Metabolic Engineering 47 (2018) 49–59

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production

Jie Zhang[a], Wenming Zong[a,b], Wei Hong[a,c], Zhong-Tian Zhang[a], Yi Wang[a,d,*]

[a] Department of Biosystems Engineering, Auburn University, Auburn, AL 36849, USA
[b] School of Engineering, Anhui Agricultural University, Hefei 230036, China
[c] Key Laboratory of Endemic and Ethnic Diseases (Guizhou Medical University), Ministry of Education, Guiyang 550000, China
[d] Center for Bioenergy and Bioproducts, Auburn University, Auburn, AL 36849, USA

ARTICLE INFO

Keywords:
Endogenous CRISPR-Cas system
CRISPR-Cas9
CRISPR-Cpf1
Clostridium tyrobutyricum
Multiplex genome editing
Butanol

ABSTRACT

Although CRISPR-Cas9/Cpf1 have been employed as powerful genome engineering tools, heterologous CRISPR-Cas9/Cpf1 are often difficult to introduce into bacteria and archaea due to their severe toxicity. Since most prokaryotes harbor native CRISPR-Cas systems, genome engineering can be achieved by harnessing these endogenous immune systems. Here, we report the exploitation of Type I-B CRISPR-Cas of *Clostridium tyrobutyricum* for genome engineering. *In silico* CRISPR array analysis and plasmid interference assay revealed that TCA or TCG at the 5′-end of the protospacer was the functional protospacer adjacent motif (PAM) for CRISPR targeting. With a lactose inducible promoter for CRISPR array expression, we significantly decreased the toxicity of CRISPR-Cas and enhanced the transformation efficiency, and successfully deleted *spo0A* with an editing efficiency of 100%. We further evaluated effects of the spacer length on genome editing efficiency. Interestingly, spacers ≤ 20 nt led to unsuccessful transformation consistently, likely due to severe off-target effects; while a spacer of 30–38 nt is most appropriate to ensure successful transformation and high genome editing efficiency. Moreover, multiplex genome editing for the deletion of *spo0A* and *pyrF* was achieved in a single transformation, with an editing efficiency of up to 100%. Finally, with the integration of the alcohol dehydrogenase gene (*adhE1* or *adhE2*) to replace *cat1* (the key gene responsible for butyrate production and previously could not be deleted), two mutants were created for n-butanol production, with the butanol titer reached historically record high of 26.2 g/L in a batch fermentation. Altogether, our results demonstrated the easy programmability and high efficiency of endogenous CRISPR-Cas. The developed protocol herein has a broader applicability to other prokaryotes containing endogenous CRISPR-Cas systems. *C. tyrobutyricum* could be employed as an excellent platform to be engineered for biofuel and biochemical production using the CRISPR-Cas based genome engineering toolkit.

1. Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) system is an RNA guided immune system in bacteria and archaea that can provide defense against foreign invaders, such as phages and plasmids (Barrangou et al., 2007; Barrangou and Marraffini, 2014; Horvath and Barrangou, 2010; Sorek et al., 2013). All the identified CRISPR-Cas systems share similar features, consisting of identical direct repeats separated by variable spacers, along with a suite of associated *cas* genes (Makarova et al., 2015; Westra et al., 2016). CRISPR-Cas systems can be classified into two classes and six types based on the signature Cas proteins and the architecture of CRISPR-*cas* loci (Makarova et al., 2011, 2015; Westra et al., 2016). A complex of multiple Cas proteins are involved in degrading the invading genetic elements in Types I, III and IV, which all belong to the Class 1 system; while Types II, V and VI in the Class 2 system can carry out the same operation by using a single large Cas protein. Among the various CRISPR-Cas systems, Type I, II, and III are the most widespread in both archaea and bacteria, and distinguished by the presence of the unique signature protein: Cas3, Cas9, and Cas10, respectively (Makarova et al., 2011, 2015). Among them, Type I systems exhibit the most diversity, and are further divided into six subtypes: I-A to I-F.

Three functional stages, termed adaptation, expression, and interference, are generally included in the development of the immunity of CRISPR-Cas systems for the defense of the potential foreign invaders

* Corresponding author at: Department of Biosystems Engineering, Auburn University, 350 Mell Street, Auburn, AL 36849, USA.
E-mail address: yiwang3@auburn.edu (Y. Wang).

https://doi.org/10.1016/j.ymben.2018.03.007
Received 7 February 2018; Received in revised form 4 March 2018; Accepted 6 March 2018
Available online 09 March 2018

(Barrangou and Marraffini, 2014; Charpentier et al., 2015; Plagens et al., 2015; Van Der Oost et al., 2014; Westra et al., 2014). During the adaptation phase, spacer sequences derived from the invading genetic elements are identified and integrated into the host genome right between the leader sequence and the first spacer, generating the new spacers of the CRISPR array. A promoter located within the CRISPR leader sequence then drives the transcription of CRISPR array (including the new spacers) to form a long precursor CRISPR RNA (crRNA) followed by the cleavage of the precursor crRNAs to make mature crRNAs. Once the invasion happens again to the host cells, a ribonucleoprotein complex (crRNP) will be formed by the mature crRNAs and specific Cas proteins to recognize the same or similar foreign genetic elements though sequence matching between the spacer on the crRNA and the protospacer on the foreign invaders, and degrade the invading DNA or RNA via interference. During the interference in Type I and Type II systems, the targeting efficiency is greatly improved if the protospacer is flanked by a short conserved sequence defined as protospacer-adjacent motif (PAM) (Bolotin et al., 2005; Deveau et al., 2008; Mojica et al., 2009). The PAM sequence is usually 2–5 nt long and located at the 5′- or 3′-end of the protospacer (Mojica et al., 2009; Shah et al., 2013). The presence of PAM sequence in the target DNA rather than in the CRISPR array of the host genome is used to discriminate 'self' and 'non-self'.

Although the Class 2 system is less abundant in the nature, their acting machineries are much simpler and more programmable. In the past few years, the *Streptococcus pyogenes* CRISPR-Cas9 (spCRISPR-Cas9) system has been engineered to be a high efficient genome editing tool that has been implemented in a broad range of organisms, such as bacteria (Huang et al., 2016; Jiang et al., 2013; Li et al., 2016), yeast (DiCarlo et al., 2013; Horwitz et al., 2015), plants (Li et al., 2013; Shan et al., 2013), mammal cells (Wang et al., 2013a), and human cells (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). Besides single gene knock-in or knock-out, successes have also been reported for multiplex genome editing (Jiang et al., 2015; Sakuma et al., 2014; Xie et al., 2015) and transcriptional regulation, including repression (Qi et al., 2013) and activation (Gilbert et al., 2013). Recently, another Class 2 CRISPR effector, Cpf1, was characterized and repurposed for genome editing (Zetsche et al., 2015, 2017). Compared to the CRISPR-Cas9 system, the CRISPR-Cpf1 system exhibited higher targeting efficiency and capability under particular circumstances (Jiang et al., 2017; Zetsche et al., 2017).

Undoubtedly, CRISPR-Cas9/Cpf1 systems are powerful genome engineering tools with which versatile genome editing purposes can be achieved. However, as a heterologous protein, in many cases, either Cas9 or Cpf1 is hard to introduce into bacteria and archaea due to their intrinsic toxicity, leading to low transformation efficiency and thus difficulty for genome editing. It has been reported that, based on the genome analysis, approximately 47% of sequenced bacteria and 87% of sequenced archaea harbor CRISPR-*cas* loci (Makarova et al., 2015). Therefore, the endogenous CRISPR-Cas system has a great potential to be repurposed for genome editing and transcriptional regulation. Through the deletion of *cas3* gene which is responsible for degrading the target DNA, the endogenous Type I-E CRISPR-Cas system in *Escherichia coli* was harnessed as a programmable gene expression regulator (Luo et al., 2015). Pyne et al. engineered the Type I-B CRISPR-Cas system in *Clostridium pasteurianum* to be an efficient genome editing tool, and successfully deleted the *cpaAIR* gene (Pyne et al., 2016).

In recent years, the genus *Clostridium* has drawn tremendous attentions as it contains various species with great potentials for the production of commodity chemicals and fuels, such as n-butanol (butanol hereafter). Butanol can be naturally produced in solventogenic clostridia through the Acetone-Butanol-Ethanol (ABE) fermentation. Although tremendous efforts have been invested on the metabolic engineering of solventogenic clostridial strains for enhanced biobutanol production, only very limited success has been achieved (Lee et al., 2016b). This is because, on one hand, there are several intrinsic byproducts in ABE fermentation including fatty acids, acetone and ethanol that are hard to eliminate; on the other, the ABE fermentation for butanol production goes through a biphasic process and is subjected to complicated metabolic regulation (Zhang et al., 2016b). Yu et al. engineered *C. tyrobutyricum* ATCC 25755 (a hyper-butyrate producer) for butanol production by inactivating the native acetate kinase (*ack*) gene or the phosphate butyryltransferase (*ptb*) gene and introducing the aldehyde/alcohol dehydrogenase (*adhE2*) from *C. acetobutylicum*, with a butanol titer of 10.0 g/L was obtained (Yu et al., 2011). Recently, the butyrate-producing metabolism of *C. tyrobutyricum* was further elucidated through whole-genome sequencing and proteomic analysis (Lee et al., 2016a). Interestingly, contradictory with the results by Yu et al. (2011), it was demonstrated that the *ptb* gene actually does not exist in *C. tyrobutyricum* and the *ack* gene can't be deleted because the deletion would lead to no endproduct and inefficient ATP generation (Lee et al., 2016a). Additionally, it was revealed that the butyrate production in *C. tyrobutyricum* is in fact dependent on the butyrate:acetate CoA transferase gene (*cat1*), which is very different from the *ptb*-butyrate kinase (*buk*) pathway for butyrate production in solventogenic clostridial strains. However, the disruption of *cat1* using mobile group II intron was unsuccessful, because the inactivation of *cat1* would likely lead to the inability of the strain for NADH oxidization (Lee et al., 2016a).

