CRISPR-Cas9[D10A] nickase-assisted base editing in solvent producer

Clostridium beijerinckii

Qi Li[1,2], François M. Seys [3], Nigel P. Minton[3], Junjie Yang[2], Yu Jiang[2], Weihong

Jiang[2], Sheng Yang[2,4*]

[1]College of Life Sciences, Sichuan Normal University, Longquan, Chengdu 610101, China

[2]Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai

Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China

[3]Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre (SBRC), School

of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK

[4] Jiangsu National Synergetic Innovation Center for Advanced Materials, Nanjing 210009, China

AUTHOR INFORMATION

Corresponding author

*E-mail: syang@sibs.ac.cn. Address: 300 Fenglin Road, Shanghai 200032, China.

Author Contributions

Qi Li performed the experiments. Qi Li and Sheng Yang wrote the manuscript. François

M. Seys, Nigel P. Minton, Junjie Yang, Weihong Jiang and Yu Jiang designed the

experiments and wrote the manuscript.

22   Conflict of Interest

23   The authors declare no competing financial interest.

24   **Abstract**

25   *Clostridium beijerinckii* is a potentially important industrial microorganism as it can

26   synthesize valuable chemicals and fuels from various carbon sources. The

27   establishment of convenient to use, effective gene tools with which the organism can

28   be rapidly modified is essential if its full potential is to be realized. Here, we developed

29   a genomic editing tool (pCBEclos) for use in *C. beijerinckii* based on the fusion of

30   cytidine deaminase (Apobec1), Cas9^{D10A} nickase and uracil DNA glycosylase inhibitor

31   (UGI). Apobec1 and UGI are guided to the target site where they introduce specific

32   base-pair substitutions through the conversion of C·G to T·A. By appropriate choice of

33   target sequence, these nucleotide changes are capable of creating missense mutation or

34   null mutations in a gene. Through optimization of pCBEclos, the system derived,

35   pCBEclos-opt, has been used to rapidly generate four different mutants in *C.*

36   *beijerinckii*, in *pyrE*, *xylR*, *spo0A* and *araR*. The efficiency of the system was such that

37   they could sometimes be directly obtained following transformation, otherwise only

38   requiring one single re-streaking step. Whilst CRISPR-Cas9 nickase systems, such as

39   pNICKclos2.0, have previously been reported in *C. beijerinckii*, pCBEclos-opt does

40   not rely on homologous recombination, a process that is intrinsically inefficient in

41   clostridia such as *C. beijerinckii*. As a consequence, bulky editing templates do not need

to be included in the knock-out plasmids. This both reduces plasmid size and makes their construction simpler, e.g., whereas the assembly of pNICKclos2.0 requires six primers for the assembly of a typical knock-out plasmid, pCBEclos-opt requires just two primers. The pCBEclos-opt plasmid established here represents a powerful new tool for genome editing in *C. beijerinckii,* which should be readily applicable to other clostridial species.

Key words: CRISPR, Cas9, nickase, base editing, *Clostridium beijerinckii*

1 Introduction

Clostridium beijerinckii, a spore-forming, solventogenic, Gram-positive bacterium, is a potentially important industrial strain as it can utilize a variety of carbon-based feedstocks to generate valuable chemicals and fuels (Chen & Blaschek, 1999; Dürre, 1998; Ezeji, Qureshi, & Blaschek, 2007; Y. Gu, Jiang, Yang, & Jiang, 2014; Jiang, Liu, Jiang, Yang, & Yang, 2015; Lee et al., 2008; Thakker, Martinez, Li, San, & Bennett, 2015). The establishment of convenient to use, effective gene tools with which the organism can be rapidly modified is essential if its full potential is to be realized. Such tools may be used both to provide an in-depth understanding of cell physiology and to enable the robust construction of engineered process organisms. Several genomic editing tools have been developed in *C. beijerinckii*. Till now, a commonly used

62 procedure is based on gene inactivation by group II introns, typified by

63 Clostron/Targetron technology (Heap et al., 2010; Heap, Pennington, Cartman, Carter,

64 & Minton, 2007; Shao et al., 2007). Here the presence of intron-encoding protein allows

65 a mobile group II intron to recognize and insert into a specific site of the genome,

66 resulting in gene disruption. Although Clostron/Targetron technology is effective, it

67 cannot achieve in-frame deletion, large fragment insertion or base editing. Moreover,

68 in common with any insertional mutagen, it can result in polar effects.

