Development of a DNA double-strand break-free base editing tool in *Corynebacterium glutamicum* for genome editing and metabolic engineering

Chen Deng [a,b], Xueqin Lv [a,b], Jianghua Li [b], Yanfeng Liu [a,b], Guocheng Du [a], Long Liu [a,b,*]

[a] *Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi, 214122, China*
[b] *Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, Wuxi, 214122, China*

ARTICLE INFO

ABSTRACT

As a traditional amino acid producing bacterium, *Corynebacterium glutamicum* is a platform strain for production of various fine chemicals. Based on the CRISPR (Clustered regularly interspaced short palindromic repeats)-Cas9 system, gene editing tools that enable base conversion in the genome of *C. glutamicum* have been developed. However, some problems such as genomic instability caused by DNA double-strand break (DSB) and off-target effects need to be solved. In this study, a DSB-free single nucleotide genome editing system was developed by construction of a bi-directional base conversion tool TadA-dCas9-AID. This system includes cytosine base editors (CBEs): activation-induced cytidine deaminase (AID) and adenine deaminase (ABEs): tRNA adenosine deaminase (TadA), which can specifically target the gene through a 20-nt single guide RNA (sgRNA) and achieve the base conversion of C-T, C-G and A-G in the 28-bp editing window upstream of protospacer adjacent motif. Finally, as a proof-of-concept demonstration, the system was used to construct a mutant library of *zwf* gene in *C. glutamicum* S9114 genome to improve the production of a typical nutraceutical *N*-acetylglucosamine (GlcNAc). The GlcNAc titer of the mutant strain K293R was increased by 31.9% to 9.1 g/L in shake flask. Here, the developed bases conversion tool TadA-dCas9-AID does not need DNA double-strand break and homologous template, and is effective for genome editing and metabolic engineering in *C. glutamicum*.

1. Introduction

CRISPR (Clustered regularly interspaced short palindromic repeats)/Cas system is an acquired immune system that is widely used in bacteria and archaea to protect against foreign DNA invasion (Makarova et al., 2013; Sampson et al., 2013). With advantages of low cost, simple construction, convenient operation and high efficiency in mediating gene fixed-point knock-in, it has been widely used in various fields, including animal model construction, gene therapy, drug research and development, food technology and biological control, and has become the most popular genome editing tool (Garneau et al., 2010; Makarova et al., 2011; Wong et al., 2017; Yang et al., 2020). However, like ZFNs and TALENs, CRISPR/Cas9 gene editing relies on the generation of DNA double-strand break (DSB), which inevitably triggers non-homologous end joining (NHEJ) repair, causes random insertion-deletion mutations, and results in higher frequency unintended base changes or off-targets cutting (LaFountaine et al., 2015; Renaud et al., 2016). Although accurate repairs can be performed in homology directed repair (HDR) by providing templates, HDR is accompanied by higher probability insertion

or deletion mutations and off-target cuts. Therefore, CRISPR/Cas9 system is more suitable for gene knockout, and how to perform base editing without introducing double-strand breaks remains a challenge (Rose et al., 2018; Salsman and Dellaire, 2017).

To overcome the above challenge, researchers have made many explorations, among which the emergence of single base editing technology brings hope for accurate and efficient base replacement. David Liu group developed a single base editor (BE) by fusion of cytosine deaminase with CRISPR/Cas9, where spCas9 was fused with cytosine deaminase to achieve single base conversion from cytosine (C) to thymine (T) within a certain mutation window (Komor et al., 2016). In addition, Keiji Nishida et al. fused cytosine deaminase from lampreys with CRISPR/Cas9 and UGI (uracil glycosylase inhibitor) to achieve 15–55% targeted mutation in mammalian cells (Nishida et al., 2016). Chang et al. fused human-derived cytosine deaminase to the C-terminus of dCas9 (dCas9-AIDx) to generate local sequence diversity through mutagenesis (Ma et al., 2016). Although the BE system can accurately convert C into T in the mutation window, there is still a part of the task of how to mutate T to C or A to G in turn. In response to this challenge, David Liu group

---

established a new single-base editing system to achieve accurate conversion of A-G (adenine base editor, ABE) (Gaudelli et al., 2017; Rees and Liu, 2018). ABE tool can convert adenine at positions 12 to 17 upstream of NGG into inosine through deamination. Inosine can be paired with cytosine, and is regarded as guanine at DNA level for code reading and replication, thus realizing mutation from A to G within the mutation window. These two single-base editing systems greatly enrich the toolbox for precise genetic modification.

Corynebacterium glutamicum is an important strain for industrial production of amino acids and is widely used in medicine, food and feed industries (Cho et al., 2017; Deng et al., 2019a; Mei et al., 2016; Zhang et al., 2017). Compared with the model strains *Escherichia coli* and *Saccharomyces cerevisiae*, the genetic modification technology of *C. glutamicum* is relatively lagging (Liu et al., 2017). Though several genome editing technologies based on CRISPR/Cas9 or Cpf1 in

C. glutamicum have been developed, these technologies are difficult to obtain mutant libraries of target genes (Cho et al., 2017; Jiang et al., 2017; Liu et al., 2017). In recent years, researches have also developed some base editing tools for *C. glutamicum*. Wang et al. fused activation-induced cytidine deaminase (AID) with Cas9 variants to achieve efficient C-T conversion on the genome of *C. glutamicum* ATCC13032 (Wang et al., 2018). Then they fused tRNA adenosine deaminase from *E. coli* (TadA) with four Cas9 variants providing 3.9-fold more target loci for gene inactivation in *C. glutamicum* ATCC13032 (Wang et al., 2019). But there is no research report that combines the two editing tools. Recently, we have used *C. glutamicum* S9114 to produce *N*-acetylglucosamine (GlcNAc) (Deng et al., 2019b), which is an effective medicine for the treatment of bone and joint diseases and can effectively alleviate bone and joint inflammation (Chen et al., 2012a, 2012b; Liu et al., 2013). However, the base editing technology of the strain *C. glutamicum* S9114

Table 1
Primers used in this study.

Primer	Sequence (5'→3')
DAF	ACAATTTCACACAGGAAACAGAATTAATTAAAAGGAGGACAACTAATGGATAAAAAGTATTCCA
DAR	CCGCCAAAACAGCCAAGCTGTCACAGGCCCAGGGTGCG
D10AF	GGCCTGGCTATCGGCACCAATTC
D10AR	GCCGATAGCCAGGCCAATGGAATAC
H840AF	GTGGACGCCATCGTCCCTCAGTC
H840AR	GACGATGGCGTCCACATCGTAATCAGAG
VectorsgRNA.F	CATTTCGAAGTGAGTTAGCGCGAATTGATCTGGTTTG
VectorsgRNA.R	TTATTGGTGCCCTTCGAAGAAGCCGCACGT
FragmentsgRNA.F	CGGTCATTTCGAAGTGAGTTAGCGCG
FragmentsgRNA.R	GGCAGTTATTGGTGCCCTTCGAAGAAGCC
sg1F	CCGGAATTCGGATTCTCGTTGGTAGGTTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
sg2F	CCGGAATTCGGAATTCCGTGAAAATGTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
sg3F	CCGGAATTCGTTCCGCATCGACCACTATTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
sg4F	CCGGAATTCGCTGCAGGCAGAAAAGATCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
sgR	TGCTCTAGAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAG
zwfseq F	CGGCGACATGAAATCGAATTAGTTC
zwfseq R	GGGTGGTGGTATCCGGAAG
cas9F	TCCAGACAATTCCGACGTGGACAAG
cas9R	GGTTGGTTTTAAACAGGAGGTCGACG
J23100F	TTGACGGCTAGCTCAGTCCTAGGTACAGTGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23102F	TTGACAGCTAGCTCAGTCCTAGGTACTGTGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23104F	TTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23105F	TTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23108F	CTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23111F	TTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23113F	CTGATGGCTAGCTCAGTCCTAGGGATTATGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23118F	TTGACGGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
J23119F	TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTAGAAGAGCGCTCTTCAGTTTTAGAGCTAGAAATAGCAAG
JR	ATAAAACGAAAGGCCCAGTCT
F23100F	TGACAGCTTATCATTTGACGGCTAGCTCAGTCC
F23102F	TTTGACAGCTTATCATTTGACAGCTAGCTCAGT
F23104F	ACAGCTTATCATTTGACAGCTAGCTCAGTC
F23105F	GACAGCTTATCATTTTACGGCTAGCTCAGTC
F23108F	ACAGCTTATCATCTGACAGCTAGCTCAGTCC
F23111F	TTGACAGCTTATCATTTGACGGCTAGCTCAGTC
F23113F	ACAGCTTATCATCTGATGGCTAGCTCAGTCC
F23118F	ACAGCTTATCATTTGACGGCTAGCTCAGTCC
F23119F	AGCTTATCATTTGACAGCTAGCTCAGTC
F231R	GAGCGTTCACCGACAAACAACAGATAAAACGAAAGGC
V231F	CCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTC
V23100R	GCTAGCCGTCAAATGATAAGCTGTCAAACCAGAT
V23102R	AGCTAGCTGTCAAATGATAAGCTGTCAAACCAGAT
V23104R	CTAGCTGTCAAATGATAAGCTGTCAAACCA
V23105R	GCTAGCCGTAAAATGATAAGCTGTCAAACCA
V23108R	AGCTAGCTGTCAGATGATAAGCTGTCAAACCAG
V23111R	CTAGCCGTCAAATGATAAGCTGTCAAACC
V23113R	GCTAGCCATCAGATGATAAGCTGTCAAACCA
V23118R	GAGCTAGCCGTCAAATGATAAGCTGTCAA
V23119R	AGCTAGCTGTCAAATGATAAGCTGTCAAACCA
zwfsg4F	**AGT**GCTGCAGGCAGAAAAGATCA
zwfsg4R	**AACT**GATCTTTTCTGCCTGCAGC
FtadaF	TCCGGATTATGCAATGAAACGGACAGCCGACGGAAG
FtadaR	ATACTTTTTATCCATTGACCCCCCGCTG
VtadaF	GGGCAGCAGCGGGGGGGTCAATGGATAAAAAGT
VtadaR	GCTGTCCGTTTCATTGCATAATCCGGAAC

