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

Chen Deng, Xueqin Lv, Jianghua Li, Yanfeng Liu, Guocheng Du, Long Liu

1. Introduction

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 swf 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 base conversions 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.
established a new single-base editing system to achieve accurate conversion of A-G (adenine base editor, ABE) (Gaudelli et al., 2017; Deng et al., 2019a; Mei et al., 2016; Zhang et al., 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 rRNA 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) (GleNac) (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 | ACAATTTCCACAGGACAGGTATATTAATGGAAGACACTAAGGTATAAGATTTCCAC |
| DAR | CGCGAGAACGGACTGTCAGGACGCCGGGGG |
| D10AF | GCCCTGCTATGAGCCCAAATCTC |
| D10AR | GCGCATGACGGCCCGGAAATATAAC |
| HF40AF | GTGAGCCGCTACCTCCGCTAGTC |
| HF40AR | GAGCTATGACCGCTACATGGTATCT |
| VectorrNA.F | CATTTCGAAAGTACTAGCCGGAATTTGATCTGTTTG |
| VectorrRNA.R | TTATTCGCTCGCTCGAGAAGACGGG |
| FragmentsrRNA.F | CCGCATTCTGAAAGTGATTTAGGGGG |
| FragmentsrRNA.R | GGGCATTATTTGCCCGCTAGAAGACCC |
| sg1F | CGGAAATCGGAATTCTGGTAGTTAGTTAGTAGCTGAGTAATAATGAGCAATAGTATA |
| sg2F | CGGAAATCGGAATTCTGGTAGTTAGTTAGTAGCTGAGTAATAATGAGCAATAGTATA |
| sg3F | GGGGATTGGTTTTAAACAGGAGGTCGACG |
| VectorsgRNA.R | TATTAGCTCGCTCGAGAAGACGGG |
| VectorsgRNA.F | CATTTCGAAAGTACTAGCCGGAATTTGATCTGTTTG |
| H840AR | GACGATGGCGTCCACATCGTAATCAGAG |
| H840AF | GTGGAGCCATCGTCCCTCAGTC |
| D10AR | GCCGATAGCCAGGCCAATGGAATAC |
| D10AF | GGCCTGGCTATCGGCACCAATTC |
| DAR | CCGGCCAAAACAGCCAAGCTGTCACAGGCCCAGGGTGCG |
| DAF | ACAATTTCCACAGGACAGGTATATTAATGGAAGACACTAAGGTATAAGATTTCCAC |

...
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 basic 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). CloneExpress 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, Hind III restriction site and SD sequence: GCTAGAAAGAAGGAGGACCCGAC (derived from pJYW-4-ecN 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 EcoR 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 Hind III/EcoR I.

The plasmid pXMJ19 was subjected to the same double digestion treatment, and the restriction enzyme digestion system contained 16 uL of fragment DNA, 1 uL of Hind III, 1 uL of EcoR I, and 2 uL of 10 × H buffer, and double distilled water was added to a total volume of 20 uL. 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 (continued)

| Name            | Description                                      | Source         |
|-----------------|--------------------------------------------------|----------------|
| pXMJ19-dCas9 2- 23105 | pXMJ19-dCas9 1 derive with sgRNA framework gene under the control of Pj23105 | This work      |
| pXMJ19-dCas9 2- 23108 | pXMJ19-dCas9 1 derive with sgRNA framework gene under the control of Pj23108 | This work      |
| pXMJ19-dCas9 2- 23111 | pXMJ19-dCas9 1 derive with sgRNA framework gene under the control of Pj23111 | This work      |
| pXMJ19-dCas9 2- 23113 | pXMJ19-dCas9 1 derive with sgRNA framework gene under the control of Pj23113 | This work      |
| pXMJ19-dCas9 2- 23118 | pXMJ19-dCas9 1 derive with sgRNA framework gene under the control of Pj23118 | This work      |
| pXMJ19-dCas9 2- 23119 | pXMJ19-dCas9 1 derive with sgRNA framework gene under the control of Pj23119 | This work      |
| pXMJ19-dCas9 2- 23100-P | pXMJ19-dCas9 2-23100 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23102-P | pXMJ19-dCas9 2-23102 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23108-P | pXMJ19-dCas9 2-23108 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23110-P | pXMJ19-dCas9 2-23110 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23111-P | pXMJ19-dCas9 2-23111 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23113-P | pXMJ19-dCas9 2-23113 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23118-P | pXMJ19-dCas9 2-23118 derive with 20N sequence and PAM site AGG | This work      |
| pXMJ19-dCas9 2- 23119-P | pXMJ19-dCas9 2-23119 derive with 20N sequence and PAM site AGG | This work      |
| TadA-TadA mutant cloned | TadA-TadA mutant cloned | This work      |
| pTYW-4-cen-C-gfjmS | pTYW-4-cen-C-gfjmS cloned, Kan' | Laboratory stock |

Cm’, chloramphenicol resistance; Kan’, kanamycin resistance.

Vectors or RNA.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 Fragmentsgra_P/Fragmentsgra_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, Enhance 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 obtain 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 EcoR1 and XbaI 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 ligations products were transformed E. coli JM109.

