1 **Markerless chromosomal gene deletion in *Clostridium beijerinckii* using**
2 **CRISPR/Cas9 system**

3 Yi Wang^{a,b,†}, Zhong-Tian Zhang^c, Seung-Oh Seo^{a,b}, Kijoong Choi^a, Ting Lu^{b,d}, Yong-Su
4 Jin^{a,b}, Hans P. Blaschek^{a,b,e,*}

5 ^aDepartment of Food Science and Human Nutrition, University of Illinois at Urbana-
6 Champaign, Urbana, IL 61801, USA

7 ^bInstitute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL
8 61801, USA

9 ^cDepartment of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL
10 61801, USA

11 ^dDepartment of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL
12 61801, USA

13 ^eThe Integrated Bioprocessing Research Laboratory (IBRL), University of Illinois at
14 Urbana-Champaign, Urbana, IL 61801, USA

15 [†]Current address: 215 Tom E. Corley Building, Biosystems Engineering Department,
16 Auburn University, Auburn, AL 36849, USA

17
18
19
20
21 ***Corresponding author:**

22 Hans P. Blaschek, The Integrated Bioprocessing Research Laboratory (IBRL)
23 University of Illinois at Urbana-Champaign,
24 1207 W Gregory Drive, Urbana, IL 61801, USA
25 Tel: 1-217-333-8224; Fax: 1-217-244-2517;
26 E-mail: blaschek@illinois.edu

ABSTRACT

The anaerobic spore-forming, gram-positive, solventogenic clostridia are notorious for being difficult to genetically engineer. Based on CRISPR/Cas9 assisted homologous recombination, we demonstrated that clean markerless gene deletion from the chromosome can be easily achieved with a high efficiency through a single-step transformation in *Clostridium beijerinckii* NCIMB 8052, one of the most prominent strains for acetone, butanol and ethanol (ABE) production. This highly efficient genome engineering system can be further explored for multiplex genome engineering purposes. The protocols and principles developed in this study provided valuable references for genome engineering in other microorganisms lacking developed genetic engineering tools.

Keywords: CRISPR/Cas9 *Clostridium* Genome engineering Homologous recombination
Metabolic engineering Butanol

Main Text:

Solventogenic clostridia are industrially significant microorganisms with great potentials for producing biofuels and biochemicals (Jones and Woods, 1986). Metabolic engineering based on designed genetic perturbations is a great strategy for studying the complex metabolism of solventogenic clostridia as well as constructing strains with desirable phenotypes. Various genetic engineering tools for the manipulation of solventogenic clostridia have recently been developed, however, many limitations still exist for efficient use of these tools. For example, the genetic tools based on mobile group II introns (ClosTron or TargeTron) have been employed by various research groups in many different strains (Heap et al., 2010; Jang et al., 2012; Wang et al., 2013). These technologies achieve disruption of genes by inserting an intron fragment at the chromosomal targeting site. One limitation of

these methods, however, is that, their efficiency heavily relies on the precise prediction of insertion sites using a specific algorithm, and it can be very challenging to find an optimal target site when the interesting insertion sequence is short (for example < 400bp). In addition, the gene knockout achieved with this method is virtually not a true deletion of gene, but rather an interruption of the gene function, whose residue activity may be variable based on the insertion site and the feature of the gene of interest (Al-Hinai et al., 2012; Wang et al., 2013). Furthermore, it has been reported that intron insertions achieved through these methods, especially for the ones at the sense orientation, may be spliced back out by the intron-encoding protein (Frazier et al., 2003); thus a complete curing of the plasmid must be achieved after the desirable mutant is obtained. Recently, a counter selection marker-assisted genetic engineering system has been developed by Al-Hinai et al (Al-Hinai et al., 2012). Allelic-exchange-based chromosomal gene deletion and integration can be obtained with the assistance of the inducible counter selection marker *mazF*, a toxin gene from *Escherichia coli*. However, with this system, an antibiotic marker is still required to be integrated into chromosome for selection purposes. Although the marker may be ‘flipped out’ by the FLP recombinase expressed on an additional plasmid (which makes the whole transformation process laborious and complicated), a remarkable scar is still left behind. Moreover, the functionality of the FLP recombinase could be dependent on the host strains, which brings up more limitations for the application of this system.