69   Traditional homologous recombination-dependent allelic exchange may also be

70 employed to edit *C. beijerinckii* genomes. Its application is reliant on the sequential

71 occurrence of single crossover and double crossover events. These occur naturally, but

72 at a very low frequency. According those cells in which the desired crossovers have

73 taken place need to be detected in the wild type population through the use of

74 appropriate selective tools. The latter have included the use of counter selection

75 markers (Al-Hinai, Fast, & Papoutsakis, 2012) or I-*Sce*I endonuclease (N. Zhang et al.,

76 2015). Their use, however, is somewhat laborious, involving numerous re-streaking of

77 colonies onto the necessary selective media, and their effectiveness can suffer from a

78 high background of false positives due to spontaneous mutants. A more effective means

79 of selecting the required double crossover mutants is to use CRISPR-Cas9 where the

80 wild type cells are eliminated on mass leaving only the desired mutant cells. In such a

81 system, typified by the previously described *C. beijerinckii* CRISPR-Cas9^{D10A} genome

82 editing tool (Li et al., 2016), all colonies obtained following transformation are in

83   essence mutants. However, whilst the use of CRISPR-Cas9 offers significant advantage

84   over the use of other counter selection markers, it remains reliant on homologous

85   recombination (HR), which is notoriously inefficient in clostridia and therefore reliant

86   on highly efficient DNA transfer. As the frequency of DNA transfer is inversely

87   proportional to plasmid size, the need to incorporate large editing templates in CRISPR-

88   Cas9 vectors for the purposes of HR compromise the system. Moreover, the inclusion

89   of an editing template in the design of the knock-out plasmid adds complexity, requiring

90   at least six primers for the assembly of the vector (Li et al., 2016). As the consequence

91   that the availability of a HR-independent *C. beijerinckii* genomic editing tool that

92   would involve fewer steps for assembly, and use relatively smaller vectors conducive

93   to high transformation frequencies, is highly desirable.

94   In recent years, the utility of CRISPR-Cas in genome editing has been extended

95   through its combination with deaminase enzymes to create a novel strategy for strain

96   engineering which is not reliant on HR. Cytidine deaminase or adenine deaminase is

97   fused to Cas9 effector protein (Cas9 nickase or dCas9) which allows its delivery to the

98   intended DNA target sites by the sgRNA/Cas9 complex. Upon delivery, the deaminase

99   converts nucleotide base pairs C·G to T·A or A·T to G·C. These conversions take place

100   in the absence of Cas9-mediated DNA double-stranded breaks (DSB) while the plasmid

101   employed do not require the relatively large editing templates associated with

102   traditional CRISPR-Cas9 genome editing vectors. To date, the base conversion activity

103   of cytidine deaminase and adenine deaminase has been used in prokaryotes (Banno,

104   Nishida, Arazoe, Mitsunobu, & Kondo, 2018; Gaudelli et al., 2017; T. Gu et al., 2018;

105   Wang, Liu, et al., 2018; Wang, Wang, et al., 2018; Zheng et al., 2018) and eukaryotes

106   (K. Kim et al., 2017; Y. B. Kim et al., 2017; Komor, Kim, Packer, Zuris, & Liu, 2016;

107   Nishida et al., 2016; Rees et al., 2017; Y. Zhang et al., 2017; Zong et al., 2017), but no

108   deaminase was applied in *Clostridium* species.

109   In this study, we established a base editing tool (pCBEclos) in *C. beijerinckkii*

110   NCIMB 8052 by the fusion of Cas9^{D10A} nickase, cytidine deaminase (rat Apobec1) and

111   uracil DNA glycosylase inhibitor (UGI) which was able to efficiently convert specific

112   C·G nucleotide base pairs in the target window sequence to T·A. In its optimized form

113   (pCBclos-opt) it proved possible to rapidly generate mutants in four different genes,

114   namely *pyrE*, *xylR*, *spo0A* and *araR*. The system does not rely on HR, a process that is

115   intrinsically inefficient in clostridia such as *C. beijerinckii*. As a consequence, bulky

116   editing templates do not need to be included in the knock-out plasmids. This both

117   reduces plasmid size and makes their construction simpler, e.g., whereas the assembly

118   of pNICKclos2.0 requires six primers for the assembly of a typical knock-out plasmid,

119   pCBEclos-opt requires just two primers.

120   To our knowledge, this is the first report of the successful application of the

121   Cas9^{D10A} nickase and deaminase mediated base editing in *Clostridium*. It represents a

122   powerful new tool for genome editing in *C. beijerinckii,* which should be readily

123   applicable to other clostridial species.

124

2 Materials and Methods

2.1 Bacterial strains, media and reagents

The bacterial strains used in this study are listed in the Supporting information, Table S1. *Escherichia coli* DH5α was used for plasmid construction and maintenance. It was grown in LB medium at 37℃, supplemented where necessary with ampicillin (100 μg/ml). *C. beijerinckii* NCIMB 8052 was used as genome editing strain, it was grown in CGM medium at 37℃ in anaerobic chamber (Thermo Forma, Inc., Waltham, MA, USA). 20 μg/ml of erythromycin was supplemented as needed for plasmid selection. For *C. beijerinckii* NCIMB 8052 *pyrE* mutant, 20 μg/l uracil was required in CGM medium.

The DNA polymerase KOD plus Neo and KOD FX (Toyobo, Osaka, Japan) were used for high fidelity DNA amplification and colony PCR, respectively. All restriction enzymes used in this study were purchased from Thermo Fisher Scientific (USA). The plasmids used in this study were assembled by ClonExpress One Step Cloning Kit (Vazyme Biotech Co., Ltd, Nanjing, China). DNA purification and plasmids extracting were performed by kits purchased from Axygen (Hangzhou, China).

