Metabolic Engineering Communications 10 (2020) e00116

Contents lists available at ScienceDirect

Metabolic Engineering Communications

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

Development of both type I–B and type II CRISPR/Cas genome editing systems in the cellulolytic bacterium *Clostridium thermocellum*

Julie E. Walker [a,b], Anthony A. Lanahan [b,c], Tianyong Zheng [b,c], Camilo Toruno [b,c], Lee R. Lynd [b,c], Jeffrey C. Cameron [a,b,d,e], Daniel G. Olson [b,c,*], Carrie A. Eckert [a,b,e,**]

[a] Renewable and Sustainable Energy Institute, University of Colorado, Boulder, CO, 80303, USA
[b] Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, 37830, USA
[c] Thayer School of Engineering, Dartmouth College, Hanover, NH, 03755, USA
[d] Department of Biochemistry, University of Colorado, Boulder, CO, 80303, USA
[e] National Renewable Energy Laboratory, Biosciences Center, Golden, USA

ARTICLE INFO

Keywords:
CRISPR
Type I–B
Cas9
Clostridium thermocellum
Thermophilic recombineering

ABSTRACT

The robust lignocellulose-solubilizing activity of *C. thermocellum* makes it a top candidate for consolidated bioprocessing for biofuel production. Genetic techniques for *C. thermocellum* have lagged behind model organisms thus limiting attempts to improve biofuel production. To improve our ability to engineer *C. thermocellum*, we characterized a native Type I–B and heterologous Type II Clustered Regularly-Interspaced Short Palindromic Repeat (CRISPR)/cas (CRISPR associated) systems. We repurposed the native Type I–B system for genome editing. We tested three thermophilic Cas9 variants (Type II) and found that GeoCas9, isolated from *Geobacillus stearothermophilus*, is active in *C. thermocellum*. We employed CRISPR-mediated homology directed repair to introduce a nonsense mutation into *pyrF*. For both editing systems, homologous recombination between the repair template and the genome appeared to be the limiting step. To overcome this limitation, we tested three novel thermophilic recombinases and demonstrated that *exo/beta* homologs, isolated from *Acidithiobacillus caldus*, are functional in *C. thermocellum*. For the Type I–B system an engineered strain, termed LL1586, yielded 40% genome editing efficiency at the *pyrF* locus and when recombineering machinery was expressed this increased to 71%. For the Type II GeoCas9 system, 12.5% genome editing efficiency was observed and when recombineering machinery was expressed, this increased to 94%. By combining the thermophilic CRISPR system (either Type I–B or Type II) with the recombinases, we developed a new tool that allows for efficient CRISPR editing. We are now poised to enable CRISPR technologies to better engineer *C. thermocellum* for both increased lignocellulose degradation and biofuel production.

1. Introduction

Members of the *Clostridium* genus have been studied for decades both as human pathogens and as industrial biocatalysts (Lynd et al., 2002; Jones et al., 2016; Xin et al., 2016; Kiu and Hall, 2018; Harris et al., 2002; Migriauli et al., 2018; Chung et al., 2014). Genetic techniques for *Clostridium* have lagged behind model organisms thus hindering advances in our understanding of their physiology and metabolism. *Clostridium thermocellum* is an obligate thermophilic and anaerobic gram positive bacterium that naturally ferments lignocellulose to ethanol and organic acids, and has recently been engineered to produce n-butanol (Tian et al., 2019; Olson et al., 2012; Lynd et al., 2005; Xu et al., 2016). The robust lignocellulose-solubilizing activity of *C. thermocellum* makes it a top candidate for consolidated bioprocessing for biofuel production (Fig. 1A). Established genetic techniques for *C. thermocellum* are laborious and introducing single nucleotide polymorphisms (SNPs) is difficult; thus efforts to increase our understanding of lignocellulose solubilization and biofuel production have been slow (Olson and Lynd,

Abbreviations: **CRISPR/Cas**, Clustered Regularly-Interspaced Short Palindromic Repeat/CRISPR associated; **sgRNA**, single guide RNA; **PAM**, protospacer adjacent motif; **HDR**, homology-directed repair; **HR**, homologous recombination; **Tm**, thiamphenicol; **5-FOA**, 5-fluoroorotic acid; **CFU**, colony forming unit; **Cas9n**, nickase Cas9; **RNP**, Cas9-sgRNA ribonucleoprotein.

* Corresponding author. Thayer School of Engineering, Dartmouth College, Hanover, NH, 03755, USA.
** Corresponding author. Renewable and Sustainable Energy Institute, University of Colorado, Boulder, CO, 80303, USA.
E-mail addresses: daniel.g.olson@dartmouth.edu (D.G. Olson), carrie.eckert@colorado.edu (C.A. Eckert).

https://doi.org/10.1016/j.mec.2019.e00116
Received 21 August 2019; Received in revised form 13 November 2019; Accepted 25 November 2019

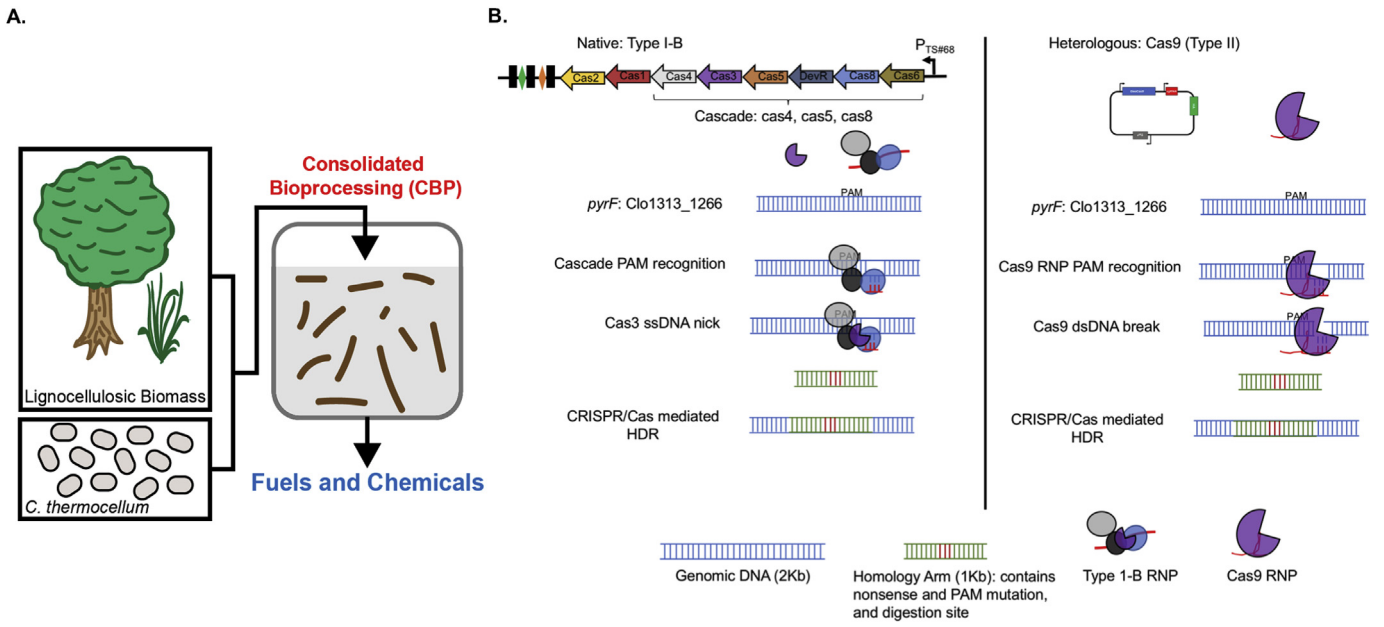


Fig. 1. CRISPR/Cas genome editing in *C. thermocellum*. A. *C. thermocellum* breaks down recalcitrant lignocellulosic biomass to produce biofuels. The robust lignocellulose activity of *C. thermocellum* make them a top candidate for consolidated bioprocessing. B. Schematic of the *C. thermocellum* native Type I–B operon (Clo1313_2969–2976) and heterologous Type II Cas9 CRISPR/Cas systems.

2012a; Mohr et al., 2013). Repurposing CRISPR (Clustered Regularly Interspaced Palindromic Repeats)/Cas (CRISPR associated) technology has revolutionized genome editing capabilities in a variety of biological systems (Gasiunas et al., 2012; Mali et al., 2013; Cong et al., 2013; Jinek et al., 2013; Hwang et al., 2013; Qi et al., 2013; Jiang et al., 2013) but has not been reported for *C. thermocellum.* To enable rapid genome engineering in *C. thermocellum* we sought to develop both a heterologous Type II and a native Type I–B CRISPR/Cas genome editing system for *C. thermocellum* (Fig. 1B).

There are six types of CRISPR systems (for a recent review see (Adli, 2018)). For the scope of our work, we will focus on outlining Type I and II CRISPR systems, each of which has a signature nuclease, Cas3 and Cas9, respectively. Both types encode *cas1* and *cas2* genes, which are involved in DNA acquisition (Nuñez et al., 2014). For Type I, the transcribed pre-crRNA is cleaved into repeat-protospacer-repeat units by Cas6 (Carte et al., 2008; Hochstrasser and Doudna, 2015). In Type II, the pre-crRNA is cleaved by RNase III and the mature crRNA is complete upon addition of the small trans-activating RNA (tracrRNA) (Deltcheva et al., 2011). Type I systems have several proteins required for the interference step while Type II systems only require the Cas9 nuclease for interference (Brouns et al., 2008; Xiao et al., 2018). In order to identify self vs. non-self during the interference stage, type I and II Cas protein(s) recognize a specific nucleotide sequence termed a Protospacer Adjacent Motif (PAM) (Mojica et al., 2009). In Type I, a group of protein(s) termed the cascade scan DNA for PAMs. Once a PAM is identified, the cascade proteins recruit the nuclease Cas3 to induce a ssDNA break (Sinkunas et al., 2011). For type II, the Cas9-sgRNA ribonucleoprotein (RNP) scans for the PAM site and induces a dsDNA break adjacent to the PAM (Brouns et al., 2008).