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) system is an immune system in bacteria and archaea that can efficiently cleave foreign DNA entering the cells, such as phage or plasmids (Sorek et al., 2013). Short DNA sequences located between the CRISPR repeats make up a CRISPR array of targets. This CRISPR array is transcribed and then processed within the repeat sequences, yielding RNA fragments called CRISPR-RNA (crRNA). The crRNA then directs the Cas nuclease to

the target site, and the presence of a specific protospacer-adjacent motif (PAM) sequence leads to a Cas-mediated cleavage of the target sequence. In type-II CRISPR/Cas systems, Cas9 forms a dual-RNA complex with crRNA and a trans-activating CRISPR-RNA (tracrRNA) that is required for Cas9 nuclease to be functional (Chylinski et al., 2013). Recently, the natural CRISPR/Cas9 system from *Streptococcus pyogenes* containing the dual-RNA (crRNA and tracrRNA) and Cas9 has been applied for genome engineering in both Gram-positive *S. pneumonia* (Jiang et al., 2013) and *Lactobacillus reuteri* (Oh and van Pijkeren, 2014) and Gram-negative *E. coli* (Jiang et al., 2013) bacteria. In all cases, the host strain needs to have an established functional recombineering system if it is not highly recombinogenic; the CRISPR/Cas9 system is primarily functional as a tool for selecting edited cells against non-edited background cells (with a possible modest induction of recombination) thereby leading to high efficiency of genome engineering. On the other hand, for the genetic engineering of eukaryotic cells, a single chimeric gRNA (with features of both crRNA and tracrRNA) along with Cas9 protein has been mostly employed, which demonstrated comparable, if not higher efficiency than the dual-RNA guiding Cas9 system (Cong et al., 2013; DiCarlo et al., 2013; Jinek et al., 2012; Mali et al., 2013).

In this study, we attempted to employ the CRISPR/Cas9 system for genetic engineering in *C. beijerinckii*, one of the most prominent solvent-producing species with various advantages over others, including being more resistant to degeneration (i.e. losing the ability to produce solvents and form spores) (Cornillot et al., 1997), and having the capabilities for consuming a broader range of substrates (Ezeji and Blaschek, 2008), etc. While the expression and functionality of natural dual-RNA guiding Cas9 system was unknown in *C. beijerinckii*, we set out to customize the chimeric gRNA guiding Cas9 system for proper expression using *C. beijerinckii* native promoters for our genetic engineering purposes. Based on our own experience and reports from the literature (Heap et al., 2010), the *spo0A*

(Cbei_1712) knockout mutant generated using group II intron based approaches displayed a severe septation defect phenotype, and because of this it can be easily identified by investigating the morphology on plates. Therefore, in this study, we selected *spo0A* as the target gene for chromosomal sequence deletion for an easy identification and demonstration of our system.

The plasmid pTJ1 (Jesse, 2003; Wang et al., 2013) has been used as the mother vector for constructing our customized CRISPR/Cas9 system. All DNA oligonucleotide sequences are listed in Table S1 in the supplemental materials. The CRISPR/Cas9 vector for *spo0A* knockout (pYWCas-spo0A-HR, Fig. 1A) was constructed using Gibson Assembly (New England Biolabs Inc., Ipswich, MA) as follows (Gibson et al., 2009). The *spoIIIE* (Cbei_0097) promoter was selected for Cas9 expression, and was amplified using primers P01 and P02 from *C. beijerinckii* NCIMB 8052 genomic DNA (gDNA). The Cas9 open reading frame (ORF) was amplified from the plasmid pMJ806 (purchased from Addgene) (Jinek et al., 2012) using primers P03 and P04. A 784bp gBlock fragment (designated gRNA-gBlk-spo0A, see Table S1) was synthesized by Integrated DNA Technologies, including a thiolase gene (Cbei_0411) terminator, a promoter of the *C. beijerinckii* small RNA gene (sCbei_5830) (Chen et al., 2011) for chimeric gRNA expression, a 20-nt guiding sequence (5'-ATAATAAAAGAATGGACAAA-3') targeting on the *spo0A* promoter region fused with the chimeric gRNA sequence (Mali et al., 2013), and a transcription terminator derived from *S. pyogenes* (Qi et al., 2013). Then these three fragments cited above (*spoIIIE* promoter, Cas9 ORF and the gBlock) were finally assembled and inserted into pTJ1 between the *ApaI* and *NotI* restriction enzyme sites, and the resultant vector was named as pYWCas-spo0A.