2.2 Plasmid construction

Cas9^{D10A} nickase and P$_{thl}$ were amplified from the plasmid pNICKclos2.0 (Li et al., 2016) by primers BE-P$_{thl}$-up/BE-P$_{thl}$-dn (apo-hm) and Cas9nclos-up/Cas9nclos-dn

144 respectively. The plasmid #73021 purchase from Addgene was used as the template to

145 amplify the Apobec1 and UGI gene by primers Apobec1-hm-up/Apobec1-hm-dn and

146 UGI-hm-up/UGI-hm-dn. The design guideline for sgRNA is as follows: 1. Choose 5'-

147 NGG-3' protospacer-adjacent motif (PAM); 2. The window area (typically from

148 positions 4 to 8 within the N20, counting the end distal PAM to the as position 1) must

149 containing C; 3. Base immediately 5' of the target C should be TC $\geq$ CC $\geq$ AC > GC

150 (Komor et al., 2016). The primers *cbei1006*-gRNA1-up1/BE-gRNA-dn were first used

151 to amplify the P$_{j23119}$-sgRNA-*pyrE* cassette from pNICKclos2.0 which was then used

152 as the template with primers *cbei1006*-gRNA1-up2/BE-gRNA-dn to produce the

153 overlapping extensions at the 5' ends of the P$_{j23119}$-sgRNA-*pyrE* cassette. Then, P$_{j23119}$-

154 sgRNA-*pyrE* cassette, P$_{thl}$, Apobec1, Cas9^{D10A} nickase and UGI were fused with

155 *Bam*HI/*Sma*I linearized pXY1 to generated plasmid pCBEclos-*cbei1006*-g1. Plasmids

156 pCBEclos-*cbei1006*-g2 and pCBEclos-*cbei1006*-g3 were derived from pCBEclos-

157 *cbei1006*-g1 by replacing the 20-bp target sequences. The construction of plasmid

158 pCBEclos-*cbei1006*-g2 has been shown here as an example. Fragment *cbei1006*-

159 gRNA2-A was amplified from plasmid pCBEclos-*cbei1006*-g1 by primers *cbei1006*-

160 gRNA2-up(A-up)/pBEclos-A-dn. Primers pBEclos-B-up/pBEclos-B-dn, pBEclos-C-

161 up/pBEclos-C-dn were used to amplify the fragments BEclos-B, BEclos-C from

162 pCBEclos-*cbei1006*-g1. *cbei1006*-gRNA2-A, BEclos-B and BEclos-C were assembled

163 together to yield plasmid pCBEclos-*cbei1006*-g2. Among them, the fragments BEclos-

164 B and BEclos-C were universal, only the fragment A (e.g. *cbei1006*-gRNA2-A) was

165 changed for each new plasmid (Fig. S1). For example, during the construction of

166 plasmid pCBEclos-*cbei1006*-g3, only the primers *cbei1006*-gRNA3-up (A-

167 up)/pBEclos-A-dn were used to amplified the fragment *cbei1006*-gRNA3-A, then this

168 fragment was fused with previously amplified fragments BEclos-B and BEclos-C to

169 generate the plasmid pCBEclos-*cbei1006*-g3.

170   Codon optimization of genes Apobec1 and UGI were performed by GenScript

171 Biotech Corp in Nanjing. Primers Apobec1-opt-up/Apobec1-opt-dn, UGI-opt-up/UGI-

172 opt-dn were adopted to amplify the optimized Apobec1 and UGI genes respectively.

173 Cas9^{D10A} nickase was amplified from the plasmid pNICKclos2.0 by primers Cas9nclos-

174 up (for opt)/Cas9nclos-dn (for opt). Apobec1, UGI and Cas9^{D10A} nickase amplified here

175 were fused with *Bam*H/*Sma*I linearized pCBEclos-*cbei1006*-g1 to generated plasmid

176 pCBEclos-*cbei1006*-g1-opt.

177   Plasmid pCBEclos-*cbei1006*-g2-opt, pCBEclos-*cbei1006*-g3-opt, pCBEclos-

178 *cbei4456*-opt, pCBEclos-*cbei2385*-g1-opt, pCBEclos-*cbei2385*-g2-opt (Addgene

179 deposits No. 118215) and pCBEclos-*cbei1712*-opt were derived from pCBEclos-

180 *cbei1006*-g1-opt by replacing the 20-bp target sequences. The construction process of

181 these plasmids was similar to the unoptimized pCBEclos series of plasmids. Here, only

182 the construction of plasmid pCBEclos-*cbei1712*-opt has been shown as an example.

183 Fragment *cbei1712*-gRNA-A was amplified from plasmid pCBEclos-*cbei1006*-g1-opt

184 by primers *cbei1712*-gRNA-up(A-up)/pBEclos-A-dn. Primers pBEclos-B-

185 up/pBEclos-B-dn, pBEclos-C-up/pBEclos-C-dn were used to amplify the fragments

186   BEclos-B-opt, BEclos-C-opt from pCBEclos-*cbei1006*-g1-opt. *cbei1712*-gRNA-A,

187   BEclos-B-opt and BEclos-C-opt were assembled together to yield plasmid pCBEclos-

188   *cbei1712*-opt. Similarly, the fragments BEclos-B-opt and BEclos-C-opt were universal

189   for constructing the optimized pCBEclos series of plasmids (Fig. S1).