Repurposing CRISPR/Cas technology allows for facile and markerless strain construction. For CRISPR/Cas genome editing using homology directed repair, all that is required is an active single guide RNA (sgRNA), a Cas complex, and a repair template (also referred to as a homology arm). CRISPR/Cas-mediated homology directed repair enables the introduction of gene deletions, SNPs, and gene knock-ins, which were previously difficult to introduce, particularly in the case of SNPs. Due to the simplicity of the Type II (Cas9) system, namely the dual tracrRNA:crRNA was engineered to be designed as a single RNA chimera (Jinek et al., 2012) thus only requiring Cas9 nuclease for genome editing, it is the most commonly exploited system for genome engineering in model and non-model organisms. However, in many organisms, expression of a heterologous Cas9 is toxic and results in low transformation efficiency (Wang et al., 2013; Jacobs et al., 2014; Liu et al., 2017; Cho et al., 2017, 2018; Lee et al., 2016). In order to overcome Cas9 toxicity it is often necessary to control *cas9*/sgRNA expression through an inducible promoter. If inducible promoters have not been characterized then another highly successful method has been to repurpose a prokaryote's native CRISPR/Cas system for genome editing (Pyne et al., 2016; Cheng et al., 2017; Li et al., 2016a).

Many prokaryotes have poor frequencies of homologous recombination so expressing bacteriophage-derived λ-Red recombinases (Beta/Exo) or Rac prophage recombinases (RecE/RecT) has been shown to increase CRISPR mediated genome editing efficiencies ten-fold or more (Jiang et al., 2013; Dong et al., 2014; Penewit et al., 2018). Recombineering machinery includes an exonuclease (Exo or RecT) and a single-stranded DNA annealing protein (Beta or RecE) that facilitates homologous recombination between the repair template and the genome. Engineering highly efficient CRISPR/Cas9 genome editing systems often requires these recombinases.

Similar to other prokaryotes, Clostridial species often have low transformation efficiencies (10^5 CFU/μg DNA for *C. thermocellum* (Olson and Lynd, 2012a)) and/or low rates of homology directed repair so establishing CRISPR/Cas systems has proven to be challenging (see (McAllister and Sorg, 2019) for an excellent, recent review on CRISPR/Cas genome editing in *Clostridium*). Nonetheless, CRISPR/Cas9 genome editing systems have successfully been established in several mesophilic *Clostridium* species (Pyne et al., 2016; Wang et al., 2015, 2016, 2017, 2018; Li et al., 2016b; Ingle et al., 2019; McAllister et al., 2017; Nagaraju et al., 2016; Huang et al., 2016). Generally, for these CRISPR/Cas systems, implementing inducible *cas9*/sgRNA expression was critical to achieve high editing efficiencies (Wang et al., 2016, 2017, 2018; McAllister et al., 2017; Wasels et al., 2017). In some *Clostridium* species, expressing wild type Cas9 (in the presence of a functional sgRNA) was toxic so a Cas9 nickase (Cas9n) variant was utilized (Li et al., 2016b, 2019; Xu et al., 2015, 2017). In Cas9n systems, one of the nuclease domains (either the HNH or RuvC domain) is inactivated thus the RNP induces ssDNA breaks which are less lethal than dsDNA breaks. In these systems, inducing ssDNA breaks still increases the rate of homology directed repair at that genomic locus; however, the cell has time

to incorporate the repair template or repair the ssDNA break before cell death occurs.

Many non-model organisms do not have a suite of characterized, inducible promoters to control a toxic, heterologous Cas9. In these prokaryotes, repurposing the native CRISPR array has been successful for genome editing (Zhang et al., 2018; Maikova et al., 2019; Pyne et al., 2016). In both systems (Type I–B or II), low homologous recombination rates can often be a limiting step for genome editing and recombineering machinery can be expressed to overcome this limitation. In addition, a "two plasmid" genome editing approach has been very successful in several organisms (Xi et al., 2015, WO2014089290A1, 2014). In this approach, the repair template is introduced first to the cell. Then, the CRISPR/Cas system is introduced to the cell and acts as a counter-selection that cuts and effectively kills all cells with genomes that are not edited. There is increasing evidence that CRISPR/Cas9 mediated genome editing is more efficient when Cas9 acts as a counter-selection post homologous recombination (Jiang et al., 2013; Mougiakos et al., 2016).

Most Cas9 systems use the most commonly utilized Cas9 variant isolated from *Streptococcus pyogenes* (SpCas9); however, SpCas9 is not active above 42 °C and cannot be used for genome editing in obligate thermophiles like *C. thermocellum* (Harrington et al., 2017). Recently, three thermophilic Cas9 variants were characterized from *Geobacillus stearothermophilus*, *Acidothermus cellulolyticus*, and *Geobacillus thermodenitrificans* T12, referred to as GeoCas9, AceCas9, and ThermoCas9, respectively (Harrington et al., 2017; Tsui et al., 2017; Mougiakos et al., 2017), that could potentially be utilized for Cas9 genome editing in obligate thermophiles like *C. thermocellum*.

For the native Type I–B system, previous work identified five CRISPR/Cas operons in *Clostridium thermocellum*; however, only CRISPR locus 5 has a fully intact Type I–B CRISPR/Cas operon (Pyne et al., 2016; Brown et al., 2014; Richter et al., 2012). Further *in silico* analysis of the Type I–B operon identified spacer-protospacer matches and putative PAM sequences (Pyne et al., 2016). In addition, *in vitro* work was done wherein the Type I–B Cascade complex from *C. thermocellum* was reconstituted and further characterized (Richter et al., 2017). Here we report the characterization of both a native Type I–B and a Type II (GeoCas9) CRISPR/Cas genome editing system in the thermophilic, industrially relevant bacterium *Clostridium thermocellum*. In addition, we present newly-developed thermophilic recombineering machinery that will be useful in other obligate thermophiles.

2. Materials and methods

2.1. Plasmids, strains and transformations

Table 1 describes the plasmids and strains used in this study. Plasmids were constructed using standard molecular biology techniques and Gibson cloning. Strains were constructed following the *C. thermocellum* strain construction protocol previously described (Olson and Lynd, 2012a), or following the CRISPR/Cas editing technology described below. All transformations were carried out using the protocol previously described (Olson and Lynd, 2012a). Briefly, overnight transformation recoveries were employed at 49–51 °C and subsequent growth was employed at 50–55 °C.

2.2. Growth and selection conditions

Generally, for growth and selection conditions we used CTFUD rich media and used concentrations for selections as previously described (Olson and Lynd, 2012a). Briefly, for experiments where *pyrF* was inactivated, 40 μg/mL of uracil was added to the medium to supplement for potential uracil auxotrophy from *pyrF* inactivation. For plasmid selection, with the exception of the second plasmid utilized in the two-step system, we used a selectable marker (gene encoding for CAT) that confers resistance to thiamphenicol. A concentration of 10 μg/mL of thiamphenicol was used for selection. When utilizing the 5-FOA counter-selectable marker the medium was supplemented with 250 μg/mL 5-FOA. For the two step/two plasmid system (described in 2.11) the second plasmid expressed a selectable marker that confers resistance to neomycin. For the Type I–B and Type II system various concentrations of neomycin had to be tested to find a concentration that was viable for cell growth but did not have background. The optimal concentration of Neomycin was 150 μg/mL for the Type I–B system and 110 μg/mL for the Type II system (data not shown). Unless otherwise noted, the concentrations described are what was used in the CTFUD rich medium, termed selective medium throughout this section.

2.3. PAM depletion assay

The CRISPR expression cassette plasmid pTY11B was amplified with ultramer pairs using Q5 polymerase. The forward ultramer in each pair contained the degenerate PAM library and spacer sequences being tested and the reverse ultramer contained the spacer expression cassette. The gel purified PCR product was Gibson assembled with the gblock XD824 and transformed into *E. coli* (NEB T7 express). An aliquot of the transformation mix was plated to determine library complexity and the remainder grown in 200 ml of LB/chloramphenicol and the plasmid DNA isolated for transformation of *C. thermocellum*. Three *E. coli* libraries were generated [pDGO180 (31 nt spacer), pDGO182 (36 nt spacer), pDGO183 (no spacer)] and used to transform *C. thermocellum* (LL1299 and LL1586) via electroporation with standard conditions. Following transformation cells recovered overnight at 50 °C and an aliquot plated to determine the number of transformants and the remainder used to inoculate a 50 ml culture of CTFUD + thiamphenicol (6 μg/ml) and grown at 55 °C for 24 h. Three independent transformations were done for each plasmid library and strain of *C. thermocellum*. Plasmid DNA was isolated and this PAM depleted library DNA used for PCR with Illumina barcode primers to amplify the PAM/spacer region. The PCR product was column purified, concentrated (Zymo Research, DNA Clean and Concentrator) and used for Illumina sequencing followed by statistical analysis.

2.4. Illumina sequencing and analysis

Genome resequencing was performed as previously described (Zhou et al., 2015). Briefly, genomic DNA was submitted to the Joint Genome Institute (JGI) for sequencing with an Illumina MiSeq instrument. Unamplified libraries were generated using a modified version of Illumina's standard protocol. 100 ng of DNA was sheared to 500 bp using a focused ultrasonicator (Covaris). The sheared DNA fragments were size selected using SPRI beads (Beckman Coulter). The selected fragments were then end repaired, A-tailed, and ligated to Illumina compatible adapters (IDT, Inc) using KAPA- Illumina library creation kit (KAPA biosystems). Libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then multiplexed into pools for sequencing. The pools were loaded and sequenced on the Illumina MiSeq sequencing platform utilizing a MiSeq Reagent Kit v2 (300 cycle) following a 2 × 150 indexed run recipe. Paired-end reads were generated, with an average read length of 150 bp and paired distance of 500 bp. Raw data were analyzed using CLC Genomics Workbench, version 8.5 (Qiagen, USA). Reads were mapped to the reference genome (NC_017992). Mapping was improved by two rounds of local realignment. The CLC Probabilistic Variant Detection algorithm was used to determine small mutations (single and multiple nucleotide polymorphisms, short insertions and short deletions). Variants occurring in less than 90% of the reads and variants that were identical to those of the wild type strain (i.e., due to errors in the reference sequence) were filtered out. The fraction of the reads containing the mutation is shown in Additional file 3. To determine larger mutations, the CLC InDel and Structural Variant algorithm was run. This tool analyzes unaligned ends of reads and annotates regions where a structural variation may have

Table 1
Description of strains and plasmids.