Since neither a desirable recombination efficiency or a functional recombineering system is available for *C. beijerinckii*, we decided to put the DNA editing template onto the same vector with the Cas9 and the gRNA sequences with the expectation that a homologous

recombination event would occur through a double-crossover event (Al-Hinai et al., 2012). Therefore, by functioning as a selection tool against non-edited cells, the CRISPR/Cas9 system was expected to confer a high efficiency for genome engineering implemented through homologous recombination. To test this hypothesis, a 2 kb fragment was amplified from *C. beijerinckii* gDNA such that 262 bp (including the 20-nt targeting sequence) covering both the *spo0A* promoter region and the beginning of the ORF was deleted with 1 kb homology flanking at both sides (Fig. S1). The fragment was inserted into the *NotI* site of pYWCas-*spo0A* to create the final vector pYWCas-*spo0A*-HR. A control vector pYWCas-*spo0A*-HRC was constructed in the same manner, but the promoter for the chimeric gRNA is missing and the 20-nt guiding sequence in pYWCas-*spo0A*-HR was replaced with a random sequence along with two *BtgZI* sites to facilitate the construction of new vectors re-targeting on other loci of the *C. beijerinckii* genome using the CRISPR/Cas9 system.

Transformation of *C. beijerinckii* NCIMB 8052 with pYWCas-*spo0A*-HR or pYWCas-*spo0A*-HRC (2 μ g for each) was performed with electroporation (Wang et al., 2013). After ~48 h of growth on tryptone–glucose–yeast extract (TGY) plates supplemented with 25 μ g/ml erythromycin, five ‘snowflake-shaped’ colonies (Fig. 1B) were observed which were tentatively identified as *spo0A* knockout mutants based on known phenotypes of the same mutant strain generated using TargeTron technology (unpublished data). Confirmation was carried out using colony PCR with a pair of primers (P09 and P10) that flank the upstream and downstream, respectively, of the 2,262 bp homologous recombination targeting region (and thus the primers can only anneal to the chromosome but not the vector). The results confirmed that all the colonies generated PCR bands with the expected deletion of the target sequence (Fig. 1D). The generated PCR product from colony PCR was gel-extracted, and additional PCR was carried out using the purified product as template with primers (P11 and P12) targeting on a shorter sequence that directly flanks the 262 bp deletion region. The

results further confirmed the desirable deletion (Fig. 1E). The PCR product was further verified by Sanger sequencing. In order to investigate whether there is a possibility that the plasmid has been integrated into the chromosome (besides the desirable homologous recombination for gene deletion), one originally identified positive colony was subjected to a re-plating on a TGY plate. Single colonies out of this re-plating were then picked and additional colony PCR was carried out with the forward primer annealing to the chromosome (P09) and the reverse primer annealing to the plasmid (P13), and thus an expected 2,337 bp band would be generated if there was a copy of plasmid integrated into the chromosome. Interestingly, eight out of the fifteen tested colonies generated the expected PCR band, while there was no band observed for the other seven, which were inferred to have no plasmid integrated (data not shown). For those seven colonies, the mutation was further confirmed by colony PCR as described above. One colony out of the seven was then picked and the hosted plasmid was cured through a subculturing process as describe previously (Wang et al., 2013).

For the transformation with the control vector, an efficiency of 2.58×10^2 cfu/ μ g-DNA was obtained. However, all the colonies showed similar morphology as the wild type on plates (Fig. 1C). Twenty-four colonies were examined using colony PCR as was the case for the mutant, and all of them demonstrated the same band size as the control using the wild type gDNA as template (data not shown).

The *spo0A* gene is considered as a global regulator gene in many spore-forming microorganisms. For example, in *Bacillus subtilis*, it has been reported that the *spo0A* gene can control the expression of up to 500 other genes (Molle et al., 2003), and what is interesting is that it can also influence the colony development (Verhamme et al., 2009). In this study, with the deletion of the promoter region and part of the *spo0A* ORF in *C. beijerinckii*, the mutant colonies demonstrated very interesting morphology on the plates. Similar phenomena have been reported in *B. subtilis* (Fujikawa and Matsushita, 1989). Some

researchers have associated this so-called ‘fractal morphology’ with chemotaxis of the *B. subtilis* cells (Aguilar et al., 2007; Ben-Jacob, 1997; Leyva et al., 2013). However, to our best knowledge, this study is the first report concerning the effect of *spo0A* gene on the colony morphology in solventogenic clostridia, although couple of studies from Papoutsakis group have previously demonstrated the possible regulation of *spo0A* gene on the motility and chemotaxis gene expression in *C. acetobutylicum* (Alsaker et al., 2004; Tomas et al., 2003). It warrants further investigation whether the deletion of *spo0A* gene can affect the expression of chemotaxis/motility genes in *C. beijerinckii*.

