CRISPR-dCas9 mediated cytosine deaminase base editing in bacillus subtilis

CRISPR-dCas9 mediated cytosine deaminase base editing in *Bacillus subtilis*

Sili Yu[1,2,3,‖], Marcus A. Price[4,‖], Yu Wang[1,2,3], Yang Liu[2,3], Yanmei Guo[2,3], Xiaomeng Ni[2,3], Susan J. Rosser[4], Changhao Bi[1,2,3]* and Meng Wang[1,2,3]*

[1]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

[2]Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, China

[3]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China

[4]Centre for Synthetic and Systems Biology and UK Centre for Mammalian Synthetic Biology, School of Biological Sciences, University of Edinburgh, UK

ABSTRACT: Base editing technology based on Clustered regularly interspaced short palindromic repeats/associated protein 9 (CRISPR/Cas9) is a recent addition to the family of CRISPR technologies. Compared with the traditional CRISPR/Cas9 technology, it does not rely on DNA double strand break and homologous recombination, and can realize gene inactivation and point mutation more quickly and simply. Herein, we first developed a base editing method for genome editing in *Bacillus subtilis* utilizing CRISPR/dCas9 (a fully nuclease-deficient mutant of Cas9 from *S. pyogenes*) and activation-induced cytidine deaminase (AID). This method achieved three and four loci simultaneous editing with editing efficiency up to 100% and 50%, respectively. Our base editing system in *B. subtilis* has a 5 nt editing window, which is similar to previous reported base editing in other microorganisms. We demonstrated that the plasmid curing rate is almost 100%, which is advantageous for multiple rounds of genome engineering in *B. subtilis*. Finally, we applied multiplex genome editing to generate a *B. subtilis* 168 mutant strain with eight inactive extracellular proteases genes in just two rounds of base editing and plasmid curing, suggesting that it is a powerful tool for gene manipulation in *B. subtilis* and industrial applications in the future.

KEYWORDS: CRISPR/dCas9, Cytidine deaminase, Genome editing, *Bacillus subtilis*

Bacillus subtilis is a well-characterized Gram-positive bacterium which has been widely applied in the production of antibiotics, medicinal proteins, industrial enzymes and biopolymers[1-5]. Because of their excellent protein secretory capability[6, 7] and generally regarded as safe (GRAS) designation, *B. subtilis* are often used as important hosts for industrial applications[8, 9].

Recently, the CRISPR/Cas9 system, which originated as a bacterial adaptive immune system, has been developed into efficient genome editing tools for various organisms[10-14], including *B. subtilis*[15-22]. However, in *B. subtilis*, a foreign DNA template is required for each genome editing target. Furthermore, a CRISPR/Cas9 induced double strand break (DSB) or nick leads to low cell survival rate, which significantly decreases the transformation efficiency and multiplex genome editing efficiency. Only three point mutations in different loci have been simultaneously edited in *B. subtilis* to date[20].

Base editing technology, based on CRISPR/Cas9, is a recent addition to the family of CRISPR technologies[23-26]. Combining the activities of the rapid and accurate genome targeting CRISPR/Cas9 system with the deaminase activity of cytidine deaminases, base editing technology can accurately convert specific Cs into Ts. Combined with bioinformatics tools, base editing technology can accurately utilize C-T mutations to generate early stop codon for gene inactivation[23-29] or targeted point mutations.

In this study, the CRISPR-dCas9-AID system was first developed for genome editing in *B. subtilis* 168. After our method optimization, a single plasmid system expressing both dCas9-AID and sgRNA could be used for programmable precise base editing in *B. subtilis* 168 with an editing efficiency up to 100%. Our base editing system in *B. subtilis* has a 5 nt editing window, which is similar to previous reported base editing in other microorganisms. Multiplex genome editing was achieved with very high efficiency (up to 100% for three loci and 50% for four loci) allowing the efficient construction, in two rounds of base editing and plasmid curing, of a strain deficient of the eight extracellular proteases secreted by *B. subtilis*.

■ RESULTS AND DISCUSSION

dCas9 is essential for successful base editing in *B. subtilis*. To test whether CRISPR/Cas9-AID enabled base editing technology can be used in *B. subtilis*,

editing plasmids containing fusion proteins of dCas9 or nCas9 (a fully or partially nuclease-deficient mutant of Cas9 from *S. pyogenes*, respectively), and AID ortholog (PmCDA1 from sea lamprey) were constructed by replacing *cas9* in pBAC0041, known to efficiently target the 5' region of the α-Amylase encoding *amyE* gene[21]. The two constructed plasmids (pBAC-dCas9-AID-*amyE*-18-target (dCas9-AID), and pBAC-nCas9-AID-*amyE*-18-target (nCas9-AID)), expressed each fusion protein under the control of the IPTG inducible P_{grac} promoter, as well as the sgRNA by the constitutive P_{veg} promoter.

Upon transformation of *B. subtilis* 168 with pBAC-dCas9-AID-*amyE*-18-target and pBAC-nCas9-AID-*amyE*-18-target, transformants were only obtained with the plasmid expressing dCas9-AID. Transformations with the control plasmids pBAC-dCas9-*amyE*-18-target and pBAC-nCas9-*amyE*-18-target, targeting the same PAM site as pBAC-dCas9-AID-*amyE*-18-target, also showed nCas9 to be lethal to the cell in the absence of AID even without induction by IPTG. On the other hand, dCas9 yielded similarly high CFUs as dCas9-AID. This is a similar result to that observed in *E. coli* by Banno *et al*.[30], but is in contrast to our previous results in the Gram positive *C. glutamicum* where nCas9-AID was used to increase base editing efficiencies[31]. The strain harboring plasmid pBAC-dCas9-AID-*amyE*-18-target was induced with 1 mM IPTG for 24 hours at 30˚C in order to determine if the targeted deamination and base editing reaction could proceed. We sequenced the target loci and found a clear double peak in the -18 position upstream from AGG PAM (Figure 1) which indicated that base editing in *B. subtilis* had occurred and is quite efficient.