In this study, firstly, we developed an efficient genome editing tool for *C. tyrobutyricum*, based on the endogenous Type I-B CRISPR-Cas system. The PAM sequences for DNA targeting purposes were identified through *in silico* CRISPR array analysis and *in vivo* plasmid interference assays. By using a lactose inducible promoter to drive the transcription of CRISPR array, multiplex genome engineering purposes have been achieved, with an editing efficiency as high as 100%. Besides, we evaluated the effects of the length of spacer sequences on the transformation efficiency and the genome editing efficiency, and further suggested the appropriate spacer length for efficient genome engineering. Secondly, using the endogenous CRISPR-Cas system, we successfully engineered *C. tyrobutyricum* for efficient butanol production. By introducing *adhE2* and meanwhile deleting *cat1*, the obtained mutant produced record high of 26.2 g/L butanol in a batch fermentation. Results from this study provides essential references and valuable guidance for harnessing endogenous CRISPR-Cas systems for genome engineering purposes. The developed mutant strain can be used as a robust workhorse for efficient biobutanol production from low-value carbon sources, and be further engineered for other valuable biochemical production.

2. Materials and methods

2.1. Bacterial strains and cultivation

All the strains used in this study are listed in Supplementary Table S1. The *E. coli* strain NEB Express (New England BioLabs Inc., Ipswich, MA) was used for general plasmid propagation. *E. coli* CA434 was employed as the donor strain for conjugation (Williams et al., 1990). All *E. coli* strains were routinely cultivated in Luria-Bertani (LB) broth or on solid LB agar plate supplemented with 30 μg/mL chloramphenicol (Cm) or 50 μg/mL kanamycin (Kan) when required. *C. tyrobutyricum* ATCC 25755 (KCTC 5387) was obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and propagated anaerobically at 37 °C in Tryptone-Glucose-Yeast extract (TGY) medium (Wang et al., 2013b). 15 μg/mL thiamphenicol (Tm), 250 μg/mL D-cycloserine, 40 mM lactose or 20 μg/mL uracil was added into the medium when required.

2.2. Identification and analysis of putative protospacer matching CRISPR spacers of C. tyrobutyricum

Nucleotide BLAST (Altschul et al., 1990; Mount, 2007) was used to analyze the CRISPR spacers of *C. tyrobutyricum*, by aligning the spacer

sequences against the existing genome sequences in the National Center for Biotechnology Information (NCBI) database. Putative protospacers were inspected for their matching with the spacers as the putative invading DNA elements, such as phage (prophage), plasmid, transposon, integrase, and so on. For the analysis, we set a maximum of 15% (a maximum of 5/34 mismatching nucleotides) for the mismatches between the putative protospacer and the corresponding CRISPR spacer of *C. tyrobutyricum* (Shariat et al., 2015).

2.3. Plasmid construction

All the plasmids and primers used in this study are listed in Supplementary Table S1 and Supplementary Table S2, respectively. The Phanta Max Super-Fidelity DNA Polymerase (Vazyme Biotech Co., Ltd., Nanjing, China) was used for the PCR to amplify DNA fragments for cloning purposes. For the attempt to delete *spo0A* gene (CTK_RS09345) in *C. tyrobutyricum* using the Type II CRISPR-Cas9 and CRISPR-Cas9 nickase (nCas9) systems derived from *S. pyogenes,* the plasmid pYW34-*BtgZ*I was chosen as the mother vector (Wang et al., 2016). This vector contains the Cas9 open reading frame (ORF) driven by the lactose inducible promoter and the chimeric gRNA sequence preceded by two *BtgZ*I sites (for easy re-targeting purpose by inserting the small RNA (sCbei_5830) promoter along with the 20-nt guiding sequence) (Wang et al., 2016). The vector pJZ23-Cas9 was constructed from pYW34-*BtgZ*I through Gibson Assembly as follows. The erythromycin (Erm) marker and CAK1 replicon of pYW34-*BtgZ*I were replaced with Cm marker and pBP1 replicon, respectively, through an *in vitro* double digestion with Cas9 nuclease following the procedure as described previously (Wang et al., 2016). The Cm marker and the pBP1 replicon were amplified from pMTL82151. The TraJ component which is essential for the conjugation was also amplified from pMTL82151 and cloned into the *Apa*I restriction site of pYW34-*BtgZ*I through Gibson Assembly, generating vector pJZ23-Cas9. To construct pJZ58-nCas9, the P*lac*-Cas9 expression cassette within pJZ23-Cas9 was replaced with the P*lac*-nCas9 expression cassette as follows. A partial fragment of the nCas9 ORF which contains the mutation (D10A) was obtained by PCR using plasmid pMJ841 (Addgene, Cambridge, MA, USA) as the template (Jinek et al., 2012). Then the partial fragment of nCas9 was fused with lactose inducible promoter (which was amplified from pYW34-*BtgZ*I) through Splicing by Overlap Extension (SOE) PCR, yielding the P*lac*-nCas9 expression cassette. The P*lac*-nCas9 expression cassette was cloned into pJZ23-Cas9 by replacing the P*lac*-Cas9 fragment between *Apa*I and *Nhe*I restriction sites, generating pJZ58-nCas9.

Based on pJZ23-Cas9 and pJZ58-nCas9, the small RNA (sCbei_5830) promoter (Wang et al., 2016) fused with the 20-nt guiding sequence (5′-GACATGCTATTGAAGTAGCG-3′) targeting on *spo0A* and two homology arms (~1 kb each) were cloned into the *BtgZ*I and *Not*I sites, respectively, as described previously (Wang et al., 2017), generating pJZ23-Cas9-*spo0A* and pJZ58-nCas9-*spo0A*.

The construction of the other vectors are included in the Supplementary Material.

2.4. Transformation of C. tyrobutyricum

Plasmids used in this study were transformed into *C. tyrobutyricum* via conjugation following published protocols with modifications (Yu et al., 2012, 2011). The donor strain *E. coli* CA434 carrying the recombinant plasmid was cultivated in LB medium supplemented with 30 μg/mL Cm and 50 μg/mL Kan. When the OD_{600} reached 1.5–2.0, about 3 mL *E. coli* CA434 cells were centrifuged and washed twice (with 1 mL fresh LB medium for each wash) to remove the antibiotics. The obtained donor cells were then mixed with 0.4 mL of the recipient culture of *C. tyrobutyricum* (which had an OD_{600} of 2.0–3.0 after an overnight growth in TGY medium). The cell mixture was spotted onto a well-dried TGY agar plate and incubated in the anaerobic chamber at 37 °C for mating purposes. After 24 h, the transconjugants were collected by washing them off the conjugation plate using one mL of TGY medium, and were then spread onto TGY plates containing 15 μg/mL Tm and 250 μg/mL D-cycloserine (for eliminating the residual *E. coli* CA434 donor cells). Transformant colonies could be generally observed after 48–96 h of incubation.

2.5. Mutant screening

The screening of mutants was performed following the protocol as described previously with modifications (Wang et al., 2017, 2016). The transformant colonies of *C. tyrobutyricum* were picked and inoculated into TGY liquid medium with addition of 15 μg/mL Tm (TGYT). The obtained cultures were then diluted serially and spread onto TGY plates supplemented with 40 mM lactose and 15 μg/mL Tm (TGYLT). The plates were incubated anaerobically at 37 °C until colonies were observed. Colony PCR (cPCR) was then performed to screen the putative mutants. When the deletion of *pyrF* is involved, 20 μg/mL uracil was added into TGYLT medium (TGYLTU) to support the growth of Δ*pyrF* strain. When shorter spacer sequence (30 bp) and shorter homology arms (~300 bp) were used for the gene deletion, a series of subculturing (1% v/v inoculum) was carried out using either TGYLT or TGYLTU liquid medium to enrich the desirable homologous recombination, before plating the culture onto the TGYLT or TGYLTU plates for selection.

2.6. Batch fermentation

Batch fermentations with various *C. tyrobutyricum* strains were carried out in 500 mL bioreactors (GS-MFC, Shanghai Gu Xin biological technology Co., Shanghai, China) with a 250 mL working volume. The fermentation medium used in this study was prepared as described previously (Zhang et al., 2017), which comprised (per liter of distilled water): 110 g glucose; 5 g yeast extract; 5 g tryptone; 3 g $(NH_4)_2SO_4$; 1.5 g K_2HPO_4; 0.6 g $MgSO_4{\cdot}7H_2O$; 0.03 g $FeSO_4{\cdot}7H_2O$, and 1 g L-cysteine. The *C. tyrobutyricum* strain was first incubated anaerobically at 37 °C in TGY medium until OD_{600} reached 1.5 and then the active seed culture was inoculated into the bioreactor at a volume ratio of 5%. The fermentation was carried out at pH 6.0 under various temperatures (20, 25, 30, 37 °C). Batch fermentations with *C. beijerinckii* NCIMB 8052 and *C. saccharoperbutylacetonicum* N1–4 under various temperatures (20, 25, 30, 35 °C) were carried out as described previously[42, 48]. Samples were taken every 12 h for the analysis.

2.7. Analytical methods

Cell growth was determined by measuring the optical density at 600 nm (OD_{600}) using a cell density meter (Ultrospec 10, Biochrom Ltd., Cambridge, England). Glucose, acetate, ethanol, butyrate and butanol concentrations in the fermentation broth were analyzed using an HPLC (Agilent 1260 series, Agilent Technologies, Santa Clara, CA, USA) equipped with a refractive index detector (RID) and an Aminex HPX-87H column (Bio-Rad, Hercules, CA, USA). 5 mM H_2SO_4 was used as the mobile phase a flow rate of 0.6 mL/min at 25 °C.

3. Results

3.1. Attempts of genome editing in C. tyrobutyricum with CRISPR-Cas9/Cpf1 systems

Recently, genome editing tools have been developed for several Gram-positive bacteria based on the Type II CRISPR-Cas9/nCas9 system derived from *S. pyogenes* (Bruder et al., 2016; Huang et al., 2016; Li et al., 2016; Pyne et al., 2016; Wang et al., 2017, 2015; Xu et al., 2015), and various Type V CRISPR-Cpf1 systems (Jiang et al., 2017; Yan et al., 2017). Therefore, firstly we sought to explore these systems for genome engineering in *C. tyrobutyricum*. The *spo0A* gene which is the master regulator for sporulation was selected as the target gene to delete. To

abate the strong toxicity of the nuclease/nickase, we constructed CRISPR-Cas9/nCas9/AsCpf1 based vectors by placing the Cas9/nCas9/AsCpf1 encoding gene under the control of a lactose inducible promoter, whereas the gRNA/crRNA were expressed from the constitutive small RNA promoter from *C. beijerinckii* (Wang et al., 2016). In addition, the homology arms for *spo0A* deletion through homologous recombination were inserted into the same plasmid (Wang et al., 2016). The resultant plasmid (pJZ23-Cas9-*spo0A*, pJZ58-nCas9-*spo0A* and pJZ60-AsCpf1-*spo0A*, respectively; Fig. 3B) was attempted to be transformed into *C. tyrobutyricum*. Although numerous attempts were implemented, no transformant could be obtained (Fig. 3C). This indicated that, due to the high toxicity of the heterologous nuclease/nickase and the limited transformation efficiency of *C. tyrobutyricum*, the genome editing is difficult to be realized in this microorganism with the CRISPR-Cas9/nCas9/Cpf1 system. It has been reported that the endogenous CRISPR-Cas system within bacteria and archaea can be harnessed for genome editing for the host microorganism (Li et al., 2015; Vercoe et al., 2013). From the genome sequence, we noticed that *C. tyrobutyricum* possesses a Type I-B CRISPR-Cas system (Lee et al., 2016a). Therefore, we next turned to exploit this endogenous CRISPR-Cas system for genome editing in *C. tyrobutyricum*.

