minipool libraries were transformed into the *B. subtilis* strain G00 for performing multiplexed *in-situ* mutagenesis, and overlapping sequencing peaks were found for the crRNA array in strains with the minipool library before selection by corresponding antibiotics (Figure 5E). After the antibiotic screening, the crRNA arrays that confer the appropriate resistance were enriched, as indicated by the pure sequencing peaks. Accordingly, the mutations occurred in the genome were consistent with those enriched crRNA arrays (Figure 5E). Subsequently, sequencing analysis was performed by selecting three single colonies in the three corresponding two-resistance agar plates. The three colonies with tetracycline-rifampicin resistance or tetracycline-spectinomycin resistance had the same mutations, while the three colonies with tetracycline-streptomycin resistance showed different mutations (Figure 5E and Supplementary Figure S16B). Moreover, we have identified a novel RpsL mutation, R99C, conferring streptomycin resistance in *B. subtilis*. This mutation is distinct from the naturally occurring K88R/E/Q/T mutations found in *B. burgdorferi*, *M. tuberculosis*, and *S. coelicolor* (corresponding to K101 in *B. subtilis*) or the K43R mutation in *M. tuberculosis* (corresponding to K55 in *B. subtilis*).

Based on the crRNA arrays enriched during the screening process, we designed a four-member crRNA array LBEL4C-Sp4 to confer resistance to the four types of antibiotics used above (Figure 5F). After conducting multiplexed *in-situ* mutagenesis with the guidance of LBEL4C-Sp4, one tetracycline-rifampicin-spectinomycin-streptomycin resistant strain, two tetracycline-rifampicin-streptomycin resistant strains, and three tetracycline-spectinomycin-streptomycin resistant strains were obtained (Figure 5G and Supplementary Figure S16C). These antibiotic-resistance mutants can be used for the development and research of new