190   **2.3 Electroporation and screening of mutant strains**

191   Plasmids were transformed into *C. beijerinckii* NCIMB 8052 using a previously

192   reported electroporation protocol (Mermelstein, Welker, Bennett, & Papoutsakis, 1992).

193   The recovered cells were spread on CGM agar supplemented with an appropriate

194   amount of erythromycin and incubated at 37℃ for approximately 2 days. The primers

195   listed in the Supplementary Table S2 were used for colony PCR, which was undertaken

196   when the transformants were visible on the CGM agar plates. Then, the PCR products

197   were extracted and sequenced to confirm the desired mutation events. For screening of

198   the *pyrE* mutants, the CGM medium was supplemented with 400 μg/l 5-fluroorotic acid

199   (5-FOA). Colony PCR was undertaken on a selection of random colonies growing on

200   the CGM agar containing 5-FOA, to confirm the expected mutation.

201   **2.4 Plasmid curing**

202   To eliminate the plasmids used in this study, mutants were first cultivated in 5 ml of

203   CGM medium without any antibiotic (T1). After growing for 12 h, 50 μl of the T1 broth

204   was used to inoculate 5 ml of fresh CGM medium and grown for 12 h until the OD_{600}

205   reached 0.8. The culture was diluted appropriately and aliquots of cells spread on a

206　nonselective CGM agar plate. The individual colonies were patch plated onto CGM

207　agar with and without erythromycin (20 μg/ml). The cells that grew on nonselective

208　medium, but were unable to grow on erythromycin CGM agar, were deemed to have

209　been cured of their plasmids.

210　**2.5 Fermentation and data analysis**

211　The fermentation of strains 8052WT, 8052*xylR*(TargeTron) (Xiao et al., 2012) and

212　8052*xylR*(BE)(*xylR* was disrupted by pCBEclos-opt) were performed anaerobically in

213　XHP_2 medium (Xiao et al., 2012) at 37℃ with xylose(60 g/l) as the carbon source for

214　72 h. 5 ml of liquid CGM was inoculated with single colony at 37 °C for about 12h,

215　then ~5% (v/v) of the inoculum was transferred into XHP2 medium for fermentation

216　when the optical density at 600 nm (OD_{600}) of the cells reached 0.8–1.0. The

217　concentrations of xylose were determined with high-performance liquid

218　chromatography (1200 series; Agilent), as described previously (Ren et al., 2010). Cell

219　density (OD_{600}) was measured using a DU730 spectrophotometer (Beckman Colter).

220

221　**3 Results**

222　**3.1 Establishment of CRISPR-Cas9^{D10A} nickase-mediated base editing system**

223　**pCBEclos-opt in *C. beijerinckii* NCIMB 8052**

224　　　To employ the deaminase-mediated base editing in *C. beijerinckii* NCIMB 8052,

225   we combined all functional components of the desired system into a single plasmid,

226   pCBEclos (Figure 1A). Transcription of the sgRNA was placed under the control of the

227   P_{j23119} promoter, and expression of the fusion protein of deaminase (rat Apobec1),

228   Cas9^{D10A} nickase and UGI under the control of the P_{thl} promoter. Cas9^{D10A} nickase

229   targets the non-edited strand and generates a nick, which promotes the use of the edited

230   strand as template for the repair of the nicked strand (Komor et al., 2016; Komor et al.,

231   2017). UGI suppresses excision of the uracil base generated by the cytosine deaminase

232   and accelerates mutagenesis (Banno et al., 2018; Komor et al., 2017) (Figure 1A). To

233   verify the desired mutation events generated via plasmid pCBEclos, DNA fragments

234   amplified by colony PCR of cells growing on counter selective media were subject to

235   Sanger sequencing for verifying the counter-selective genes; and colony PCR and

236   sequencing were directly performed from the transformants for non-selectable genes

237   (Figure 1B).

238   The *pyrE* gene (*cbei1006*) encoding orotate phosphoribosyltransferase was

239   selected as the first target gene in *C. beijerinckii* NCIMB 8052. Inactivation of the *pyrE*

240   gene leads to uracil auxotrophy and to resistance to the uracil analog 5-FOA (Ehsaan et

241   al., 2016; Tripathi et al., 2010), making such mutants readily distinguishable from wild

242   type cells. Accordingly, the plasmid pCBEclos-*cbei1006*-g1 carrying the spacer that

243   targets the *pyrE* gene was electroporated into *C. beijerinckii* NCIMB 8052 and a total

244   of 55 transformants from those obtained on CGM media supplemented with

245   erythromycin. To establish if any of these transformants were mutants, a total of 20

246 randomly selected colonies were subject to colony PCR and the amplified DNA

247 fragment subject to Sanger sequencing. All of the sequence reads obtained were wild

248 type. To ascertain whether mutant cells were present within the population, all of the

249 55 primary transformants were patch plated onto CGM agar media supplemented with

250 5-FOA. Of these, 22 were found to be resistant to 5-FOA. However, even after an

251 extended period of time, these colonies grew poorly (Figure 2A). Further screening of

252 a randomly selected 9 representatives of these 22 clones by Sanger sequencing of the

253 DNA fragment amplified by colony PCR indicate that all 9 contained the expected

254 mutational change (Figure 2B).