3.2. In silico analysis of the Type I-B CRISPR-Cas system of C. tyrobutyricum

Based on the genome sequence (Lee et al., 2016a), two CRISPR arrays were identified located at two different loci within the *C. tyrobutyricum* genome. The first CRISPR array (Array1) contains 17 spacers (length: 34–38 nt) flanked by direct repeat sequences of 30 nt (5′-ATTGAACCTTAACATGAGATGTATTTAAAT-3′). However, no putative Cas-encoding gene was found at the upstream or downstream of Array1. The second CRISPR array (Array2) was comprised of eight spacers (length: 34–38 nt) flanked by direct repeat sequences of 30 nt (5′-GTTGAACCTTAACATGAGATGTATTTAAAT-3′; which is only one nucleotide different from that of Array1) (Fig. 1A & Supplementary Fig. S1B). A core *cas* gene operon (*cas6-cas8b-cas7-cas5-cas3-cas4-cas1-cas2*) was found at the upstream of Array2, indicating that this CRISPR-Cas system belongs to the Type I-B subtype (Fig. 1A) (Makarova et al., 2015).

The CRISPR-Cas system is known as an immune system, and its spacer sequences are typically derived from the invading genetic elements during the ‘adaptation’ stage (Barrangou and Marraffini, 2014; Charpentier et al., 2015; Plagens et al., 2015; Van Der Oost et al., 2014; Westra et al., 2014). Therefore, we set out to analyze all the 25 spacer sequences specified in Array1 and Array2 using Nucleotide BLAST

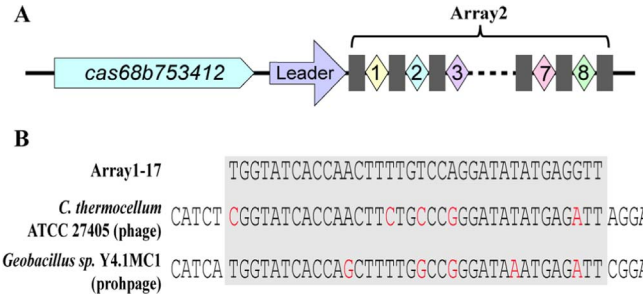


Fig. 1. Characterization of the Type I-B CRISPR-Cas system in *C. tyrobutyricum*. (A) Structure of the central Type I-B CRISPR-Cas locus in the genome of *C. tyrobutyricum*. The central CRISPR-Cas locus possesses a representative Type I-B *cas* operon including *cas6-cas8b-cas7-cas5-cas3-cas4-cas1-cas2* (briefly *cas68b753412*) and the Array2 containing 8 distinct spacers (diamonds) separated by 30-nt direct repeats (rectangles). The transcription of Array2 is driven by a promoter within the leader sequence. (B) Identification of putative protospacer matches via *in silico* analysis of *C. tyrobutyricum* CRISPR spacers. Spacer-protospacer mismatches are indicated with red color. Only five nt of the 5′- and 3′-end adjacent sequences are provided. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Altschul et al., 1990; Mount, 2007), aiming to elucidate whether any spacer sequence matches the putative invading DNA elements, including phage (prophage), plasmid, transposon, and integrase. In order to determine the putative protospacers, a mismatch of less than 15% (5/34 mismatching nucleotides or less) was defined (Shariat et al., 2015). Among all the 25 spacers in the CRISPR-Cas system of *C. tyrobutyricum*, only one spacer sequence (the 17th spacer within Array 1, Array1–17: 5′-TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT-3′) hit (with five mismatches) the putative protospacers found in phage sequence from *C. thermocellum* and prophage sequence from *Geobacillus thermoglucosidasius* (Fig. 1B).

3.3. Identification of protospacer adjacent motif (PAM) sequences

Plasmid transformation interference assay was carried out to test the activity of the Type I-B CRISPR-Cas system of *C. tyrobutyricum* and meanwhile identify the putative PAM sequences (Pyne et al., 2016). The plasmid employed in interference assay contains a protospacer for the DNA targeting purpose and a 5-nt putative PAM sequence located at the 5′- or 3′-end of the protospacer which is essential for the recognition by the Type I CRISPR-Cas system (Table 1). Though the spacer Array1–17 was the only spacer found to match the invading DNA elements, there was no adjacent Cas-encoding genes associated with Array1 discovered. Therefore, additionally we decided to employ another spacer (Array2-1: GCATTCAGACTTGCAACTGTAACTCCCTAGTA CTCCCC) derived from Array2 as the protospacer for the plasmid interference purpose. The 5-nt sequences derived from the upstream or downstream of identified putative protospacers were tested as putative PAM sequences (Fig. 1B and Table 1). Our *in silico* analysis revealed that the *C. tyrobutyricum* CRISPR array possessed high homology to the CRISPR array of *C. pasteurianum*, for both the leader and direct repeat sequences (Supplementary Fig. S1) (Pyne et al., 2016). Therefore, we hypothesized that the CRISPR-Cas system of *C. tyrobutyricum* may share the same or similar PAM with that of *C. pasteurianum*. Hence, the PAM sequences for *C. pasteurianum* Type I-B CRISPR-Cas system were also employed in the plasmid interference assay (Table 1) (Pyne et al., 2016). Altogether, 14 interference plasmids were constructed by combining different protospacer and PAMs (Table 1).

Since both a protospacer and PAM sequence have been included on the interference plasmid, there would be no transformants (the specific plasmid is cleaved and eliminated; we define this as the ‘interference response’) if the CRISPR-Cas system is functional with a particular combination of the protospacer and PAM. As shown in Table 1, no matter what PAM sequences were employed, there was no interference response observed when the protospacer Array1–17 was used. This result suggests that Array1 which does not have an adjacent *cas* gene operon may be silent in the genome of *C. tyrobutyricum*, or it was possibly derived from a gene transfer event which was unrelated to the development of the CRISPR-Cas immunity system in *C. tyrobutyricum*. While combinations of protospacer Array2-1 with 5′ adjacent PAM sequences 5′-CATCA-3′ or 5′-TTTCA-3′, derived from *C. tyrobutyricum* and *C. pasteurianum* respectively, successfully triggered the interference response (Table 1). Plasmids contained combinations of Array2-1 (as the protospacer) and other PAMs were transformed efficiently into *C. tyrobutyricum* (Table 1). These results indicated that Array2 along with the associated core *cas* gene operon in *C. tyrobutyricum* is active and highly functional. Furthermore, the specific PAM sequence located at the 5′-end of the protospacer is essential for the target recognition of Cas proteins.

We used 5-nt PAM sequences in the plasmid transformation interference assay on the basis that most identified PAMs within various microorganisms vary between 2 and 5 nt (Shah et al., 2013). However, it is noteworthy that the two functional PAM sequences contain a conserved 3-nt sequence 5′-TCA-3′ which may play the critical role for the target recognition for *C. tyrobutyricum* Type I-B CRISPR-Cas system. To test our hypothesis, various PAMs (5′-NTCA-3′ with point mutations

Table 1
Effect of different combinations of protospacers and PAM sequences on the transformation efficiency.

Plasmid	5′ PAM	Protospacer[a]	3′ PAM	Transformation efficiency (x 10^2 CFU/mL donor)[b]
pMTL82151	–	–	–	4.9 ± 0.6
pIF-1	CATCT	TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT		4.2 ± 0.8
pIF-2− 2	CATCA	TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT		3.7 ± 0.4
pIF-3		TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT	AGGAT	4.8 ± 0.1
pIF-4		TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT	CGGAT	4.2 ± 0.7
pIF-	AATTG	TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT		3.9 ± 0.5
pIF− 6	TTTCA	TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT		3.3 ± 0.4
pIF-7	TATCT	TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT		5.1 ± 0.2
pIF-8	CATCT	GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC		3.8 ± 0.3
pIF-9	CATCA	GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC		0 ± 0
pIF-10		GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC	AGGAT	3.5 ± 0.9
pIF-11		GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC	CGGAT	4.1 ± 0.1
pIF-12	AATTG	GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC		4.0 ± 0.7
pIF-13	TTTCA	GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC		0 ± 0
pIF-14	TATCT	GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC		3.9 ± 0.9

[a] Two kinds of protospacers were used here: Array1–17(TGGTATCACCAACTTTTGTCCAGGATATATGAGGTT) and Array2-1 (GCATTCAGACTTGCAACTGTAACTCCCTAGTACTCCCC).
[b] Values are average ± standard deviation based on at least two independent replicates.