Strain	Organism	Description	Accession Number	Reference or source
DSM1313	*Clostridium thermocellum*	wild type	CP002416	DSMZ
LL1299	*Clostridium thermocellum*	DSM1313 ΔhptΔ0478		
LL1584	*Clostridium thermocellum*	LL1299 Peno-Cas operon	SRP164871	This study
LL1585	*Clostridium thermocellum*	LL1299 P#0815-Cas operon	SRP164872	This study
LL1586	*Clostridium thermocellum*	LL1299 PTsac#0068-Cas operon	SRP164875	This study
LL1587	*Clostridium thermocellum*	LL1299 PTsac#2130-Cas operon	SRP164873	This study
LL1588	*Clostridium thermocellum*	LL1299 PTsac#0530-Cas operon	SRP164874	This study
61	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas5 mutation	SRP164861	This study
62	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas3 mutation	SRP164864	This study
63	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas8 mutation	SRP164863	This study
67	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas5 mutation	SRP164866	This study
70	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas5 mutation	SRP164865	This study
71	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas8 mutation	SRP164867	This study
72	*Clostridium thermocellum*	LL1586 one-step escape mutant, cas8 mutation	SRP164868	This study
Plasmids	**Description**	**crRNA/spacer sequence**		**Reference or Source**
pJEW54	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$non-target_sgRNA	TTCCCTGGTACCTAGGAACCCG		This study
pJEW55	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$sgRNA#1	gggcatatttgttctggtcaag		This study
pJEW56	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$sgRNA#2	tcgtttcttttttccgtctgcaa		This study
pJEW57	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$sgRNA#3	CGGGTTGACTGTCAGGGCATCC		This study
pJEW63	$P_{Tsac0068}$ThermoCas9; $P_{Clo1313_2638}$non-target_sgRNA	TGTCATAGCGCTAGATCCGG		This study
pJEW64	$P_{Tsac0068}$ThermoCas9; $P_{Clo1313_2638}$sgRNA#1	gcatatttgttctggtcaag		This study
pJEW70	$P_{Tsac0068}$AceCas9; $P_{Clo1313_2638}$non-target_sgRNA	CTTCCGCTGAGACTCCCCTTACAC		This study
pJEW71	$P_{Tsac0068}$AceCas9; $P_{Clo1313_2638}$sgRNA#1	ggattagaccctaaaattgaatat		This study
pJEW72	$P_{Tsac0068}$AceCas9; $P_{Clo1313\_2638}$ sgRNA#2	caagatacgggttgactgtcaggg		This study
pJEW68	$P_{Tsac0068}$GeoCas9; $P_{Clo1313\_2638}$ non-target_sgRNA_HA	TTCCCTGGTACCTAGGAACCCG		This study
pJEW69	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$_sgRNA#1_HA	gggcatatttgttctggtcaag		This study
pJEW84	$P_{Tsac0068}$GeoCas9n; $P_{Clo1313\_2638}$ non-target_sgRNA_HA	TTCCCTGGTACCTAGGAACCCG		This study
pJEW85	$P_{Tsac0068}$GeoCas9n; $P_{Clo1313_2638}$sgRNA#1_HA	gggcatatttgttctggtcaag		This study
pJEW112	HA			This study
pJEW106	$P_{Tsac530}$A.caldusBeta_Exo; HA			This study
pJEW107	$P_{Tsac530}$C.stercorariumRecT_RecE; HA			This study
pJEW108	$P_{Tsac530}$GeobacillusRecT_RecE; HA			This study
pJEW136	$P_{Tsac530}$A.caldusBeta_Exo; HA_500bp			This study
pJEW137	$P_{Tsac530}$A.caldusBeta_Exo; HA_50bp			This study
pJEW117	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$non-target_sgRNA	TTCCCTGGTACCTAGGAACCCG		This study
pJEW111	$P_{Tsac0068}$GeoCas9; $P_{Clo1313_2638}$sgRNA#1	gggcatatttgttctggtcaag		This study
pTY11B	$P_{Clo1313\_1194}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY21B	$P_{Clo1313\_1194}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY32B	$P_{Clo1313\_1194}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY62B	$P_{Clo1313\_1194}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY11C	$P_{Clo1313\_2638}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY21C	$P_{Clo1313\_2638}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY32C	$P_{Clo1313\_2638}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pTY62C	$P_{Clo1313\_2638}$	ATAATGACATTTATGGTACTGTTGTGGTAATAGACGA		This study
pDGO185N	pyrF targeting, deletion	CAAGTTTCATAAAACACCCTCATGCCTTCAAGGCCGT		This study
pDGO186N	pyrF targeting, deletion	TGAGATTGTTGCGGAGTATGTTGAATCATGGGGTGAA		This study
pDGO187N	pyrF target, deletion, no sgRNA control	none		This study
pDGO186N–S1_nheI	pyrF target, stop codon insertion, target sgRNA	TGAGATTGTTGCGGAGTATGTTGAATCATGGGGTGAA		This study
pDGO186N–S1_CS	pyrF non-target control, stop codon insertion, non-target control sgRNA	CTTGAAGGCATGAGGGTGTTTTATGAAACTTGCAAAT		This study
pDGO186NX-S1_nheI	pyrF target, stop codon insertion	none		This study
pDGO186NXR-S1_nheI	pyrF target, stop codon insertion, recombinase	none		This study
pDGO186N-CS3neo	target sgRNA	TGAGATTGTTGCGGAGTATGTTGAATCATGGGGTGAA		This study
pDGO186N-ContS-neo	control sgRNA	CTTGAAGGCATGAGGGTGTTTTATGAAACTTGCAAAT		This study

occurred, which are called breakpoints. Since the read length averaged 150 bp and the minimum mapping fraction was 0.5, a breakpoint can have up to 75 bp of sequence data. The resulting breakpoints were filtered to eliminate those with fewer than ten reads or less than 20% "not perfectly matched." The breakpoint sequence was searched with the Basic Local Alignment Search Tool (BLAST) algorithm (Altschul et al., 1990) for similarity to known sequences. Pairs of matching left and right breakpoints were considered evidence for structural variations such as transposon insertions and gene deletions. Raw data is available from the JGI Sequence Read Archive.

2.5. Cas promoter characterization

LL1004, LL1299, LL1584, LL1585, LL1586, LL1587, LL1588 were grown in CTFUD at 55 °C and cells were collected at different stages of logarithmic growth and RNA prepared. 1 ml of bacterial culture was pelleted and lysed by digestion with lysozyme (15 mg/ml) and proteinase K (20 mg/ml). RNA was isolated with an RNAeasy minikit (Qiagen#74104) and digested with TURBO DNase (Applied Biosystems) to remove contaminating DNA. cDNA was synthesized from 500 ng of RNA using the iScript cDNA synthesis kit (BioRad). qPCR reactions for each sample were performed in triplicate using cDNA with SsoFast EvaGreen

Supermix (BioRad) at an annealing temperature of 55 °C to determine Cas6 (Clo1313_2976) and recA (Clo1313_1163) RNA levels. A gblock (IDT) containing the *cas6* and *recA* amplicons was used to generate a standard curve for each amplicon and *cas6* RNA levels were normalized to *recA* RNA levels. RNA levels for *cas6* and *recA* at different growth stages for each strain were combined and averaged to determine *cas6* RNA levels and relative promoter strength.

2.6. Single guide RNA design

The terminology single guide RNA (sgRNA) is used throughout the text to describe the gRNA for both the Type I–B and Type II system. For the Type I–B system this refers to the crRNA that is processed to target *pyrF*. For the Type II system this refers to the single RNA chimera that is composed of the crRNA and the tracrRNA.

Type I–B. Using the plasmid toxicity assay and PAM depletion library results, we designed guide RNAs by identifying a strong PAM site on the sense strand (PAM sequence: TTA) in the *pyrF* locus (Clo1313_1266) and then used the 37 base pair sequence immediately downstream as a target spacer (termed target sgRNA). As a control we identified a 37 base pair sequence on the sense strand in the *pyrF* locus without a predicted PAM site (PAM sequence: GGC) and used this for the non-target spacer (termed non-target sgRNA). In general, 37bp spacers that were immediately downstream of strong predicted PAM sites were used for the target sgRNA; non-target sgRNA were the same length but lacked predicted PAM sites.

Type II. Single guide RNAs were designed using Benchling software. Briefly, using the CRISPR tool in Benchling we assigned the guide type to single guide and the reference genome to ASM876v1 (*Clostridium acetobutylicum* ATCC 824). The guide length was set to 22, 24, or 20 and the PAM site was set to 5′-NNNNCRAA-3′, 5′-NNNCC-3′, or 5′-NNNNCNAA-3′ for GeoCas9, AceCas9, and ThermoCas9, respectively. Only sgRNAs with 40–60% GC content were selected for further testing.

2.7. CRISPR/cas toxicity assay

A toxicity assay was used to demonstrate functional DNA cutting for both the Type I–B and Type II systems. Differences in the systems required slightly different assay design (described below).

Type I–B. Two versions of the toxicity assay were used. The first characterized plasmid transformation efficiency following transformation of strains LL1299 or LL1586 (strain LL1299 with a strong promoter driving the Type I–B *cas* operon) with a plasmid containing a spacer expression cassette targeting various PAM/spacer combinations on the same plasmid. For this, the previously characterized autonomously replicating plasmid pMU102 was modified with a spacer expression cassette that used the Clo1313_1194 promoter to express the 37 bp spacer, 30 bp repeats, and terminator from CRISPR locus 4 adjacent to Clo1313_1653. The plasmid also contained the same spacer with various adjacent 5′ PAM sequences. Transformations were employed as previously described and plated on rich medium with thiamphenicol. All transformations were done in biological duplicates and toxicity or percent of cells killed was calculated using the equation: (CFUs no PAM/spacer sgRNA + CFUs PAM/spacer sgRNA)/(CFUs no PAM/spacer sgRNA) * 100.

The second assay characterized bacterial survival following transformation of strains LL1299 or LL1586 by a plasmid with or without (i.e., no sgRNA) a sgRNA targeting the Clo1313 *pyrF* locus. The plasmid used above was modified to include a ~1 kb pyrF repair template (repair template) and express a sgRNA targeting the *pyrF* locus (the control omitted the sgRNA from the CRISPR expression cassette). All transformations were done in biological duplicates and toxicity results or percent cells killed was calculated using the equation: (CFUs no sgRNA + CFUs pyrF sgRNA)/(CFUs no sgRNA) * 100.