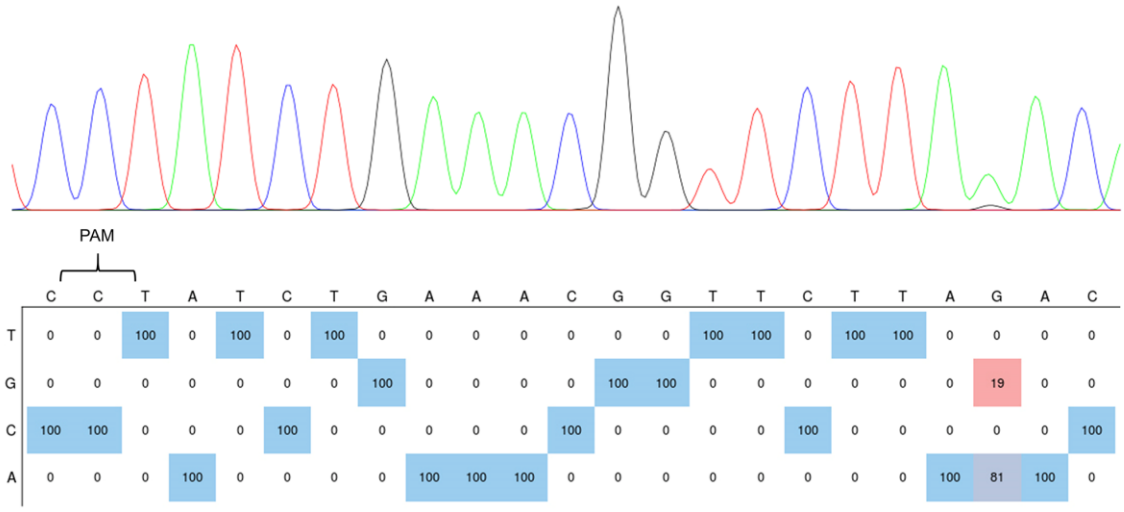


Figure 1. Programmable base editing by CRISPR/dCas9-AID in *B. subtilis*. Detectable editing double peak at -18 C upstream of the AGG PAM. The editing efficiency was estimated at 81% following analysis of sanger sequencing results (https://moriaritylab.shinyapps.io/editr_v10/)[32].

Plasmid curing. Because multiple rounds of sequential genome editing is inevitable for most strain engineering works, a plasmid curing system was realized by replacing the original replication protein, RepA, with the temperature-sensitive RepF[33] to produce plasmid pBAC-dCas9-AID-*amyE*-18-target^{TS}. This plasmid was used to transform *B. subtilis* 168 and transformants were selected for growth. After induction, the sequencing results showed similar editing efficiency to pBAC-dCas9-AID-*amyE*-18-target at 88%. Sequencing-confirmed, successfully edited strains were then cultured in LB medium at 37°C, 220 rpm for 24 hours without antibiotic for plasmid curing. Plasmid curing efficiency was determined to be approaching 100% (Figure S1), which is very advantages for complex genome engineering in *B. subtilis*.

Base editing method optimization. To further investigate the base editing efficiencies at different positions within the protospacer region, a target sequence containing 5 consecutive cytosines from positions -16 to -20 upstream of the PAM, termed hereafter as 5Cs, was selected from the genome of *B. subtilis* 168. A 5Cs editing plasmid, pBAC-dCas9-AID-*5Cs*-target^{TS}, was constructed and used to transformed into *B. subtilis* 168. After 24 hours of growth and induction with 1 mM IPTG at 30°C, we sequenced the target loci and found the editing efficiencies were quite low. The editing efficiency of -17 position was less than 20% and -18 position was less than 60%. In order to improve the editing efficiency for further applications, we first studied the effect of IPTG concentrations ranging from 0, 0.1, 0.25, 0.5 and 1.0 mM. The results showed that only -17 and -18 positions were edited and the editing efficiency didn't change significantly (Figure S2). Therefore, the final concentration of IPTG was set at 0.1 mM in the following experiments.

We also observed that the editing efficiency does not improve further after 6 hours of induction in 0.1 mM IPTG at 30°C (Figure 2a). As it has been reported that multiple passages could improve the editing efficiency[31], after 6 hours of induction, the culture was transferred into fresh LB with antibiotic and IPTG at the same initial OD_{600}. The editing efficiency following the 2nd passage could increase to 39% for -17 position and 78% for -18 position (Figure 2b).

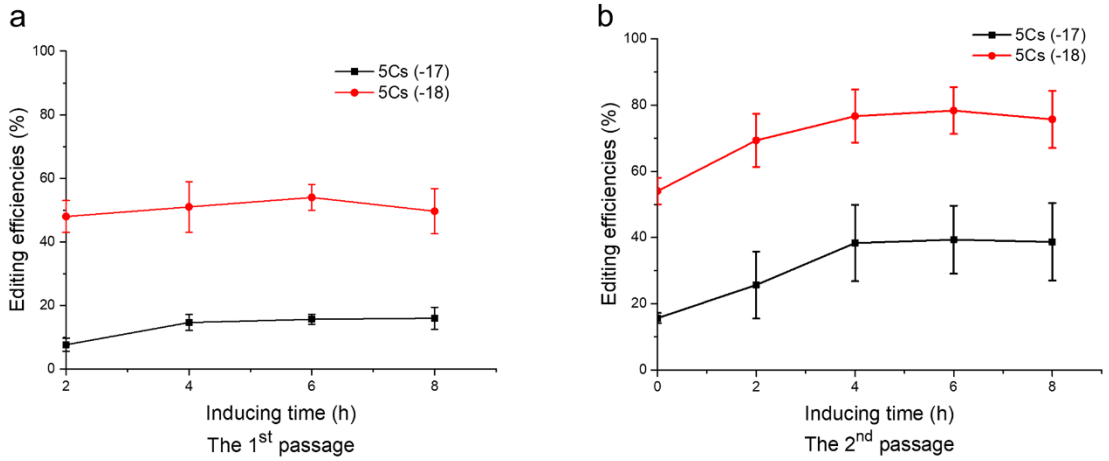


Figure 2. Adjusting culture conditions to optimize editing efficiency. Influence of induction time on the editing efficiency for (a) the 1st passage and (b) the 2nd passage.