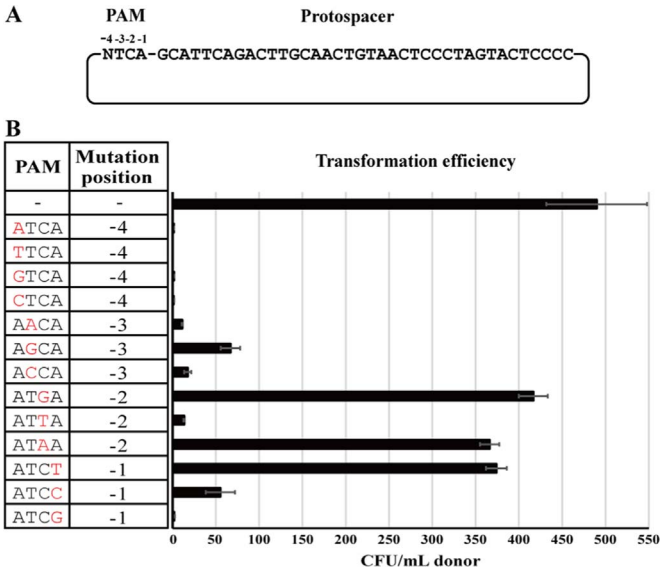


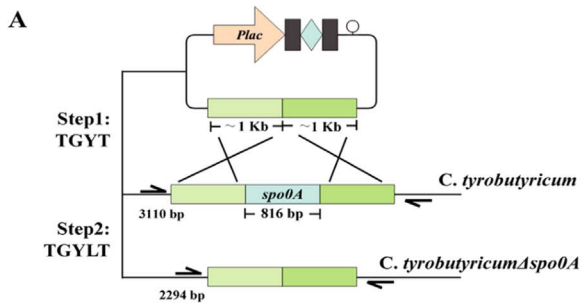
Fig. 2. Identification of protospacer adjacent motif (PAM) sequences of the Type I-B CRISPR-Cas system in *C. tyrobutyricum*. (A) The map of plasmids used in systematic mutagenesis assays. Mutation positions were indicated on the PAM sequence. Array2-1 was used as the protospacer. (B) Variant PAM sequences used in the assay and their corresponding transformation efficiencies. The plasmid pMTL82151 (PAM, -; Mutation position, -) was used as the control. Data are based on at least two independent replicates.

at different positions) built upon 5′-TCA-3′ were systematically evaluated for their functionality (Fig. 2). As shown in Fig. 2B, significant differences in the transformation efficiency were observed with plasmids containing different PAMs (along with Array2-1 as the protospacer). All the plasmids contained point mutations at position –4 triggered the interference response, suggesting that the first three nucleotides (5′-TCA-3′) encompass the core PAM sequence. When 'T' located at position –3 was mutated, only slightly increased transformation efficiency was obtained, indicating that the nucleotide on position –3 had a minor effect on target recognition. Nevertheless, high transformation efficiency (comparable to the control plasmid pMTL82151) was observed when 'C' located at position –2 was mutated to 'G' or 'A' or 'A' located at position –1 was mutated to 'T'. The transformation efficiency was slightly increased (compared to that with 5′-NTCA-3′) when 'A' located at position –1 was mutated to 'C', while 'TCG' kept the similar level of transformation efficiency with 5′-NTCA-3′. These data demonstrated that, for the appropriate function of the PAM sequence, pyrimidine nucleotides ('C' and 'T'), rather than purine nucleotides ('G' and 'A'), are more preferable at the position –2, and conversely, purine nucleotides are better options than pyrimidine nucleotides at the position –1. Overall, 3-nt sequences 5′-TCA-3′ (TCA) and 5′-TCG-3′ (TCG) (also written as TCR collectively for both) which led to an approximately 1,000-fold drop in plasmid transformation efficiency (compared to the control plasmid pMTL82151, Fig. 2B) were concluded to be the functional PAM sequences of the Type I-B CRISPR-Cas system in *C. tyrobutyricum*.

3.4. Development of an inducible CRISPR-Cas system for genome editing in *C. tyrobutyricum*

Known that the endogenous Type I-B CRISPR-Cas system of *C. tyrobutyricum* was functional and had high interference activity against plasmids possessing proper protospacer and PAM sequences, we then attempted to engineer this system to be a genome editing tool for *C. tyrobutyricum*. Two parts are required for such a system: 1) a synthetic CRISPR expression cassette, containing a spacer targeting on the specific genome sequence; 2) gene editing cassette, comprised of a pair of homology arms to achieve homologous recombination (Fig. 3A and Supplementary Fig. S2) (Wang et al., 2016). The *spo0A* gene was selected as the first target gene to be deleted. The 816 bp *spo0A* ORF contains a total of 28 potential PAMs (TCR) including 24 TCA and 4 TCG. One of the PAM (TCA) along with its downstream protospacer sequence (38-nt *spo0A* spacer1) was selected as the target site. Plasmid pJZ69-leader-38*spo0A*, comprised of a synthetic CRISPR expression cassette and a *spo0A* editing cassette (upstream and downstream homology arms, ~1 kb each), was constructed to delete the *spo0A* gene in *C. tyrobutyricum* (Supplementary Fig. S2). In the synthetic CRISPR expression cassette, the native CRISPR leader and terminator sequences were used to drive the transcription of synthetic CRISPR array which contained the 38-nt *spo0A* spacer1 flanked by 30-nt direct repeat sequences (Supplementary Fig. S2). Conjugation was carried out. However, no transformants were obtained with pJZ69-leader-38*spo0A*, although the expected transformation efficiency was obtained with pMTL82151 as the control. Many attempts have been conducted, and the results were consistently the same (data not shown). Therefore, it seemed like, even with the endogenous CRISPR-Cas system, the instant expression could be highly toxic to the cells and thus no transformants could be obtained.

Generally, the leader sequence of the CRISPR array contains a promoter for CRISPR array transcription and a regulatory signal for the uptake of new spacer-repeat elements (Arslan et al., 2014; Erdmann and Garrett, 2012; Li et al., 2014; Pougach et al., 2010; Pul et al., 2010; Yosef et al., 2012). In this study, however, for the genome editing



B

Plasmid	Promoter for *Cas* gene or RNA guide	gRNA or spacer
pJZ23-Cas9-*spo0A*	P*lac* for Cas9	20 nt gRNA
pJZ58-nCas9-*spo0A*	P*lac* for nCas9	20 nt gRNA
pJZ60-AsCpf1-*spo0A*	P*lac* for AsCpf1	23 nt crRNA
pJZ69-leader-38*spo0A*	Leader for CRISPR array	38 nt *spo0A* spacer1
pJZ74-P*lac*-10*spo0A*	P*lac* for CRISPR array	10 nt *spo0A* spacer1
pJZ74-P*lac*-20*spo0A*	P*lac* for CRISPR array	20 nt *spo0A* spacer1
pJZ74-P*lac*-30*spo0A*	P*lac* for CRISPR array	30 nt *spo0A* spacer1
pJZ74-P*lac*-38*spo0A*	P*lac* for CRISPR array	38 nt *spo0A* spacer1
pJZ74-P*lac*-50*spo0A*	P*lac* for CRISPR array	50 nt *spo0A* spacer1
pJZ75-P*lac*-38*spo0A*	P*lac* for CRISPR array	38 nt *spo0A* spacer2
pJZ76-P*ara*-38*spo0A*	P*ara* for CRISPR array	38 nt *spo0A* spacer1

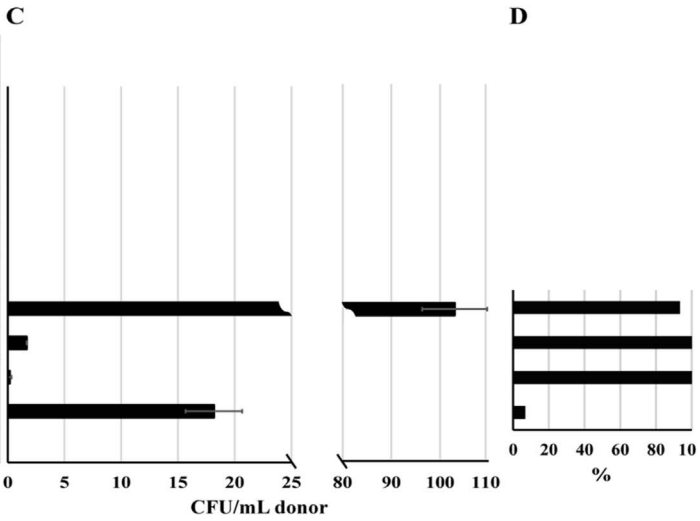


Fig. 3. Markerless genome editing in *C. tyrobutyricum* using the endogenous Type I-B CRISPR-Cas system. (A) Schematic illustrating the work flow of *spo0A* gene deletion using the lactose inducible CRISPR-Cas system. The lactose inducible promoter was used to driven the transcription of synthetic CRISPR array. ~1 kb upstream and downstream homology arms were used for the deletion of *spo0A* gene. Two screening steps are involved here. In the first step, the plasmid was transformed into *C. tyrobutyricum* under the selection of Tm. In the second step, lactose was applied to induce the transcription of synthetic CRISPR array and eliminate the wild type background cells, thus selecting for the desirable mutant. Pairs of half arrows and the numbers in the figure indicate the cPCR target regions and the PCR amplicon sizes, respectively. (B) Various plasmids carrying the CRISPR-Cas9/nCas9/AsCpf1 and Type I-B CRISPR-Cas systems were tested for the deletion of *spo0A*. Promoters and the length of spacers were optimized for the CRISPR-Cas system in order to improve the transformation efficiency and editing efficiency. *Plac*, lactose inducible promoter; *Para*, arabinose inducible promoter. (C) The transformation efficiency of different plasmids. Data are based on at least two independent replicates. (D) The genome editing efficiency of different plasmids that can be transformed into *C. tyrobutyricum*. Fifteen colonies of each transformant were picked and screened for mutation. The editing efficiency were calculated as the ratio of the number of *spo0A* mutants to the total of fifteen colonies.

purposes, only the promoter function of the leader sequence is needed. In order to reduce the toxicity of endogenous CRISPR-Cas system, a lactose inducible promoter and an arabinose inducible promoter were evaluated for the transcription of the synthetic CRISPR array in place of the native leader sequence (Fig. 3A and B) (Wang et al., 2016). The resultant plasmids pJZ74-P*lac*-38*spo0A* and pJZ76-P*ara*-38*spo0A* were tried to transform into *C. tyrobutyricum*. Transformants were generated with pJZ74-P*lac*-38*spo0A*, with an overall transformation efficiency of 1.7 CFU/mL donor (Fig. 3C); while the transformation with pJZ76-P*ara*-38*spo0A* was failed, suggesting that the arabinose inducible promoter was less stringent than the lactose inducible promoter for the expression of the CRISPR array in *C. tyrobutyricum* (Fig. 3C). As a control (or as a means to further confirm the appropriate PAM sequence), a 38-nt *spo0A* spacer2 (corresponding PAM: TCT) was employed to replace the 38-nt *spo0A* spacer1 in pJZ74-P*lac*-38*spo0A*, generating plasmid pJZ75-P*lac*-38*spo0A*. Results demonstrated that the transformation efficiency with pJZ75-P*lac*-38*spo0A* (~ 18.2 CFU/mL donor) increased more than an order of magnitude compared to that with pJZ74-P*lac*-38*spo0A* (Fig. 3C). The obtained transformants (with either pJZ74-P*lac*-38*spo0A* or pJZ75-P*lac*-38*spo0A*) were cultivated in TGYT medium, and then spread onto TGYLT plates to induce the expression of the synthetic CRISPR array. Colony PCR was carried out with randomly picked colonies to screen the *spo0A* deletion mutants. Results showed that one out of fifteen (6.7%) of the tested colonies was *spo0A* deletion mutant (Δ*spo0A*) from the transformants with pJZ75-P*lac*-38*spo0A* (Fig. 3D). While all tested colonies were Δ*spo0A* mutants from the transformants with pJZ74-Plac-38*spo0A*, representing an editing efficiency of 100% (Fig. 3D). These results confirmed our above conclusion concerning the PAM sequence: the targeting efficiency of TCA is much higher than TCT. The Δ*spo0A* mutant was further verified by Sanger sequencing (data not shown). Collectively, we proved that with the inducible endogenous CRISPR-Cas system, efficient genome editing could be achieved in *C. tyrobutyricum*.