255 Our hypothesis to explain this observation is that the initial transformant colonies

256 are composed of a mixture of wild type and mutant cells in which the former vastly

257 predominate. The ratio of mutant to wild type is such that the mutant sequence cannot

258 be detected as it is swamped by the wild type. As an additional consequence, the

259 transformants grow poorly when initially plated on agar media containing 5-FOA as

260 they are predominately wild type cells, which are sensitive to this uracil analog.

261 Additionally, vector maintenance and constitutive expression of the codon un-

262 optimized fusion protein might also be responsible for reduced growth. Thus, our initial

263 base editing system (designated pCBEclos) although functional in *C. beijerinckii*

264 NCIMB 8052, was deemed relatively inefficient.

265 We also selected two other target sites (Fig. S2A) within the *pyrE* gene to test

266 pCBEclos. Similar to pCBEclos-*cbei1006*-g1, 20 out of 22 transformants obtained with

267   plasmid pCBEclos-*cbei1006*-g2 were found to grow on CGM agar containing 5-FOA,

268   albeit weakly. In contrast, no cells resistant to 5-FOA were obtained with plasmid

269   pCBEclos-*cbei1006*-g3 (Fig. S2B). 9 out of those 20 clones obtained from pCBEclos-

270   *cbei1006*-g2 which grew on 5-FOA medium were all shown by Sanger sequencing of

271   amplified PCR products, to contained the desired mutations (Fig. S2C). The results

272   indicated that the selection of different target sites on the same gene was not a fruitful

273   way to improve the efficiency of the initially established base editing plasmid

274   pCBEclos.

275   As the Addgene-derived Apobec1 and UGI genes used in the pCBEclos plasmid

276   system were optimized for expression in human cells, they may not be well expressed

277   in *Clostridium*. This could explain the poor efficiency of pCBEclos. Accordingly, we

278   elected to optimize the Apobec1 and UGI codons used based on *C. beijerinckii* NCIMB

279   8052 genome codon usage. The humanized components on plasmid pCBEclos-

280   *cbei1006*-g1 were thereafter replaced with the *Clostridium* optimized Apobec1 and

281   UGI genes to generate plasmid pCBEclos-*cbei1006*-g1-opt. Following the procedure

282   showed in Figure 1B, plasmid pCBEclos-*cbei1006*-g1-opt was electroporated into *C.*

283   *beijerinckii* NCIMB 8052 and transformed cells plated onto CGM agar supplemented

284   with erythromycin. The transformation frequencies obtained equated to 18.2 CFU/μg

285   DNA. Sanger sequencing of the amplified DNA obtained by colony PCR of six

286   randomly selected transformants revealed that three of them contained the desired

287   mutational changes. However, the reads obtained comprised a mixture of wild type and

288 mutant reads in the target region (Figure 3A). These cells were therefore re-streaked

289 once onto fresh CGM agar plates and two of single colonies tested again by Sanger

290 sequencing of the PCR amplified product. All of the purified colonies appeared to be

291 clean mutants with no detectable wild type sequence (Figure 3C). In parallel to the

292 above, 49 primary pCBEclos-*cbei1006*-g1-opt transformants were patch plated onto

293 CGM agar containing 5-FOA. On the basis of their growth, 46 out of the 49 colonies

294 were found to be resistant to 5-FOA. Moreover, in this case the growth observed was

295 vigorous, in contrast to the poor growth previously obtained when using the un-

296 optimized pCBEclos system (Figure 3B). The new base editing tool was designation

297 the pCBEclos-opt system. In contrast to pCBEclos, clones containing the desired C·G

298 to T·A mutations obtained simply by plating cells electroporated with the pCBEclos-

299 opt system onto CGM media containing erythromycin. The detection of the desired

300 mutants using the pCBEclos system requires subsequent screening of primary

301 transformants on selective media (Figure 3D). Moreover, the ratio of positive 5-FOA

302 resistant colonies was improved by about 2-folds via pCBEclos-opt system, compared

303 to the previous pCBEclos system (Figure 3D).

304   Successive rounds of base editing require that the initially used editing plasmid is

305 cured from the cell if an additional mutation is required. In order to test the efficiency

306 of plasmid curing, the edited *C. beijerinckii* NCIMB 8052 containing the expected *pyrE*

307 mutation was cultured in nonselective liquid CGM that was supplemented with

308 exogenous uracil (20 µg/l). After two subcultures, clonal populations were isolated by

309   plating to single colonies on nonselective CGM plates and these single colonies were

310   patch plated onto CGM agar with and without erythromycin supplementation. The

311   result showed that all 56 colonies could grow on the nonselective CGM medium, but

312   they were sensitive to erythromycin (Fig. S3). These data indicated that curing of

313   plasmid pCBEclos-*cbei1006*-g1 from the cells took place with 100% efficiency after

314   only two subcultures.

315   **3.2 Expansion of the pCBEclos-opt system to further genes in *C. beijerinckii***

316   **NCIMB 8052.**

317   After demonstrating the functionality of cytidine deaminase-based gene editing on

318   the *pyrE* gene, we further expanded the pCBEclos-opt system to other genes in *C.*

319   *beijerinckii* NCIMB 8052 (namely: araR or *cbei4456*, encoding a GntR family

320   transcriptional regulator; *xylR or cbei2385*, encoding the transcriptional regulator of

321   xylose metabolism; and *spo0A* or *cbei1712*, encoding response regulator receiver

322   protein).