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

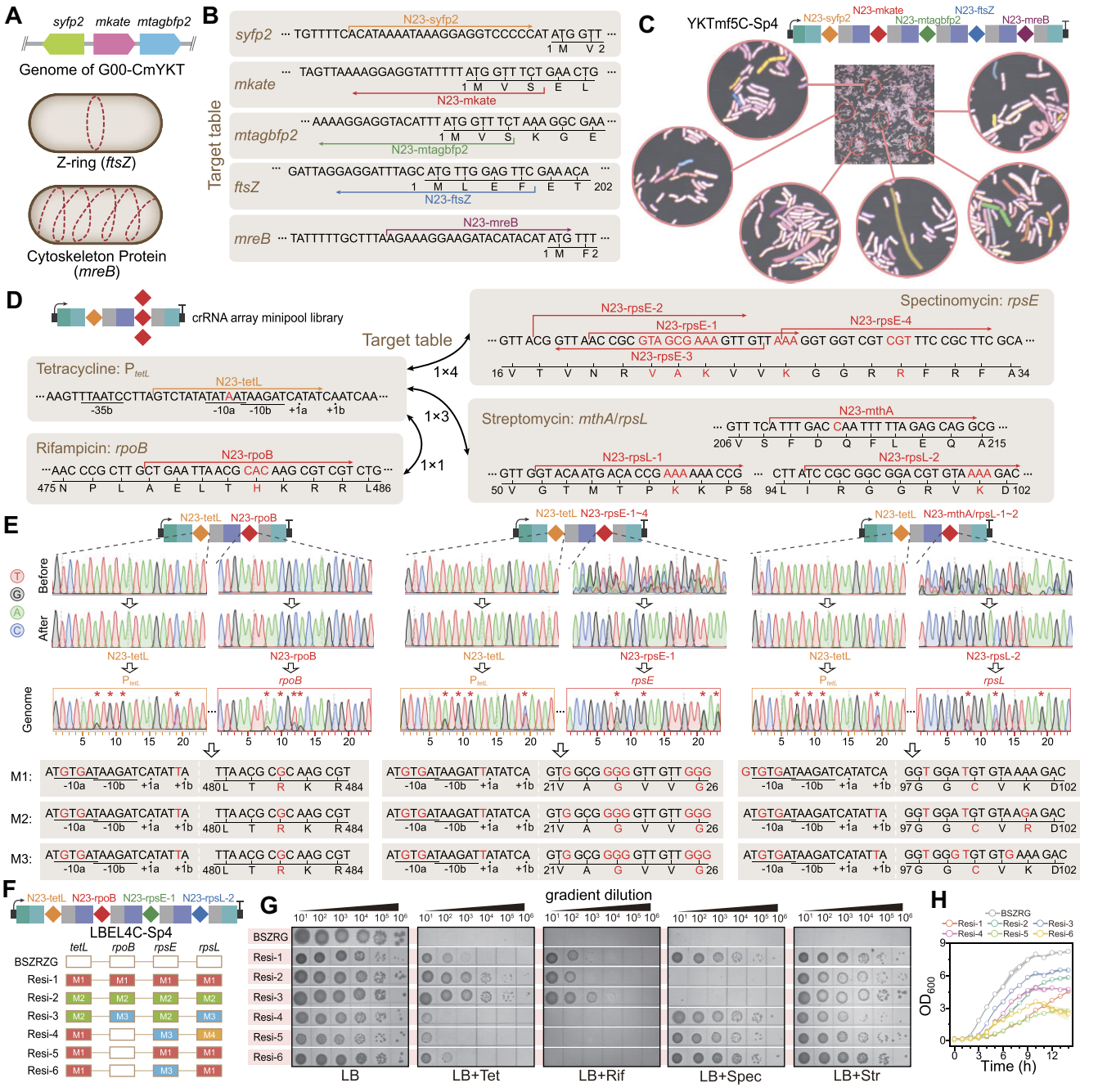


Figure 5. Multiplexed *in-situ* mutagenesis for cell physiology reprogramming. (**A**) To validate MultiduBE in generic phenotypic diversification, strain G00-CmYKT was constructed by integrating three fluorescent proteins (sYFP2, mKate, and mtagBFP2) into the genome of strain G00. For cellular morphology reprogramming, FtsZ (tubulin to form Z-ring at the midcell) and MreB (cytoskeleton protein) were selected as the targets for MultiduBE. (**B**) The target table containing designed crRNAs for verifying MultiduBE in generic phenotypic diversification. (**C**) The five-member crRNA array YKTmc5C-Sp4 guided MultiduBE in generic phenotypic diversification. After treatment by MultiduBE, confocal images of strain G00-CmYKT were obtained by merging the brightfield, green (500–550 nm), red (570–616 nm) and blue (429–474 nm) channels. (**D**) The target table containing designed crRNAs for antibiotic (tetracycline, rifampicin, spectinomycin, or streptomycin) resistant mutant generation. Nucleotides marked in red indicate potential mutation sites based on previous studies. (**E**) Sequencing analysis of the crRNA arrays and the genomic targets. (**F**) Based on the crRNA arrays enriched during the screening process, a four-member crRNA array LBEL4C-Sp4 was designed to confer resistance to the four types of antibiotics used. M1–M3 represent the corresponding mutations for each gene, and details can be found in Supplementary Figure S16C. (**G**) Approximately 10^8 cells were diluted with distilled water, and 5 μl aliquots of the cell suspension were spotted onto the agar plate with specific antibiotics. (**H**) Growth curves of the wild and mutated strains. Data are presented as mean values ± SD from three independent biological replicates (n = 3). Lines indicate the mean, and shaded areas represent the SD.

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

antibiotics. Moreover, the fitness cost of resistance, which would allow susceptible bacteria to outcompete resistant bacteria if the selective pressure from antibiotics is reduced (57), was determined from the reduced bacterial growth rate (Figure 5H).

Multiplexed *in-situ* mutagenesis for metabolic regulation

Surfactin is a lipopeptide biosurfactant synthesized by a non-ribosomal peptide synthetase (NRPS) in *Bacillus* species, and it is widely used in the fields of biopharmaceuticals, environmental remediation, oilfield exploitation, cosmetics, and daily necessities (58). As shown in Figure 6A, we selected five genes as the targets for surfactin synthesis overproduction based on previous studies (58–60), including *srfAA* (encoding surfactin synthetase, a critical speed limiting step), *remA* (encoding the regulatory protein of biofilm formation-related genes, associated with cell fitness), *spoIVB* (encoding serine protease, related to spore formation), *lcfB* (encoding long-chain fatty-acid-CoA ligase, providing precursor fatty acyl-CoA), and *ilvB* (encoding acetolactate synthase large subunit, providing precursor branched-chain amino acids). The targets for crRNAs include the negative regulation sites in the promoters of *srfAA*, *lcfB*, and *ilvB*; the 5′- UTR and NCS of *remA*; as well as the core region between -10 and -35 in the promoter of *spoIVB* (Figure 6A).

The 4′-phosphopantetheinyl transferase for activating surfactin synthetase encoded by *sfp* is inactive in *B. subtilis* 168 (59), so we first performed a reverse mutation on the *sfp* gene in strain G600 (*B. subtilis* 168Δ*epr*::XylR-P_{xylA}-*comKS,* Δ*trpC2*::*trpC0,* Δ*gudB*::*gudB*$^{+}$, *aprE*^{0}, *nprE*^{0}, *bpr*^{0}, *mpr*^{0}, *nprB*^{0}) (27), generating the strain Sur-s0. Then, a five-member crRNA array sur5C-Sp4 was constructed and used for the multiplexed *in-situ* mutagenesis in the genes related to surfactin synthesis in the strain Sur-s0 (Figure 6B). Based on a colorimetric method (31), 12 mutant strains (Sur-s1∼s12) were screened (Figure 6B), and the mutations in the target genes were analyzed after eliminating the MultiduBE plasmid (Supplementary Table S4). Although editing occurred at all five targets guided by sur5C-Sp4 (Supplementary Figure S17A), the maximum number of mutations obtained involved only three genes. Finally, strains G600, Sur-s0, Sur-s1, Sur-s11, and Sur-s12 were selected for shake-flask production of surfactin. As shown in Figure 6C, in comparison to the starting strain Sur-s0 that could synthesize 702.7 mg/l surfactin, the surfactin titer of the strain Sur-s1 (with mutations in three other genes *srfAA*, *remA* and *spoIVB*) improved by 42.0% to 997.5 mg/l. Furthermore, the changes in the expression of these three mutated genes in Sur-s1 were also consistent with the improvement of surfactin production (Supplementary Figure S17B). The growths of these five strains were similar except for Sur-s11, which showed impaired growth and reduced surfactin titer (Figure 6C). Given that the probability of occurrence decreases with the combination of mutations, more efficient screening methods are required for combinatorial mutations involving a greater number of genes in this situation.

Riboflavin (also known as vitamin B2) is a heat-resistant, water-soluble vitamin that is essential for humans as it is a direct precursor for the coenzymes flavin adenine (FAD) and flavin mononucleotide (FMN), which are crucial for cellular processes (61). As shown in Figure 6D, the synthesis of riboflavin in *B. subtilis* is subjected negative feedback regulation with GTP and FMN by the guanine and FMN riboswitches, respectively (61). In addition, riboflavin kinase, encoded by the essential gene *ribC*, can convert riboflavin to FMN, which not only consumes riboflavin but also further strengthens the feedback inhibition of the FMN riboswitch (61). Moreover, glucose-6-phosphate dehydrogenase, encoded by *zwf*, is a key step in the pentose phosphate pathway, which is also known to affect riboflavin synthesis (61). Therefore, we selected the guanine riboswitch in front of *purE*, the FMN riboswitch in front of ribD, *ribC*, and *zwf* as mutagenic targets for enhancing riboflavin synthesis in the wild strain G600. As shown in Figure 6E, crRNAs were designed to target the key points of the two riboswitches and RibC (marked in red), as well as the 5′-UTR and NCS of Zwf. Moreover, the crRNA targeting P_{tet} for tetracycline resistance described earlier was set to observe the induced base editing without corresponding antibiotic selection pressure.