has not been reported yet. In this work, we constructed a base conversion tool that can simultaneously achieve A-G and C-T/G in *C. glutamicum* S9114 strain. This new method can diversify specific DNA sequences in the cell and realize multiple mutations of single nucleotides in the genome. This method is also simple to operate and can introduce a variety of single base transformations into the gene sequence. Moreover, a method for rapidly ligating single guide RNA (sgRNA) targeting 20 nt protospacer (20N) on the expression vectors was also developed, and this greatly improves the efficiency of linking various sgRNA target sequences to vectors. Finally, by single-base editing of *zwf* gene in *C. glutamicum* S9114 genome, mutant strains capable of producing higher GlcNAc titer were obtained. Overall, we provide a good base conversion tool for genome editing of *C. glutamicum*, and this will be useful for the metabolic engineering of *C. glutamicum* for production of various fine chemicals and nutraceuticals.

2. Materials and methods

2.1. Bacterial strains, plasmids, and materials

The primers used in this study are listed in Table 1. All strains and plasmids used in this study are listed in Table 2. *E. coli* JM109 was used as the host for plasmid construction and was cultured on LB medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl). PrimeSTAR HS DNA polymerase, restriction endonucleases, PCR reagents, Genomic Extraction Kit, and DNA purification kit were purchased from Takara (Dalian, China). T4 DNA Ligase Kit was purchased from Thermo Fisher Scientific (Shanghai, China). Restriction Enzymes were purchased from New England Biolabs (Beijing, China). ClonExpress II, One Step Cloning Kit was purchased from Vazyme (Nanjing, China). Annealing Buffer for DNA Oligos (5X) was purchased from Beyotime Biotechnology (Shanghai, China).

2.2. Plasmid construction

The pXMJ19 plasmid was used as a skeleton structure for the plasmid construction in the single base editing system. Codon optimization and gene synthesis were carried out using the *Streptococcus pyogenes* Cas9 gene provided on NCBI as a template. At the same time, *Hin*d III restriction site and SD sequence: GCTAGAAAGAAGGAGGACCCGAC (derived from pJYW-4-ceN plasmid (Deng et al., 2019a) were added at the 5′ end of the Cas9 gene. At the same time, a linker composed of SGGS, *Homo sapiens* activation-induced cytidine deaminase gene and restriction site *Eco*R I were added at 3′ end of the Cas9 gene. The whole genetic framework Cas9-linker-AID was synthesized by GENEWIZ biotechnology co., ltd (Suzhou, China), and the synthetic gene was PCR-amplified with primer DAF/DAR to obtain a Cas9-AID fusion expression cassette with a double restriction site *Hin*d III/*Eco*R I.

The plasmid pXMJ19 was subjected to the same double digestion treatment, and the restriction enzyme digestion system contained 16 μl of fragment DNA, 1 μl of *Hin*d III, 1 μl of *Eco*R I, and 2 μl of 10 × H buffer, and double distilled water was added to a total volume of 20 μl. The 1% agarose gel electrophoresis was carried out to recover the digested product. The pXMJ19-Cas9-AID recombinant plasmid was confirmed by DNA sequencing. The obtained recombinant vector pXMJ19-cas9-AID was used as a template, and two pairs of primers (D10AF/D10AR and H840AF/H840AR) were designed. PCR conditions were: pre-denaturation at 98 °C for 3 min; denaturation at 98 °C for 30 s; annealing at 55 °C for 30 s; extension at 72 °C for 5 min. The reaction was carried out for 30 cycles and the final extension was 72 °C for 10 min. The PCR product was recovered using a DNA purification kit, and the Cas9 gene was mutated to dCas9 which would lose cleavage activity but still retain DNA binding ability by site-directed mutagenesis, and the recombinant vector pXMJ19-dcas9-AID was confirmed by DNA sequencing.

Construction of sgRNA expression framework: the extracted plasmid pXMJ19-dcas9 was used as template, and primers VectorsgRNA.F/

Table 2
Strains and plasmids used in this study.

Name	Description	Source
Strains		
C. glutamicum S9114 ΔnagA-ΔgamA-Δldh	*C. glutamicum* S9114 derivate: ΔnagA-ΔgamA-Δldh	Laboratory stock
E. coli JM109		Laboratory stock
M1	*C. glutamicum* S9114 ΔnagA-ΔgamA-Δldh embedding *zwf* mutant P286L	This work
M2	*C. glutamicum* S9114 ΔnagA-ΔgamA-Δldh embedding *zwf* mutants P286L, A287G	This work
M3	*C. glutamicum* S9114 ΔnagA-ΔgamA-Δldh embedding *zwf* mutant P286S	This work
M4	*C. glutamicum* S9114 ΔnagA-ΔgamA-Δldh embedding *zwf* mutant L290stop	This work
M5	*C. glutamicum* S9114 ΔnagA-ΔgamA-Δldh embedding *zwf* mutants P286L, K293E	This work
M6	*C. glutamicum* S9114 ΔnagA-ΔgamA-Δldh embedding *zwf* mutant K293R	This work
Plasmids		
pJYW-4-ceN	pJYW-4 derivate, Ptac-*GNA1*,Kan^r	Laboratory stock
pXMJ19	Cloning vector, Cm^r	Laboratory stock
pXMJ19-Cas9-AID	pXMJ19 derivate with Cas9 and AID cloned	This work
pXMJ19-dCas9-AID	pXMJ19-Cas9-AID with D10A and H840A mutation sites	This work
pFST-porb	pFST with sgRNA expression framework cloned	Laboratory stock
pXMJ19-dCas9 1	pXMJ19-dCas9-AID with sgRNA expression framework cloned	This work
pXMJ19-dCas9 1-*zwf*1	pXMJ19-dCas9 1 with PAM1 of gene *zwf*	This work
pXMJ19-dCas9 1-*zwf*2	pXMJ19-dCas9 1 with PAM2 of gene *zwf*	This work
pXMJ19-dCas9 1-*zwf*3	pXMJ19-dCas9 1 with PAM3 of gene *zwf*	This work
pXMJ19-dCas9 1-*zwf*4	pXMJ19-dCas9 1 with PAM4 of gene *zwf*	This work
pMD19-T		Laboratory stock
pMD19-T Vector23100	pMD19-T derivate with sgRNA framework gene under the control of Pj23100	This work
pMD19-T Vector23102	pMD19-T derivate with sgRNA framework gene under the control of Pj23102	This work
pMD19-T Vector23104	pMD19-T derivate with sgRNA framework gene under the control of Pj23104	This work
pMD19-T Vector23105	pMD19-T derivate with sgRNA framework gene under the control of Pj23105	This work
pMD19-T Vector23108	pMD19-T derivate with sgRNA framework gene under the control of Pj23108	This work
pMD19-T Vector23111	pMD19-T derivate with sgRNA framework gene under the control of Pj23111	This work
pMD19-T Vector23113	pMD19-T derivate with sgRNA framework gene under the control of Pj23113	This work
pMD19-T Vector23118	pMD19-T derivate with sgRNA framework gene under the control of Pj23118	This work
pMD19-T Vector23119	pMD19-T derivate with sgRNA framework gene under the control of Pj23119	This work
pXMJ19-dCas9 2-23100	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23100	This work
pXMJ19-dCas9 2-23102	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23102	This work
pXMJ19-dCas9 2-23104	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23104	This work

Table 2 (*continued*)

Name	Description	Source
pXMJ19-dCas9 2-23105	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23105	This work
pXMJ19-dCas9 2-23108	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23108	This work
pXMJ19-dCas9 2-23111	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23111	This work
pXMJ19-dCas9 2-23113	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23113	This work
pXMJ19-dCas9 2-23118	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23118	This work
pXMJ19-dCas9 2-23119	pXMJ19-dCas9 1 derivate with sgRNA framework gene under the control of Pj23119	This work
pXMJ19-dCas9 2-23100-P	pXMJ19-dCas9 2-23100 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23100-P	pXMJ19-dCas9 2-23100 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23102-P	pXMJ19-dCas9 2-23102 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23104-P	pXMJ19-dCas9 2-23104 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23105-P	pXMJ19-dCas9 2-23105 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23108-P	pXMJ19-dCas9 2-23108 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23111-P	pXMJ19-dCas9 2-23111 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23113-P	pXMJ19-dCas9 2-23113 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23118-P	pXMJ19-dCas9 2-23118 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 2-23119-P	pXMJ19-dCas9 2-23119 derivate with 20N sequence and PAM site AGG	This work
pXMJ19-dCas9 3	pXMJ19-dCas9 2-23119-P derivate with TadA-TadA mutant cloned	This work
pTYW-4-ceN-C.glglmS	pJYW-4-ceN derivate with C.glglmS cloned, Kanʳ	Laboratory stock

Cmʳ, chloramphenicol resistance; Kanʳ, kanamycin resistance.