2.4. Single base editing in C. glutamicum S9114

The recombinant plasmids pXMJ19-dCas9 1-zwf, pXMJ19-dCas9 1-zwf2, pXMJ19-dCas9 1-zwf3, pXMJ19-dCas9 1-zwf4 with four 20-nt sgRNA of the target gene zwf, which have been correctly sequenced, are separately transformed into the C. glutamicum S9114 ΔnagA-Δgama-Aldh competence constructed in the previous study (Deng et al., 2019b), and then transferred to the LBB (Luria–Bertiari broth with 18.5 g/L brain heart infusion) plate with 10 μg/μl chloramphenicol and 0.02 mM iso-propyl-β-D-thiogalactosidase (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 OD600 = 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-zwf4 as templates, primers J23100F/JR, J23102F/JR, J23104F/JR, J23105F/JR, J23107F/JR, J23109F/JR, J23111F/FR, J23113F/FR, J23115F/FR, and J23119F/FR were used to amplify sgRNA fragments with p231 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 Vector23111, pMD19-T Vector23113, pMD19-T Vector23115, and pMD19-T Vector23118. Using pMD19-T Vector23100, pMD19-T Vector23102, pMD19-T Vector23104, pMD19-T Vector23105, pMD19-T Vector23108, pMD19-T Vector23111, pMD19-T Vector23113, pMD19-T Vector23115, 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 and pXMJ19-dCas9 2-23119 were respectively digested with SapI 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 zwfsg4F, and 20 μl primer zwfsg4R. 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 SapI 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 ddH2O 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 2-23105, 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 XbaI restriction enzyme.

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

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-C( Gu et al., 2017). The selected target gene in C. glutamicum, which controls the major competition reactions of GlcNAc was S9114 for base editing experiments, and sgRNA, we used the online tool Mfold to predict the secondary structure of sgRNA, we used the online tool Mfold to predict the secondary structure of sgRNA (http://unafold.rna.albany.edu/?r=mfold) (Wiese and Hendriks, 2006).

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-C-eC- C.gltGSmS 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; KH2PO4, 1.0; (NH4)2SO4, 0.5; MgSO4, 0.5; CaCO3, 20.0; and FeSO4, 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 OD562 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.

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 OD962 of the fermentation broth was determined by UV-visible spectrophotometer (UVmini-1240, Shimadzu Corporation, Japan), and the calculation formula based on cell OD962 and dry cell weight (DCW) was: DCW = 0.6495 × OD962–2.7925 (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 H2SO4 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 ptc-cas9-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.cshl.edu/) 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 P231 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).
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, 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, and pXMJ19-dCas9 2-23119, respectively. The relative strengths of Pj23100, Pj23102, Pj23104, Pj23105, Pj23108, Pj23111, Pj23113, 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-Spal, 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 SapI. 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).
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, and pXMJ19-dCas9 2-23118, 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 pXMJ19-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.
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.)
In order to realize the application of base editing tools in metabolic engineering, we tried to construct a mutant library on the genome of *C. glutamicum* 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; 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 chose AID because the new AID-mediated mutagenesis technology has unique features compared to the other cytosine deaminase: (1) the redundant sequence downstream of the promoter had a great effect of sgRNA, and can effectively improve and expand the scope of AID; (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 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 CRISPR/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. 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 Ptc. (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 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

This work was financially supported by the National Natural Science Foundation of China (31622001, 31671845, 31600068), and the 111 Project (111-2-06).

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

Cheng Deng: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, 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.

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

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

Banno, S., Kuscu, C., Kac, F., 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 AID-mediated mutagenesis (TAM) enables efficient CRISPR/Cas9-based genome editing in mammalian cells. Nat. Methods 13 (12), 1029–1035.

C. Deng et al. Metabolic Engineering Communications 11 (2020) e00135

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

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

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

LaFontaine, 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, D., Gu, C., 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, S., Yang, X., Cheng, D., 2016. Targeted AID-mediated mutagenesis (TAM) enables efficient CRISPR/Cas9-based genome editing 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.L., 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.L., Koonin, E.V., 2013. The basic building blocks and evolution of CRISPR/Cas machinery. 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 strain 5914. Gene 579 (2 Pt A), 615–622.

Mol, C.D., Arvai, A.S., Sanderson, R.J., Slupphaug, G., Kawakami, K., Kool, H.E., Moobath, 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), aaa7829.

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

Peng, X.C., Xue, X., Sun, W., Deng, C., Li, 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 the transcriptome of living cells. Nat. Rev. Genet. 19 (10), 779–788.

Renaud, J.W., Boix, C., Charpentier, M., De Cian, A., 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), 539–540.

Salsman, J., Dellaire, G., 2017. Precision genome editing in the CRISPR era. Biochem. J. 474, 181–187.

Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Yen, Y.L., Weiss, D.S., 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature 479 (7344), 254–257.

Shen, M.W., Arbab, M., Hsu, Y.J., Worstell, D., Culbertson, S.J., Shah, A., Cassa, C.A., Liao, D.R., Gifford, D.K., Sherwood, R.L., 2018. Predictable and precise template-free genome editing in bacteria. Nat. Biotechnol. 36 (7), 626–631.

Shen, X., Liu, Y., Liu, J., Gao, Q., Zou, Z., Wang, Y., Qin, P., Li, J., Guo, W., Chen, Z., Zhou, L., Zeng, F., Gao, C., 2019. Cytosine, but not adenosine, base editors induce genome-wide off-target mutations in rice. Science 364 (6437), 292–295.

Shi, A., El-Sayed, A., Abril, F., Bhi, S., Afzal, Z., Norman, D.J., Ali, G.S., 2017. A highly efficient ligand-independent cloning system for CRISPR/Cas9 based genome editing in plants. Plant Methods 13, 86.

S.111

S.111
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.