Type II. The *Thermoanaerobacterium saccharolyticum* promoter #68 (ABC transporter) was used for expression of Cas9 and the previously characterized *C. thermocellum* promoter #2638 with a disrupted ribosome binding site was used to transcribe the sgRNA. The Cas9 variant and sgRNA were placed on the previously characterized autonomously replicating vector (pNW33N) to transform into *C. thermocellum* strain DSM1313 or LL1299.

For the toxicity assay, a non-target sgRNA and one to three target sgRNAs (termed target sgRNA) for each Cas9 variant were transformed to DSM1313. All target sgRNAs targeted the *pyrF* chromosomal coding region (Clo1313_ 1266). Transformations were employed as previously described and plated on CTFUD rich medium with 10 μg/mL thiamphenicol. All transformations were done in biological duplicates or triplicates and when more than one sgRNAs was tested then the combined average of CFUs for all sgRNAs was calculated. Toxicity results or percent cells killed was calculated using the equation: (CFUs non-target sgRNA - CFUs target sgRNA)/(CFUs non-target sgRNA) * 100.

2.8. One step CRISPR/Cas genome editing in C. thermocellum

For one-step genome editing, the plasmid for each system was transformed to strain LL1299 and recovered in medium overnight at 51 °C. The recovered transformation was plated on medium with thiamphenicol and uracil and incubated for 3–5 days at 53 °C. If transformants were observed, ten colonies were pooled and grown up overnight in selective medium. To increase homologous recombination between the repair template and the genome, the cultures were passaged at a 1:20 dilution in selective medium. Passaging was repeated a total of five times. On the fifth passage, cultures were grown to mid-exponential, serially diluted and plated on selective medium ± 5-FOA. This was done for two or more biological replicates.

The effect of CRISPR (CRISPR fold change) for the one step method was determined using these equations: X (increase in 5-FOA^R phenotype) = (CFUs on Tm + 5-FOA for target sgRNA or non-target sgRNA)/(CFUs on Tm target sgRNA or non-target sgRNA); CRISPR fold-change for one step: (X target sgRNA)/(X non-target sgRNA).

Type I–B. The plasmid used in the toxicity assay was modified to include a ~1 kb repair template targeting the *pyrF* locus and introducing a stop codon at amino acid position #196 in *pyrF* and NheI digestion site in the region immediately 3′ of a predicted PAM sequence. One version of the plasmid expressed a sgRNA targeting the *pyrF* locus at the site with the predicted PAM sequence and a second version expressed a control sgRNA that targeted the *pyrF* locus at a different site without a predicted PAM sequence. The one step plasmid with a target or non-target sgRNA was transformed into strain LL1586 (LL1299 with promoter Tsac_0068 driving *cas* operon expression) and recovered in CTFUD overnight at 50 °C.

Type II. The plasmid used in the toxicity assay (Cas9 and non-target sgRNA/sgRNA#1) with the addition of a *pyrF* repair template introducing a stop codon at amino acid position 173 in *pyrF*, PAM mutation, and an EcoRI digest site was used for a one step genome editing approach (pJEW68 and pJEW69). In addition, a nickase GeoCas9 variant was constructed via Q5 site-directed mutagenesis (New England Biolabs'; catalog #: E0554S) wherein the catalytic domain was mutated at position H582A. The nickase Cas9 variant was subcloned into the vector containing the *pyrF* repair template and either a non-target or target sgRNA (pJEW84 and pJEW85, respectively) and tested using the one step genome editing approach.

2.9. Identification of thermophilic recombineering machinery

To identify thermophilic recombineering machinery, we searched all of the publicly available microbial genomes in the JGI IMG database using the "Genome Search by Metadata Category tool." We selected all genomes annotated as "thermophile" and searched for the presence of putative phage-type endonuclease (TIGRfam: TIGR03033; 27 hits). *recT* hits (21) and *bet* hits (1) were filtered for. Only genomes in which the *recT* or *beta* was directly adjacent to the *recE* or *exo*, respectively, were selected for further analysis. For *recT/recE* we filtered for bacteria that are gram positive, have an optimal temperature of at least 60 °C (optimal growth temperature of *C. thermocellum*) and are facultative anaerobes or anaerobes. The only *beta/exo* hit from *Acidithiobacillus caldus* (gene numbers: Atc1291 and Atc1292) and two *recT/recE* hits from *Clostridium stercorarium* (gene numbers: Cst0375 and Cst0374) and *Geobacillus* sp. (gene numbers: Geo2951 and Geo2953) were selected for testing.

2.10. Characterization of thermophilic recombinases in C. thermocellum

A codon optimized Gblock (Eurofins) was synthesized for each recombineering gene (see gene numbers above) and assembled via Gibson assembly on an autonomously replicating plasmid (pNW33N) containing a 1 kb *pyrF* repair template and thiamphenicol resistance. The *Thermoanaerobacterium saccharolyticum* promoter #530 (30S ribosomal protein S6) was used to drive expression of the recombineering machinery. Plasmids with and without the recombineering machinery were transformed to LL1299 and plated on medium with 6 μg/mL thiamphenicol and uracil. Transformants were observed for all plasmids (pJEW112, pJEW106, pJEW108) with the exception of the plasmid expressing *recE/T* from *Clostridium stercorarium* (pJEW107). Approximately 30 colonies were pooled from the transformation, grown up to mid-exponential phase, and serial dilutions for each culture were plated on selective medium ± 5-FOA. The percentage of colonies exhibition homology-directed repair was determined by the number of colonies in the presence or absence of 5-FOA.

2.11. A two plasmid CRISPR/Cas genome editing system in C. thermocellum

For both systems (Type I–B and Type II), the repair template (repair template) +/− *Atc* recombineering machinery plasmids were transformed into strain LL1299. 10–30% of the recovered transformation was plated on medium with thiamphenicol and uracil and incubated for 3–5 days at 53 °C. ~30 colonies were pooled and grown overnight in selective medium. To increase homologous recombination between the repair template and the genome, the cultures were passaged up to three times at a 1:20 dilution in selective medium. Genome editing was verified by restriction digest of a PCR amplicon covering the edit location. After two to three passages the culture was grown in 50 mL selective medium. The "killing vector" was transformed into a strain of LL1299 that already contained the repair template plasmid. 50–100% of the recovered transformation was plated on medium with Neomycin, uracil, +/− 5-FOA. A subset of colonies were picked into selective medium (Type I–B) or medium supplemented with uracil (Type II). No neomycin selection was used when growing colonies for the Type II system due to technical difficulties with the Neomycin selection that are further described in the discussion.

The effect of CRISPR (CRISPR fold change) for the two-plasmid method was determined using the equation below. Briefly, primers were designed wherein one primer annealed to the *pyrF* locus outside of the repair template and the second annealed to the repair template. Restriction enzyme, NheI or EcoRI for Type I–B and GeoCas9, respectively, was added to the PCR reactions and products were resolved on a 1% agarose gel. Note, the one-step and two-step protocol experimental differences required us to calculate CRISPR fold-change in slightly different ways.

CRISPR fold-change for two step system was calculated by: (% of checked edited colonies for the target sgRNA verified by restriction digest)/(% of checked edited colonies for non-target gRNA verified by restriction digest).

Type I–B. The plasmid pDGO186N–S1_nheI used for the one step protocol was modified and used for step-one of the two step protocol. The sgRNA expression cassette was removed leaving just the *pyrF* repair template in pDGO186NX-S1_nheI and in the case of pDGO186NXR-S1_nheI; the *A. caldus* exo/*beta* recombineering machinery was inserted and expressed in addition to the *pyrF* repair template. For step two of the protocol, pDGO186N–S1_nheI was modified by removing the *pyrF* repair template to generate sgRNA expression plasmids ('killing' and control). The thiamphenicol resistance gene was replaced with the neomycin resistance gene. pDGO186N-CS3neo expressed the same target sgRNA as used in the one step protocol and pDGO186N-ContSneo the same control (non-target sgRNA) as in the one step protocol.

Type II. The same plasmids described in the "characterization of thermophilic recombineering machinery" section were used during step-one of the two step protocol. Briefly, one plasmid contained the 1 kb *pyrF* repair template (pJEW112) and another plasmid contained the 1 kb *pyrF* repair template and expressed *A. caldus* exo/beta recombineering machinery (pJEW106). The "killing vector" was similar to the plasmid previously described in the "CRISPR/Cas toxicity assay" section with the exception that the vector used for two step editing conferred neomycin resistance rather than thiamphenicol resistance. The "killing vector" conferred neomycin resistance and expressed GeoCas9 and a non-target or target sgRNA (pJEW117 and pJEW111 respectively).

3. Results & discussion

3.1. Type I–B Cas operon and PAM identification

Of the five CRISPR/Cas operons in *Clostridium thermocellum* only CRISPR locus 5 has a fully intact Type I–B CRISPR/Cas operon (Supplemental Table 1) (Brown et al., 2014; Richter et al., 2012). We began repurposing the Type I–B system at CRISPR locus 5 by first identifying the PAM sequence. Our preliminary approach to PAM identification was to try to identify the original sequence of the invading elements that ended up as CRISPR spacers in *C. thermocellum*, by BLAST search, as described by Pyne (Pyne et al., 2016). From this analysis, we found the following PAM sequences (5' - 3'): CAGTTA, AATCCA, TTGTTA, AGGTTA, AGTTA, GATCA, and TAGTT. We tested these spacers using plasmid pTY#B/C (see Table 1), where the synthetic CRISPR array was driven by either the Clo1313_1194 promoter or the Clo1313_2638 promoter (Supplemental Fig. 1) (Olson et al., 2015). The synthetic CRISPR array on the plasmid is designed to target a second copy of the spacer elsewhere on the plasmid, such that a plasmid with a functional PAM sequence will cause itself to be cut by Cas3 and will be unable to provide antibiotic resistance to the cell. Thus, we hypothesized that functional PAM sequences would lead to reduced transformation efficiency. In the presence of a strong promoter (Clo1313_1194 or 2638 [69]), and with PAM sequences 5′-CAGTTA-3′ and 5′-GATCA-3′ we observed a 10-fold decrease in transformation efficiency (from 5×10^4 to $<5 \times 10^3$ CFU/ug DNA), indicating that CRISPR-mediated interference was happening.