VectorsgRNA.R were used to linearize plasmid pXMJ19-dcas9. PCR conditions were: denaturation at 98 °C for 30s, annealing at 55 °C for 30s, extension at 72 °C for 5min, reaction for 30 cycles, extension for 10 min at 72 °C and reaction for 30 cycles, and finally, extension for 10 min at 72 °C. PCR products were recovered by DNA purification kit to obtain the linearized pXMJ19-dcas9-AID plasmid. Using the primer Fragmentsgrna.F/Fragmentsgrna.R and plasmid pFST-porb as template, PCR was carried out according to the following conditions: denaturation at 98 °C for 30s; Annealing at 55 °C for 30s; Extension at 72 °C for 1min, reaction for 30 cycles; Extension at 72 °C for 10min. The PCR product was recovered by DNA purification kit and the entire sgRNA expression framework gene was amplified from plasmid pFST-porb.

Ligation was carried out using the ClonExpress II One Step Cloning Kit, and the linearized vector obtained by PCR and the target gene fragment with the vector homologous end were recovered and mixed in a molar ratio of 3:1.5 x CE II Buffer 4 μl, Exnase II 2 μl, and ddH₂O was added to make the total volume of the ligation system to 20 μl. The reaction was carried out at 37 °C for 30 min and lowered to 4 °C for incubation. Then, 10 μl of the ligation system was transformed to *E. coli* JM109 competent cells (The preparation method of competent cells is described in the instructions of the Takara *E. coli* competent kit). The correct transformants of the colony PCR were selected for sequencing verification, and the recombinant expression vector pXMJ19-dcas9-AID-sgRNA (pXMJ19-dCas9 1) was obtained, and a single-base editing plasmid of all-in-one was constructed.

2.3. sgRNA design

For dCas9 system, it is necessary to design a RNA oligonucleotide containing 20 base pairs, which can target any genome site and is an important part of CRISPR Cas9 system. Many factors need to be considered when designing gRNA sequences. In order to find a suitable gRNA sequence, the online gRNA design tool (https://chopchop.cbu.uib.no/) was used to design the gRNA of the target gene *zwf*. The NGG sequence in which the first bit of protospacer adjacent motif (PAM) in gRNA library is A, T, C, and G was selected as PAM sequence. Using plasmid pFST-porb as template and primers sg1F/sgR, sg2F/sgR, sg3F/sgR, sg4F/sgR, PCR amplification was carried out to obtain four sgRNA containing 20N sequence. PCR products were recovered by 2% agarose gel, and the same volume of isopropanol was added to mix in the sample. The recovered product and plasmid pXMJ19-dCas9 1 were digested with *Eco*R I and *Xba*I for 2h, respectively, and then agarose gel was recovered. The T4 DNA Ligase was used to mix the fragments with the vector in proportion and incubate at 22 °C for 10 min, and 10 μl of the ligation products were transformed to *E. coli* JM109.

2.4. Single base editing in C. glutamicum S9114

The recombinant plasmids pXMJ19-dCas9 1-*zwf*1, pXMJ19-dCas9 1-*zwf*2, pXMJ19-dCas9 1-*zwf*3, pXMJ19-dCas9 1-*zwf*4 with four 20-nt sgRNA of the target gene *zwf*, which have been correctly sequenced, are separately transformed into the *C. glutamicum* S9114 Δ*nagA*-Δ*gamA*-Δ*ldh* competence constructed in the previous study (Deng et al., 2019b), and then transferred to the LBB (Luria–Bertani broth with 18.5 g/L brain heart infusion) plate with 10 μg/μl chloramphenicol and 0.02 mM isopropyl-β-D-thiogalactoside (IPTG). After a large number of colonies were grown for 48 h, some colonies were selected, and primers zwfseqF/R were designed to carry out high-throughput sequencing of gene sequences near genome mutation sites to verify mutation effect.

Elimination of pXMJ19-dCas9 1-*zwf* plasmid: the strain containing pXMJ19-dCas9 1-*zwf* was cultured in 5 ml LBB medium at 30 °C until OD₅₆₂ = 1, and 0.01 mM IPTG was added for 10 h. The bacterial liquid was streaked onto a plate without antibiotic and IPTG for overnight culture. Primer cas9F/cas9R was used to verify whether the plasmid remained in the single colony. At the same time, they were transferred to antibiotic-free plate and chloramphenicol (10 μg/ml) resistant plate respectively to verify the plasmid elimination effect. The strain that cannot grow on chloramphenicol resistant plates but can grow on antibiotic-free plates is the mutant strain that eliminates pXMJ19-dCas9 1-*zwf* plasmid.

2.5. Promoter optimization of sgRNA

Using recombinant plasmids pXMJ19-dCas9 1-*zwf*4 as templates, primers J23100F/JR, J23102F/JR, J23104F/JR, J23105F/JR, J23108F/JR, J23111F/JR, j23111 F/JR, J23118F/JR, and J23119F/JR were used to amplify sgRNA fragments with pJ231 series plasmids, and the fragments were ligated to pMD19-T vector after recovery and purification. T vector ligation method was as follows: pMD19-T Vector 1 μl, target fragment 2 μl, ddH₂O 2 μl, then 5 μl of Solution I, reacted at 16 °C for 30 min, and then transformed into *E. coli* JM109. The correct recombinant plasmids were named pMD19-T Vector23100, pMD19-T Vector23102, pMD19-T Vector23104, pMD19-T Vector23105, pMD19-T Vector23108, pMD19-T Vector23111, pMD19-T Vector23113, pMD19-T Vector23118, and pMD19-T Vector23119. Using pMD19-T Vector23100, pMD19-T Vector23102, pMD19-T Vector23104, pMD19-T Vector23105, pMD19-T Vector23108, pMD19-T Vector23111, pMD19-T Vector23113, pMD19-T Vector23118, pMD19-T Vector23119 as templates, primers F23100F/F231R, F23102F/F231R, F23104F/F231R, F23105F/F231R, F23108F/F231R, F23111F/F231R, F23113F/F231R, F23118F/F231R, F23119F/F231R were used for introducing a series of different promoter sequences upstream sgRNA fragment via PCR reaction.

Then ClonExpress II One Step Cloning Kit was adopted to respectively connect the fragments with corresponding expression vectors pXMJ19-dcas9-sgRNA to obtain a series of plasmids reconstructed with various sgRNA promoters: pXMJ19-dCas9 2-23100, pXMJ19-dCas9 2-23100, pXMJ19-dCas9 2-23102, pXMJ19-dCas9 2-23104, pXMJ19-dCas9 2-23105, pXMJ19-dCas9 2-23108, pXMJ19-dCas9 2-23111, pXMJ19-dCas9 2-23113, pXMJ19-dCas9 2-23118, pXMJ19-dCas9 2-23119.

2.6. Development of efficient mutation method for targeting sequence 20N

Plasmids pXMJ19-dCas9 2-23100, pXMJ19-dCas9 2-23100, pXMJ19-dCas9 2-23102, pXMJ19-dCas9 2-23104, pXMJ19-dCas9 2-23105, pXMJ19-dCas9 2-23108, pXMJ19-dCas9 2-23111, pXMJ19-dCas9 2-23113, pXMJ19-dCas9 2-23118 and pXMJ19-dCas9 2-23119 were respectively digested with *Sap*I restriction enzymes. The synthetic reaction system of DNA double-stranded complex with cohesive terminus contained: Annealing Buffer for DNA Oligos (5X) 10 μl, 20 μl primer *zwf*sg4F, and 20 μl primer *zwf*sg4R. The reaction conditions were as follows: keep 98 °C for 2 min, and then the temperature was reduced to 4 °C at a speed of 0.1 °C/s, and finally, the temperature was kept at 4 °C for 2 min. PCR instrument can be used for the reaction in this process. The double-stranded DNA fragments with cohesive terminus and linearized plasmids digested by *Sap*I restriction enzyme were ligated.

The ligation method was as follows: 1 μl of T4 ligase, 2 μl of T4 buffer, linearized plasmid pXMJ19-dCas9 2-23119 100 ng, diluted 10 times of DNA double-stranded fragment 1 μl, and ddH$_2$O was added to a total volume of 20 μl. The mixed system was then incubated for 10 min at 22 °C. Then 10 μl of the ligation system was transferred to *E. coli* JM109. A series of recombinant plasmids with 20N sequence and PAM site AGG were obtained: pXMJ19-dCas9 2-23100-P, pXMJ19-dCas9 2-23100-P, pXMJ19-dCas9 2-23100-P, pXMJ19-dCas9 2-23100-P, pXMJ19 -dCas9 2-23100-P, pXMJ19-dCas9 2-23100-P, pXMJ19-dCas9 2-23100-P, pXMJ19-dCas9 2-23100-P, pXMJ19-dCas9 2-23119-P.

2.7. Construction of multidirectional base editing plasmid containing both TadA and AID

The fusion framework TadA-TadA mutant-3xGGS-XTEN-3xGGS composed by TadA and the TadA mutant and its intermediate linker were synthesized according to the full DNA sequence of the plasmid NG-ABEmax provided at http://www.addgene.org/124163/. The homologous arms sequences of the vector pXMJ19-dCas9 2-23119-P were added to both ends of the synthesized gene TadA-TadA mutant-3xGGS-XTEN-3xGGS using the primer FtadaF/FtadaR. The primers VtadaF/VtadaR were used to linearize the pXMJ19-dCas9 2-23119-P plasmid. The linearized plasmid and the TadA-TadA mutant-3xGGS-XTEN-3xGGS fusion gene frame carrying the vector homology arms were ligated with T4 ligase to construct the recombinant vector pXMJ19-dCas9 3, which contained both TadA and AID. Similarly, plasmid pXMJ19-dCas9 3 was transformed into *C. glutamicum* S9114 for base editing experiments, and different transformants were selected for high throughput sequencing to verify base editing effects.