C. beijerinckii has a very low homologous recombination efficiency, making it almost impossible to isolate positive mutants out of the background non-edited cells, as indicated by the transformation with the control vector in this study. With the assistance of CRISPR/Cas9 for selection against non-edited cells, we achieved a markerless gene deletion with a high efficiency. Further exploration of this highly efficient genetic engineering system for multiplex genome engineering purposes (including large fragment deletion, gene knock-in, and even single nucleotide modification, etc.) is underway in our laboratory. Notably, the principles we developed in this study can be easily applied for the genome engineering on other microorganisms.

Acknowledgements

This work was supported by Department of Energy (DOE) grant #2011-01219 to HPB. We thank Dr. Roderick I. Mackie for sharing his lab facilities. We also thank Mr. Wenyan Jiang (Dr. Luciano A. Marraffini group at The Rockefeller University), Dr. Esteban Toro (Dr. Adam P. Arkin group at UC-Berkeley), Dr. Jason Peters (Dr. Carol Gross group at UC-San Francisco), and Dr. Martin Jinek (Dr. Jennifer Doudna group at UC-Berkeley) for their helpful discussions.

Fig.1 Deletion of gene fragment from the chromosome of *Clostridium beijerinckii* using CRISPR/Cas9 system. (A) The customized CRISPR/Cas9 vector targeting on *spo0A* gene (Cbei_1712) in *C. beijerinckii* achieved a markerless gene fragment deletion via a homologous recombination event. The boxed regions (with black dash lines) show the approximate regions of homology (~1 kb each) incorporated in the CRISPR/Cas9 vector. The region covered by the red dash line box indicated the gene sequence that has been deleted. Putative deletion mutants were screened by two rounds of PCRs with the indicated primers. *spoIIE*-P: the promoter of *spoIIE* gene (Cbei_0097); Cas9: the *Streptococcus pyogenes* Cas9 ORF amplified from the plasmid pMJ806 (purchased from Addgene) (Jinek et al., 2012); *thl*-T: terminator of thiolase gene (Cbei_0411); *sRNA*-P: the promoter of the small RNA gene (sCbei_5830) (Chen et al., 2011); gRNA: the chimeric gRNA (Mali et al., 2013), with a 20-nt guiding sequence (5'-ATAATAAAAGAATGGACAAA-3') targeting on the *spo0A* promoter region; tracrRNA-T: the transcription terminator derived from *S. pyogenes* (Qi et al., 2013); CAK1: the *C. beijerinckii* Gram-positive replicon; Erm^r: the erythromycin resistant marker. (B) *spo0A* mutant colonies (re-spread from the original colony) showed snowflake-like shape. (C) The colonies from the control transformation (re-spread from the original colony) showed round shape (similar to the wild type morphology). (D) Colony PCR with primers flanking the upstream and downstream respectively of the 2,262 bp homologous recombination targeting region (thus only annealing to the chromosome but not the vector) confirmed the expected deletion. Lanes M: 1 kb DNA marker, with the numbers on the left corresponding to the marker length in kb (New England Biolabs Inc., Ipswich, MA); Lane 1: directly picked the *spo0A* mutant colony as template (PCR amplicon size: 2,767 bp); Lane 2: using the *C. beijerinckii* 8052 wild type genomic DNA as template (PCR amplicon size: 3,029 bp); (E) PCR with primers directly flanking the expected deletion region confirmed the desirable deletion of the gene fragment. Lanes M: 100 bp DNA marker, with the numbers on the left

227 corresponding to the marker length in bp (New England Biolabs Inc., Ipswich, MA); Lane 1:
228 using the gel-extraction purified colony PCR product as template (PCR amplicon size: 317
229 bp); Lane 2: using the *C. beijerinckii* 8052 wild type genomic DNA as template (PCR
230 amplicon size: 579 bp).