To further characterize the specific PAM sequence, we performed a PAM depletion assay, where we built a library of all 4096 possible 6 bp PAM sequences, transformed them into *C. thermocellum*, and identified which sequences were subsequently depleted from the library. We found that the base pairs at the −6, −5 and −4 positions were unimportant. Of the remaining 81, 3 bp PAM sequences, 10 showed strong depletion with both spacers (Supplemental Table 2) and can be summarized by the degenerate sequences 5′-TTN-3′ and 5′-TNA-3'.

3.2. Toxicity assay to determine CRISPR/Cas activity

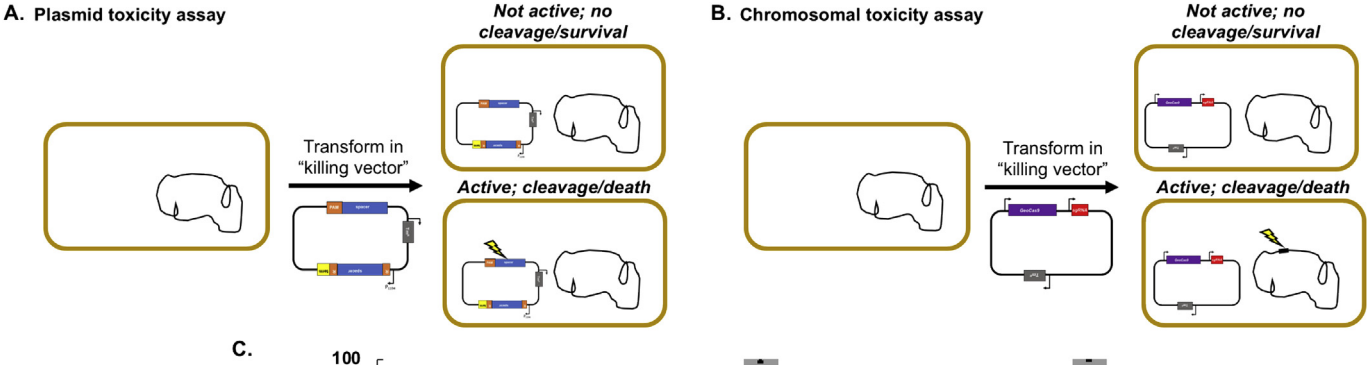
Like most bacteria, *C. thermocellum* does not encode for proteins responsible for non-homologous end joining; thus a single or double-strand DNA break caused by an active Cas:sgRNA RNP complex cannot be repaired under normal circumstances (Pitcher et al., 2007). We used a toxicity assay wherein the Cas:sgRNA RNP complex targets either a plasmid or the chromosome for cleavage (Fig. 2). If the RNP is active then cleavage will occur destroying a plasmid that confers resistance to an antibiotic (Fig. 2A), or the chromosome (Fig. 2B) will be targeted for a ss/dsDNA break causing cell death.

We first tested the activity of the Type I–B system by designing a plasmid that expresses the sgRNA (repeat-spacer-repeat) such that the RNP targets the same plasmid. For all cases, we used the PAM sequence 'TTA' (full PAM sequence: 5′-CAGTTA-3′) and tested several spacers targeting the plasmid. Constructs without a spacer were used as a negative control. When the endogenous *cas* operon promoter, and spacer array promoter were used we observed no cell killing. We hypothesized that expression of the sgRNA was too low for killing so we used a strong constitutive promoter, Clo1313_1194 (Olson et al., 2015), to drive expression of the sgRNA. Swapping out the native spacer array promoter for the stronger promoter Clo1313_1194 resulted in 88% cell killing (Fig. 2C).

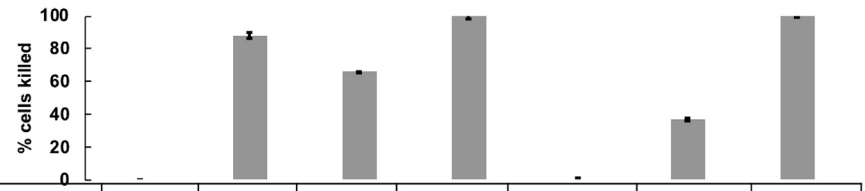
Once we confirmed the Type I–B system was active we next determined the activity of the RNP when targeting the chromosome rather than a plasmid. A spacer targeting the chromosomal *pyrF* locus (Clo1313_1266) was tested using the same system as described above (i.e., native promoter to drive *cas* expression and Clo1313_1194 promoter to drive sgRNA expression), and 65.8% cell killing was observed (Fig. 2C). Ideally, 100% cell killing is desired so there is minimal to no background when employing genome editing as Cas cleavage should act as a counter-selection. To further improve the efficiency of CRISPR-mediated DNA cleavage, we tested five promoters for upregulation of the *cas* operon (Clo1313_2969–2976). These promoters included two from *C. thermocellum* (enolase and 0815 [69]), and three from *Thermoanaerobacterium saccharolyticum* (Tsac_0068, 0530, 2130). Of these, the Clo1313_0815 and Tsac_0068 promoters gave the most reliable results (Supplemental Fig. 2). We chose the Tsac_0068 promoter because it is not present in the *C. thermocellum* genome (and thus is not a target for unwanted homologous recombination). The resulting strain was named LL1586 and was able to achieve 99.8% cell killing in our toxicity assay.

Three thermophilic Cas9 variants were recently characterized termed ThermoCas9, AceCas9, and GeoCas9 (Harrington et al., 2017; Tsui et al., 2017; Mougiakos et al., 2017). We employed a toxicity assay to test the activity of all three thermophilic Cas9 variants in *C. thermocellum*. Similar to the design of the Type I–B system, we used the Tsac_0068 promoter to drive expression of Cas9 and the strong *C. thermocellum* promoter Clo1313_2638 to drive expression of the sgRNA. For this assay, we targeted the *pyrF* locus. Two to three sgRNAs were designed for each thermophilic Cas9 variant and one non-target sgRNA was designed as a negative control. For ThermoCas9 no cell killing was observed, for AceCas9 37% cell killing was observed, and for GeoCas9 100% cell killing was observed (Fig. 2C). We are not sure why ThermoCas9 or AceCas9 did not work at all or as effectively as GeoCas9. We speculate that the Cas9 variant could be misfolding in *C. thermocellum* or the sgRNAs tested may not be effective sgRNAs. Nonetheless, for GeoCas9 all three sgRNAs tested resulted in 100% cell death whereas ~150 CFUs were obtained for the non-target sgRNA control. Based on these results, the GeoCas9 variant was selected for evaluating CRISPR/Cas9-mediated homology-directed repair (HDR).

In summary, the toxicity assay demonstrated a highly active engineered Type I–B CRISPR/Cas system and a highly active Type II GeoCas9 system (Table 2). Making the Type I–B system functional required identifying the correct PAM sequence and placing strong constitutive



C.



System	Type I-B	Type I-B	Type I-B	Type I-B	ThermoCas9	AceCas9	GeoCas9
Plasmid Target	+	+	-	-	-	-	-
Chromosome Target	-	-	+	+	+	+	+
cas promoter	Native	Native	Native	Tsac_0068	Tsac_0068	Tsac_0068	Tsac_0068
gRNA promoter	Native	Clo_1194	Clo_1194	Clo_1194	Clo_2638	Clo_2638	Clo_2638

*Cells killed = (CFUs NT–CFUs target)/(CFUs NT). Negative values were assigned as 0.

Fig. 2. Toxicity Assay. A. A toxicity assay wherein the RNP targets the transformed plasmid. If the RNP is active, then cleavage of the plasmid will occur, destroying the plasmid conferring resistance to the selection. B. A toxicity assay wherein the RNP targets the chromosome. If the RNP is active, then cleavage of the chromosome will occur, resulting in cell death. C. Table describing the results of the toxicity assay for various systems tested.

Table 2
Key differences between the native Type I–B and the Type II CRISPR systems.

Feature	Native type I–B	Type II
Organism	*Clostridium thermocellum*	*Geobacillus stearothermophilus*
Nuclease	Cas3	Cas9
PAM	5′-TTN-3′, 5′-TCD-3′	5′-NNNNCRAA-3′
guide RNA	97 nt, spacer (ranges from 31 to 42 nt), with two flanking repeats (30 nt each)	140 nt, crispr/spacer RNA sequence (22 nt) and tracrRNA (118 nt) fused together to form a sgRNA
Holoenzyme	4 proteins (cas4, cas5, cas8 and cas3) in complex with a sgRNA	One protein (Cas9) in complex with a sgRNA

promoters upstream of both the chromosomal *cas* operon and the plasmid-encoded sgRNA. Making the Type II system functional required identification of a thermostable Cas9 protein and putting strong promoters in front of both the *cas9* gene and sgRNA. We are now poised to introduce a repair template for CRISPR/Cas mediated homology directed repair.

3.3. A one-step CRISPR/Cas genome editing system is effective but not efficient

We targeted the *pyrF* gene (Clo1313_1266) for inactivation because the resulting mutants can easily be identified by their resistance to the toxic uracil analog 5-fluoroorotic acid (5-FOA). Furthermore, deletion of *pyrF* has no fitness defect for *C. thermocellum* when the media is supplemented with uracil (Tripathi et al., 2010). Based on previous reports, we expect some background level of 5-FOA resistance due to spontaneous *pyrF* point mutations; however, this should be the same for the non-target and target sgRNAs, and thus not interfere with measurements of CRISPR effectiveness/fold-change. For both systems, a *pyrF* repair template was designed containing a nonsense mutation to inactivate *pyrF* function, a unique restriction site to facilitate verification of the modification, and a mutation in either the PAM or spacer region to prevent subsequent editing by the RNP complex. The repair template was included on the same plasmid as the sgRNA.

For the Type II GeoCas9 system, ~50 CFUs were observed from the transformation with the non-target sgRNA; however, no CFUs were observed when the target sgRNA was transformed (Table 3). We predicted that GeoCas9 was inducing dsDNA breaks and killing cells before homology-directed repair could occur. Other CRISPR/Cas genome editing systems have overcome the high toxicity of Cas9-mediated DNA cleavage by controlling *cas9* gene expression with an inducible promoter. To attempt to overcome the Cas9 toxicity, we used a recently characterized thermophilic riboswitch (Marcano-Velazquez et al., 2019) to control *geoCas9* expression; however, no CFUs were observed (data not shown). We predict that the Cas9 RNP is extremely potent, and thus even leaky expression of Cas9 from the inducible riboswitch is enough to kill cells.