As shown in Figure 6F, a five-member crRNA array rib5C-Sp4 was constructed in the MultiduBE plasmids pWLT-duBE-1a and pWLT-duBE-2b to promote the multiplexed *in-situ* mutagenesis for the emergence of riboflavin overproducer from strain G600 that do not overproduce riboflavin. After being rewashed with medium without aTC, fluorescence-activated droplet sorting (FADS) was conducted to screen and select the mutants with high fluorescence, which correlates with a high extracellular riboflavin titer in the droplets (Figure 6F and Supplementary Figure S17C). The cells in the collected droplets were spread on agar plates containing kanamycin, and single colonies were picked and screened by fluorescence measurement in 96-well plates (Figure 6G). A total of 14 mutated strains (Rib-s1∼s14) with high fluorescence were obtained, and the mutations in the target genes were analyzed after eliminating the MultiduBE plasmid (Supplementary Table S5). Although tetracycline was not used for selection, all 14 strains had mutations in the P_{tet} site. The mutations observed in the selected strains were more diverse than those in the surfactin synthesis regulation, with all five targets in strain Rib-S12 undergoing mutations. This highlights the importance of an efficient screening strategy for obtaining multi-site combinatorial mutants with low probabilities. Subsequently, five mutant strains (Rib-s4, Rib-s8, Rib-s11, Rib-s12 and Rib-s14) were selected for shake-flask culture, and a control strain Rib-s0 was constructed by replacing P_{ribD} and the FMN riboswitch with a strong constitutive promoter P_{veg} (Figure 6H). All five mutated strains could produce more riboflavin than that of the control strain Rib-s0 (21.5 mg/l), and the riboflavin titer of Rib-s12 (344.8 mg/l) was 16-fold that of the control (Figure 6H). Moreover, the increased expression of the synthetic pathways (encoded by the *pur* operon, the *rib* operon, and *zwf*) and the reduced enzyme activity of the riboflavin consumption pathway (encoded by *ribC*) in Rib-s12 correspond with the riboflavin overproduction phenotype (Supplementary Figure S17D).

Discussion

The ability to generate *in-situ* mutagenesis in multiple genomic loci combinatorially is a powerful way for the study of complex biological functions. In this study, we combined the multiplexed capabilities of dCas12a with the activities of two deaminases to create MultiduBE, which can simultaneously mutate C and A (or G and T in the opposite strand)