323   Accordingly, *C. beijerinckii* NCIMB 8052 was transformed with plasmid

324   pCBEclos-*cbei4456*-opt encoding a sgRNA that targets *araR*. In this case all 3

325   transformants obtained harbored the desired C·G to T·A mutation. However, as with

326   *pyrE*, all three represented a mixed population composed of the wild type and desired

327   mutant (Figure 4A). One pure colony harboring the desired mutation could be isolated

328   after single-round re-streaking of one of the transformants (Figure 4B).

16

329   The plasmid pCBEclos-*cbei2385*-g1-opt targeting *xylR* was transformed in *C.*

330   *beijerinckii* NCIMB 8052 and yielded 3 transformants that were screened by colony

331   PCR and Sanger sequencing. The sequencing results showed that 2 transformants were

332   mixtures (Figure 4C), while the last colony was wild-type. A pure mutant could be

333   obtained by single-round re-streaking one of the mixed colonies on the CGM agar

334   (Figure 4D).

335   In the case of the *xylR* gene, further improvements in mutagenesis efficiency were

336   sought by changing the target sequence of pCBEclos-opt. Accordingly, the 20-bp spacer

337   on plasmid pCBEclos-*cbei2385*-g1-opt was replaced to yield pCBEclos-*cbei2385*-g2-

338   opt. The latter was found to be electroporated into 8052 with efficiency of 29.8 CFU/μg

339   DNA. Sanger sequencing of the colony PCR product of five randomly selected

340   transformants showed that one of them was a pure mutant (Figure 5A), one was a

341   mixture and the other three were wild type. As previously, re-streaking of a mixed clone

342   onto CGM agar and subsequent testing of individual single colonies easily allowed the

343   isolation of a pure mutant (Fig. S4A). Thus, in contrast to the previous plasmid targeting

344   *xylR*, pCBEclos-*cbei2385*-g1-opt, changing the target site to that present in pCBEclos-

345   *cbei2385*-g2-opt allowed the direct isolation of a pure mutant. As *xylR* was inactivated

346   via TargeTron previously (Xiao et al., 2012) and it was related to xylose consumption.

347   In order to test the fermentation phenotype of *xylR* mutant we obtained by pCBEclos-

348   opt, we fist cleared the plasmid pCBEclos-*cbei2385*-g2-opt with efficiency of 34/39 to

349   obtain the strain 8052*xylR*(BE) (Fig. S5A). Then, 8052WT, 8052*xylR*(BE) as well as

350 8052*xylR*(TargeTron) (Xiao et al., 2012) were cultured in XHP2 medium containing 60

351 g/l xylose for 72h. The results showed that the fermentation phenotype of 8052(BE)

352 was close to the 8052*xylR*(TargeTron), both mutants consumed 10% more xylose than

353 strain 8052WT (Fig. S5B).

354 Attempting to edit *spo0A*, plasmid pCBEclos-*cbei1712*-opt was introduced into

355 8052, and transformants were obtained with transformation of 110.6 CFU/μg DNA. 2

356 out of 6 random picked colonies were pure mutated colonies, 3 were mixed colonies,

357 and 1 was pure wild type (Figure 5B). Pure mutated strains could be isolated from all

358 3 mixed colonies (Fig. S4B).

359

360 **4 Discussion**

361 Genome editing tools based on CRISPR-Cas9 systems traditionally introduce a

362 DSB at a specific locus under the guidance of a sgRNA. During the repair of the DSB,

363 precise genome editing can be achieved in the presence of a donor DNA template by

364 exploiting the host's homologous recombination mechanisms. Based on this principle,

365 CRISPR-Cas9 mediated genome editing has been widely used in bacteria. However,

366 some bacteria have inefficient HR system and lack a functional non-homologous end

367 joining (NHEJ) repair pathway, which prevents the repair of Cas9-mediated DSBs and

368 results in cell death. Therefore, it is necessary to establish HR independent genome

369 editing tools in such bacteria. *C. bejerinckii* NCIMB 8052 is one of these bacteria

370 lacking an effective DSB repair pathway. One such HR independent tool available in

371 *C. bejerinckii* NCIMB 8052 is the group II intron-based gene inactivation, but it is as

372 precise as Cas9-mediated genome editing and it has polar effects.

373   In this study, we first established a CRISPR-mediated base editing tool pCBEclos

374 in *Clostridium* by the fusion of Apobec1, Cas9^{D10A} nickase and UGI. The conversion

375 of C·G to T·A at the target sites were realized via pCBEclos in *C. bejerinckii* NCIMB

376 8052. We initially established the pCBEclos system by directly applying Apobec1 and

377 UGI obtained from Addgene. However, the pCBEclos plasmid was inefficient and it

378 required selective medium to screen the edited strains, such as culturing the *pyrE*

379 mutants on 5-FOA plates. This pCBEclos system with poor efficiency is not suitable

380 for genes that do not exhibit a selectable phenotype. Fortunately, the base editing

381 efficiency was greatly improved after the optimization of Apobec1 and UGI, and the

382 desired mutants of *pyrE*、*xylR*、*spo0A* or *araR* could be directly detected in the

383 transformants of *C. bejerinckii* NCIMB 8052 via this optimized pCBEclos-opt system.