Another way to overcome Cas9 toxicity is by using a nickase Cas9 (Cas9n) variant wherein one of the cleavage sites is mutated thus inducing ssDNA breaks rather than dsDNA breaks. A GeoCas9n variant, GeoCas9^{H582A}, was tested using the one step system. For the GeoCas9n, when plated without 5-FOA, ~200 CFUs were observed with the non-target sgRNA and ~150 CFUs were observed from the target sgRNA (Table 3). For the type I–B system, using the engineered strain, LL1586, ~18,000 CFUs were observed for the non-target sgRNA and ~7 CFUs were observed for the target sgRNA. However, none of the colonies screened (in either system) showed signs of editing. In addition, neither system showed phenotypic evidence of CRISPR-based *pyrF* inactivation (i.e. an increase in 5-FOA resistant colonies).

Table 3
One-step CRISPR/Cas genome editing results.

Metric	Spacer	Type I–B	Type II	Type II (nickase)
Transformation efficiency (CFU/μg DNA)	Non-target	18,000 ± 7000	55 ± 7	200 ± 70
	Target	7 ± 4.2	0 ± 0	150 ± 14
Percent increase observed in 5-FOA^R phenotype	Non-target	11 ± 2.9	n/a	0.03 ± 0.01
	Target	14 ± 3.4	n/a	0.21 ± 0.07
CRISPR fold change[a]		1.3 ± 0.13	n/a	7.3 ± 0.19

**CRISPR fold change based off of 5-FOA^R = X for target gRNA/X for non-target gRNA.

[a] Percent increase observed in 5-FOA^R phenotype/X= (Cfus on Tm + 5FOA)/(Cfus on Tm).

It has been noted in CRISPR/Cas genome editing for other Clostridial organisms that serial transfer in liquid media increased the number of edited genomes (Xu et al., 2015, WO2014089290A1, 2014). We speculate that this is required due to insufficient homology-directed repair. To increase the occurrence of homology-directed repair we employed serial transfers. After five rounds of serial transfer, we measured CRISPR-based editing based on 5-FOA resistance. For both Type I–B and Type II editing efficiencies were very low for both target and non-target sgRNA. CRISPR fold-change was 7.3-fold above background (i.e., non-target sgRNA) for the Cas9n system, and 1.3-fold above background for the Type I–B system (Table 3).

In both systems, CRISPR/Cas increased genome editing efficiencies; however, the overall editing efficiency was too low for practical use as a genome editing tool. In addition, for the Cas9n system, the non-target sgRNA and target sgRNA had similar transformation efficiencies showing that ssDNA nicks cannot be used as a counter-selectable marker. In order to use the Type II (Cas9) system as a counter-selection we predicted it would be necessary to use the WT Cas9 which makes dsDNA breaks. On the other hand, for the Type I–B system the transformation efficiency for the target sgRNA was ~10,000x lower than the non-target sgRNA, indicating that the Type I–B system can be used as an effective counter-selection.

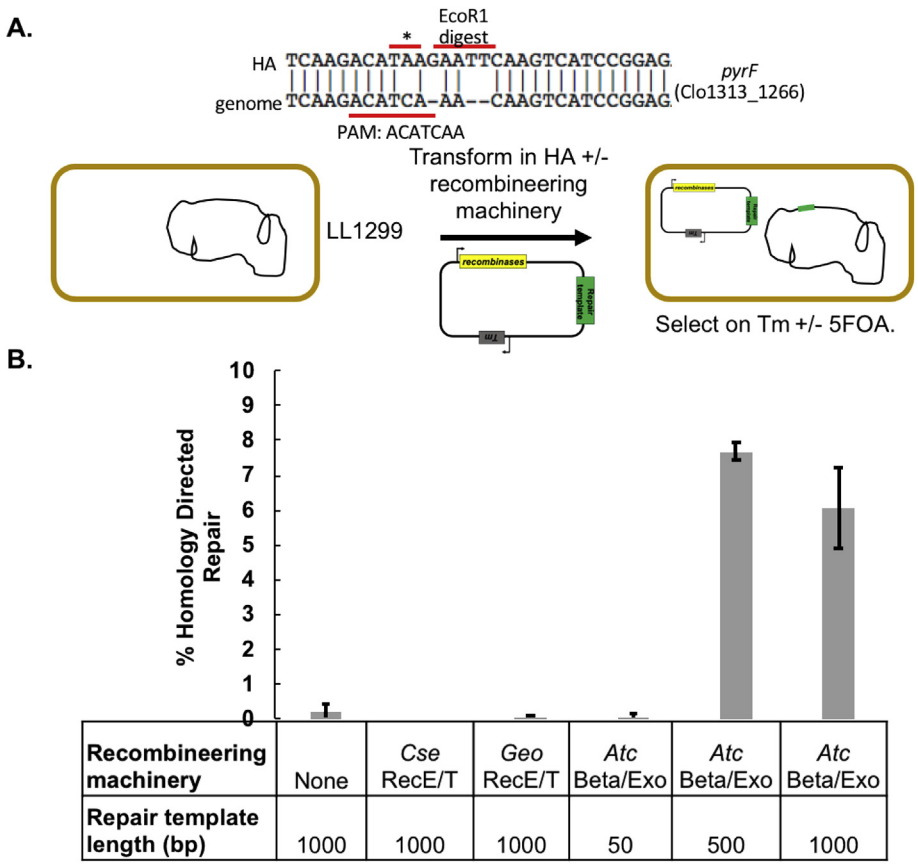
3.4. Type I–B escape mutants correlated with point mutations in cas genes

To better understand the appearance of the Type I–B escape mutants (i.e. colonies transformed with a target sgRNA, but that still had the wild type sequence at the *pyrF* locus), we resequenced total DNA (genomic and plasmid) from seven mutants. In all cases, we observed a mutation in one of the genes in the *cas* operon (specifically *cas3*, *cas5* or *cas8*) that would be expected to inactivate the encoded protein (Supplemental Table 3). We did not observe any mutations in the plasmid sequence. Although the plasmid replicon is known to support copy numbers of 10–1000 in *C. thermocellum* (Olson and Lynd, 2012b), our resequencing data unexpectedly showed that plasmid copy number was approximately one.

3.5. A recombineering system overcomes limitations of homology-directed repair in C. thermocellum

The combination of high cutting efficiency (from the killing assay) and low genome editing efficiency (from the one-step editing protocol) indicated that homology directed repair was the rate limiting step. To improve the rate of homologous recombination, we identified thermophilic recombinases (e.g. *beta/exo* from λ-Red or *recE/recT* from Rac prophage). Three recombineering machinery pairs were tested in *C. thermocellum*: *beta/exo* isolated from *Acidithiobacillus caldus* (*Atc*), *recE/recT* isolated from *Clostridium stercorarium* (*Cse*) and *recE/recT* isolated from *Geobacillus* sp (*Geo*) (Fig. 3).

Recombineering plasmids containing the same *pyrF* repair template used for CRISPR/Cas9 genome editing and expressing each pair of recombineering genes were transformed into *C. thermocellum* (strain LL1299). Interestingly, no colonies were obtained for *Cse recE/T*, indicating toxicity of *C. stercorarium* recombinases in *C. thermocellum*. For the other pairs of recombinases, sufficient numbers of colonies were obtained to assay for changes in homology-directed repair (HDR). ~30



Recombineering machinery	None	*Cse* RecE/T	*Geo* RecE/T	*Atc* Beta/Exo	*Atc* Beta/Exo	*Atc* Beta/Exo
Repair template length (bp)	1000	1000	1000	50	500	1000

Fig. 3. Recombineering machinery. A. Pictorial description of the repair template used to test the recombineering machinery and schematic for experimentally determining percent homology directed repair. B. Percent homology directed repair was determined using recombineering machinery isolated from three thermophilic organisms. *Cse* denotes *Clostridium stercorarium, Geo* denotes *Geobacillus* sp, and *Atc* denotes *Acidithiobacillus caldus.* Various repair template lengths (1000, 500 and 50bp) were tested for the active recombineering machinery isolated from *A. caldus.*

colonies were pooled, subcultured, and plated with and without 5-FOA to measure changes in the rate of *pyrF* mutation due to HDR. The *Geo recE/T* recombinases did not show any change in HDR (Fig. 3B). By contrast, the *Atc* recombinases showed a 35-fold increase in HDR compared to baseline homologous recombination. Further experiments showed that for *Atc*, the length of the repair template can be decreased from 1000 bp to 500 bp with no effect on HDR (Fig. 3B). We expect this technology will be transferable to other industrially relevant thermophilic organisms.

3.6. Efficient genome editing using a two-step CRISPR/Cas, recombineering system

Due to the low editing efficiency of the one step CRISPR/Cas system, we devised a two-step genome editing approach similar to ones described previously (WO2014089290A1, 2014). The first step involves transforming cells with a "repair" plasmid containing the repair template and (optionally) recombinases. A certain fraction of the population will incorporate the repair template onto the genome. The second step involves transforming the cells with a second plasmid containing the CRISPR machinery (sgRNA, and *cas9* for the Type II system). The second "killing" plasmid is used as a counterselection to kill cells that were not edited by the first plasmid (Fig. 4A).

In *C. thermocellum*, two antibiotic resistance markers (*cat* and *neo*) (Olson and Lynd, 2012a) and two origins of replication (pNW33N (Olson and Lynd, 2012a) and pBAS2 (Groom et al., 2016)) have been characterized. However, the two plasmids have different temperature sensitivities that makes them incompatible (data not shown). Thus, we used the pNW33N origin of replication for both plasmids but were able to maintain both of them simultaneously by using different antibiotic resistance markers for each one: *cat* for the "repair" plasmid and *neo* for the "killing" plasmid.

After transformation with the "repair" plasmid, cells were grown for about 13 generations (3 passages of 1:20 dilution) to allow time for recombination to occur. Since we are using *pyrF* as a reporter, we could then measure the frequency of repair either by plating on 5-FOA (which would measure the sum of *pyrF* mutations due to homologous recombination and due to random point mutations) (Fig. 3) or by transforming with a "killing" vector and measuring changes in transformation efficiency (which would measure the sum of *pyrF* mutations due to homologous recombination and "escape" mutations in other parts of the CRISPR machinery) or by doing both. This third option (transforming with a "killing" vector and selection on 5-FOA) resulted in no transformants, so we proceeded with the second option ("killing" vector only), and subsequently verified the presence of the target mutation by PCR and restriction digest.