2.8. Application of base editing tool for improving GlcNAc production

After high-throughput sequencing, six strains producing missense mutations and nonsense mutations, M1, M2, M3, M4, M5, M6, were selected from the mutant library obtained by base editing of *zwf* gene in *C. glutamicum* strain S9114. They were prepared as *C. glutamicum* competent cells, and the recombinant expression vector pTYW-4-ceN-*C.glglmS* constructed in the previous work was transformed into these competent cells (Deng et al., 2019b).

The seed culture solution stored in the glycerol tube was inoculated into LBB broth. The seed medium contained the following components (g/L): glucose, 25.0; corn syrup, 20.0; KH$_2$PO$_4$, 1.0; (NH$_4$)$_2$SO$_4$, 0.5; and urea, 1.25. The fermentation medium contained the following components (g/

L): glucose, 100.0; corn syrup, 10.0; KH$_2$PO$_4$, 1.0; (NH$_4$)$_2$SO$_4$, 20.0; MgSO$_4$, 0.5; CaCO$_3$, 20.0; and FeSO$_4$, 0.18. The final pH of the seed and fermentation medium was adjusted to 7.0. A total of 25 mg/L kanamycin was added to all media for detection of transformants or for recombinant culture. A single colony grown on the LBB plate was inoculated into the seed culture medium. After 16–18 h of culture, the bacterial culture medium of the seed culture was inoculated into the fermentation medium at an initial OD$_{562}$ of 1.6 and grown for 72 h. After centrifugation at 10 000×*g* for 10 min, the supernatant was used to determine GlcNAc titer.

2.9. Quantification of cell density and GlcNAc

The OD$_{562}$ of the fermentation broth was determined by UV–visible spectrophotometer (UVmini-1240, Shimadzu Corporation, Japan), and the calculation formula based on cell OD$_{562}$ and dry cell weight (DCW) was: DCW = $0.6495 \times$ OD$_{562}$–2.7925 (Yin et al., 2013). The concentrations of GlcNAc in the fermentation broth were measured with high-performance liquid chromatography (HPLC) equipped with an HPX-87H column (Bio-Rad) and a refractive index detector. HPLC was carried out with 5 mM H$_2$SO$_4$ as the mobile phase with a flow rate of 0.6 ml/min at 35 °C (Gu et al., 2017).

2.10. Statistical analysis

All data were expressed as mean ± SD. Differences were determined by 2-tailed Student's t-test between two groups, or one-way ANOVA followed by post-hoc Tukey's test for multiple groups. Statistical significance is indicated as * for $p < 0.05$ and ** for $p < 0.01$.

3. Results

3.1. Construction of cytosine deaminase base editing tool in C. glutamicum

In this study, pXMJ19 plasmid was used as the framework structure, and ptac-dCas9-AID expression cassette and Ptrc-sgRNA expression cassette were inserted into the pXMJ19 plasmid to express dCas9-AID fusion protein and sgRNA (Fig. 1A), respectively.

Avoiding the off-target effect of Cas9 is a critical step in the design of sgRNA. In this study, online gRNA design tool (https://chopchop.cbu.uib.no/) was used to design the target sequence. The selected target gene was *zwf*, which controls the major competition reactions of GlcNAc synthesis (pentose phosphate pathway) (Gu et al., 2019). Four PAM sequences with the highest efficiency and the first bases A, T, C and G were screened out as binding sequences. The results show that only when PAM sequence was AGG, the mutation efficiency was 4% at C6 position, while if PAM sequence was the other sequence, the mutation efficiency was almost zero (Fig. 1B).

To further optimize the single-base editing tool in *C. glutamicum*, we analyzed the DNA sequence of plasmid pXMJ19-dCas9 1 and found that a 46 bp redundant sequence existed in the downstream of the Ptrc promoter transcription start site. This partial sequence will transcribe before the sgRNA and affect the secondary structure of sgRNA (Fig. 1A). To investigate the effect of the extra 46 bp on the secondary structure of sgRNA, we used the online tool Mfold to predict the secondary structure of sgRNA (http://unafold.rna.albany.edu/?q=mfold) (Wiese and Hendriks, 2006).

The RNA structure prediction shows that the RNA transcribed from the 46 bp sequence of this segment was likely to be complementary to the mRNA formed from the 20N sequence to form double-stranded RNA, which made the 20N sequence unable to combine with the genome because it is hidden, resulting in lower mutation efficiency (Fig. 1D). Only when the 20N sequence was taken as the transcription initiation sequence, the probability of being hidden will be reduced (Fig. 1E). Therefore, we used Pj231 series promoters, which have been proved to be suitable for sRNA synthesis in *C. glutamicum*, to avoid the formation of the redundant sequence (Liu et al., 2017; Noh et al., 2019).
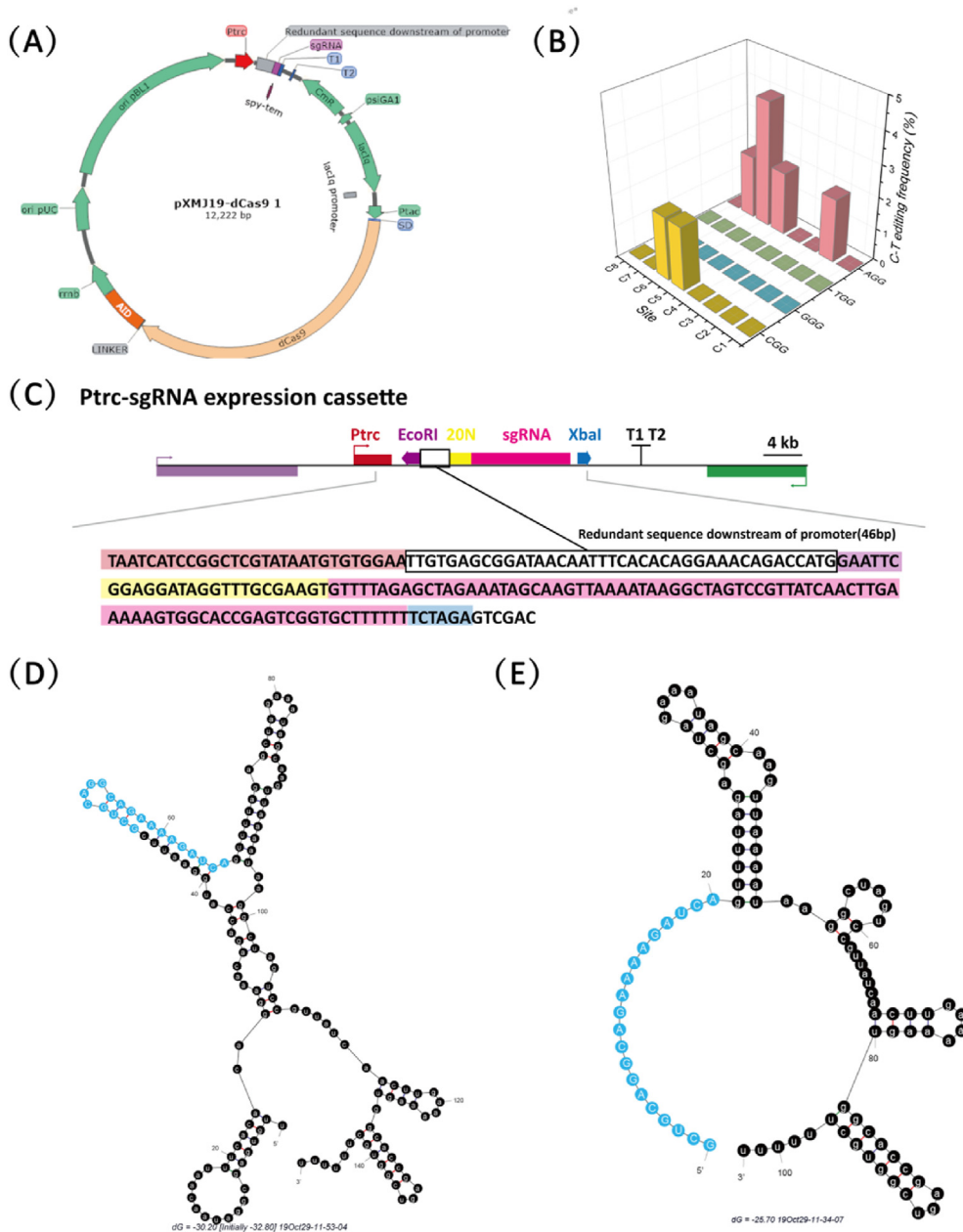
Fig. 1. The base editing effect using pXMJ19-dCas9 1 plasmid in *C. glutamicum*. (A) Structure and schematic diagram of plasmid pXMJ19-dCas9 1 used for single base editing. (B) Single base editing efficiency of different 20 PAMS. (C) Structure and sequence of the Ptrc-sgRNA expression cassette. (D) Secondary structure of the gRNA formed by the excess sequence downstream of the Ptrc promoter with 20N and the sgRNA scaffold. The base of the 20 nt target sequence is highlighted by a blue sphere. (E) 20 nt target sequence and the secondary structure of the gRNA formed by the downstream sgRNA scaffold. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Promoter optimization of sgRNA in cytosine single base editing tool and development of high efficiency 20N target linking strategy

To remove the effect of redundant sequences on sgRNA, we selected a series of constitutive promoters including Pj23100, Pj23102, Pj23104, Pj23105, Pj23108, Pj23111, Pj23113, Pj23118, and Pj23119 to replace the original Ptrc promoter, yielding a series of recombinant single base editing vectors: pXMJ19-dCas9 2-23100, pXMJ19-dCas9 2-23102, pXMJ19-dCas9 2-23104, pXMJ19-dCas9 2-23105, pXMJ19-dCas9 2-23108, pXMJ19-dCas9 2- 23111, pXMJ19-dCas9 2-23113, pXMJ19-dCas9 2-23118, and pXMJ19-dCas9 2-23119, respectively. The relative strengths of Pj23100, Pj23102, Pj23104, Pj23105, Pj23108, Pj23111, Pj23113, Pj23118, and Pj23119 were 1, 0.86, 0.72, 0.24, 0.51, 0.58, 0.01, 0.56 and 1, respectively (http://parts.igem.org/Promoters/Catalog/Anderson).