384 Furthermore, the loss of plasmid pCBEclos-*cbei1006*-g1-opt after gene editing was

385 achieved with efficiency of 100% after only two subcultures, allowing for successive

386 rounds of base editing. When mixed colonies of wild-type cells and mutants were

387 obtained, pure colonies harboring the desired mutation could be isolated by subsequent

388 re-streaking of the mixed colonies. Targeting a different locus within the *xylR* allowed

389 to isolate pure colonies of the desired genotype without the need of a re-streak. This

390 improvement in mutagenesis efficiency might reinforce the hypothesis of Komor et al

391 (2016), that the base immediately 5' and 3' of the target C may result in the different

392 editing efficiency.

393 The whole process of Cas9^{D10A} nickase-mediated base editing, including

394 electroporation, editing, identification and plasmid curing, only took five days. Unlike

395 the pNICKclos2.0 system we established previously, pCBEclos-opt does not rely on

396 homologous recombination, and as such DNA repair templates are not required when

397 using this system to edit gene. Therefore, the assembly of pCBEclos-opt is easier than

398 pNICKclos2.0, requiring only two primers instead of six. PCR amplification is

399 performed to obtain the part A that contains the new 20-bp target sequence, then this

400 part A is fused with the universal part B and C to generate the new plasmid (Fig. S1)

401 using ClonExpress One Step Cloning Kit. Its high genome editing efficiency of and the

402 simplicity of its assembly make pCBEclos-opt a useful genome editing tool in

403 *Clostridium*. If mutagenesis efficiency can be improved, a plasmid library of

404 pCBEclos-opt containing sgRNAs targeting each gene in *C. bejerinckii* NCIMB 8052

405 could be used to produce a mutant library that could be selected against a desired

406 phenotype.

407 In summary, this study is the first report that successfully applied Cas9^{D10A}

408 nickase-mediated base editing tool in *Clostridium*. A similar strategy would likely be

409 effective in other *Clostridium* strains. The base editing plasmid pCBEclos-opt we

410 established here will accelerate the metabolic engineering of *Clostridium* for the

411 optimization of chemicals and solvents in the future.

412

Acknowledgements

417

Supporting Information

Table S1: Plasmids and strains used in this study.

Table S2: Oligonucleotides used in this study.

Figure S1: Schematic for construction of pCBEclos series plasmids.

Figure S2: Target sites on *pyrE* gene were changed to test the pCBEclos system.

Figure S3: Clearance of plasmid pCBEclos-*cbei1006*-g1-opt.

Figure S4: Purification of the mixed *xylR* and *spo0A* mutants on plates.

Figure S5: Clearance of plasmid pCBEclos-*cbei2385*-g2-opt and the xylose

consumption of 8052WT, 8052xylR(TargeTron) and 8052xylR(BE).

427

References

Al-Hinai, M. A., Fast, A. G., & Papoutsakis, E. T. (2012). Novel system for efficient

430   isolation of *Clostridium* double-crossover allelic exchange mutants enabling

431   markerless chromosomal gene deletions and DNA integration. *Appl Environ*

432   *Microbiol, 78*(22), 8112-8121. doi:10.1128/AEM.02214-12

433 Banno, S., Nishida, K., Arazoe, T., Mitsunobu, H., & Kondo, A. (2018). Deaminase-

434   mediated multiplex genome editing in *Escherichia coli*. *Nat Microbiol*

435   doi:10.1038/s41564-017-0102-6

436 Chen, C. K., & Blaschek, H. P. (1999). Acetate enhances solvent production and

437   prevents degeneration in *Clostridium beijerinckii* BA101. *Appl Microbiol*

438   *Biotechnol, 52*(2), 170-173. doi:10.1007/s002530051504

439 Dürre, P. (1998). New insights and novel developments in clostridial

440   acetone/butanol/isopropanol fermentation. *Appl Microbiol Biotechnol, 49*(6),

441   639-648.

442 Ehsaan, M., Kuit, W., Zhang, Y., Cartman, S. T., Heap, J. T., Winzer, K., & Minton, N.

443   P. (2016). Mutant generation by allelic exchange and genome resequencing of

444   the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnol*

445   *Biofuels, 9*, 4. doi:10.1186/s13068-015-0410-0

446 Ezeji, T., Qureshi, N., & Blaschek, H. P. (2007). Butanol production from agricultural

447   residues: Impact of degradation products on *Clostridium beijerinckii* growth

448   and butanol fermentation. *Biotechnol Bioeng, 97*(6), 1460-1469.

449   doi:10.1002/bit.21373

450 Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D.

451   I., & Liu, D. R. (2017). Programmable base editing of A*T to G*C in genomic

452   DNA without DNA cleavage. *Nature, 551*(7681), 464-471.

453   doi:10.1038/nature24644

454   Gu, T., Zhao, S., Pi, Y., Chen, W., Chen, C., Liu, Q., . . . Ji, Q. (2018). Highly efficient

455   base editing in *Staphylococcus aureus* using an engineered CRISPR RNA-

456   guided cytidine deaminase. *Chem Sci, 9*(12), 3248-3253.