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

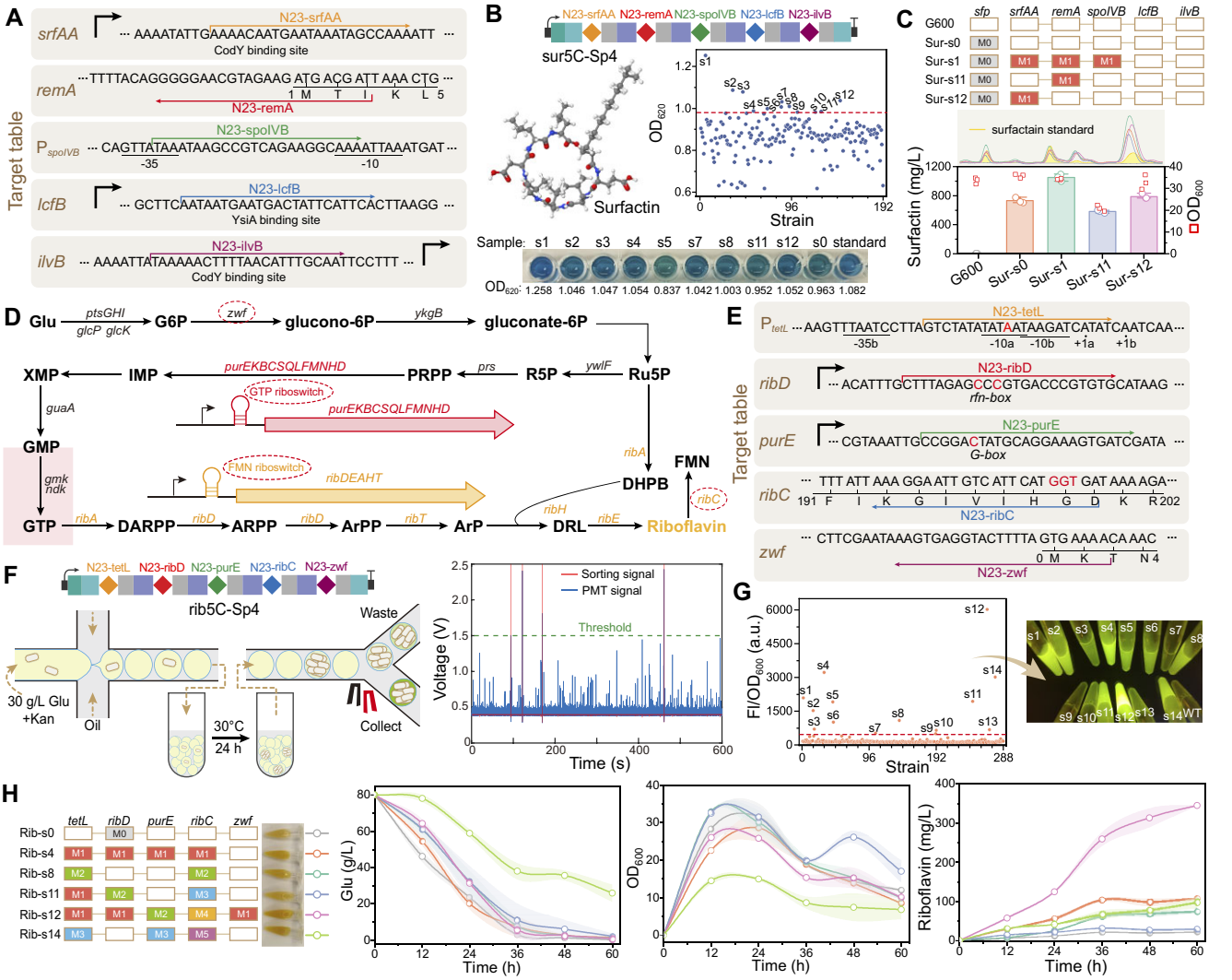


Figure 6. Multiplexed *in-situ* mutagenesis for metabolic regulation. (**A**) The target table containing designed crRNAs for surfactin synthesis regulation. (**B**) A five-member crRNA array sur5C-Sp4 was constructed and utilized for multiplexed *in-situ* mutagenesis on genes related to surfactin synthesis. Using a colorimetric method, 12 mutated strains (Sur-s1∼s12) were screened. (**C**) Strains G600, Sur-s0, Sur-s1, Sur-s11 and Sur-s12 were selected for shake-flask production of surfactin. M1 represents the corresponding mutation for each gene, and details can be found in Supplementary Table S4. Data are presented as mean values ± SD from three independent biological replicates ($n = 3$). The circles represent individual data points of surfactin titer, and the squares represent individual data points of OD_{600}. The chromatogram of the surfactin standard (including four surfactin isoforms) is shown in yellow with a light shadow, while the surfactin chromatograms of selected strains are shown in the colors corresponding to the circles. (**D**) The synthetic pathway and regulation manner of riboflavin in *B. subtilis*. (**E**) The target table containing designed crRNAs for the emergence of the riboflavin overproducer. Nucleotides marked in red indicate potential mutation sites based on previous studies. (**F**) A five-member crRNA array rib5C-Sp4 was constructed to facilitate multiplexed *in-situ* mutagenesis for the emergence of riboflavin overproducer from strain G600. Fluorescence-activated droplet sorting (FADS) was conducted to screen the mutations overproducing riboflavin. (**G**) The isolated single colonies were picked and screened by fluorescence measurement in 96-well plates. (**H**) Five mutated strains (Rib-s4, Rib-s8, Rib-s11, Rib-s12 and Rib-s14) were selected for shake-flask culture. The strain Rib-s0, constructed by replacing P_{ribD} and FMN riboswitch with a strong constitutive promoter P_{veg}, was used as a control. M1–M5 represent different mutations for each gene, and details can be found in Supplementary Table S5. Data are presented as mean values ± SD from three independent biological replicates ($n = 3$). Lines indicate the mean and shaded areas represent the SD.

bases at multiple targets (Figure 1). MultiduBE, encoded on a single plasmid, allows for multiplexed *in-situ* mutagenesis, relying only on trans-action guided by specific crRNA arrays and therefore it does not require further genomic manipulation. Despite the possibility of off-targeting or unexpected mutations (Supplementary Table S1), the successful generation of mutant strains exhibiting desired properties (Figures 5, 6) confirms the efficacy of MultiduBE in directed mutagenesis. In comparison to other transcriptional regulatory processes, such as CRISPR interference and activation (CRISPRi/a) (8,19), MultiduBE stands out by inducing mutations at the genome level, providing greater stability and robustness to the resulting modifications. This genome-level modification can also serve as the foundation for subsequent dynamic regulation systems based on CRISPRi/a. In contrast to the previously used nCas9-based CBE or dual-function BE integrated into the genome (62,63), our method employs an all-in-one plasmid with general applicability in both *E. coli* and *B. subtilis* (Figure 3 and Supplementary Figure S14), which ensures better genetic stability by eliminating the MultiduBE plasmid after mutagenesis. For the initial construction and optimization of MultiduBE, we employed a rapid and

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

cost-effective analysis process using Sanger sequencing. However, Sanger sequencing cannot detect low-frequency mutations and reveal detailed components of the mutations. Therefore, we also used NGS for a more detailed mutation analysis, which further revealed the diversity of *in-situ* mutations generated by MultiduBE (Figure 4). Although we observed low-frequency BEOP, these mutations were still beneficial for diversity, as we observed in *ribC* for the overproduction of riboflavin (Supplementary Table S5).

In previous BE works based on Cas9, the nickase nCas9 (D10A) is often employed as it improves editing efficiency by an average of threefold (64). Owning to the nick produced in the non-edited target DNA strand by nCas9, DNA replication and repair will primarily use the edited DNA strand without nick as the template. As illustrated in Supplementary Figure S18A, we investigated multiplex editing performance of the nCas9-based duBE by modifying a previously reported system (63) and compared it with the dCas12a-based duBE. The sgRNA array constructed with multiple promoters showed high instability, with the majority experiencing recombination and fragment deletion after transformation into *B. subtilis* (Supplementary Figure S18B,C). Conversely, the tRNA-based sgRNA array demonstrated relatively better stability, and fragment deletion was only observed in some of the single colonies after transformation into *B. subtilis* but not in cultures after induction (Supplementary Figure S18C,D). Although the nCas9-based duBE possesses a higher activity than that based on dCas12a, the editing efficiency was significantly reduced when incorporating these sgRNAs into the five-member tRNA-sgRNA array, especially losing editing in target 3 (Supplementary Figure S18D). In comparison, the stability of crRNA arrays proved significantly superior to sgRNA arrays, with no instances of fragment deletion observed in both single colonies after transformation into *B. subtilis* and cultures after induction (Supplementary Figure S18C,E). Moreover, the property of nCas9 also creates diversity base conversions except for C > T and A > G (65–67). For instance, uracil-DNA glycosylase (UNG), which will remove the intermedium U base produced by nCas9 and cytosine deaminase, can initiate base-excision repair (BER) and allows the construction of BE enabled C > A and C > G conversions (66,67). After replacing nCas9 with dCas9, only the C > T conversion can be found, approving the specific DNA damage caused by nCas9 is necessary for UNG-initiated C > A and C > G conversions. Similarly, the BE enabled A > C and A > T editing was also built by coupling nCas9 and adenine deaminase with an N-methylpurine DNA glycosylase (MPG), which can excise the intermedium hypoxanthine (I) base and initiates BER (65). Cas12a contrary to Cas9, possesses an RNase activity, which is beneficial for the multi-targeting of MultiduBE (Figure 1). However, the nickase version of Cas12a that solely cleaves the target DNA strand does not exist (the cleavage of the target strand by the Nuc domain relies on the non-target strand cleavage by the RuvC domain) (68), thereby limiting both the editing efficiency (Figure 3 and Figure 4) and diversity of the Cas12a-based BE. This could explain why certain TadA variants, such as TadA7.10, TadA8.17, TADAC-3.1, and TadA-dual, commonly used with nCas9, showed no activity when fused with dCas12a (Figure 2F and Supplementary Figure S6E). In the future, engineered nickase version of Cas12a that could only cleave the target DNA strand may be built by coupling and modifying the endonuclease cleavage domain (69), which will further improve the editing efficiency of BE based on dCas12a.