In addition, long primers needed to be synthesized when the target gene specific binding sequence 20N was linked to the vector in the previous studies (Doench et al., 2014). Moreover, difficulties in gel extraction of short DNA sequence, low recovery rate and long manipulation time of using restriction endonuclease constrained the efficiency in the whole process. Therefore, we designed a method for rapid ligation of the vector with the specific sequence 20N in *C. glutamicum*. We added Type IIS Restriction Enzymes site-*Spa*I, which recognize asymmetric DNA sequences and cleave outside their recognition sequences, upstream and downstream of the 20N sequence that needs to be inserted into the vector (Arakawa, 2016; Khan et al., 2017; Pan et al., 2014). In addition, the corresponding double-stranded fragment was obtained by fusing a reverse complementary primer with the 20N sequence (Fig. 2). A linearized vector containing two cohesive ends was obtained by digesting the vector with *Sap*I. The linearized vector with cohesive terminus and DNA double-stranded complex obtained by PCR reaction were ligated with T4 ligase, and after 10 min, it was transformed into *E. coli* JM109 to obtain 20N replaced recombinant plasmid (Fig. 2).
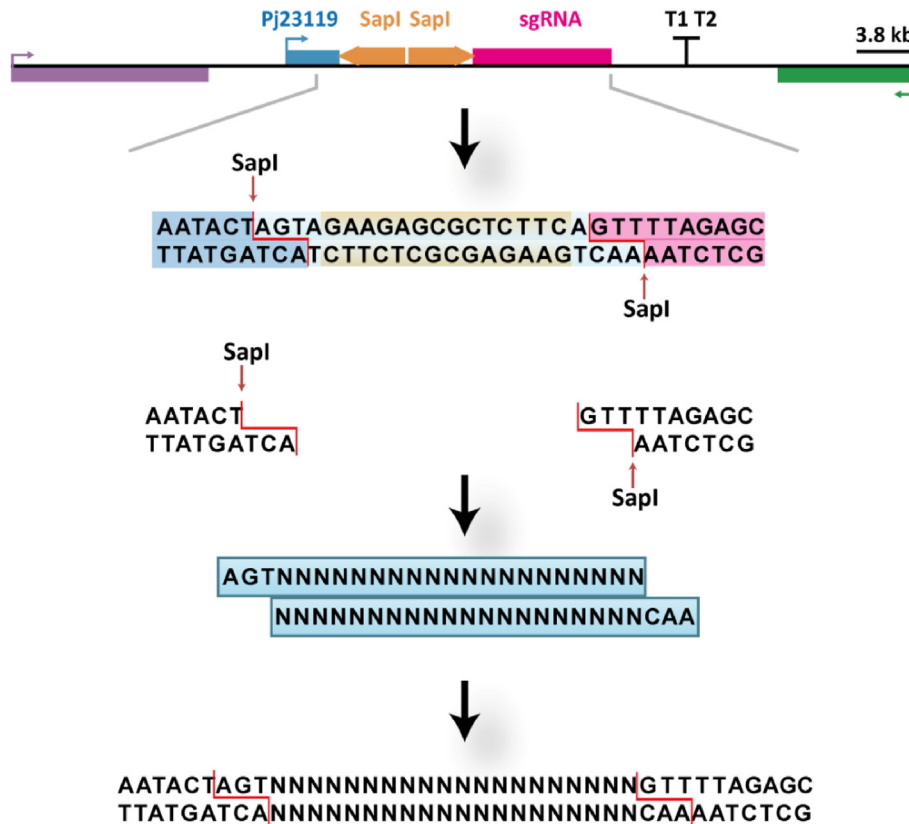
Fig. 2. Type IIS Restriction Enzymes *Sap*I-mediated efficient replacement strategy for 20 nt target sequence. The red dotted line portion is the *Sap*I restriction site. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.3. Different promoters for controlling expression of sgRNA led to different base mutation effects

The 20N target sequence high-efficiency ligation method was used to connect the PAM site AGG with the highest mutation efficiency together with the upstream 20N sequence to vectors pXMJ19-dCas9 2-23100, pXMJ19-dCas9 2-23102, pXMJ19-dCas9 2-23104, pXMJ19-dCas9 2-23105, pXMJ19-dCas9 2-23108, pXMJ19-dCas9 2-23111, pXMJ19-dCas9 2-23113, pXMJ19-dCas9 2-23118, and pXMJ19-dCas9 2-23119, which contained different promoters for driving sgRNA expression. These vectors were transformed into *C. glutamicum* S91114 for single base editing. The results of high-throughput sequences show that the modified series of tools resulted in three types of mutations in the genome: C-G, C-T and base deletion. The mutation range can be detected as far as 28bp upstream of PAM site. The overall mutation effect mediated by pXMJ19-dCas9 2-23119-P was the best, with a maximum mutation efficiency of 30% at C6 position (Fig. 3A–J), and the second was PXMJ 19-DCAS92-23100-P. However, no mutation was detected in pXMJ19-dCas9 2-23113-P strain. Therefore, pXMJ19-dCas9 2-23119-P plasmid was subsequently selected for further modification.

3.4. Addition of adenosine deaminase TadA enables base conversion of A-T

In the above studies, we have realized some mutations such as C-T, C-G and base deletion in the genome of *C. glutamicum* S9114. To widen the range of mutations in base editing, adenosine deaminase from *E. coli* that can deaminate adenine and mutate to guanine was introduced into the base mutation tool. We used the DNA sequence of TadA-TadA mutant fusion adenosine deaminase on the NG-ABE max plasmid (Huang et al., 2019) used in the ABE system to fuse the adenine deaminase fusion sequence upstream of the dCas9 gene to construct Ptac-TadA-dCas9-AID

expression cassette, thereby obtaining the recombinant vector pXMJ19-dCas9 3 (Fig. 4A). Similarly, after transformation into *C. glutamicum* S9114 for single-base editing experiments, the sequencing results show that the adenine of A3-A9 sites produced an A-T conversion in the mutation window, which produced a maximum mutation efficiency of 16% at A8 site (Fig. 4B), indicating that TadA-TadA mutant can also exert a deamination effect in the *C. glutamicum* S9114 strain and realize the conversion A-G on the genome.

3.5. Application of single base editor for improved GlcNAc production by C. glutamicum S9114

Due to the existence of common precursors glucose-6-P and fructose-6-P, pentose phosphate pathway is an important competitive pathway in the synthesis of GlcNAc. Therefore, regulating the expression of gene *zwf* encoding glucose 6-phosphate dehydrogenase is beneficial to the production of GlcNAc. Since pentose phosphate pathway is necessary for cell growth, *zwf* gene cannot be knocked out. Therefore, the *zwf* gene which controls the primary rate-limiting step in pentose phosphate pathway was selected as the editing target (Fig. 5A) to obtain the mutant S9114 strain with improved GlcNAc titer.