We evaluated all combinations of CRISPR targeting (target vs. non-target control), CRISPR system (Type I–B vs II), and recombinase (±), in biological duplicates (Supplemental Table 4 and Supplemental Fig. 3). In the absence of a recombinase, we would expect <1% of the cells to survive the "killing" plasmid (Fig. 3). Instead, we see 20–50% survival for the Type I–B system and >100% for the Type II system. In the presence of a recombinase, we would expect 5–10% of the cells to survive the "killing" plasmid. Again, we see much higher survival, 10–30% for the Type I–B system and 50–60% for the Type II system. One possible explanation for the increased survival is a large contribution from escape mutants. Another possibility is that the "killing" plasmids actually interact with the "repair" plasmids to stimulate recombination.

There are two important metrics to consider when evaluating the usefulness of a gene editing tool. One metric is the total number of correct transformants, the other is the fraction of correct transformants (Fig. 4B). Ideally, both would be achieved simultaneously, however there are some cases where one or the other is more important. In the absence of both recombinases and targeting sgRNAs, only a single colony in one replicate was correct. This low rate of background recombination is expected, and

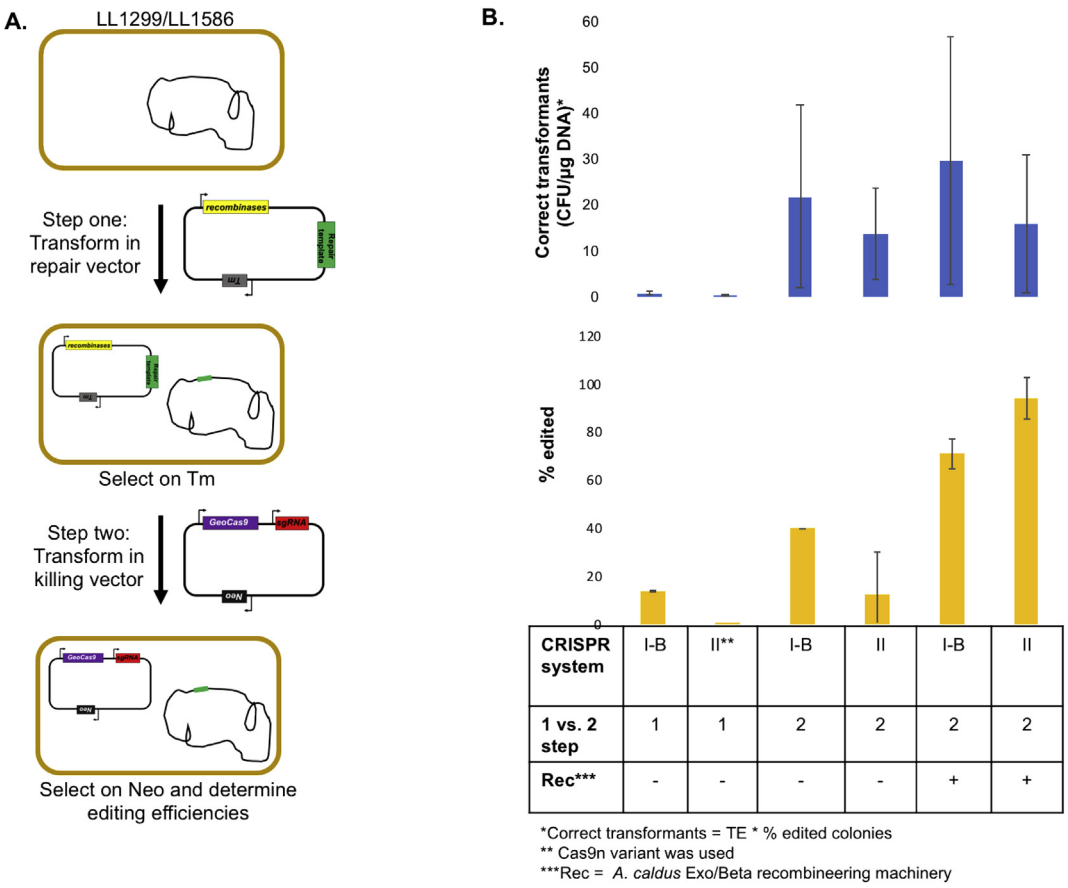


Fig. 4. Two step system results. A. Schematic for two-step CRISPR/Cas genome editing. B. Summary of correct transformants (blue bars) and percent genomes edited (yellow bars) for the target sgRNA for various CRISPR/Cas systems tested. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

much too low to be useful for genetic engineering. Introducing recombinases resulted in a large improvement with an average of 14–18 correct transformants (per µg DNA). In the complete system, with both recombinases and the target sgRNA, the Type I–B system had 27 correct colonies, and the Type II system had 16 correct colonies. However, the Type I–B system only had 71% correct, whereas it was 94% for the Type II system (Fig. 4B).

Given the large variability of transformation efficiency, the relatively small differences between these numbers is likely due to random variation, however, to maximize the number of correct colonies (e.g. because of a need to identify a rare mutant), the Type I–B system might be preferred. On the other hand, for minimization of colony screening, the Type II system might be preferred, due to its higher fraction of correct colonies. Other considerations include the choice of PAM sequence. The Type I–B system allows for easy multiplexing and its T-rich, 3 nucleotide PAM sequence is found frequently in the *C. thermocellum* genome. The Type II GeoCas9 has a more complex PAM sequence that could limit the choice of editing locations.

While this tool is effective it does have limitations that we are currently optimizing. The transformation efficiency is low even for the non-target sgRNA constructs. The maximum transformation efficiency of *C. thermocellum* is about 10^5 CFU/ug DNA (Olson and Lynd, 2012a), however for the 2-step system, the second transformation frequently had an efficiency of 10^1–10^2 CFU/ug DNA. We speculate that this results from transforming two plasmids with the same origin of replication (pNW33N). The somewhat worse transformation efficiency for the Type II system seems to be related to its overall toxicity. This is a frequent problem with heterologous expression of *cas9* genes and is thought to be due to off-target cutting. The higher level of background colonies in the Type I–B system (compared to the Type II system) is probably due to mutations in the *cas* operon. The Type II system avoids the problem of *cas9* mutations because the *cas9* is on a multi-copy plasmid, along with the sgRNA.

In addition to the low transformation efficiency with the two-plasmid system, the Neomycin selection has historically been an unreliable selectable marker in *C. thermocellum,* since the *neo* gene only confers a 2 to 4-fold increase in Neomycin resistance. Based on the high killing during the toxicity assay we speculate the high background observed during the two-plasmid system is due to the weak neomycin selection. Nonetheless, by combining the thermophilic CRISPR/Cas system (either Type I–B or Type II) with the recombinases, we developed a new tool that allows for efficient CRISPR/Cas genome editing.

4. Conclusions

In this study, we report the development of the heterologous Type II and native Type I–B CRISPR/Cas genome editing systems in the industrially relevant thermophile *C. thermocellum*. We found that both the Type II GeoCas9 and engineered Type I–B system are extremely efficient at CRISPR mediated killing. However, a one-step CRISPR/Cas genome editing system resulted in very low genome editing outputs. We concluded that the one-step system is too inefficient to use as a tool, likely due to low homologous recombination in *C. thermocellum*. To overcome this apparent limitation in homologous recombination, we performed detailed characterization of novel, thermophilic recombineering machinery and found that expressing Exo/Beta homologs from *A. caldus* greatly increases homologous recombination in *C. thermocellum*. By combining a thermophilic CRISPR system (either Type I–B or Type II)

with recombinases via a two-step protocol, we developed a new tool that allows for efficient CRISPR/Cas genome editing. Overall, we improved genome editing efficiency from 14% to 70% for the Type I–B system and from 0.21% to 94% for the Type II system. Both systems enable strains to be constructed in about half the time (2 weeks) compared to traditional *C. thermocellum* strain construction (4 weeks). Our work offers efficient editing tools that combine CRISPR/Cas and recombineering for rapid genome editing in *C. thermocellum*. These tools may also be useful for genome editing in other thermophilic microbes.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Lee Lynd is a co-founder of the Enchi Corporation, which has a financial interest in *C. thermocellum*. Work within was filed as a provisional patent, Provisional Patent Application No. 62/896,555 titled *Novel Recombineering Machinery to Increase Homology Directed Genome Editing in Thermophilic Microbes*.

Acknowledgements

We thank Timothy Chapman for this assistance characterizing the promoter insertion strains. We thank Michael Pyne for his assistance identifying Type I–B PAM sequences by BLAST search. We thank Adam Guss for his insightful discussions and technical advice. We thank Emily Freed for her critical reading of the manuscript and insightful discussions. We thank Yanhao Jiang for his assistance in molecular cloning for the Type II constructs. We would also like to thank members of the Cameron Lab, Lynd Lab, and Gill Lab for their support. This work was supported by the Center for Bioenergy Innovation. The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. Genome resequencing was performed by the Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, and is supported by the Office of Science of the U.S. Department of Energy under contract number DE-AC02–05CH11231.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mec.2019.e00116.