5 nt Editing window in *B. subtilis*. The d/nCas9-AID editing system was previously reported to target -16 to -20 positions upstream of the PAM sequence[24, 26, 30]. Our results following editing of 5Cs showed that only -17 and -18 positions have heightened editing efficiencies, while those at other positions were substantially lower. Following the sequencing of ten single colonies, obtained after spreading the edited culture on LB plates, the average editing efficiency of each single colony was calculated to represent the overall editing efficiency. We found that 5Cs (-18) position had the highest editing efficiency of 76%, while 5Cs (-17) position was second highest at 54%. The editing efficiency at 5Cs (-19) position was only 4%, while 5Cs (-16) and 5Cs (-20) positions had little or no editing.

To investigate the discrepancy, we constructed five editing plasmids targeting five different loci on the *amyE* gene with individual cytosine at the -16 to -20 positions within the protospacer region. Following the sequencing of ten single colonies for each locus, the results were similar to those editing efficiencies obtained at the 5Cs loci, as the editing efficiency at the *amyE* (-18) locus was still the highest at 100% and *amyE* (-17) locus was 51%. The results from the 5Cs locus found that the editing efficiency at position (-19) was higher than that for position (-20). One difference here was that the efficiency at the *amyE* (-20) locus was higher than for the *amyE* (-19) locus (20% and 13% respectively). This is most likely to have been due to the sequence dependency for the efficiency of each editing target. The *amyE* (-16) locus had the lowest editing efficiency of only 2% (Figure 3). We

therefore conclude that the dCas9-AID editing system in *B. subtilis* retains the 5nt editing window observed in other organisms, with the caveat that editing efficiency at the -16 position is extremely low.

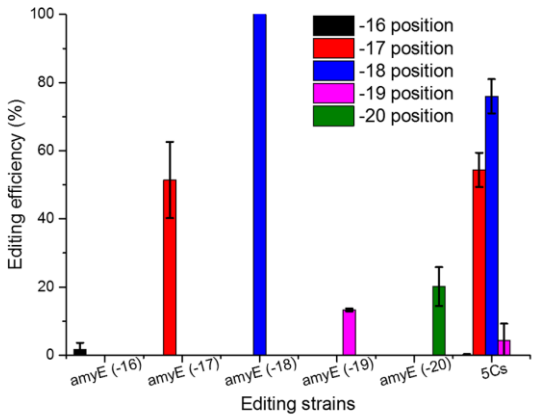


Figure 3. Target window for *B. subtilis* 168 base editing. Editing strains contained plasmids targeting individual -16 to -20 C positions or consecutive 5Cs. 10 randomly selected colonies were analyzed to calculate the overall editing efficiency of different positions.

Multiplex genome editing. Multiplex genome editing can greatly reduce the work load during genome engineering processes through the simultaneous targeting of multiple loci. However, it is much harder to achieve in prokaryotes[34] due to their inferior DNA repair system when DSB are introduced in comparison to eukaryotes. Alternatively, base editing technology represents an excellent option for multiplex genome editing because no DSB is introduced. We first tested multiplex base editing in *B. subtilis* by simultaneously expressing three gRNAs to edit *amyE* (-18), *scoC* (-18) and 5Cs. A three gene editing plasmid, pBAC-dCas9-AID-5Cs-scoC-1-amyE-18-3targets^{TS}, was constructed and used to transformed into *B. subtilis* 168. We then performed our optimized base editing method and mixed culture sequencing showed that the three genes could be edited at the same time. After plating an IPTG induced culture on LB plates containing chloramphenicol and IPTG, ten single colonies were picked and the targeted regions were sequenced. Surprisingly, the editing efficiency was 100% for all three genes (Figure 4), which is much higher than the previously reported multiplex genome editing in *C. glutamicum*[31], *E. coli*[35] and *Streptomyces*[29].

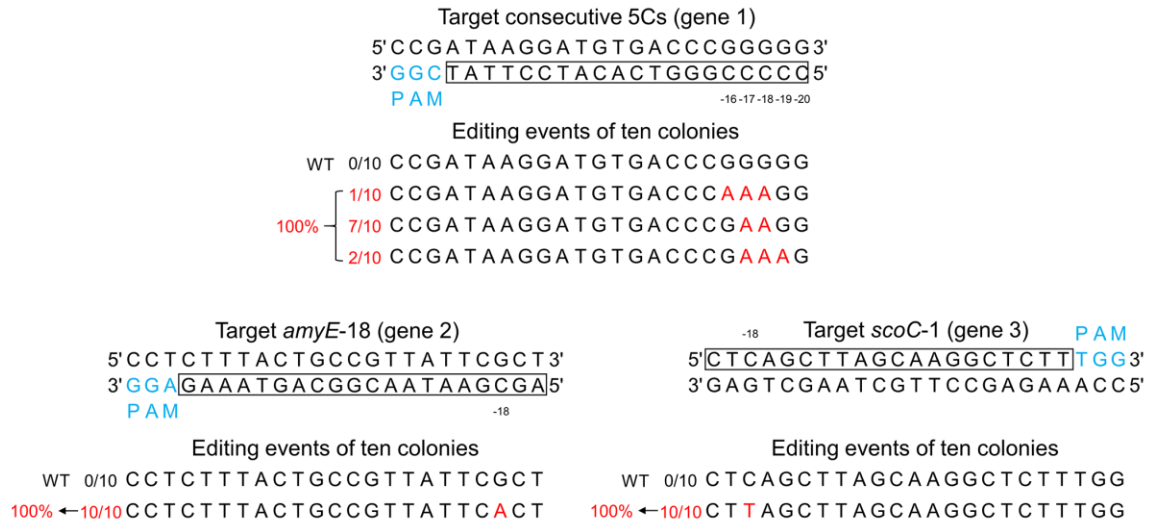


Figure 4. Multiplex genome editing in *B. subtilis* 168. Three sgRNAs were expressed to simultaneously target the consecutive 5Cs (-16 to -20), the transcriptional regulator gene *scoC*-1 (-18) and *amyE*-18 (-18). Editing events of ten colonies were counted. PAM sites are shown in blue, while the modified bases are shown in red.