3.5. Effects of spacer length on transformation efficiency and genome editing efficiency

In *C. tyrobutyricum* genome, a total of 25 spacer sequences were identified in Array1 and Array2 with lengths ranging from 34 to 38 nt. In order to mimic the feature of the native Type I-B CRISPR array, the 38-nt *spo0A* spacer1 was employed to develop the genome editing platform for the deletion of *spo0A*. However, it is reasonable to question whether the length of the spacer has an effect on the transformation efficiency and genome editing efficiency of the CRISPR-Cas genome engineering platform. To answer this question, we replaced the 38-nt *spo0A* spacer1 in plasmid pJZ74-P*lac*-38*spo0A* with 10 nt, 20 nt, 30 nt, and 50 nt of *spo0A* spacer1 (while the PAM sequence TCA was kept the same), yielding pJZ74-P*lac*-10*spo0A*, pJZ74-P*lac*-20*spo0A*, pJZ74-P*lac*-30*spo0A*, and pJZ74-P*lac*-50*spo0A*, respectively (Fig. 3B). Surprisingly, no transformant was obtained after several attempts with pJZ74-P*lac*-10*spo0A* or pJZ74-P*lac*-20*spo0A*. This might be because the shorter spacer sequences (≤ 20 nt) led to severe off-target effects which killed the host cells (Fig. 3C). However, when 30-nt, 38-nt and 50-nt spacers were used, transformation efficiencies of 103.0 CFU/mL donor, 1.7

CFU/mL donor and 0.2 CFU/mL donor were obtained, respectively (Fig. 3C). The longer spacers can bind more tightly to the target and thus increase the self-targeting activity of the endogenous CRISPR-Cas system, which may contribute to the decreased transformation efficiency. The genome editing efficiency was also assessed for the transformants obtained with pJZ74-P*lac*-30*spo0A*, pJZ74-P*lac*-38*spo0A* or pJZ74-P*lac*-50*spo0A*. Interestingly, colonies of various sizes were observed for the transformants harboring pJZ74-P*lac*-30*spo0A* on the TGYLT plates, while the colonies from the other two transformants appeared homogeneous in sizes (Supplementary Fig. S3). Then, large and small colonies of the transformant harboring pJZ74-P*lac*-30*spo0A* were picked separately to screen for the Δ*spo0A* mutant, and editing efficiencies of 93.3% and 13.3% were obtained, respectively (Supplementary Fig. S3B). The different genome editing efficiency for large and small colonies might be due to the low self-targeting activity of the endogenous CRISPR-Cas system when 30-nt spacer was employed. In this case, some of the host cells could survive from the selection of the endogenous CRISPR-Cas system, but their growth was still inhibited. So, as shown in Supplementary Fig. S3B, most of the small colonies were wild type cells with growth inhibited, whereas most of the large colonies were mutant cells without growth interference because their target site for the CRISPR-Cas system had been eliminated. On the other hand, the editing efficiencies of transformants obtained with pJZ74-P*lac*-38*spo0A* or pJZ74-P*lac*-50*spo0A* were both 100% (Fig. 3D).

3.6. Multiplex genome engineering

As described above, single gene deletion was achieved with high efficiency using the inducible endogenous CRISPR-Cas system. Here, we further explored this system for multiplex genome editing in *C. tyrobutyricum*. The *pyrF* gene encoding the enzyme orotidine 5-phosphate decarboxylase (involved in the *de novo* pyrimidine biosynthesis) together with the *spo0A* gene were selected as targets to delete. In order to have the CRISPR-Cas system target onto two loci at the same time, we simply put two spacers targeting on *spo0A* and *pyrF* respectively onto the same CRISPR array insulated by three direct repeats (Fig. 4A). Considering that the longer spacer is more toxic to the host cells as we demonstrated above, 30 nt was used for both spacers (*spo0A* spacer1 and *pyrF* spacer3). In addition, as we noticed that, with the increase of the plasmid size, the transformation efficiency decreases dramatically (especially when the vector size > 10 kb; data not shown), we used shorter homology arms for the deletion of both genes (two homology arms for the deletion of each gene, with the length of each arm is ~300 bp), to keep the final vector size < 9 kb (Fig. 4A). Control plasmids pJZ77-P*lac*-30*spo0A* and pJZ77-P*lac*-30*pyrF* were also constructed for deleting *spo0A* and *pyrF* individually by using the same corresponding modules (spacer and homology arms) in pJZ77-P*lac*-30*spo0A*/30*pyrF* for deleting *spo0A* and *pyrF*, respectively. The three plasmids were successfully transformed into *C. tyrobutyricum*, and the resulting transformants were then spread onto TGYLTU plates. However, no mutant was detected (47 colonies from each transformant were screened with cPCR) for any of the three transformants, which was not surprising considering the reduced editing efficiency when shorter spacers and homology arms were used. In order to enrich the desirable homologous recombination, a series of subculturing was performed in TGYLTU liquid medium. Then mutant screening was performed with cPCR for the 8th and 15th generations of the subculture. For the 8th generation, for *spo0A* and *pyrF* deletion respectively, editing efficiencies of 59.6% and 40.4% were obtained with the one-spacer approach (using pJZ77-P*lac*-30*spo0A* and pJZ77-P*lac*-30*pyrF*, respectively), while editing efficiencies of 53.2% and 31.9% were obtained with the two-spacer approach (using pJZ77-P*lac*-30*spo0A*/30*pyrF*) (Fig. 4B). In addition, double deletion was also detected with the two-spacer approach, but at a much lower rate (6.4%) (Fig. 4B). For the 15th generation, up to 100% editing efficiencies were observed for *spo0A* and *pyrF* deletion with both one-spacer and two-spacer approaches, which meant that as high as 100% editing efficiency for the double deletion was also achieved with the two-spacer approach (Fig. 4C).

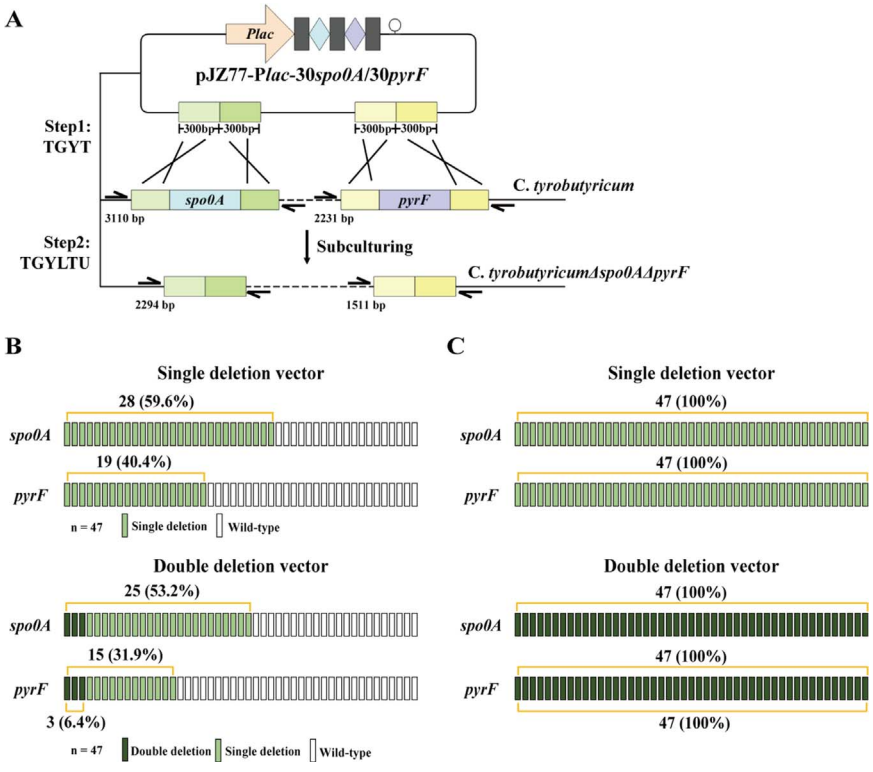


Fig. 4. Multiplex gene editing in *C. tyrobutyricum* using the inducible endogenous Type I-B CRISPR-Cas system. (A) Schematic illustrating the work flow of *spo0A* and *pyrF* double deletion using lactose inducible CRISPR-Cas system. A total of ~1.2 kb upstream and downstream homology arms of *spo0A* and *pyrF* (~300 bp each) were employed in the double deletion vector. Two 30-nt spacers targeting on *spo0A* and *pyrF* genes, respectively, were used. The screening procedure of double deletion was similar with that for single deletion (Fig. 3A), except that a series of subculturing was required before plating the culture on the TGYLTU plates. Pairs of half arrows and the numbers in the figure indicate the cPCR target regions and the PCR amplicon sizes, respectively. Detection of gene deletion events was carried out at the 8th (B) and 15th (C) generations during the subculturing. Single deletion vectors pJZ77-P*lac*-30*spo0A* and pJZ77-P*lac*-30*pyrF* were used as controls. 47 colonies of each transformant were picked and screened for mutations. The open rectangle, light green rectangle, and dark green rectangle represent wild type strain, single deletion mutant of *spo0A* or *pyrF*, and double deletion mutant, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