Previous studies on multiplexed cytosine or adenine base editing using Cas12a focused predominantly on the synthetic BE protein (20–24), paying less attention to the impact of the crRNA array. Here, we designed relatively long separators and extended the DR as needed to match the crRNA array assembly process (Supplementary Figure S1). These modifications of the crRNA array significantly increased the multi-target editing efficiency in both *E. coli* and *B. subtilis* (Figure 4 and Supplementary Figure S16). We also developed a two-round PCR based crRNA array assembly strategy, which requires only the synthesis of an 82-nt forward primer containing DR+, N23, and 32-nt Sp4 for editing new targets (Supplementary Figure S1). This novel strategy enables the production of large-scale crRNA arrays, where an oligo pool is synthesized on a DNA chip, leading to random multi-target genome-wide mutagenesis. The *mthA* mutation in *B. subtilis* can confer low-level resistance to streptomycin (49), so we mixed the crRNA targeting *mthA* with crRNAs targeting *rpsL*, a previously validated target in other hosts (56), into a crRNA minipool library for streptomycin-resistant mutant generation (Figure 5D and Supplementary Figure S16A). With the selective pressure, novel mutations of *rpsL* conferring streptomycin resistance were found for *B. subtilis* (Figure 5E). The robust correlation between genomic mutations and the crRNAs implies that the crRNA array could serve as trackers for mutation targets in the genome. Additionally, for the multiplexed *in-situ* mutagenesis guided by the crRNA array rib5C-Sp4, the mutation in P_{tetL} was found in all the selected riboflavin overproducing strains (Figure 6H and Supplementary Table S5), even when the corresponding selective pressure tetracycline was not exerted. This result suggests that tetracycline resistance can be employed to preclude unmutated strains. However, as demonstrated in metabolic regulation (Figure 6), an efficient screening method is also essential for acquiring specific mutations from a vast combinatorial mutant library. Recent researches in genetically encoded biosensors have provided excellent opportunities in this regard (70).

Some limitations should be considered in the context of multiplexed mutagenesis using MultiduBE. For example, the challenge in constructing crRNA arrays is increased when a large number of genes are selected as the targets simultaneously. Additionally, while MultiduBE showed no substrate motif preference, its high activity was confined to a specific editing window (protospacer positions 7–15) (Supplementary Figure S19). To potentially expand the editing window, liquid-liquid phase separation (LLPS) could be explored, as it has been shown to improve the efficiency of CRISPRa in mammalian cells and mice (71). As shown in Supplementary Figure S20, we also try to broaden the targeting scope to the non-canonical PAM (ncPAM ≠ TTV) by constructing the broad-spectrum duBE (duBEbs), incorporating a broad-spectrum variant of dCas12a (72). However, duBEbs-1a, although capable of generating base editing in all targets with a five-member crRNA array containing two ncPAM crRNAs, exhibited reduced editing activity compared to duBE-1a. Furthermore, in Target 3 with a ncPAM GCTC, duBE-1a also possessed a certain level of activity (Supplementary Figure S20C), which is consistent with the previous study (72). Future work will involve searching for broad-spectrum mutations that better match duBE. Simultaneously, it is crucial to consider the trade-

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

off between off-target effects and PAM preference, as stringent PAM recognition is known to reduce genome-wide off-target effects (73).

In conclusion, we revealed that the specific combination mode of cytosine deaminase, adenine deaminase, and dCas12a would guarantee the activities of all three proteins simultaneously, which enables multiplexed targeting and dual base editing. Furthermore, we found that interference in the crRNA array was a key factor that affects the performance of MultiduBE, and that introducing the synthetic separator sequence into the crRNA array significantly improved the multiplexed editing efficiency in both *E. coli* and *B. subtilis*. Finally, we demonstrated the value of MultiduBE for multiplexed *in-situ* mutagenesis, taking cell physiology reprogramming and metabolic regulation as examples. Given that mutations in the genome can be rapidly tracked using the crRNA array, MultiduBE will also be useful for genome-scale mutagenesis and multi-gene interaction analysis.

Data availability

Reads from NGS of high-throughput targeted amplicon sequencing, whole genome resequencing, RNA-seq, and miRNA sequencing are submitted at Sequence Read Archive (PRJNA967093). The sequences of plasmid and crRNA array with the necessary annotations used in this study are uploaded to Benchling, and the according URLs are shown in Supplementary Note S2 and Supplementary Note S4. Raw gel images, unedited micrographs, and source data generated by the BEAT software are deposited in Zenodo (10.5281/zenodo.10631289). Flow cytometry data is deposited at FlowRepository (FR-FCM-Z6AW).

Supplementary data

Supplementary Data are available at NAR Online.

Funding

National Key R&D Program of China [2020YFA0908300, 2018YFA0900300]; National Natural Science Foundation of China [32021005, 32070085, 31930085, 32300064]; China National Postdoctoral Program for Innovative Talents [BX2021113]; Natural Science Foundation of Jiangsu Province [BK20221083]; Fundamental Research Funds for the Central Universities [USRP52019A, JUSRP221013, JUSRP124034, and JUSRP121010]. Funding for open access charge: National Natural Science Foundation of China.

Conflict of interest statement

None declared.