231

232

Fig. 1

- Aguilar, C., Vlamakis, H., Losick, R., Kolter, R., 2007. Thinking about *Bacillus subtilis* as a multicellular organism. *Curr. Opin. Microbiol.* 10, 638-643.
- Al-Hinai, M.A., Fast, A.G., Papoutsakis, E.T., 2012. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. *Appl. Environ. Microbiol.* 78, 8112-8121.
- Alsaker, K.V., Spitzer, T.R., Papoutsakis, E.T., 2004. Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. *J. Bacteriol.* 186, 1959-1971.
- Ben-Jacob, E., 1997. From snowflake formation to growth of bacterial colonies II: Cooperative formation of complex colonial patterns. *Contemp. Phys.* 38, 205-241.
- Chen, Y., Indurthi, D.C., Jones, S.W., Papoutsakis, E.T., 2011. Small RNAs in the genus *Clostridium*. *mBio* 2, e00340-00310.
- Chylinski, K., Le Rhun, A., Charpentier, E., 2013. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* 10, 726-737.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823.
- Cornillot, E., Nair, R.V., Papoutsakis, E.T., Soucaille, P., 1997. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *J. Bacteriol.* 179, 5442-5447.
- 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. *Nucl. Acids Res.* 41, 4336-4343.
- Ezeji, T., Blaschek, H.P., 2008. Fermentation of dried distillers' grains and solubles (DDGS) hydrolysates to solvents and value-added products by solventogenic clostridia. *Bioresour. Technol.* 99, 5232-5242.
- Frazier, C.L., San Filippo, J., Lambowitz, A.M., Mills, D.A., 2003. Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. *Appl. Environ. Microbiol.* 69, 1121-1128.
- Fujikawa, H., Matsushita, M., 1989. Fractal growth of *Bacillus subtilis* on agar plates. *J. Phys. Soc. Jpn.* 58, 3875-3878.
- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343-345.
- Heap, J.T., Kuehne, S.A., Ehsaan, M., Cartman, S.T., Cooksley, C.M., Scott, J.C., Minton, N.P., 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J. Microbiol. Methods* 80, 49-55.
- Jang, Y.S., Lee, J.Y., Lee, J., Park, J.H., Im, J.A., Eom, M.H., Lee, J., Lee, S.H., Song, H., Cho, J.H., Seung, D.Y., Lee, S.Y., 2012. Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. *mBio* 3, 1-9.
- Jesse, T.W., 2003. Genetic characterization and manipulation of solvent-producing clostridia. PhD dissertation. Department of Food Science and Human Nutrition. University of Illinois at Urbana-Champaign, Urbana, IL.
- Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, L.A., 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotech.* 31, 233-239.
- 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.
- Jones, D.T., Woods, D.R., 1986. Acetone-butanol fermentation revisited. *Microbiol. Rev.* 50, 484-524.
- Leyva, J., Málaga, C., Plaza, R.G., 2013. The effects of nutrient chemotaxis on bacterial aggregation patterns with non-linear degenerate cross diffusion. *Statistical Mechanics and its Applications* 392, 5644-5662.
- 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.
- Molle, V., Fujita, M., Jensen, S.T., Eichenberger, P., González - Pastor, J.E., Liu, J.S., Losick, R., 2003. The *Spo0A* regulon of *Bacillus subtilis*. *Mol. Microbiol.* 50, 1683-1701.
- Oh, J.H., van Pijkeren, J.P., 2014. CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucl. Acids Res.* 42, e131.
- 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.
- Sorek, R., Lawrence, C.M., Wiedenheft, B., 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* 82, 237-266.
- Tomas, C.A., Alsaker, K.V., Bonarius, H.P.J., Hendriksen, W.T., Yang, H., Beamish, J.A., Paredes, C.J., Papoutsakis, E.T., 2003. DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5. *J. Bacteriol.* 185, 4539-4547.
- Verhamme, D.T., Murray, E.J., Stanley-Wall, N.R., 2009. DegU and Spo0A jointly control transcription of two loci required for complex colony development by *Bacillus subtilis*. *J. Bacteriol.* 191, 100-108.

295 Wang, Y., Li, X., Milne, C.B., Janssen, H., Lin, W., Phan, G., Jin, Y.S., Price, N.D., Blaschek, H.P., 2013.
296 Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of
297 acid production pathways in *Clostridium beijerinckii*. Appl. Environ. Microbiol. 79, 5853-5863.
298