Inactivation of Eight Extracellular Protease Genes. *B. subtilis* 168 is a model strain with important industrial applications particularly in enzyme production. *B. subtilis* 168 has eight extracellular proteases, the genes for which are *mpr*, *nprB*, *vpr*, *wprA*, *aprE*, *nprE*, *epr* and *bpr*. These proteases will degrade secreted heterologous proteins, leading to decreased enzyme production titers during industrial applications. The modified *B. subtilis* 168 strain, WB800, where all eight proteases had been knocked out, was previously constructed and has been shown to enhance the stability of the secreted heterologous proteins because of extracellular protease deficient[8, 9]. However, the construction of WB800 required extensive time and resources, performing eight rounds of sequential deletions to remove each extracellular protease gene. This time intensive sequential genome editing process, while somewhat improved, has persisted with the development of CRISPR-Cas9-Mediated Multi-Round Genome Editing system[36].

To demonstrate the rapid genome engineering ability of base editing in *B. subtilis*, we inactivated these eight genes in just two rounds of multiplex base editing. We first chose four genes (*mpr*, *nprB*, *vpr* and *wprA*) for the introduction of early stop codons. After we obtained the strain where these four genes had been inactivated (Figure S3a), we removed the editing plasmid, utilizing the temperature sensitive origin of replication. The cured strain, BS△1-4Pro, was then transformed with another editing plasmid targeting the last

four protease genes (*aprE*, *nprE*, *epr* and *bpr*). After editing (Figure S3b) and plasmid curing, we obtained the eight extracellular proteases deficient strain BS△1-8Pro (Figure S4). Using this strain, a protease activity assay was performed using FTC (fluorescein isothiocyante)-casein as a substrate provided by the Pierce Fluorescent Protease Assay Kit as reported[36]. The protease assay showed that the extracellular activity of inactive strain BS△1-8Pro was significantly decreased to a similar level as that of the WB800 control strain (Figure 5).

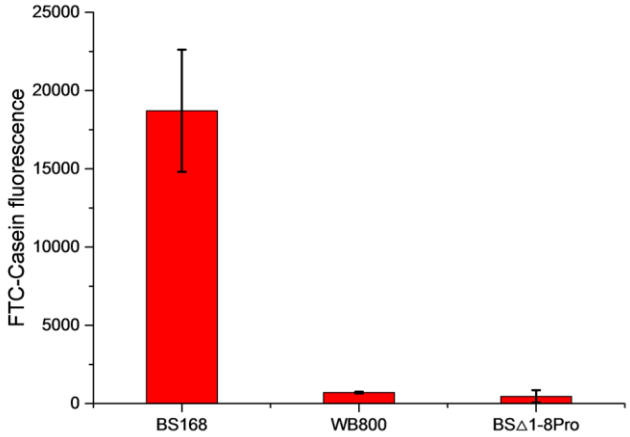


Figure 5. Protease activity assay. The protease activity of inactive strain BS△1-8Pro is compared to the wild type strain (*B. subtilis* 168) and the WB800 control strain.

Genome scale analysis of CRISPR/dCas9-AID targets for gene inactivation. The CRISPR/dCas9-AID system can create TGA (opal), TAG (amber), and TAA (ochre) stop codons by editing the CGA (Arg), CAG (Gln), and CAA (Gln) codons respectively on the coding strand, or by targeting the CCA on the noncoding strand of a target gene to edit one or both Gs to As within a TGG (Trp) codon. To assess the genome-scale coverage of this gene inactivation method in *B. subtilis*, we performed bioinformatics analysis of the *B. subtilis* 168 genome[37]. 15337 unique gRNAs were obtained, which can target 3480 of the total 4106 genes (84.7%) to introduce early stop codons (Figure. 6a). In addition, these codons are uniformly distributed within the coding region (Figure. 6b). We compared the reported essential genes with targetable genes in *B. subtilis* 168 genome[38], suggesting that 85.1% of nonessential genes (3266 genes) are editable and the remaining 569 non-essential genes are not editable using current CRISPR/dCas9-AID technique reported here (Figure. 6c).

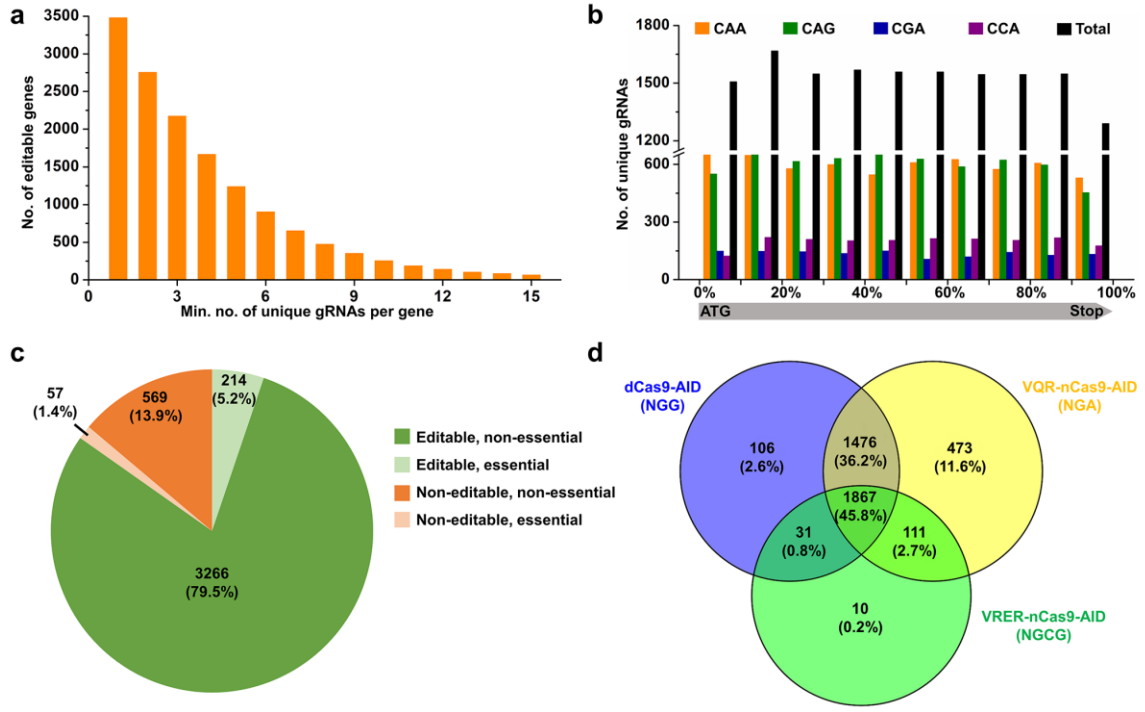


Figure 6. CRISPR/dCas9(D10A&H840A)-AID targets for gene inactivation in *B. subtilis* 168 genome. (a) Number of editable genes with at least one unique gRNA. (b) Relative positions of the CRISPR/dCas9-AID targets for gene inactivation. (c) Statistics of editable genes with at least one targetable codon for gene inactivation in *B. subtilis* 168 genome. (d) Number of genes that cannot be edited by CRISPR/ dCas9-AID but can be edited by Cas9 variants (VQR and VRER)[26].