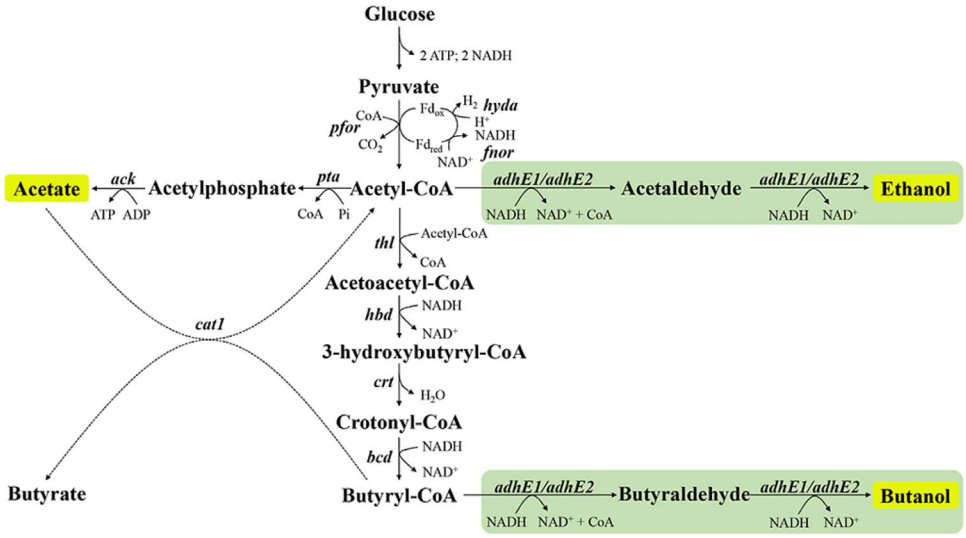


Fig. 5. Schematic diagram of the metabolic pathway of Δ*cat1*::*adhE1* and Δ*cat1*::*adhE2*. The major products of the two mutants are shown in yellow boxes, and the ethanol and butanol biosynthesis pathways which are absent in the wild type strain are shown in green boxes. The butyrate biosynthesis pathway which is disrupted from the wild type strain is shown with dotted lines. Key genes in the pathway: *pfor*, pyruvate::ferredoxin oxidoreductase; *hyda*, hydrogenase; *fnor*, ferredoxin NAD^+ oxidoreductase; *pta*, phosphotransacetylase; *ack*, acetate kinase; *thl*, thiolase; *hbd*, beta-hydroxybutyryl-CoA dehydrogenase; *crt*, crotonase; *bcd*, butyryl-CoA dehydrogenase; *cat1*, butyrate: acetate coenzyme A transferase; *adhE1/adhE2*, aldehyde-alcohol dehydrogenase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.7. Engineer *C. tyrobutyricum* for butanol production

C. tyrobutyricum is a hyper-butyrate producer (Fu et al., 2017), indicating that the metabolic pathway from glucose to butyryl-CoA is highly favorable (Fig. 5). Therefore, using the high efficient endogenous CRISPR-Cas system, we attempted to engineer the *C. tyrobutyricum* for hyper-butanol production. Two aldehyde/alcohol dehydrogenase genes (*adhE1* and *adhE2*) which can convert butyryl-CoA to butanol were chosen to introduce into *C. tyrobutyricum*. In order to drive more metabolic flux towards C4 products, the *pta-ack* operon which was responsible for acetate formation was initially selected to be deleted or replaced by *adhE1/adhE2* (Fig. 5). However, none of the attempts was successful (data not shown), suggesting that the *pta-ack* operon was vital for *C. tyrobutyricum* metabolism and thus cannot be deleted. This result is coincident with that was reported by Lee and co-workers (Lee et al., 2016a).

In *C. tyrobutyricum*, *cat1* is the essential gene for butyrate biosynthesis, and the *ptb-buk* operon as seen in solventogenic clostridial strains does not exist (Lee et al., 2016a) (Fig. 5). Therefore, we hypothesized that the deletion of *cat1* could possibly eliminate butyrate production, and thus the introduction of *adhE1/adhE2* can lead to the conversion of butyryl-CoA for enhanced butanol production. However, it was previously reported that the disruption of *cat1* was not achievable (with the mobile group II intron), because the inactivation of *cat1* could likely lead to the inability of the strain for NADH oxidization (Lee et al., 2016a). Here, based on the high efficient CRISPR-Cas system for genome engineering, we attempted to delete the *cat1* gene or replace it with *adhE1* or *adhE2*. Similar as the previous report, the deletion of *cat1* was fruitless despite numerous attempts, however the replacement of *cat1* with *adhE1* or *adhE2* was successful, yielding mutants Δ*cat1*::*adhE1* and Δ*cat1*::*adhE2*, respectively. As the control, the recombinants WT(pJZ98-P*cat1-adhE1*) and WT(pJZ98-P*cat1-adhE2*) were also obtained by introducing the plasmid-based *adhE1* and *adhE2* (driven by the *cat1* promoter) overexpression vectors into *C. tyrobutyricum*. Initial batch fermentations were carried out at 37 °C (the optimum temperature for the cell growth of *C. tyrobutyricum*). Results demonstrated that acetate (14.8 g/L), ethanol (9.7 g/L) and butanol (8.7 g/L) were the major products with a low level of butyrate (1.3 g/L) produced for the control strain WT(pJZ98-P*cat1-adhE1*) (Table 2 and Supplementary Fig. S4A). However, for WT(pJZ98-P*cat1-adhE2*), acetate (6.9 g/L), ethanol (7.4 g/L) and butyrate (34.1 g/L) were the major products, with only a small amount of butanol (2.0 g/L) was produced (Table 2 and Supplementary Fig. S5A). As we expected, with the butyrate biosynthesis pathway replaced with the butanol producing pathway, mutants Δ*cat1*::*adhE1* and Δ*cat1*::*adhE2* produced negligible butyrate (0.6–0.8 g/L) but high levels of butanol (15.0 g/L). In addition, significant amounts of acetate (15.1–20.8 g/L) and ethanol (5.2–5.3 g/L) were also produced by the two mutants (Table 2 and Supplementary Figs. S4B and S5B).

3.8. Enhance butanol production by carrying out fermentation at low temperatures

It is well known that the limited butanol tolerance of the host is a major bottleneck for butanol production in microorganisms (Peabody and Kao, 2016). Recent studies showed that lower temperature could alleviate the alcohol toxicity and thus increase the alcohol production (Knoshaug and Zhang, 2009; Ramio-Pujol et al., 2015; Zhang et al., 2016a). Therefore, batch fermentations were further carried out at 30, 25 and 20 °C with Δ*cat1*::*adhE1* and Δ*cat1*::*adhE2*, respectively. As seen in Table 2 and Supplementary Figs. S4 and S5, the acetate production was kept at the similar levels at different temperatures. However, the production of ethanol and butanol was significantly increased at these lower temperatures. Butanol titers for Δ*cat1*::*adhE1* and Δ*cat1*::*adhE2* obtained at 20 °C were 21.4 and 26.2 g/L, respectively, which increased by 42.7% and 74.7%, respectively, compared with that obtained at 37 °C. While the total BE (butanol and ethanol) production of Δ*cat1*::*adhE1* and Δ*cat1*::*adhE2* reached the maximum of 35.6 and 38.2 g/L, respectively at 25 °C.

4. Discussion

Within the past few years, CRISPR-Cas, the adaptive immune system from bacteria and archaea, has been repurposed for versatile genome editing and transcriptional regulation in various species. However, so far, the majority of such applications are based on the Type II CRISPR-Cas9 system derived from *S. pyogenes*. Due to the unique feature of the chromosome of prokaryotic cells, the expression of the heterologous Cas9 is highly toxic, thus leading to poor transformation efficiency and failure of genome editing (Wang et al., 2016). Recently, the type V CRISPR-Cpf1 system has also been exploited for genome editing purposes (Zetsche et al., 2015). It has advantages over the CRISPR-Cas9 system due to its smaller size of the effector protein (Cpf1) and the more compact RNA guide (crRNA) (Zetsche et al., 2015). Although the toxicity of Cpf1 is much lower than that of Cas9 as demonstrated in specific strains, remarkable decrease in transformation efficiency is still observed with the expression of Cpf1 in the host (Ungerer and Pakrasi, 2016). Therefore, it is challenging to carry out genome editing with CRISPR-Cas9/Cpf1 systems in microorganisms with low DNA transformation efficiencies. On the other hand, the endogenous CRISPR-Cas

Table 2
Summary of fermentation results for *C. tyrobutyricum* mutants at various temperatures[a].

Strain	Temperature (°C)	Glucose consumption (g/L)	Acetate (g/L)	Butyrate (g/L)	Ethanol (g/L)	Butanol (g/L)	Total BE (g/L)	BE yield (g/g of glucose)
WT(pJZ98-P*cat1-adhE1*)	37	79.1	14.8	1.3	9.7	8.7	18.4	0.23
WT(pJZ98-P*cat1-adhE2*)	37	109.0	6.9	34.1	7.4	2.0	9.4	0.09
Δ*cat1*::*adhE1*	37	87.1	20.8	0.6	5.3	15.0	20.3	0.23
Δ*cat1*::*adhE2*	37	75.5	15.1	0.8	5.2	15.0	20.2	0.27
Δ*cat1*::*adhE1*	30	109.6	22.5	0.8	14.3	17.2	31.5	0.29
Δ*cat1*::*adhE2*	30	96.7	12.3	1.3	10.8	21.1	31.9	0.33
Δ*cat1*::*adhE1*	25	111.9	22.8	1.3	16.6	19.0	35.6	0.32
Δ*cat1*::*adhE2*	25	109.4	13.9	1.8	12.8	25.4	38.2	0.35
Δ*cat1*::*adhE1*	20	111.9	21.8	1.6	10.4	21.4	31.8	0.28
Δ*cat1*::*adhE2*	20	112.2	15.2	2.4	8.9	26.2	35.1	0.31

[a] The fermentation profiles are provided in Supplementary Figs. S4 and S5; values are based on at least two independent replicates.

system has been harnessed for genome editing and transcriptional regulation in a few species of bacteria (Gomaa et al., 2014; Luo et al., 2015; Vercoe et al., 2013) and archaea (Li et al., 2015; Zebec et al., 2014). In this case, when the endogenous CRISPR-Cas machinery is employed, the transformation efficiency and thus the chance of success for genome editing are normally much higher than the case when heterologous CRISPR-Cas9/Cpf1 systems are used (Pyne et al., 2016). In this work, after many unsuccessful attempts for genome editing with the CRISPR-Cas9 or CRISPR-AsCpf1 systems, we successfully repurposed the Type I-B CRISPR-Cas system of *C. tyrobutyricum* as an efficient genome editing tool for this microorganism.