References

1. Korry,B.J., Lee,S.Y.E., Chakrabarti,A.K., Choi,A.H., Ganser,C., Machan,J.T. and Belenky,P. (2021) Genotoxic agents produce stressor-specific spectra of spectinomycin resistance mutations based on mechanism of action and selection in *Bacillus subtilis*. *Antimicrob. Agents Chemother.*, **65**, e00891-21.
2. Halperin,S.O., Tou,C.J., Wong,E.B., Modavi,C., Schaffer,D.V. and Dueber,J.E. (2018) CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature*, **560**, 248–252.
3. Esvelt,K.M., Carlson,J.C. and Liu,D.R. (2011) A system for the continuous directed evolution of biomolecules. *Nature*, **472**, 499–503.
4. Ravikumar,A., Arzumanyan,G.A., Obadi,M.K.A., Javanpour,A.A. and Liu,C.C. (2018) Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell*, **175**, 1946–1957.
5. Tian,R., Rehm,F.B.H., Czernecki,D., Gu,Y., Zürcher,J.F., Liu,K.C. and Chin,J.W. (2024) Establishing a synthetic orthogonal replication system enables accelerated evolution in *E. coli*. *Science*, **383**, 421–426.
6. Moore,C.L., Papa,L.J.I. and Shoulders,M.D. (2018) A processive protein chimera introduces mutations across defined DNA regions *in vivo*. *J. Am. Chem. Soc.*, **140**, 11560–11564.
7. Molina,R.S., Rix,G., Mengiste,A.A., Álvarez,B., Seo,D., Chen,H., Hurtado,J.E., Zhang,Q., García-García,J.D., Heins,Z.J., *et al.* (2022) *In vivo* hypermutation and continuous evolution. *Nat. Rev. Methods Primer*, **2**, 37.
8. Wu,Y., Liu,Y., Lv,X., Li,J., Du,G. and Liu,L. (2020) Applications of CRISPR in a microbial cell factory: from genome reconstruction to metabolic network reprogramming. *ACS Synth. Biol.*, **9**, 2228–2238.
9. 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.
10. Grünewald,J., Zhou,R., Lareau,C.A., Garcia,S.P., Iyer,S., Miller,B.R., Langner,L.M., Hsu,J.Y., Aryee,M.J. and Joung,J.K. (2020) A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nat. Biotechnol.*, **38**, 861–864.
11. Sakata,R.C., Ishiguro,S., Mori,H., Tanaka,M., Tatsuno,K., Ueda,H., Yamamoto,S., Seki,M., Masuyama,N., Nishida,K., *et al.* (2020) Base editors for simultaneous introduction of C-to-T and A-to-G mutations. *Nat. Biotechnol.*, **38**, 865–869.
12. Zhang,X., Zhu,B., Chen,L., Xie,L., Yu,W., Wang,Y., Li,L., Yin,S., Yang,L., Hu,H., *et al.* (2020) Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.*, **38**, 856–860.
13. Li,C., Zhang,R., Meng,X., Chen,S., Zong,Y., Lu,C., Qiu,J.-L., Chen,Y.-H., Li,J. and Gao,C. (2020) Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat. Biotechnol.*, **38**, 875–882.
14. Xie,J., Huang,X., Wang,X., Gou,S., Liang,Y., Chen,F., Li,N., Ouyang,Z., Zhang,Q., Ge,W., *et al.* (2020) ACBE, a new base editor for simultaneous C-to-T and A-to-G substitutions in mammalian systems. *BMC Biol.*, **18**, 131.
15. Shelake,R.M., Pramanik,D. and Kim,J.-Y. (2023) Improved dual base editor systems (iACBEs) for simultaneous conversion of adenine and cytosine in the bacterium *Escherichia coli*. *mBio*, **14**, e02296-22.
16. Yuan,Q. and Gao,X. (2022) Multiplex base- and prime-editing with drive-and-process CRISPR arrays. *Nat. Commun.*, **13**, 2771.
17. 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.
18. Wu,Y., Li,Y., Jin,K., Zhang,L., Li,J., Liu,Y., Du,G., Lv,X., Chen,J., Ledesma-Amaro,R., *et al.* (2023) CRISPR–dCas12a-mediated genetic circuit cascades for multiplexed pathway optimization. *Nat. Chem. Biol.*, **19**, 367–377.
19. Shaw,W.M., Studená,L., Roy,K., Hapeta,P., McCarty,N.S., Graham,A.E., Ellis,T. and Ledesma-Amaro,R. (2022) Inducible expression of large gRNA arrays for multiplexed CRISPRai applications. *Nat. Commun.*, **13**, 4984.
20. Li,X., Wang,Y., Liu,Y., Yang,B., Wang,X., Wei,J., Lu,Z., Zhang,Y., Wu,J., Huang,X., *et al.* (2018) Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.*, **36**, 324–327.
21. Wang,X., Ding,C., Yu,W., Wang,Y., He,S., Yang,B., Xiong,Y.-C., Wei,J., Li,J., Liang,J., *et al.* (2020) Cas12a base editors induce efficient and specific editing with low DNA damage response. *Cell Rep.*, **31**, 107723.