In this study, we report the first development of a deaminase base editing technique within *B. subtilis*. Without foreign donor DNA and the introduction of a double-strand break, a single plasmid system expressing both dCas9-AID and sgRNA could be used for programmable, precise base editing in *B. subtilis* 168 at high efficiencies. Although our base editor has a 5 nt editing window, only -17 and -18 positions were edited with very high efficiencies up to 100%. While positions -16, -19 and -20 could only be edited with low efficiencies in contrast to our previous work in *C. glutamicum*[31] where the editing efficiencies of -16 to -20 positions were all more than 50%. It would be very interesting to investigate the reason behind the differences in editing window between different organisms. Using the approach reported here, we suggest avoiding the targeting of positions -16, -19 and -20, if possible, in *B. subtilis.*

We also achieved very efficient multiplex genome editing in *B. subtilis* 168 with our method. Interestingly, we noted that it is much easier to obtain successfully multiplex edited single colonies when a 2nd passage of cells are induced on an LB plate with IPTG,

rather than liquid LB media (Table S1 and Table S2). It is possible that liquid media represents a competitive growth environment that results in multiplex edited strains having a slight disadvantage in terms of growth, leading to lowered probabilities to obtain multiplex edited single colonies.

The 100% efficiency of a triple sites editing experiment was significantly higher than 23.3% observed previously in *C. glutamicum*[31]. It is possible that this difference is due to dCas9 being used here, while nCas9 is used for editing of *C. glutamicum*, which, as noted in this study, can have a significant impact on cell survival rate. Extremely high multiplex genome editing efficiency is advantageous in complex genome engineering works and it would be very interesting to further increase targeting loci numbers to see its limitation.

Off-target effects are a universal drawback of all genome editing tools. We performed genome sequencing of the eight extracellular protease deficient strain BS△1-8Pro and found there were a total of nineteen extra single nucleotide variants (SNVs) besides the eight targeted proteases, including nine missense variants, seven synonymous variants and three intragenic variants (Table S3). All nineteen extra mutations are either C-T or G-A, which might be attributed to the deaminase activity of AID. However we cannot rule out the chance that these mutations arose spontaneously during the lengthy experiment – an effect perhaps exacerbated by dCas-AID over-expression. We analyzed these extra mutations sites using cas-offinder[39] and found none to coincide with any theoretical off-target sites of the eight sgRNAs that were used to target each protease (data not shown). Therefore, these extra mutations are likely the results of random AID activity or spontaneous mutations occurring during the experiment. As a comparison, in a reported base-editing[31] application in *C. glutamicum*, nine extra mutations were found in a strain, after one round of base editing, where only a single site was targeted. In our case, the final extracellular protease deficient strain went through two rounds of base editing targeting eight individual sites.

In the field of metabolic engineering, random mutations in microorganisms are inevitable and not as critical as in mammalian cells so long as they do not affect cell growth and the desired phenotype. In the future, we can reduce the adverse effect of AID by using more stringent inducible promoters and reducing the induction time appropriately to limit the chance of introducing random mutations, while still obtaining the desired mutations.

Target coverage is another important issue for our base editor in *B. subtilis*. Bioinformatics analysis indicates that 569 predicted non-essential genes in *B. subtilis* 168 cannot be inactivated by our current base editor (Fig. 6c). However, by using just two Cas9 variants with different PAM sequences (VQR-Cas9 (NGA), VRER-Cas9 (NGCG))[26], which have been proven to be quite efficient in editing the genome of *C. glutamicum* , the non-editable genes can be reduced to only 32 in *B. subtilis* 168 (Fig. 6d). Further development of this base editor tool via testing with different Cas9 variants[40-42] and other CRISPR effector proteins, such as the recently reported dMAD7 variant[40-43] would significantly improve its efficiency and flexibility.

■ METHODS

Microorganisms and Culture Medium. The bacterial strains used in this study are listed in Table 1. *E. coli* DH5α and *E. coli* DB3.1 were used to construct the recombinant plasmids and cultivated aerobically at 37°C in Luria–Bertani (LB) broth. *B. subtilis* cells were also cultured in LB broth, LB agar. When required, the antibiotics were supplemented with ampicillin (100 μg/mL) for E. *coli* or chloramphenicol (5 μg/mL) for *B. subtilis*. To express dCas9-AID protein, 0-1.0 mM final concentration of IPTG (isopropyl-β-D-1-thiogalactopyranoside) was added in the medium.

Plasmids and gRNAs design. The initial pBAC0041 amyE target plasmid was modified to express *dcas9-AID* (pBAC-dCas9-AID-*amyE*-18-target) or *ncas9-AID* (pBAC-nCas9-AID-*amyE*-18-target) by removing *cas9* between *Bam*HI and *Xba*I sites, following by amplification of *dcas9-AID* and *ncas9-AID* with primers containing a *Bsa*I site. These were digested with *Bsa*I, yielding compatible overhangs to the digested pBAC0041 backbone, and the fragments were ligated with T4 DNA ligase.

To facilitate rapid and simple construction of different gRNAs expression plasmids, a Golden Gate assembly strategy was used. Two *Bsa*I sites were removed from pBAC-dCas9-AID-*amyE*-18-target and the 20 bp target sequence was replaced with a *Bsa*I-*ccdB*-*Bsa*I cassette to obtain pBAC-dCas9-AID-gRNA::*ccdB* plasmid. Then a pair of 24 bp primers were annealed to form a double-strand DNA (dsDNA) with cohesive ends and the dsDNA was ligated to the *Bsa*I digested pBAC-dCas9-AID-gRNA::*ccdB* plasmid backbone using a Golden Gate assembly reaction to generate the editing plasmid. For

multiplex genome editing plasmid construction, repeated *Bsa*I-target-gRNAscaffold-terminator-promoter-target-*Bsa*I with different targets and complementary cohesive ends were connected with the pBAC-dCas9-AID-gRNA::*ccdB* plasmid backbone by Golden Gate. To obtain the temperature sensitive pBAC-dCas9-AID-gRNA::*ccdB*^{TS} plasmid that could be easily cured after base editing, the encoding gene of a temperature-sensitive original replication protein RepF was amplified from pKSV7 plasmid then used to replace the original replication protein gene *repA*. Plasmids used in this study are listed in Table 2 and primers used for gRNAs design are listed in Table 3.