References

WO2014089290A1, 2014. Crispr-based genome modification and regulation - Google Patents. at. https://patents.google.com/patent/WO2014089290A1.
Adli, M., 2018. The CRISPR tool kit for genome editing and beyond. Nat. Commun. 9, 1911.
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.
Brouns, S.J.J., et al., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964.
Brown, S.D., et al., 2014. Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of Clostridium autoethanogenum and analysis of CRISPR systems in industrial relevant Clostridia. Biotechnol. Biofuels 7, 40.
Carte, J., Wang, R., Li, H., Terns, R.M., Terns, M.P., 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev. 22, 3489–3496.
Cheng, F., et al., 2017. Harnessing the native type I-B CRISPR-Cas for genome editing in a polyploid archaeon. J. Genet. Genom. 44, 541–548.
Cho, J.S., et al., 2017. CRISPR/Cas9-coupled recombineering for metabolic engineering of Corynebacterium glutamicum. Metab. Eng. 42, 157–167.
Cho, S., et al., 2018. High-level dCas9 expression induces abnormal cell morphology in Escherichia coli. ACS Synth. Biol. 7, 1085–1094.
Chung, D., Cha, M., Guss, A.M., Westpheling, J., 2014. Direct conversion of plant biomass to ethanol by engineered Caldicellulosiruptor bescii. Proc. Natl. Acad. Sci. U.S.A. 111, 8931–8936.
Cong, L., et al., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.
Deltcheva, E., et al., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602–607.
Dong, H., Tao, W., Gong, F., Li, Y., Zhang, Y., 2014. A functional recT gene for recombineering of Clostridium. J. Biotechnol. 173, 65–67.
Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U.S.A. 109, E2579–E2586.
Groom, J., et al., 2016. Promiscuous plasmid replication in thermophiles: use of a novel hyperthermophilic replicon for genetic manipulation of Clostridium thermocellum at its optimum growth temperature. Metab. Eng. Commun. 3, 30–38.
Harrington, L.B., et al., 2017. A thermostable Cas9 with increased lifetime in human plasma. Nat. Commun. 8, 1424.
Harris, L.M., Welker, N.E., Papoutsakis, E.T., 2002. Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in Clostridium acetobutylicum ATCC 824. J. Bacteriol. 184, 3586–3597.
Hochstrasser, M.L., Doudna, J.A., 2015. Cutting it close: CRISPR-associated endoribonuclease structure and function. Trends Biochem. Sci. 40, 58–66.
Huang, H., et al., 2016. CRISPR/Cas9-Based efficient genome editing in Clostridium ljungdahlii, an autotrophic gas-fermenting bacterium. ACS Synth. Biol. 5, 1355–1361.
Hwang, W.Y., et al., 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. 31, 227–229.
Ingle, P., et al., 2019. Generation of a fully erythromycin-sensitive strain of Clostridioides difficile using a novel CRISPR-Cas9 genome editing system. Sci. Rep. 9, 8123.
Jacobs, J.Z., Ciccaglione, K.M., Tournier, V., Zaratiegui, M., 2014. Implementation of the CRISPR-Cas9 system in fission yeast. Nat. Commun. 5, 5344.
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.
Jinek, M., et al., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.
Jinek, M., et al., 2013. RNA-programmed genome editing in human cells. Elife 2, e00471.
Jones, S.W., et al., 2016. CO2 fixation by anaerobic non-photosynthetic mixotrophy for improved carbon conversion. Nat. Commun. 7, 12800.
Kiu, R., Hall, L.J., 2018. An update on the human and animal enteric pathogen Clostridium perfringens. Emerg. Microb. Infect. 7, 141.
Lee, Y.J., Hoynes-O'Connor, A., Leong, M.C., Moon, T.S., 2016. Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. Nucleic Acids Res. 44, 2462–2473.
Li, Y., et al., 2016. Harnessing Type I and Type III CRISPR-Cas systems for genome editing. Nucleic Acids Res. 44, e34.
Li, Q., et al., 2016. CRISPR-based genome editing and expression control systems in Clostridium acetobutylicum and Clostridium beijerinckii. Biotechnol. J. 11, 961–972.
Li, Q., et al., 2019. CRISPR-Cas9D10A nickase-assisted base editing in the solvent producer Clostridium beijerinckii. Biotechnol. Bioeng. 116, 1475–1483.
Liu, J., et al., 2017. Development of a CRISPR/Cas9 genome editing toolbox for Corynebacterium glutamicum. Microb. Cell Factories 16, 205.
Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66, 506–577 table of contents.
Lynd, L.R., van Zyl, W.H., McBride, J.E., Laser, M., 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr. Opin. Biotechnol. 16, 577–583.
Maikova, A., Kreis, V., Boutserin, A., Severinov, K., Soutourina, O., 2019. Using an endogenous CRISPR-Cas system for genome editing in the human pathogen Clostridium difficile. Appl. Environ. Microbiol. 85.
Mali, P., et al., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823–826.
Marcano-Velazquez, J.G., Lo, J., Nag, A., Maness, P.-C., Chou, K.J., 2019. Developing riboswitch-mediated gene regulatory controls in thermophilic bacteria. ACS Synth. Biol. 8, 633–640.
McAllister, K.N., Sorg, J.A., 2019. CRISPR genome editing systems in the genus Clostridium: a timely advancement. J. Bacteriol. https://doi.org/10.1128/JB.00219-19.
McAllister, K.N., Bouillaut, L., Kahn, J.N., Self, W.T., Sorg, J.A., 2017. Using CRISPR-Cas9-mediated genome editing to generate C. difficile mutants defective in selenoproteins synthesis. Sci. Rep. 7, 14672.
Migriauli, I., et al., 2018. Epidemiology of clostridium difficile infection in hospitalised pediatric patients in Georgia. Georgian Med. News 172–176.
Mohr, G., et al., 2013. A targetron system for gene targeting in thermophiles and its application in Clostridium thermocellum. PLoS One 8, e69032.
Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., Almendros, C., 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology (Read.) 155, 733–740.
Mougiakos, I., Bosma, E.F., de Vos, W.M., van Kranenburg, R., van der Oost, 2016. J. Next generation prokaryotic engineering: the CRISPR-Cas toolkit. Trends Biotechnol. 34, 575–587.
Mougiakos, I., et al., 2017. Characterizing a thermostable Cas9 for bacterial genome editing and silencing. Nat. Commun. 8, 1647.
Nagaraju, S., Davies, N.K., Walker, D.J.F., Köpke, M., Simpson, S.D., 2016. Genome editing of Clostridium autoethanogenum using CRISPR/Cas9. Biotechnol. Biofuels 9, 219.
Nuñez, J.K., et al., 2014. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. Nat. Struct. Mol. Biol. 21, 528–534.
Olson, D.G., Lynd, L.R., 2012. Transformation of Clostridium thermocellum by electroporation. Methods Enzymol. 510, 317–330.
Olson, D.G., Lynd, L.R., 2012. Computational design and characterization of a temperature-sensitive plasmid replicon for gram positive thermophiles. J. Biol. Eng. 6, 5.

Olson, D.G., McBride, J.E., Shaw, A.J., Lynd, L.R., 2012. Recent progress in consolidated bioprocessing. Curr. Opin. Biotechnol. 23, 396–405.
Olson, D.G., et al., 2015. Identifying promoters for gene expression in Clostridium thermocellum. Metab. Eng. Commun. 2, 23–29.
Penewit, K., et al., 2018. Efficient and scalable precision genome editing in Staphylococcus aureus through conditional recombineering and CRISPR/Cas9-Mediated counterselection. mBio 9.
Pitcher, R.S., Brissett, N.C., Doherty, A.J., 2007. Nonhomologous end-joining in bacteria: a microbial perspective. Annu. Rev. Microbiol. 61, 259–282.
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., et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183.
Richter, H., et al., 2012. Characterization of CRISPR RNA processing in Clostridium thermocellum and Methanococcus maripaludis. Nucleic Acids Res. 40, 9887–9896.
Richter, H., Rompf, J., Wiegel, J., Rau, K., Randau, L., 2017. Fragmentation of the CRISPR-Cas type I-B signature protein Cas8b. Biochim. Biophys. Acta Gen. Subj. 1861, 2993–3000.
Sinkunas, T., et al., 2011. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J. 30, 1335–1342.
Tian, L., et al., 2019. Metabolic engineering of Clostridium thermocellum for n-butanol production from cellulose. Biotechnol. Biofuels 12, 186.
Tripathi, S.A., et al., 2010. Development of pyrF-based genetic system for targeted gene deletion in Clostridium thermocellum and creation of a pta mutant. Appl. Environ. Microbiol. 76, 6591–6599.
Tsui, T.K.M., Hand, T.H., Duboy, E.C., Li, H., 2017. The impact of DNA topology and guide length on target selection by a cytosine-specific Cas9. ACS Synth. Biol. 6, 1103–1113.
Wang, H., et al., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153, 910–918.
Wang, Y., et al., 2015. Markerless chromosomal gene deletion in Clostridium beijerinckii using CRISPR/Cas9 system. J. Biotechnol. 200, 1–5.
Wang, Y., et al., 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.
Wang, S., Dong, S., Wang, P., Tao, Y., Wang, Y., 2017. Genome editing in Clostridium saccharoperbutylacetonicum N1-4 with the CRISPR-Cas9 system. Appl. Environ. Microbiol. 83.
Wang, S., et al., 2018. Genome engineering of Clostridium difficile using the CRISPR-Cas9 system. Clin. Microbiol. Infect. 24, 1095–1099.
Wasels, F., Jean-Marie, J., Collas, F., López-Contreras, A.M., Lopes Ferreira, N., 2017. A two-plasmid inducible CRISPR/Cas9 genome editing tool for Clostridium acetobutylicum. J. Microbiol. Methods 140, 5–11.
Xi, L., Schmidt, J.C., Zaug, A.J., Ascarrunz, D.R., Cech, T.R., 2015. A novel two-step genome editing strategy with CRISPR-Cas9 provides new insights into telomerase action and TERT gene expression. Genome Biol. 16, 231.
Xiao, Y., Luo, M., Dolan, A.E., Liao, M., Ke, A., 2018. Structure basis for RNA-guided DNA degradation by cascade and Cas3. Science 361.
Xin, B., et al., 2016. Co-utilization of glycerol and lignocellulosic hydrolysates enhances anaerobic 1,3-propanediol production by Clostridium diolis. Sci. Rep. 6, 19044.
Xu, T., et al., 2015. Efficient genome editing in Clostridium cellulolyticum via CRISPR-Cas9 nickase. Appl. Environ. Microbiol. 81, 4423–4431.
Xu, Q., et al., 2016. Dramatic performance of Clostridium thermocellum explained by its wide range of cellulase modalities. Sci. Adv. 2, e1501254.
Xu, T., Li, Y., He, Z., Van Nostrand, J.D., Zhou, J., 2017. Cas9 nickase-assisted RNA repression enables stable and efficient manipulation of essential metabolic genes in Clostridium cellulolyticum. Front. Microbiol. 8, 1744.
Zhang, J., Zong, W., Hong, W., Zhang, Z.-T., Wang, Y., 2018. Exploiting endogenous CRISPR-Cas system for multiplex genome editing in Clostridium tyrobutyricum and engineer the strain for high-level butanol production. Metab. Eng. 47, 49–59.
Zhou, J., et al., 2015. Physiological roles of pyruvate ferredoxin oxidoreductase and pyruvate formate-lyase in Thermoanaerobacterium saccharolyticum JW/SL-YS485. Biotechnol. Biofuels 8, 138.