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

22. Kempton,H.R., Love,K.S., Guo,L.Y. and Qi,L.S. (2022) Scalable biological signal recording in mammalian cells using Cas12a base editors. *Nat. Chem. Biol.*, **18**, 742–750.
23. Chen,F., Lian,M., Ma,B., Gou,S., Luo,X., Yang,K., Shi,H., Xie,J., Ge,W., Ouyang,Z., *et al.* (2022) Multiplexed base editing through Cas12a variant-mediated cytosine and adenine base editors. *Commun. Biol.*, **5**, 1163.
24. Gaillochet,C., Peña Fernández,A., Goossens,V., D'Halluin,K., Drozdzecki,A., Shafie,M., Van Duyse,J., Van Isterdael,G., Gonzalez,C., Vermeersch,M., *et al.* (2023) Systematic optimization of Cas12a base editors in wheat and maize using the ITER platform. *Genome Biol.*, **24**, 6.
25. Cheng,Y., Zhang,Y., Li,G., Fang,H., Sretenovic,S., Fan,A., Li,J., Xu,J., Que,Q. and Qi,Y. (2023) CRISPR-Cas12a base editors confer efficient multiplexed genome editing in rice. *Plant Commun.*, **4**, 100601.
26. Wu,Y., Chen,T., Liu,Y., Tian,R., Lv,X., Li,J., Du,G., Chen,J., Ledesma-Amaro,R. and Liu,L. (2020) Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in *Bacillus subtilis*. *Nucleic Acids Res.*, **48**, 996–1009.
27. Li,Y., Wu,Y., Liu,Y., Li,J., Du,G., Lv,X. and Liu,L. (2022) A genetic toolkit for efficient production of secretory protein in *Bacillus subtilis*. *Bioresour. Technol.*, **363**, 127885.
28. Wu,Y., Liu,Y., Lv,X., Li,J., Du,G. and Liu,L. (2020) CAMERS-B: cRISPR/Cpf1 assisted multiple-genes editing and regulation system for *Bacillus subtilis*. *Biotechnol. Bioeng.*, **117**, 1817–1825.
29. Xu,L., Liu,Y. and Han,R. (2019) BEAT: a python program to quantify base editing from sanger sequencing. *CRISPR J.*, **2**, 223–229.
30. 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.
31. Yang,H., Yu,H. and Shen,Z. (2015) A novel high-throughput and quantitative method based on visible color shifts for screening *Bacillus subtilis* THY-15 for surfactin production. *J. Ind. Microbiol. Biotechnol.*, **42**, 1139–1147.
32. Sun,G., Wu,Y., Huang,Z., Liu,Y., Li,J., Du,G., Lv,X. and Liu,L. (2023) Directed evolution of diacetylchitobiose deacetylase via high-throughput droplet sorting with a novel, bacteria-based biosensor. *Biosens. Bioelectron.*, **219**, 114818.
33. You,J., Yang,C., Pan,X., Hu,M., Du,Y., Osire,T., Yang,T. and Rao,Z. (2021) Metabolic engineering of *Bacillus subtilis* for enhancing riboflavin production by alleviating dissolved oxygen limitation. *Bioresour. Technol.*, **333**, 125228.
34. Mack,M., van Loon,A.P.G.M. and Hohmann,H.-P. (1998) Regulation of riboflavin biosynthesis in *Bacillus subtilis* is affected by the activity of the flavokinase/flavin adenine dinucleotide synthetase encoded by *ribC*. *J. Bacteriol.*, **180**, 950–955.
35. Liu,X., Liu,X., Zhou,C., Lv,J., He,X., Liu,Y., Xie,H., Wang,B., Lv,X., Tang,L., *et al.* (2021) Engineered FnCas12a with enhanced activity through directional evolution in human cells. *J. Biol. Chem.*, **296**, 100394.
36. Meliawati,M., Schilling,C. and Schmid,J. (2021) Recent advances of Cas12a applications in bacteria. *Appl. Microbiol. Biotechnol.*, **105**, 2981–2990.
37. Bandyopadhyay,A., Kancharla,N., Javalkote,V.S., Dasgupta,S. and Brutnell,T.P. (2020) CRISPR-Cas12a (Cpf1): a versatile tool in the plant genome editing tool box for agricultural advancement. *Front. Plant Sci.*, **11**, 584151.
38. Li,A., Mitsunobu,H., Yoshioka,S., Suzuki,T., Kondo,A. and Nishida,K. (2022) Cytosine base editing systems with minimized off-target effect and molecular size. *Nat. Commun.*, **13**, 4531.
39. Cheng,T.-L., Li,S., Yuan,B., Wang,X., Zhou,W. and Qiu,Z. (2019) Expanding C–T base editing toolkit with diversified cytidine deaminases. *Nat. Commun.*, **10**, 3612.
40. Yan,D., Ren,B., Liu,L., Yan,F., Li,S., Wang,G., Sun,W., Zhou,X. and Zhou,H. (2021) High-efficiency and multiplex adenine base editing in plants using new TadA variants. *Mol. Plant*, **14**, 722–731.
41. Neugebauer,M.E., Hsu,A., Arbab,M., Krasnow,N.A., McElroy,A.N., Pandey,S., Doman,J.L., Huang,T.P., Raguram,A., Banskota,S., *et al.* (2023) Evolution of an adenine base editor into a small, efficient cytosine base editor with low off-target activity. *Nat. Biotechnol.*, **41**, 673–685.
42. Lam,D.K., Feliciano,P.R., Arif,A., Bohnuud,T., Fernandez,T.P., Gehrke,J.M., Grayson,P., Lee,K.D., Ortega,M.A., Sawyer,C., *et al.* (2023) Improved cytosine base editors generated from TadA variants. *Nat. Biotechnol.*, **41**, 686–697.
43. Zetsche,B., Heidenreich,M., Mohanraju,P., Fedorova,I., Kneppers,J., Degennaro,E.M., Winblad,N., Choudhury,S.R., Abudayyeh,O.O., Gootenberg,J.S., *et al.* (2017) Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.*, **35**, 31–34.
44. Magnusson,J.P., Rios,A.R., Wu,L. and Qi,L.S. (2021) Enhanced Cas12a multi-gene regulation using a CRISPR array separator. *eLife*, **10**, e66406.
45. Grünewald,J., Zhou,R., Garcia,S.P., Iyer,S., Lareau,C.A., Aryee,M.J. and Joung,J.K. (2019) Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature*, **569**, 433–437.
46. Galinier,A., Foulquier,E. and Pompeo,F. (2021) Metabolic control of cell elongation and cell division in *Bacillus subtilis*. *Front. Microbiol.*, **12**, 697930.
47. Blair,J.M.A., Webber,M.A., Baylay,A.J., Ogbolu,D.O. and Piddock,L.J.V. (2015) Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.*, **13**, 42–51.
48. Maughan,H., Galeano,B. and Nicholson,W.L. (2004) Novel *rpoB* mutations conferring rifampin resistance on *Bacillus subtilis*: global effects on growth, competence, sporulation, and germination. *J. Bacteriol.*, **186**, 2481–2486.
49. Tojo,S., Kim,J.-Y., Tanaka,Y., Inaoka,T., Hiraga,Y. and Ochi,K. (2014) The *mthA* mutation conferring low-level resistance to streptomycin enhances antibiotic production in *Bacillus subtilis* by increasing the S-adenosylmethionine pool size. *J. Bacteriol.*, **196**, 1514–1524.
50. Bechhofer,D.H. and Stasinopoulos,S.J. (1998) *tetA* (L) mutants of a tetracycline-sensitive strain of *Bacillus subtilis* with the polynucleotide phosphorylase gene deleted. *J. Bacteriol.*, **180**, 3470–3473.
51. Ilina,E., Malakhova,M., Bodoev,I., Filimonova,A., Oparina,N. and Govorun,V. (2013) Mutation in ribosomal protein S5 leads to spectinomycin resistance in *Neisseria gonorrhoeae*. *Front. Microbiol.*, **4**, 51849.
52. Chen,S.-C., Hu,L.-H., Zhu,X.-Y. and Yin,Y.-P. (2020) Gonococcal urethritis caused by a multidrug resistant *Neisseria gonorrhoeae* strain with high-level resistance to spectinomycin in China. *Emerg. Microbes Infect.*, **9**, 517–519.
53. He,X., Miao,V. and Baltz,R.H. (2005) Spectinomycin resistance in *rpsE* mutants is recessive in *Streptomyces roseosporus*. *J. Antibiot.*, **58**, 284–288.
54. Criswell,D., Tobiason,V.L., Lodmell,J.S. and Samuels,D.S. (2006) Mutations conferring aminoglycoside and spectinomycin resistance in *Borrelia burgdorferi*. *Antimicrob. Agents Chemother.*, **50**, 445–452.
55. Sun,Y.-J., Luo,J.-T., Wong,S.-Y. and Lee,A.S.G. (2010) Analysis of *rpsL* and *rrs* mutations in Beijing and non-Beijing streptomycin-resistant *mycobacterium tuberculosis* isolates from Singapore. *Clin. Microbiol. Infect.*, **16**, 287–289.
56. Hosaka,T., Xu,J. and Ochi,K. (2006) Increased expression of ribosome recycling factor is responsible for the enhanced protein synthesis during the late growth phase in an antibiotic-overproducing *Streptomyces coelicolor* ribosomal *rpsL* mutant. *Mol. Microbiol.*, **61**, 883–897.
57. Andersson,D.I. and Hughes,d.,D. (2010) Antibiotic resistance and its cost: is it possible to reverse resistance?*Nat. Rev. Microbiol.*, **8**, 260–271.