Base editing in *B. subtilis.* *B. subtilis* 168 was transformed with 2 μg pBAC-dCas9-AID-gRNAtarget^{TS} plasmids by the Spizizen natural competency method[44]. The transformed cells were spread on LB agar plates supplemented with chloramphenicol and grown at 30°C. Single colonies were selected from the plate and grown in LB medium containing chloramphenicol, overnight at 30°C with agitation. The overnight culture was diluted to an initial OD_{600} of 0.1 in 3 mL LB medium containing chloramphenicol and 0.1 mM IPTG. To perform multiple passages to promote increased editing efficiencies, following 8 h of growth and inducing at 30°C with agitation, the cultures were transferred into fresh 3 mL LB medium with chloramphenicol and 0.1 mM IPTG or were spread on LB agar plates containing chloramphenicol and 0.1 mM IPTG. These were subsequently grown further for 24 h at 30°C.

Targeted genomic regions were PCR amplified from these cultures, which contain a mixed population of edited or non-edited cells, to verify the editing events by sanger sequencing. Through the analysis of sequencing chromatographs online (https://moriaritylab.shinyapps.io/editr_v10/), the editing efficiency of the mixed population could be obtained preliminarily. Further calculation of editing efficiency was determined by sequencing the targeted genomic regions of single colonies obtained after spreading the editing cultures on LB agar. Ten colonies were picked and the target region was amplified by PCR for sequencing. The editing efficiency at each protospacer position was established for each colony and the average between these colonies were used to represent the overall editing efficiency.

Plasmid curing. To obtain plasmid-free strains, edited *B. subtilis* containing pBAC-dCas9-AID-gRNAtarget^{TS} plasmids were incubated at 37°C in antibiotic-free LB

medium for 24 h. Subsequently, the cultures were diluted and spread on LB with and without chloramphenicol separately. After incubation at 30°C for 48 h, the loss rate of plasmids pBAC-dCas9-AID-gRNAtarget^{TS} was calculated by counting the number of colonies formed. The single colonies which had grown on LB plates without chloramphenicol were streaked on new LB plates and grown at 30°C with and without antibiotic to confirm plasmid loss.

Whole genome sequencing and analysis. Genomic DNA from base edited strains was extracted using the Wizard Genomic DNA Purification Kit (Promega (Beijing) Biotech Co., China). Library construction and genome sequencing was performed by Novogene (Beijing, China) using Illumina NovaSeq 6000 sequencing platform. Quality assurance of the output was analyzed using FastQC software (v.0.10.1) and NGSQC Toolkit software (v.2.3.3). BWA alignment software (v.0.7.17) and SAM tools software (v.1.9) were used for alignment and variant calling, respectively. Variations were annotated using the SnpEff software (v.4.3i). Through analysis with the cas-offinder online tool (http://www.rgenome.net/cas-offinder/), we can determine if potential off-target positions of each of the eight gRNAs coincide with the sequenced SNPs.

Table 1. Strains used in this study

Strains	Description	Source
E. coli DH5α	For plasmids construction	Laboratory stock
E. coli DB3.1	For plasmids construction	Laboratory stock
B. subtilis 168	Wide type strain	Laboratory stock
BS-dCas-AID-18	*B. subtilis* 168 containing plasmid pBAC-dCas9-AID-*amyE*-18-target	This study
BS-dCas-18	*B. subtilis* 168 containing plasmid pBAC-dCas9-*amyE*-18-target	This study
BS-*amyE*-16 to BS-*amyE*-20	*B. subtilis* 168 containing plasmid pBAC-dCas9-AID-*amyE*-16-target^{TS} to pBAC-dCas9-AID-*amyE*-20-target^{TS}	This study
BS-*5Cs*	*B. subtilis* 168 containing plasmid pBAC-dCas9-AID-*5Cs*-target^{TS}	This study
BS-*5Cs*-*scoC*-1-*amyE*-18	*B. subtilis* 168 containing plasmid pBAC-dCas9-AID-*5Cs*-*scoC*-1-*amyE*-18-3targets^{TS}	This study
BS-*mpr*-4-*nprB*-3-*vpr*-1-*wprA*-5	*B. subtilis* 168 containing plasmid pBAC-dCas9-AID-*mpr*-4-*nprB*-3-*vpr*-1-*wprA*-5-4targets^{TS}	This study
BS-*aprE*-1-*nprE*-1-*epr*-1-*bpr*-4	*B. subtilis* 168 containing plasmid pBAC-dCas9-AID-*aprE*-1-*nprE*-1-*epr*-1-*bpr*-4-4targets^{TS}	This study
BS△1-4Pro	*B. subtilis* 168 △*mpr*△*nprB*△*vpr*△*wprA*	This study
BS△1-8Pro	*B. subtilis* 168 △*mpr*△*nprB*△*vpr*△*wprA* △*aprE*△*nprE*△*epr*△*bpr*	This study

Table 2. Plasmids used in this study

Plasmids	Description	Source
pKSV7	For *repF* gene amplification	Laboratory stock
pBAC0041	Targeting *amyE* gene -18C position	[21]
pBAC-d/nCas9-AID-*amyE*-18-target	Targeting *amyE* gene -18C position	This study
pBAC-d/nCas9-*amyE*-18-target	Control plasmids targeting *amyE* gene -18C position	This study
pBAC-dCas9-AID-gRNA::ccdB^{TS}	Modified temperature-sensitive *ccdB* plasmid	This study
pBAC-dCas9-AID-*amyE*-X-target^{TS}	Targeting different regions of *amyE* gene	This study
pBAC-dCas9-AID-*5Cs*-target^{TS}	Targeting consecutive 5Cs from -16 to -20 in B. *subtilis* 168 genome	This study
pBAC-dCas9-AID-*5Cs*-*scoC*-1-*amyE*-18-3targets^{TS}	Targeting *amyE*, *scoC and 5Cs* simultaneously	This study
pBAC-dCas9-AID-*mpr*-4-*nprB*-3-*vpr*-1-*wprA*-5-4targets^{TS}	Targeting *mpr*, *nprB*, *vpr and wprA* 4 genes simultaneously	This study
pBAC-dCas9-AID-*aprE*-1-*nprE*-1-*epr*-1-*bpr*-4-4targets^{TS}	Targeting *aprE*, *nprE*, *epr* and *bpr* 4 genes simultaneously	This study