The *zwf* mutant library obtained by the single base editing tool pXMJ19-dCas9 3 was analyzed and screened, and 6 mutant strains which produced missense mutations and nonsense mutations in base editing window were selected: M1 (P286L) M2 (P286L, A287G), M3 (P286S), M4 (L290stop), M5 (P286L, K293E), and M6 (K293R) (Fig. 5B). They were prepared into competent cells, and then were transformed with the expression vector pJYW-4-*ceN-Cglglms*, to produce GlcNAc in *C. glutamicum* S9114. The final GlcNAc titer of mutant strain M6 was increased to 9.1 g/L, which was 31.9% higher than that of the control strain (Fig. 5C). However, the GlcNAc titer of the double mutant strain M2 and the strain M4 inserted with the stop codon TAG decreased significantly.
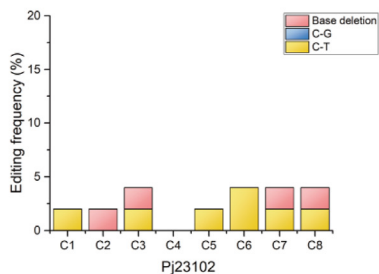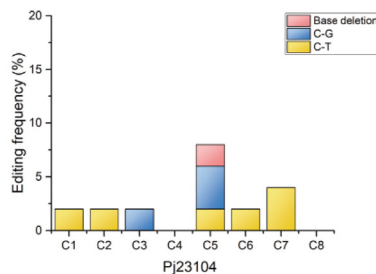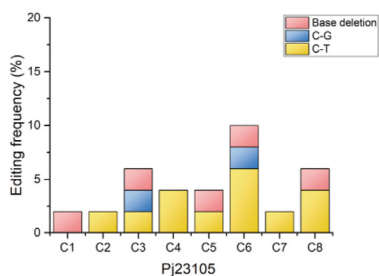
(A) TGC₁C₂AGC₃GC₄AGC₅TGC₆AGGC₇AGAAAAGATC₈AAGG
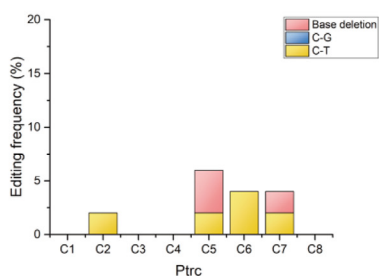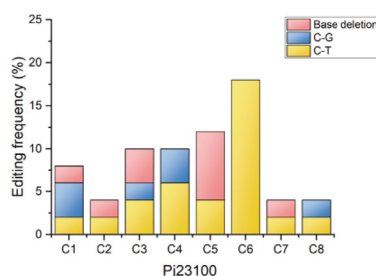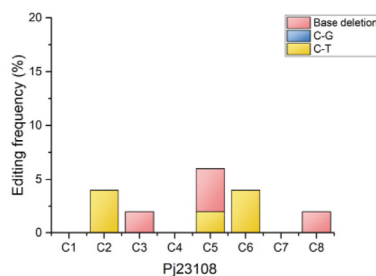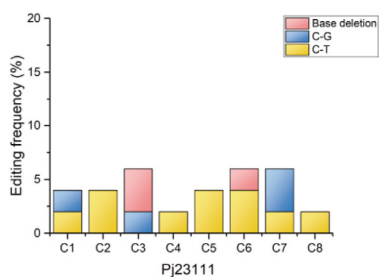
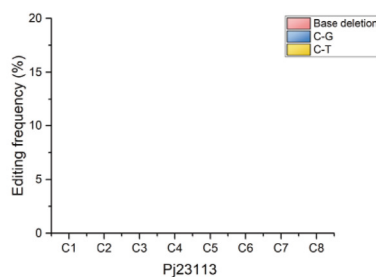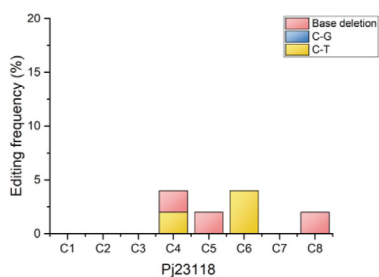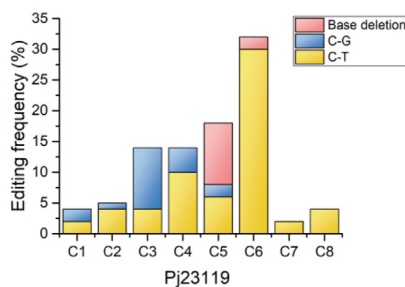**Fig. 3.** Base editing efficiency with different gRNA promoters. (A) Single base editing efficiency using the Ptrc-sgRNA expression cassette. (B) Single base editing efficiency using the Pj23100-sgRNA expression cassette. (C) Single base editing efficiency using the Pj23102-sgRNA expression cassette. (D) Single base editing efficiency using the Pj23104-sgRNA expression cassette. (E) Single base editing efficiency using the Pj23105-sgRNA expression cassette. (F) Single base editing efficiency using the Pj23106-sgRNA expression cassette. (G) Single base editing efficiency using the Pj23111-sgRNA expression cassette. (H) Single base editing efficiency using the Pj23113-sgRNA expression cassette. (I) Single base editing efficiency using the Pj23118-sgRNA expression cassette. (J) Single base editing efficiency using the Pj23119-sgRNA expression cassette. Different base conversion efficiencies are represented by histograms of different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
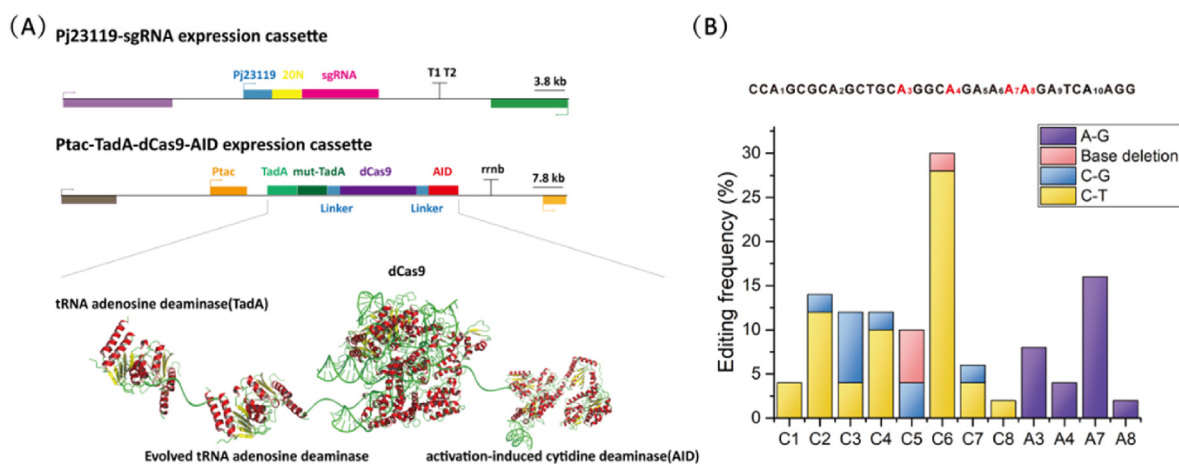
Fig. 4. Double deaminase single base editing strategy. (A) The N of dCas9 in this system is ligated with tRNA adenosine deaminase (TadA) and Evolved tRNA adenosine deaminase, and the C-terminus is linked to activation-induced cytidine deaminase (AID) and is commonly initiated by the promoter Ptrc. (B) Single base editing efficiency after addition of TadA. The purple histogram shows the A-G base conversion efficiency. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In order to realize the application of base editing tools in metabolic engineering, we tried to construct a mutant library on the genome of *zwf* and screen positive mutant strains with higher GlcNAc titer from the mutant library. In addition, we also found that single base substitution cannot only change the genetic code (missense mutation) in the open reading frame, but also produce the stop codon in advance. If the stop codon (TAA, TGA and TAG) appears in advance, gene knockout can be realized (Billon et al., 2017; Kuscu et al., 2017).

4. Discussion

This work aims to construct a single base editing system that can be used for genome editing of *C. glutamicum*. Activation-induced cytosine deaminase, specifically expressed by central B cells, is the molecular basis for high frequency mutations in somatic cells (Banno et al., 2018; Wang et al., 2018). AID deamination of cytosine to uracil results in U-G base mismatch. U-G base mismatch has two repair paths: first, if mismatch is not repaired in time, mutation of C-T and G-A will occur during DNA replication; Second, U-G mismatch was excised by uracil-N-glycosylase (UNG) to produce pyrimidine-free sites, and template-independent mononucleotide mutations were introduced by base excision repair (Assefa et al., 2012; Shen et al., 2018).

Based on this, Chang et al. fused human cytosine deaminase to the C-terminus of dCas9 (dCas9-AIDx) to form TAM (Targeted Aid-Mediated Mutagenesis), which can produce local sequence diversity through mutagenesis (Ma et al., 2016). By the fusion of dCas9 protein and AID, dCas9 can recruit AID into the targeting sequence under the targeting effect of sgRNA, and can effectively improve and expand the scope of AID by increasing the local concentration of AID. High frequency mutation is realized for specific genes in non-B cells to generate genetic diversity. In order to realize base editing in glutamic acid-producing strain *C. glutamicum* S9114, we first fused dCas9 with AID, then further optimized PAM targets, and studied the effect of sgRNA secondary structure on gene editing efficiency of *C. glutamicum* S9114. The experimental results show that when the PAM is AGG, a base editing tool has highest editing efficiency, which may be due to the CRISPR-Cas9-based tools in *C. glutamicum* have PAM codon preference (Peng et al., 2017), and PAM preference is different in different strains. At the same time, the secondary structure of sgRNA also has a great influence on the efficiency of base editing. Optimizing the secondary structure of sgRNA can improve the efficiency of base editing.

In order to connect multiple groups of targeted gene binding sequences to the tool plasmid at the same time, we designed an efficient seamless connection method of 20N sequence. Compared with the

previous DNA ligation method, the method is more time-saving and accurate, and the *Spa*I site was excised by enzyme digestion and does not appear on the recombinant vector. This method can be applied to any case where a shorter sequence is replaced on multiple parallel plasmids. New England Biolabs (UK) Ltd. currently offers more than 45 Type IIS restriction enzymes, which can be selected according to specific conditions (https://international.neb.com/tools-and-resources/selection-charts/type-iis- restriction-enzymes).

Furthermore, by modifying the sgRNA promoter, the mutation efficiency by the base editing tool was significantly improved, indicating that the redundant sequence downstream of the promoter had a great influence on its transcribed gRNA. The combination of the AID gene and dCas9 was chosen because the new AID-mediated mutagenesis technology has unique features compared to the other cytosine deaminase: (1) cytosine can be randomly transformed to the other three bases in the absence of UGI; (2) The activity of dCas9-AID fusion protein only depends on sgRNA and is independent of the primary sequence recognized by AID; (3) dCas9-AID can induce mutation of multiple cytosines at the same time (Yoshikawa, 2002). In addition, this method does not rely on DNA cutting and repair, which can greatly improve the efficiency of precise editing of single base and avoid double strand breaks of DNA. If the base editing is expected to result in a single C-T base conversion, adding UGI can be considered. As a phage protein, UGI protects its genome from the repair of host UNG when it invades *E. coli*. Therefore, UGI is an inhibitor of UNG, which prevents UNG from resecting the U produced by AID and cannot complete downstream repair (Jin et al., 2019; Mol et al., 1995). After the addition of UGI, the mutation direction is more singular, leaving only C-T and G-A mutations. In summary, the UGI protein can be selectively inserted according to different needs.