In silico analysis of the CRISPR array in *C. tyrobutyricum* identified only one spacer sequence that can match protospacers from phage (prophage) of *Clostridium* and *Geobacillus* (Fig. 1B). However, we hypothesized that, due to the possible horizontal transferring property of CRISPR-Cas loci between closely-related species, the Type I-B CRISPR-Cas systems from different *Clostridium* species could be very similar and share similar/same PAMs and direct repeat sequences (Pyne et al., 2016; Rath et al., 2015). Indeed, our subsequent *in silico* analysis demonstrated high homology between the CRISPR array in *C. tyrobutyricum* and that in *C. pasteurianum* (Supplementary Fig. S1) (Pyne et al., 2016). Therefore, the three PAM sequences from *C. pasteurianum* along with the putative PAMs identified in *C. tyrobutyricum* were employed to assess the activity of the endogenous CRISPR-Cas system of *C. tyrobutyricum*. The *in vivo* plasmid interference assay revealed that the Cas protein in *C. tyrobutyricum* had high affinity to the 5′ adjacent PAM sequences TCA and TCG (Fig. 2B). These results verified our hypothesis that the Type I-B CRISPR-Cas system from *C. tyrobutyricum* shares the same PAM sequence (TCA) as that in *C. pasteurianum*, as well as those in *Clostridium tetani* and *C. thermocellum* (Pyne et al., 2016).

In attempt for the genome editing with the endogenous CRISPR-Cas system, initially, the native leader sequence was used as the promoter to drive the transcription of the synthetic CRISPR array. However, no transformants were obtained, likely due to the toxicity of the endogenous CRISPR-Cas system when it was instantly expressed. Then, a lactose inducible promoter was employed to replace the leader sequence to drive the expression of the CRISPR-Cas system, resulting in an overall transformation efficiency of 1.7 CFU/mL donor (Fig. 3C). This transformation efficiency is still low, but is enough to enable us to obtain desirable mutants with a high editing efficiency. With this, we demonstrated that the inducible expression of the endogenous CRISPR-Cas array is achievable (although the configuration of the original native leader sequence for the CRISPR array regulation was complex) and innovative to realize efficient genome editing in the host microorganism. It is also worthwhile to point out that, the same inducible promoter was tried to drive the expression of Cas9, nCas9 or AsCpf1 proteins to achieve genome editing for the same microorganism, however no successful transformation can be achieved with any of the plasmids containing these heterologous nuclease (or nickase) proteins. This confirmed that the toxicity of the endogenous CRISPR-Cas system is much lower than that of heterologous CRISPR-Cas9/nCas9/AsCpf1 systems and thus more implementable for genome editing purposes (Fig. 3B and C).

Although the markerless genome engineering platform was developed and high editing efficiency could be obtained, the transformation efficiency was still low which would restrict the application of the genome editing platform in *C. tyrobutyricum*. The length of spacers identified from the CRISPR Arrary1 and Arrary2 are not all the same (ranging from 34 to 38 nt). We reasoned that the length of the spacer might have an impact on the transformation efficiency and/or genome editing efficiency. Therefore, various lengths of spacers were systematically evaluated in the developed CRISPR-Cas system in the context for *spo0A* deletion. Results indicated that, the transformation was not successful when the spacer $\leq$ 20 nt was used, suggesting possible severe off-target effects (Fig. 3C). Spacers ranging from 30 to 50 nt can be used for targeting purposes for the successful genome editing. Comparatively, when shorter spacers were used, the genome editing efficiency was slightly decreased (for 30-nt spacer, an editing efficiency of 93.3% was obtained based on the large colonies), but meanwhile the transformation efficiency was dramatically enhanced (by approximately 500-fold for the 30-nt spacer). Therefore, depending on the different genome editing purposes, one can make a tradeoff between the transformation efficiency and genome editing efficiency by using a spacer of an appropriate length. Briefly, based on the above results for the deletion of *spo0A* particularly, spacers of 30–38 nt seems good options. However, this could be target site specific and/or strain specific. To our best knowledge, this is the first report describing the effect of the spacer length in the endogenous CRISPR-Cas system on the off-target effect, transformation efficiency and also genome editing efficiency. It should be pointed out that, one of the advantages with such an endogenous CRISPR-Cas system comparing to the type II CRISPR-Cas9 system for genome editing is that, the employment of the longer spacer sequence (30–38 nt vs. 20 nt for the spCRISPR-Cas9) can abate the potential off-target effect. Apparently, the longer the spacer sequence is, the more specific the targeting of the crRNA. For eukaryotic cells, the off-target effect can lead to unspecific mutations on the chromosome, which is highly problematic for various applications. This won't occur for the prokaryotic cells due to their inefficient endogenous nonhomologous end-joining (NHEJ) capability for the automatic DNA repairing. However, the off-target effect in prokaryotic cells can lead to cell death and thus failure of genome editing.

CRISPR-Cas9-mediated multiplex genome editing has been reported by several groups (Cong et al., 2013; Jiang et al., 2015; Sakuma et al., 2014; Westbrook et al., 2016; Xie et al., 2015). However, in order to edit multiple loci in a single step, multiple plasmids or a large construct carrying multiple guide RNAs are required to be delivered into the cells.

Recently, the CRISPR-Cpf1 system has also been engineered to edit several loci of the chromosome at the same time by using a single CRISPR array that encoded multiple spacer sequences (Zetsche et al., 2017). Whereas, the toxicity issues of Cas9/Cpf1 are still obstacles for the employment of these machineries in the species with low DNA transformation efficiencies. In this study, multiplex genome editing was achieved by using the endogenous CRISPR-Cas system of *C. tyrobutyricum* (Fig. 4A). A synthetic CRISPR array carrying two spacers was used for the chromosome targeting to delete *spo0A* and *pyrF* simultaneously, yielding an editing efficiency of up to 100% (Fig. 4C). To date, this is the first success for multiplex genome editing in microorganisms with underdeveloped genome engineering tools such as *Clostridium*.

C. tyrobutyricum is a natural hyper-butyrate producer, which has been engineered for butanol production previously (Yu et al., 2011). Yu et al. (2011) reported that, with the deletion of *ack* or *ptb* along with the overexpression of *adhE2*, the *C. tyrobutyricum* mutant could produce up to 10 g/L butanol. However, based on the genome sequence and results by Lee et al. (2016a), as well as the results in this study, the *ptb* gene is not present in the genome of *C. tyrobutyricum* and the *ack* gene can't be deleted because the elimination of acetate production pathway would lead to no endproduct and inefficient ATP generation (Fig. 5). Based on the results from Lee et al. (2016a), the *cat1* gene was actually the essential gene for butyrate production in *C. tyrobutyricum*, and the deletion of *cat1* was not achievable either. In this study, based on the developed CRISPR-Cas genome engineering system, we successfully replaced the *cat1* gene with *adhE1/adhE2*. In this way, the butyrate production in *C. tyrobutyricum* was almost eliminated and the microorganism was converted into a hyper-butanol producer (Fig. 5). Previous studies have demonstrated that the lower temperature is beneficial to enhance the butanol tolerance of host strains (Baer et al., 1987; Knoshaug and Zhang, 2009), which may be because of the change of cell membrane composition and fluidity under lower temperatures (Baer et al., 1987). Therefore, here fermentations for butanol production with the *C. tyrobutyricum* mutant were further carried out at lower temperatures. At 20 °C, the butanol production in the mutant Δ*cat1*::*adhE2* reached 26.2 g/L in a regular batch fermentation. To the best of our knowledge, this is the highest butanol production that has ever been reported in a batch fermentation. We also investigated the butanol production of *C. beijerinckii* NCIMB 8052 and *C. saccharoperbutylacetonicum* N1–4 at lower temperatures (Supplementary Table S3), to confirm whether carrying out fermentations at low temperatures is a broadly applicable mechanism to achieve high butanol production with other strains as well. Although the butanol production of these two solventogenic clostridia was increased at lower temperatures, the increment was far lower than that obtained with mutant Δ*cat1*::*adhE2*, indicating that *C. tyrobutyricum* has much greater potential and thus is a more favorable host for butanol production. Furthermore, there is no acetone production in the fermentation of Δ*cat1*::*adhE2* as seen in the ABE fermentation; butanol, ethanol and acetate are the only primary endproducts. This on one hand simplifies the downstream recovery process; on the other, these endproducts could be further upgraded to high-value biochemicals (such as diesel, esters, etc.) through chemical or biochemical processes (Gangadwala et al., 2003; Zhang et al., 2017).

Overall, the inducible endogenous CRISPR-Cas platform developed herein extends the existing genetic toolbox for genome editing. The protocol developed in this study can be readily adapted to other microorganisms that carry active endogenous CRISPR-Cas systems. Finally, the mutant Δ*cat1*::*adhE2* generated using the endogenous CRISPR-Cas system exhibited great potential for economic biobutanol production.

Acknowledgements

We thank Dr. Hans Blaschek (University of Illinois at Urbana-Champaign) for providing pYW34-*BtgZI*, Dr. Mike Young (Aberystwyth University, UK) for providing *E. coli* CA434, and Dr. Nigel Minton (University of Nottingham, UK) for providing pMTL82151. Auburn University has filed a patent application covering the work described in this article. The application names Y.W. and J.Z. as inventors.