Table 3. Primers used for gRNAs design

Primers	Sequence
TGamyE-16-F	AAACGGCTCCAATGATTCGGATTT
TGamyE-16-R	ATGTAAATCCGAATCATTGGAGCC
TGamyE-17-F	AAACGGCATTGATCGTGCCTGTCA
TGamyE-17-R	ATGTTGACAGGCACGATCAATGCC
TGamyE-18-F	AAACCTTTACTGCCGTTATTCGCT
TGamyE-18-R	ATGTAGCGAATAACGGCAGTAAAG
TGamyE-19-F	AAACCAGGTCTGGAAAAGAAAAGA
TGamyE-19-R	ATGTCAGGTCTGGAAAAGAAAAGA
TGamyE-20-F	AAACTGATTCCCTTCCTTTACTTG
TGamyE-20-R	ATGTCAAGTAAAGGAAGGGAATCA
ApBsaAmyE18-F	GGTCTCCAAACCTTTACTGCCGTTATTCGCTACATTTATTGTACAACACGAGCCCATTTTT
ApBsascoC1hf-F	GGTCTCCCTTGCTAAGCTGAGACATTTATTGTACAACACGAGCCCATTTTTG
ApBsascoC1hf-R	GGTCTCGCAAGGCTCTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC
ApBsa5Cs-R	GGTCTCGATGTCCCCCGGGTCACATCCTTATGTTTTAGAGCTAGAAATAGC
ApBsahfnprB3-F	GGTCTCCCTTTTCACTTGTTTACATTTATTGTACAACACGAGCCCATTTTT
ApBsampr4-R	GGTCTCGATGTTCAGGTTTCCGCCCCTTATGGTTTTAGAGCTAGAAATAGC
ApBsahfvpr1-F	GGTCTCCCAATGCTGCCTCGTACATTTATTGTACAACACGAGCCCATTTTT
ApBsahfnprB3-R	GGTCTCGAAAGTTTCTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC
ApBsawprA5-F	GGTCTCCAAACTTTCCTGATTTCTTCTGATCACATTTATTGTACAACACGAGCCCATTTTT
ApBsahfvpr1-R	GGTCTCGATTGCATTTGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC

Primers	Sequence
ApBsahfnpr E1-F	GGTCTCCACAGGCGCATATCGACATTTATTGTACAACACGAGCCCATTTT TG
ApBsaaprE1 -R	GGTCTCGATGTGCAATTTAAGTATGTTAACGGTTTTAGAGCTAGAAATA GC
ApBsahfepr 1-F	GGTCTCCCGCTGCTGTTTTGCACATTTATTGTACAACACGAGCCCATTTT T
ApBsahfnpr E1-R	GGTCTCGAGCGAGAAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTC
ApBsabpr4- F	GGTCTCCAAACACCATCCTCATATGTCTGACACATTTATTGTACAACACG AGCCCATTTTT
ApBsahfepr 1-R	GGTCTCGAGCGAGAAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTC

■ AUTHOR INFORMATION

Corresponding Authors

*Email: wangmeng@tib.cas.cn (M.W.)

bi_ch@tib.cas.cn (C.B.)

Author Contributions

‖S.Y. and M.A.P. have contributed equally to this work. S.Y., M.A.P., Y.L and Y.G designed and performed the experiments and wrote the manuscript. M.W., C.B. and S.J.R oversaw the study. Y.W. and X.N performed the genome scale analysis and NGS data analysis. All authors edited the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

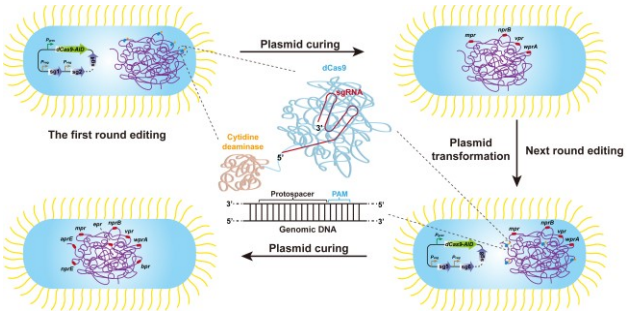
This work was supported by the National Key R&D Program of China (2018YFA0901900 and 2018YFA0902900), the National Natural Science Foundation of China (31970063), the Key Project of Chinese Academy of Sciences (QYZDB-SSW-SMC012), the Biological Resources Programme of Chinese Academy of Sciences (KFJ-BRP-009), a Newton Fund PhD placement program grant, ID 352639434, under the UK-China Joint Research and Innovation Partnership Fund, and a PhD studentship for M.A.P. from The Industrial Biotechnology Innovation Center (IBioIC).

■ REFERENCES

(1) Wang, T., et al., Natural products from Bacillus subtilis with antimicrobial properties. *Chinese Journal of Chemical Engineering* **2015,** *23* (4), 744-754.