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

58. Xia,L. and Wen,J. (2022) Available strategies for improving the biosynthesis of surfactin: a review. *Crit. Rev. Biotechnol.*, **43**, 1111–1128.
59. Wu,Q., Zhi,Y. and Xu,Y. (2019) Systematically engineering the biosynthesis of a green biosurfactant surfactin by *Bacillus subtilis* 168. *Metab. Eng.*, **52**, 87–97.
60. Wang,M., Yu,H., Li,X. and Shen,Z. (2020) Single-gene regulated non-spore-forming *Bacillus subtilis*: construction, transcriptome responses, and applications for producing enzymes and surfactin. *Metab. Eng.*, **62**, 235–248.
61. You,J., Pan,X., Yang,C., Du,Y., Osire,T., Yang,T., Zhang,X., Xu,M., Xu,G. and Rao,Z. (2021) Microbial production of riboflavin: biotechnological advances and perspectives. *Metab. Eng.*, **68**, 46–58.
62. Hao,W., Cui,W., Cheng,Z., Han,L., Suo,F., Liu,Z., Zhou,L. and Zhou,Z. (2021) Development of a base editor for protein evolution via *in situ* mutation *in vivo*. *Nucleic Acids Res.*, **49**, 9594–9605.
63. Hao,W., Cui,W., Suo,F., Han,L., Cheng,Z. and Zhou,Z. (2022) Construction and application of an efficient dual-base editing platform for *Bacillus subtilis* evolution employing programmable base conversion. *Chem. Sci.*, **13**, 14395–14409.
64. Komor,A.C., Kim,Y.B., Packer,M.S., Zuris,J.A. and Liu,D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, **533**, 420–424.
65. Tong,H., Wang,X., Liu,Y., Liu,N., Li,Y., Luo,J., Ma,Q., Wu,D., Li,J., Xu,C., *et al.* (2023) Programmable A-to-Y base editing by fusing an adenine base editor with an N-methylpurine DNA glycosylase. *Nat. Biotechnol.*, **41**, 1080–1084.
66. Zhao,D., Li,J., Li,S., Xin,X., Hu,M., Price,M.A., Rosser,S.J., Bi,C. and Zhang,X. (2021) Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.*, **39**, 35–40.
67. Kurt,I.C., Zhou,R., Iyer,S., Garcia,S.P., Miller,B.R., Langner,L.M., Grünewald,J. and Joung,J.K. (2021) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.*, **39**, 41–46.
68. Yamano,T., Nishimasu,H., Zetsche,B., Hirano,H., Slaymaker,I.M., Li,Y., Fedorova,I., Nakane,T., Makarova,K.S., Koonin,E.V., *et al.* (2016) Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell*, **165**, 949–962.
69. Aman,R., Syed,M.M., Saleh,A., Melliti,F., Gundra,S.R., Wang,Q., Marsic,T., Mahas,A. and Mahfouz,M.M. (2024) Peptide nucleic acid-assisted generation of targeted double-stranded DNA breaks with T7 endonuclease I. *Nucleic Acids Res.*, gkae148.
70. Yu,W., Xu,X., Jin,K., Liu,Y., Li,J., Du,G., Lv,X. and Liu,L. (2023) Genetically encoded biosensors for microbial synthetic biology: from conceptual frameworks to practical applications. *Biotechnol. Adv.*, **62**, 108077.
71. Ma,S., Liao,K., Li,M., Wang,X., Lv,J., Zhang,X., Huang,H., Li,L., Huang,T., Guo,X., *et al.* (2023) Phase-separated DropCRISPRa platform for efficient gene activation in mammalian cells and mice. *Nucleic Acids Res.*, **51**, 5271–5284.
72. Chen,Z., Sun,J., Guan,Y., Li,M., Lou,C. and Wu,B. (2021) Engineered DNase-inactive Cpf1 variants to improve targeting scope for base editing in E. coli. *Synth. Syst. Biotechnol.*, **6**, 326–334.
73. Zhou,J., Chen,P., Wang,H., Liu,H., Li,Y., Zhang,Y., Wu,Y., Paek,C., Sun,Z., Lei,J., *et al.* (2022) Cas12a variants designed for lower genome-wide off-target effect through stringent PAM recognition. *Mol. Ther.*, **30**, 244–255.

Downloaded from https://academic.oup.com/nar/article/52/8/4739/7639551 by guest on 20 May 2024

Received: July 21, 2023. **Revised:** March 13, 2024. **Editorial Decision:** March 15, 2024. **Accepted:** March 19, 2024
© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com