Although cytosine deaminase can accurately convert C in the mutation window to T, it cannot reversely mutate T to C or A to G. Recently, David Liu group reported a new single-base editing system ABE which can convert A to G by adenosine deaminase and CRIPSR/Cas9 (Hu et al., 2018; Ryu et al., 2017). After 7 rounds of evolution and transformation, they successfully obtained a TadA mutant capable of efficiently deaminating DNA adenine directly, and constructed an ABE system capable of efficiently inducing A > G mutation by fusion with nCas9 (D10A). There are many factors affecting the editing efficiency of TadA, such as the composition of PAM site and 20N sequence, the length of linker between TadA and TadA mutant, etc (Koblan et al., 2018). In order to obtain a more accurate single base editing tool, cytosine deaminase and adenosine deaminase edited by single base can be mutated and optimized in the future, thus obtaining a high-precision single base editing tool which can completely eliminate RNA off-target and maintain DNA editing activity.
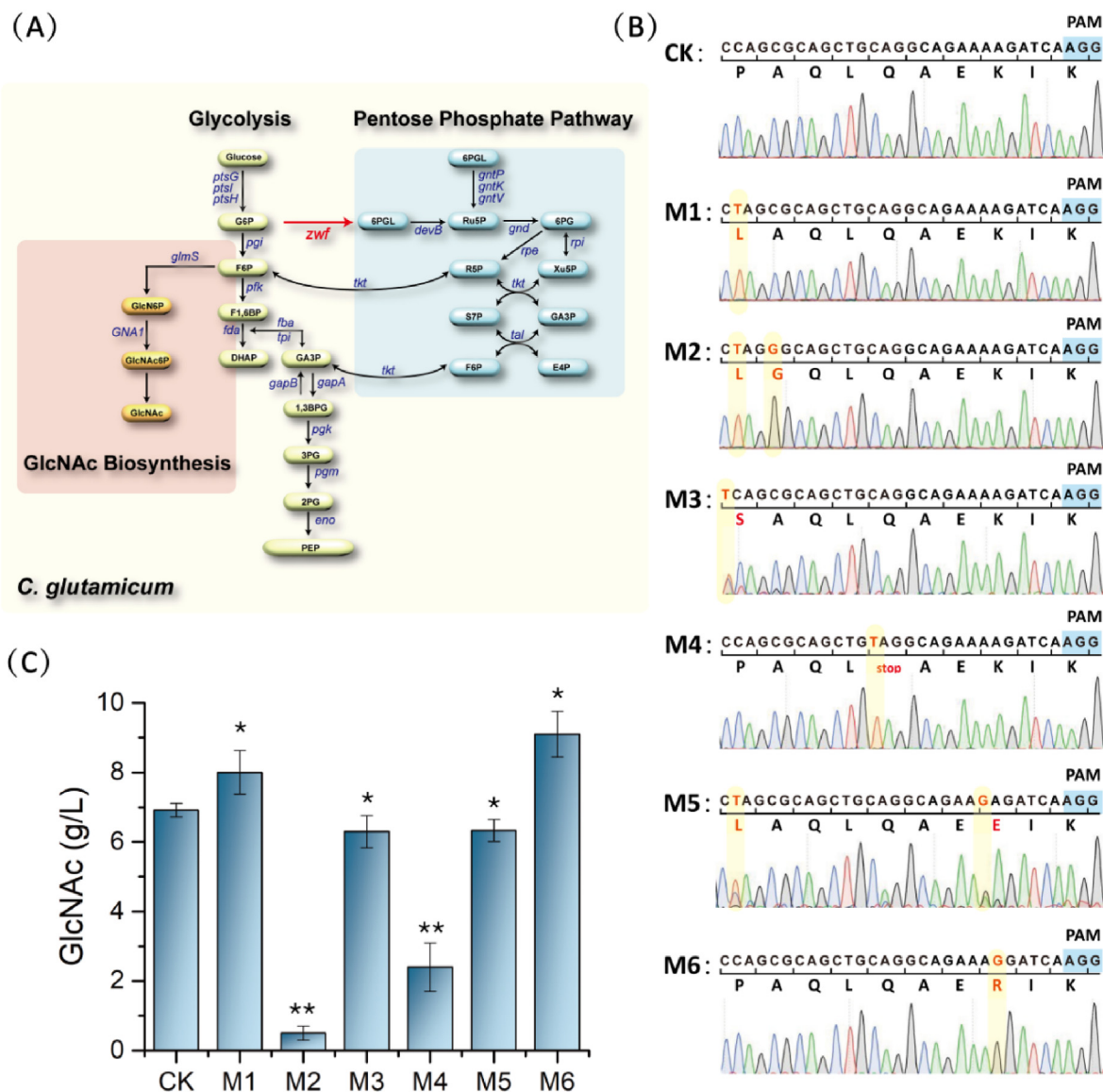
Fig. 5. The *zwf* gene in the *C. glutamicum* S9114 strain was modified using base editing strategy. (A) Synthetic pathway of GlcNAc in *C. glutamicum*. The *zwf* gene is highlighted in red, and the pentose phosphate pathway linked by the gene is highlighted by blue square. (B) 6 mutant strains generated by single-base editing of *zwf* using the Ptac-TadA-dCas9-AID expression cassette. The mutated bases are highlighted by red letters, and the PAM sequences are highlighted by blue squares. Among them, the M4 strain produced a premature stop codon in *zwf*. (C) GlcNAc titer in 6 single-mutant strains. All data were the average of three independent studies with standard deviations. All data were expressed as mean ± SD. Differences were determined by 2-tailed Student's t-test between two groups, or one-way. ANOVA followed by post-hoc Tukey's test for multiple groups. Statistical significance is indicated as * for p < 0.05 and ** for p < 0.01 relative to control strain S9114 Δ*nagA*-Δ*gamA*-Δ*ldh*, respectively. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In this study, we combined pyrimidine base conversion tool and purine base conversion tool to introduce a variety of base mutations into the genome of *C. glutamicum* S9114, thus greatly improving the efficiency of single base mutation on the genome of *C. glutamicum*.

Changes in one or several bases (substitutions or deletions) lead to changes in the coding frame, resulting in changes in the structure of the translated protein, which is conducive to the construction of mutant libraries in *C. glutamicum*, while it is difficult to construct mutant libraries in *C. glutamicum* genome in other studies in the past. In terms of functional screening, the single-base edited dCas9-AID system can convert C-T or C-G, and greatly enhances the genetic diversity of target sequences. Here the TadA-dCas9-AID system overcomes the limitation that the previous dCas9-AID system cannot realize A-G conversion, and can generate more new amino acid mutations in the target gene during the screening process. The combination of two mutation tools produces a wider mutation library from which a gain-of-function mutant can be obtained. However, wild-type Cas9 system can only realize enhanced mutation through HDR, and it is almost impossible to realize large-scale functional screening in *C. glutamicum*. The base editing tool developed in this study can led to point mutations/premature termination, so it can greatly improve the efficiency of base substitution in *C. glutamicum* and can be used for the metabolic engineering of *C. glutamicum*.

In summary, the single base editing tool obtained in this study greatly increased the editable site on the genome of *C. glutamicum* and improved the gene editing efficiency of *C. glutamicum*. The development of a single base editor provides an idea for site-directed editing and modification of key nucleotide variants in the *C. glutamicum* genome. In the future, we will further optimize the base editor by optimizing sgRNA design and using a more rigorous CRISPR/Cas system or by protein engineering to improve base editing efficiency and expand mutation types, and try to further narrow the mutation window and even change only a single base without affecting the surrounding sequence.

Funding

Declaration of competing interest

The authors declare that they have no conflicts of interest.

The research performed did not involve human participants and/or animals.

CRediT authorship contribution statement

Chen Deng: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing - original draft. **Xueqin Lv:** Investigation, Methodology, Project administration, Validation, Visualization, Writing - original draft. **Jianghua Li:** Resources, Software, Supervision. **Yanfeng Liu:** Conceptualization, Data curation, Formal analysis, Writing - review & editing, Supervision. **Guocheng Du:** Resources, Software, Supervision. **Long Liu:** Conceptualization, Data curation, Formal analysis, Writing - review & editing, Supervision.

References

Arakawa, H., 2016. A method to convert mRNA into a gRNA library for CRISPR/Cas9 editing of any organism. Sci. Adv. 2 (8), e1600699.

Assefa, N.G., Niiranen, L., Willassen, N.P., Smalas, A., Moe, E., 2012. Thermal unfolding studies of cold adapted uracil-DNA N-glycosylase (UNG) from Atlantic cod (Gadus morhua). A comparative study with human UNG. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 161 (1), 60–68.

Banno, S., Nishida, K., Arazoe, T., Mitsunobu, H., Kondo, A., 2018. Deaminase-mediated multiplex genome editing in *Escherichia coli*. Nat. Microbiol. 3 (4), 423–429.