Funding

This work was supported by the United States Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) Hatch program.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2018.03.007.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., Pul, Ü., 2014. Detection and characterization of spacer integration intermediates in type I-E CRISPR–Cas system. Nucleic Acids Res. 42, 7884–7893.
Baer, S.H., Blaschek, H.P., Smith, T.L., 1987. Effect of butanol challenge and temperature on lipid composition and membrane fluidity of butanol-tolerant *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 53, 2854–2861.
Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709–1712.
Barrangou, R., Marraffini, L.A., 2014. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Mol. Cell 54, 234–244.
Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S.D., 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151, 2551–2561.
Bruder, M.R., Pyne, M.E., Moo-Young, M., Chung, D.A., Chou, C.P., 2016. Extending CRISPR-Cas9 technology from genome editing to transcriptional engineering in the genus *Clostridium*. Appl. Environ. Microbiol. 82, 6109–6119.
Charpentier, E., Richter, H., van der Oost, J., White, M.F., 2015. Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiol. Rev. 39, 428–441.
Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.
Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., Moineau, S., 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J. Bacteriol. 190, 1390–1400.
DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343.
Erdmann, S., Garrett, R.A., 2012. Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. Mol. Microbiol. 85, 1044–1056.
Fu, H., Yu, L., Lin, M., Wang, J., Xiu, Z., Yang, S.T., 2017. Metabolic engineering of *Clostridium tyrobutyricum* for enhanced butyric acid production from glucose and xylose. Metab. Eng. 40, 50–58.
Gangadwala, J., Mankar, S., Mahajani, S., Kienle, A., Stein, E., 2003. Esterification of acetic acid with butanol in the presence of ion-exchange resins as catalysts. Ind. Eng. Chem. Res. 42, 2146–2155.
Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–451.
Gomaa, A.A., Klumpe, H.E., Luo, M.L., Selle, K., Barrangou, R., Beisel, C.L., 2014. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. mBio 5, e00928–13.
Horvath, P., Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science 327, 167–170.
Horwitz, A.A., Walter, J.M., Schubert, M.G., Kung, S.H., Hawkins, K., Platt, D.M., Hernday, A.D., Mahatdejkul-Meadows, T., Szeto, W., Chandran, S.S., 2015. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. Cell Syst. 1, 88–96.
Huang, H., Chai, C., Li, N., Rowe, P., Minton, N.P., Yang, S., Jiang, W., Gu, Y., 2016. CRISPR/Cas9-based efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas-fermenting bacterium. ACS Synth. Biol. 5, 1355–1361.
Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, L.A., 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. 31, 233–239.
Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., Yang, S., 2015. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. Appl. Environ. Microbiol. 81, 2506–2514.
Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., Sun, B., Chen, B., Xu, X., Li, Y., 2017. CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. Nat. Commun. 8, 15179.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.
Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., Doudna, J., 2013. RNA-programmed genome editing in human cells. elife 2, e00471.
Knoshaug, E.P., Zhang, M., 2009. Butanol tolerance in a selection of microorganisms. Appl. Biochem. Biotechnol. 153, 13–20.
Lee, J., Jang, Y.-S., Han, M.-J., Kim, J.Y., Lee, S.Y., 2016a. Deciphering *Clostridium tyrobutyricum* metabolism based on the whole-genome sequence and proteome analyses. mBio 7, e00743–16.
Lee, S.H., Yun, E.J., Kim, J., Lee, S.J., Um, Y., Kim, K.H., 2016b. Biomass, strain engineering, and fermentation processes for butanol production by solventogenic clostridia. Appl. Microbiol. Biotechnol. 100, 8255–8271.
Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., Sheen, J., 2013. Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat. Biotechnol. 31, 688–691.
Li, M., Wang, R., Zhao, D., Xiang, H., 2014. Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. Nucleic Acids Res. 42, 2483–2492.
Li, Q., Chen, J., Minton, N.P., Zhang, Y., Wen, Z., Liu, J., Yang, H., Zeng, Z., Ren, X., Yang, J., 2016. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. Biotechnol. J. 11, 961–972.
Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., Chen, L., Liang, Y.X., She, Q., 2015. Harnessing Type I and Type III CRISPR-Cas systems for genome editing. Nucleic Acids Res. 44, e34.
Luo, M.L., Mullis, A.S., Leenay, R.T., Beisel, C.L., 2015. Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. Nucleic Acids Res. 43, 674–681.
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., 2011. Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. 9, 467–477.
Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J.M., Terns, R.M., Terns, M.P., White, M.F., Yakunin, A.F., Garrett, R.A., van der Oost, J., Backofen, R., Koonin, E.V., 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. 13, 722–736.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823–826.
Mojica, F., Diez-Villasenor, C., Garcia-Martinez, J., Almendros, C., 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155, 733–740.
Mount, D.W., 2007. Using the basic local alignment search tool (BLAST). CSH Protoc. pdb. top17.
Peabody, G.L., Kao, K.C., 2016. Recent progress in biobutanol tolerance in microbial systems with an emphasis on *Clostridium*. FEMS Microbiol. Lett. 363 (fnw017).
Plagens, A., Richter, H., Charpentier, E., Randau, L., 2015. DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. FEMS Microbiol. Rev. 39, 442–463.
Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K.A., Djordjevic, M., Wanner, B.L., Severinov, K., 2010. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. Mol. Microbiol. 77, 1367–1379.
Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., Wagner, R., 2010. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. Mol. Microbiol. 75, 1495–1512.
Pyne, M.E., Bruder, M.R., Moo-Young, M., Chung, D.A., Chou, C.P., 2016. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. Sci. Rep. 6, 25666.
Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183.
Ramio-Pujol, S., Ganigue, R., Baneras, L., Colprim, J., 2015. Incubation at 25 degrees C prevents acid crash and enhances alcohol production in *Clostridium carboxidivorans* P7. Bioresour. Technol. 192, 296–303.
Rath, D., Amlinger, L., Rath, A., Lundgren, M., 2015. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie 117, 119–128.
Sakuma, T., Nishikawa, A., Kume, S., Chayama, K., Yamamoto, T., 2014. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci. Rep. 4, 5400.
Shah, S.A., Erdmann, S., Mojica, F.J., Garrett, R.A., 2013. Protospacer recognition motifs: mixed identities and functional diversity. RNA Biol. 10, 891–899.
Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.L., 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31, 686–688.
Shariat, N., Timme, R.E., Pettengill, J.B., Barrangou, R., Dudley, E.G., 2015. Characterization and evolution of *Salmonella* CRISPR-Cas systems. Microbiology 161, 374–386.
Sorek, R., Lawrence, C.M., Wiedenheft, B., 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. Annu. Rev. Biochem. 82, 237–266.
Ungerer, J., Pakrasi, H.B., 2016. Cpf1 is a versatile tool for CRISPR genome editing across diverse species of Cyanobacteria. Sci. Rep. 6, 39681.
Van Der Oost, J., Westra, E.R., Jackson, R.N., Wiedenheft, B., 2014. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. 12, 479–492.
Vercoe, R.B., Chang, J.T., Dy, R.L., Taylor, C., Gristwood, T., Clulow, J.S., Richter, C., Przybilski, R., Pitman, A.R., Fineran, P.C., 2013. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. PLoS Genet. 9, e1003454.
Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013a. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153, 910–918.
Wang, S., Dong, S., Wang, P., Tao, Y., Wang, Y., 2017. Genome editing in *Clostridium saccharoperbutylacetonicum* N1-4 using CRISPR-Cas9 system. Appl. Environ. Microbiol. 83, e00233–17.
Wang, Y., Li, X., Milne, C.B., Janssen, H., Lin, W., Phan, G., Hu, H., Jin, Y.S., Price, N.D., Blaschek, H.P., 2013b. Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in *Clostridium beijerinckii*. Appl. Environ. Microbiol. 79, 5853–5863.
Wang, Y., Zhang, Z.T., Seo, S.O., Choi, K., Lu, T., Jin, Y.S., Blaschek, H.P., 2015. Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. J. Biotechnol. 200, 1–5.
Wang, Y., Zhang, Z.T., Seo, S.O., Lynn, P., Lu, T., Jin, Y.S., Blaschek, H.P., 2016. Bacterial genome editing with CRISPR-Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. ACS Synth. Biol. 5, 721–732.
Westbrook, A.W., Moo-Young, M., Chou, C.P., 2016. Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. Appl. Environ. Microbiol. 82, 4876–4895.
Westra, E.R., Buckling, A., Fineran, P.C., 2014. CRISPR-Cas systems: beyond adaptive immunity. Nat. Rev. Microbiol. 12, 317–326.
Westra, E.R., Dowling, A.J., Broniewski, J.M., van Houte, S., 2016. Evolution and ecology of CRISPR. Annu. Rev. Ecol. Evol. S 47, 307–331.
Williams, D.R., Young, D.I., Young, M., 1990. Conjugative plasmid transfer from *Escherichia coli* to *Clostridium acetobutylicum*. Microbiology 136, 819–826.
Xie, K., Minkenberg, B., Yang, Y., 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. USA 112, 3570–3575.
Xu, T., Li, Y., Shi, Z., Hemme, C.L., Li, Y., Zhu, Y., Van Nostrand, J.D., He, Z., Zhou, J., 2015. Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. Appl. Environ. Microbiol. 81, 4423–4431.
Yan, M.Y., Yan, H.Q., Ren, G.X., Zhao, J.P., Guo, X.P., Sun, Y.C., 2017. CRISPR-Cas12a–assisted recombineering in bacteria. Appl. Environ. Microbiol. AEM (00947-17).
Yosef, I., Goren, M.G., Qimron, U., 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. Nucleic Acids Res. 40, 5569–5576.
Yu, M., Du, Y., Jiang, W., Chang, W.L., Yang, S.T., Tang, I.C., 2012. Effects of different replicons in conjugative plasmids on transformation efficiency, plasmid stability, gene expression and n-butanol biosynthesis in *Clostridium tyrobutyricum*. Appl. Microbiol. Biotechnol. 93, 881–889.
Yu, M., Zhang, Y., Tang, I.C., Yang, S.T., 2011. Metabolic engineering of *Clostridium tyrobutyricum* for n-butanol production. Metab. Eng. 13, 373–382.
Zebec, Z., Manica, A., Zhang, J., White, M.F., Schleper, C., 2014. CRISPR-mediated targeted mRNA degradation in the archaeon *Sulfolobus solfataricus*. Nucleic Acids Res. 42, 5280–5288.
Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F., 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759–771.
Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S., 2017. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol. 35, 31–34.
Zhang, J., Taylor, S., Wang, Y., 2016a. Effects of end products on fermentation profiles in *Clostridium carboxidivorans* P7 for syngas fermentation. Bioresour. Technol. 218, 1055–1063.
Zhang, J., Wang, S., Wang, Y., 2016b. Biobutanol Production from Renewable Resources: Recent Advances 1 Elsevier Inc.
Zhang, Z.T., Taylor, S., Wang, Y., 2017. In situ esterification and extractive fermentation for butyl butyrate production with *Clostridium tyrobutyricum*. Biotechnol. Bioeng. 114, 1428–1437.