(2) Widner, B., et al., Hyaluronic acid production in Bacillus subtilis. *Appl Environ Microbiol* **2005,** *71* (7), 3747-52.
(3) Ouattara, H. G., et al., Regulation of the synthesis of pulp degrading enzymes in Bacillus isolated from cocoa fermentation. *Food Microbiol* **2017,** *63*, 255-262.
(4) Song, Y., et al., High-Efficiency Secretion of beta-Mannanase in Bacillus subtilis through Protein Synthesis and Secretion Optimization. *J Agric Food Chem* **2017,** *65* (12), 2540-2548.
(5) Ling, M., et al., Combinatorial promoter engineering of glucokinase and phosphoglucoisomerase for improved N-acetylglucosamine production in Bacillus subtilis. *Bioresour Technol* **2017,** *245* (Pt A), 1093-1102.
(6) Chen, J., et al., High-level intra- and extra-cellular production of D-psicose 3-epimerase via a modified xylose-inducible expression system in Bacillus subtilis. *J Ind Microbiol Biotechnol* **2016,** *43* (11), 1577-1591.
(7) Chen, J., et al., A novel strategy for protein production using non-classical secretion pathway in Bacillus subtilis. *Microb Cell Fact* **2016,** *15*, 69.
(8) Westers, L., et al., Secretion of functional human interleukin-3 from Bacillus subtilis. *J Biotechnol* **2006,** *123* (2), 211-24.
(9) Phuong, N. D., et al., Production of XynX, a large multimodular protein of thermoanaerobacterium sp., by protease-deficient Bacillus subtilis strains [corrected]. *Appl Biochem Biotechnol* **2012,** *168* (2), 375-82.
(10) DiCarlo, J. E., et al., Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res* **2013,** *41* (7), 4336-43.
(11) Cobb, R. E., et al., High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system. *ACS Synth Biol* **2015,** *4* (6), 723-8.
(12) Shan, Q., et al., Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* **2013,** *31* (8), 686–688.
(13) Jiang, Y., et al., Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* **2015,** *81* (7), 2506-14.
(14) Mali, P., et al., RNA-guided human genome engineering via Cas9. *Science* **2013,** *339* (6121), 823-6.
(15) Westbrook, A. W., et al., Development of a CRISPR-Cas9 Tool Kit for Comprehensive Engineering of Bacillus subtilis. *Appl Environ Microbiol* **2016,** *82* (16), 4876-95.
(16) Zhang, K., et al., Multigene disruption in undomesticated Bacillus subtilis ATCC 6051a using the CRISPR/Cas9 system. *Sci Rep* **2016,** *6*, 27943.
(17) So, Y., et al., A Highly Efficient CRISPR-Cas9-Mediated Large Genomic Deletion in Bacillus subtilis. *Front Microbiol* **2017,** *8*, 1167.
(18) Altenbuchner, J., Editing of the Bacillus subtilis Genome by the CRISPR-Cas9 System. *Appl Environ Microbiol* **2016,** *82* (17), 5421-7.
(19) Burby, P. E.; Simmons, L. A., CRISPR/Cas9 Editing of the Bacillus subtilis Genome. *Bio Protoc* **2017,** *7* (8).
(20) Liu, D., et al., Development and characterization of a CRISPR/Cas9n-based multiplex genome editing system for Bacillus subtilis. *Biotechnol Biofuels* **2019,** *12*, 197.
(21) Price, M. A., et al., CRISPR-Cas9 In Situ engineering of subtilisin E in Bacillus subtilis. *PLoS One* **2019,** *14* (1), e0210121.
(22) Watzlawick, H.; Altenbuchner, J., Multiple integration of the gene ganA into the Bacillus subtilis chromosome for enhanced beta-galactosidase production using the CRISPR/Cas9 system. *AMB Express* **2019,** *9* (1), 158.
(23) Gaudelli, N. M., et al., Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **2017,** *551* (7681), 464-471.
(24) Komor, A. C., et al., Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **2016,** *533* (7603), 420-4.

(25) Liu, Z., et al., Improved base editor for efficient editing in GC contexts in rabbits with an optimized AID-Cas9 fusion. *FASEB J* **2019,** *33* (8), 9210-9219.
(26) Kim, Y. B., et al., Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol* **2017,** *35* (4), 371-376.
(27) Jiang, W., et al., BE-PLUS: a new base editing tool with broadened editing window and enhanced fidelity. *Cell Res* **2018,** *28* (8), 855-861.
(28) Tan, J., et al., Expanding the genome-targeting scope and the site selectivity of high-precision base editors. *Nat Commun* **2020,** *11* (1), 629.
(29) Zhao, Y., et al., Multiplex genome editing using a dCas9-cytidine deaminase fusion in Streptomyces. *Sci China Life Sci* **2019**.
(30) Nishida, K., et al., Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **2016,** *353* (6305).
(31) Wang, Y., et al., MACBETH: Multiplex automated Corynebacterium glutamicum base editing method. *Metab Eng* **2018,** *47*, 200-210.
(32) Kluesner, M. G., et al., EditR: A Method to Quantify Base Editing from Sanger Sequencing. *CRISPR J* **2018,** *1*, 239-250.
(33) Villafane, R., et al., Replication control genes of plasmid pE194. *J Bacteriol* **1987,** *169* (10), 4822-9.
(34) Choi, K. R.; Lee, S. Y., CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnol Adv* **2016,** *34* (7), 1180-1209.
(35) Banno, S., et al., Deaminase-mediated multiplex genome editing in Escherichia coli. *Nat Microbiol* **2018,** *3* (4), 423-429.
(36) Lim, H.; Choi, S. K., Programmed gRNA Removal System for CRISPR-Cas9-Mediated Multi-Round Genome Editing in Bacillus subtilis. *Front Microbiol* **2019,** *10*, 1140.
(37) Wang, Y., et al., Expanding targeting scope, editing window, and base transition capability of base editing in Corynebacterium glutamicum. *Biotechnol Bioeng* **2019,** *116* (11), 3016-3029.
(38) Kobayashi, K., et al., Essential Bacillus subtilis genes. *Proc Natl Acad Sci U S A* **2003,** *100* (8), 4678-83.
(39) Bae, S., et al., Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **2014,** *30* (10), 1473-5.
(40) Zetsche, B., et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **2015,** *163* (3), 759-71.
(41) Burstein, D., et al., New CRISPR-Cas systems from uncultivated microbes. *Nature* **2017,** *542* (7640), 237-241.
(42) Yang, H., et al., PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease. *Cell* **2016,** *167* (7), 1814-1828 e12.
(43) Price, M. A., et al., Expanding and understanding the CRISPR toolbox for Bacillus subtilis with MAD7 and dMAD7. *Biotechnol Bioeng* **2020**.
(44) Anagnostopoulos, C.; Spizizen, J., Requirements for Transformation in Bacillus subtilis. . *Journal of Bacteriology* **1961,** *81*, 741-746.



For Table of Contents Only