Billon, P., Bryant, E.E., Joseph, S.A., Nambiar, T.S., Hayward, S.B., Rothstein, R., Ciccia, A., 2017. CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. Mol. Cell 67 (6), 1068–1079 e4.

Chen, X., Liu, L., Li, J., Du, G., Chen, J., 2012a. Improved glucosamine and *N*-acetylglucosamine production by an engineered *Escherichia coli* via step-wise regulation of dissolved oxygen level. Bioresour. Technol. 110, 534–538.

Chen, X., Liu, L., Li, J., Du, G., Chen, J., 2012b. Optimization of glucose feeding approaches for enhanced glucosamine and *N*-acetylglucosamine production by an engineered *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 39 (2), 359–365.

Cho, J.S., Choi, K.R., Prabowo, C.P.S., Shin, J.H., Yang, D., Jang, J., Lee, S.Y., 2017. CRISPR/Cas9-coupled recombineering for metabolic engineering of *Corynebacterium glutamicum*. Metab. Eng. 42, 157–167.

Deng, C., Lv, X., Li, J., Liu, Y., Du, G., Amaro, R.L., Liu, L., 2019a. Synthetic repetitive extragenic palindromic (REP) sequence as an efficient mRNA stabilizer for protein production and metabolic engineering in prokaryotic cells. Biotechnol. Bioeng. 116 (1), 5–18.

Deng, C., Lv, X., Liu, Y., Li, J., Lu, W., Du, G., Liu, L., 2019b. Metabolic engineering of *Corynebacterium glutamicum* S9114 based on whole-genome sequencing for efficient *N*-acetylglucosamine synthesis. Synth. Syst. Biotechnol. 4 (3), 120–129.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E., 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat. Biotechnol. 32 (12), 1262–1267.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468 (7320), 67–71.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., Liu, D.R., 2017. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551 (7681), 464–471.

Gu, Y., Deng, J., Liu, Y., Li, J., Shin, H.D., Du, G., Chen, J., Liu, L., 2017. Rewiring the glucose transportation and central metabolic pathways for overproduction of *N*-acetylglucosamine in *Bacillus subtilis*. Biotechnol. J. 12 (10), 1700268.

Gu, Y., Lv, X., Liu, Y., Li, J., Du, G., Chen, J., Rodrigo, L.A., Liu, L., 2019. Synthetic redesign of central carbon and redox metabolism for high yield production of *N*-acetylglucosamine in *Bacillus subtilis*. Metab. Eng. 51, 59–69.

Huang, T.P., Zhao, K.T., Miller, S.M., Gaudelli, N.M., Oakes, B.L., Fellmann, C., Savage, D.F., Liu, D.R., 2019. Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. Nat. Biotechnol. 37 (6), 626–631, 2019.

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., Liu, D.R., 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556 (7699), 57–63.

Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., Sun, B., Chen, B., Xu, X., Li, Y., Wang, R., Yang, S., 2017. CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. Nat. Commun. 8, 15179.

Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.-L., Zhang, F., Gao, C., 2019. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science 364 (6437), 292–295.

Khan, A.A., El-Sayed, A., Akbar, A., Mangravita-Novo, A., Bibi, S., Afzal, Z., Norman, D.J., Ali, G.S., 2017. A highly efficient ligation-independent cloning system for CRISPR/Cas9 based genome editing in plants. Plant Methods 13, 86.

Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., Liu, D.R., 2018. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. 36 (9), 843–846.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., Liu, D.R., 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533 (7603), 420–424.

Kuscu, C., Kuscu, C., Tufan, T., Yang, J., Adli, M., 2017. CRISPR-STOP: gene silencing through base editing-induced nonsense mutations. Nat. Methods 14, 710–712.

LaFountaine, J.S., Fathe, K., Smyth, H.D., 2015. Delivery and therapeutic applications of gene editing technologies ZFNs, TALENs, and CRISPR/Cas9. Int. J. Pharm. 494 (1), 180–194.

Liu, J., Wang, Y., Lu, Y., Zheng, P., Sun, J., Ma, Y., 2017. Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. Microb. Cell Factories 16 (1), 205.

Liu, L., Liu, Y., Shin, H.D., Chen, R., Li, J., Du, G., Chen, J., 2013. Microbial production of glucosamine and *N*-acetylglucosamine: advances and perspectives. Appl. Microbiol. Biotechnol. 97 (14), 6149–6158.

Ma, Y., Zhang, J., Yin, W., Zhang, Z., Song, Y., Chang, X., 2016. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. Nat. Methods 13 (12), 1029–1035.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J., Koonin, E.V., 2011. Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. 9 (6), 467–477.

Makarova, K.S., Wolf, Y.I., Koonin, E.V., 2013. The basic building blocks and evolution of CRISPR-CAS systems. Biochem. Soc. Trans. 41 (6), 1392–1400.

Mei, J., Xu, N., Ye, C., Liu, L., Wu, J., 2016. Reconstruction and analysis of a genome-scale metabolic network of *Corynebacterium glutamicum* S9114. Gene 575 (2 Pt 3), 615–622.

Mol, C.D., Arvai, A.S., Sanderson, R.J., Slupphaug, G., Kavli, B., Krokan, H.E., Mosbaugh, D.W., Tainer, J.A., 1995. Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. Cell 82 (5), 701–708.

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K.Y., Shimatani, Z., Kondo, A., 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 353 (6305), aaf8729.

Noh, M., Yoo, S.M., Yang, D., Lee, S.Y., 2019. Broad-spectrum gene repression using scaffold engineering of synthetic sRNAs. ACS Synth. Biol. 8 (6), 1452–1461.

Pan, X., Wan, B., Li, C., Liu, Y., Wang, J., Mou, H., Liang, X., 2014. A novel whole genome amplification method using type IIS restriction enzymes to create overhangs with random sequences. J. Biotechnol. 184, 1–6.

Peng, F, Wang, X, Sun, W, Dong, G, Yang, Y, Liu, X, Bai, Z, 2017. Efficient gene editing in *Corynebacterium glutamicum* using the CRISPR/Cas9 system. Microb. Cell. Fact. 16 (1), 201.

Rees, H.A., Liu, D.R., 2018. Base editing: precision chemistry on the genome and transcriptome of living cells. Nat. Rev. Genet. 19 (12), 770–788.

Renaud, J.B., Boix, C., Charpentier, M., De Cian, A., Cochennec, J., Duvernois-Berthet, E., Perrouault, L., Tesson, L., Edouard, J., Thinard, R., Cherifi, Y., Menoret, S., Fontaniere, S., de Croze, N., Fraichard, A., Sohm, F., Anegon, I., Concordet, J.P., Giovannangeli, C., 2016. Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. Cell Rep. 14 (9), 2263–2272.

Rose, J.C., Stephany, J.J., Wei, C.T., Fowler, D.M., Maly, D.J., 2018. Rheostatic control of Cas9-mediated DNA double strand break (DSB) generation and genome editing. ACS Chem. Biol. 13 (2), 438–442.

Ryu, S., Koo, T., Kim, K., Lim, K., Baek, G., Kim, S., Kim, H., Kim, D., Lee, H., Chung, E., 2017. Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. Nat. Biotechnol. 36 (6), 536–539.

Salsman, J., Dellaire, G., 2017. Precision genome editing in the CRISPR era. Biochem. Cell. Biol. 95 (2), 187–201.

Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.L., Weiss, D.S., 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature 497 (7448), 254–257.

Shen, M.W., Arbab, M., Hsu, J.Y., Worstell, D., Culbertson, S.J., Krabbe, O., Cassa, C.A., Liu, D.R., Gifford, D.K., Sherwood, R.I., 2018. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature 563 (7733), 646–651.

Wang, Y., Liu, Y., Liu, J., Guo, Y., Fan, L., Ni, X., Zheng, X., Wang, M., Zheng, P., Sun, J., Ma, Y., 2018. MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. Metab. Eng. 47, 200–210.

Wang, Y., Liu, Y., Li, J., Yang, Y., Ni, X., Cheng, H., Huang, T., Guo, Y., Ma, H., Zheng, P., Wang, M., Sun, J., Ma, Y., 2019. Expanding targeting scope, editing window, and base transition capability of base editing in *Corynebacterium glutamicum*. Biotechnol. Bioeng. 116 (11), 3016–3029.

Wiese, K.C., Hendriks, A., 2006. Comparison of P-RnaPredict and mfold algorithms for RNA secondary structure prediction. Bioinformatics 22 (8), 934–942.

Wong, L., Engel, J., Jin, E., Holdridge, B., Xu, P., 2017. YaliBricks, a versatile genetic toolkit for streamlined and rapid pathway engineering in *Yarrowia lipolytica*. Metab. Eng. Commun. 5, 68–77.

Yang, Z., Edwards, H., Xu, P., 2020. CRISPR-Cas12a/Cpf1-assisted precise, efficient and multiplexed genome-editing in *Yarrowia lipolytica*. Metab. Eng. Commun. 10, e00112.

Yin, L., Shi, F., Hu, X., Chen, C., Wang, X., 2013. Increasing l-isoleucine production in *Corynebacterium glutamicum* by overexpressing global regulator Lrp and two-component export system BrnFE. J. Appl. Microbiol. 114 (5), 1369–1377.

Yoshikawa, K., 2002. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. Science 296 (5575), 2033–2036.

Zhang, B., Yu, M., Zhou, Y., Li, Y., Ye, B.C., 2017. Systematic pathway engineering of *Corynebacterium glutamicum* S9114 for L-ornithine production. Microb. Cell Factories 16 (1), 158.