457   doi:10.1039/c8sc00637g

458   Gu, Y., Jiang, Y., Yang, S., & Jiang, W. (2014). Utilization of economical substrate-

459   derived carbohydrates by solventogenic clostridia: pathway dissection,

460   regulation and engineering. *Curr Opin Biotechnol, 29*, 124-131.

461   Heap, J. T., Kuehne, S. A., Ehsaan, M., Cartman, S. T., Cooksley, C. M., Scott, J. C., &

462   Minton, N. P. (2010). The ClosTron: Mutagenesis in Clostridium refined and

463   streamlined. *J Microbiol Methods, 80*(1), 49-55.

464   doi:10.1016/j.mimet.2009.10.018

465   Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P., & Minton, N. P. (2007). The

466   ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J*

467   *Microbiol Methods, 70*(3), 452-464. doi:10.1016/j.mimet.2007.05.021

468   Jiang, Y., Liu, J., Jiang, W., Yang, Y., & Yang, S. (2015). Current status and prospects

469   of industrial bio-production of n-butanol in China. *Biotechnol Adv, 33*(7), 1493-

470   1501. doi:10.1016/j.biotechadv.2014.10.007

471   Kim, K., Ryu, S.-M., Kim, S.-T., Baek, G., Kim, D., Lim, K., . . . Kim, J.-S. (2017).

472   Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol,*

473   *35*(5), 435-+. doi:10.1038/nbt.3816

474   Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., & Liu, D. R. (2017).

475   Increasing the genome-targeting scope and precision of base editing with

476   engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol, 35*(4), 371-+.

477   doi:10.1038/nbt.3803

478   Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016).

479   Programmable editing of a target base in genomic DNA without double-

480   stranded DNA cleavage. *Nature, 533*(7603), 420-+. doi:10.1038/nature17946

481   Komor, A. C., Zhao, K. T., Packer, M. S., Gaudelli, N. M., Waterbury, A. L., Koblan,

482   L. W., . . . Liu, D. R. (2017). Improved base excision repair inhibition and

483   bacteriophage Mu Gam protein yields C:G-to-T: A base editors with higher

484   efficiency and product purity. *Sci. Adv, 3*(8). doi:10.1126/sciadv.aao4774

485   Lee, S. Y., Park, J. H., Jang, S. H., Nielsen, L. K., Kim, J., & Jung, K. S. (2008).

486   Fermentative butanol production by Clostridia. *Biotechnol Bioeng, 101*(2), 209-

487   228. doi:10.1002/bit.22003

488   Li, Q., Chen, J., Minton, N. P., Zhang, Y., Wen, Z., Liu, J., . . . Yang, S. (2016). CRISPR-

489   based genome editing and expression control systems in *Clostridium*

490   *acetobutylicum* and *Clostridium beijerinckii. Biotechnol J, 11*(7), 961-972.

491   doi:10.1002/biot.201600053

492   Mermelstein, L. D., Welker, N. E., Bennett, G. N., & Papoutsakis, E. T. (1992).

Expression of cloned homologous fermentative genes in *Clostridium*
acetobutylicum ATCC 824. *Nat Biotechnol, 10*(2), 190-195.

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., . . . Kondo,
A. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate
adaptive immune systems. *Science, 353*(6305). doi:10.1126/science.aaf8729

Rees, H. A., Komor, A. C., Yeh, W.-H., Caetano-Lopes, J., Warman, M., Edge, A. S. B.,
& Liu, D. R. (2017). Improving the DNA specificity and applicability of base
editing through protein engineering and protein delivery. *Nat commun, 8*,
15790-15790. doi:10.1038/ncomms15790

Ren, C., Gu, Y., Hu, S., Wu, Y., Wang, P., Yang, Y., . . . Jiang, W. (2010). Identification
and inactivation of pleiotropic regulator CcpA to eliminate glucose repression
of xylose utilization in Clostridium acetobutylicum. *Metab Eng*, 12(5), 446-454.
doi: 10.1016/j.ymben.2010.05.002

Shao, L., Hu, S., Yang, Y., Gu, Y., Chen, J., Yang, Y., . . . Yang, S. (2007). Targeted
gene disruption by use of a group II intron (targetron) vector in *Clostridium
acetobutylicum*. *Cell Res, 17*(11), 963-965. doi:10.1038/cr.2007.91

Thakker, C., Martinez, I., Li, W., San, K. Y., & Bennett, G. N. (2015). Metabolic
engineering of carbon and redox flow in the production of small organic acids.
J Ind Microbiol Biotechnol, 42(3), 403-422. doi:10.1007/s10295-014-1560-y

Tripathi, S. A., Olson, D. G., Argyros, D. A., Miller, B. B., Barrett, T. F., Murphy, D.
M., . . . Caiazza, N. C. (2010). Development of *pyrF*-based genetic system for

514   targeted gene deletion in *Clostridium thermocellum* and creation of a pta mutant.

515   *Appl Environ Microbiol, 76*(19), 6591-6599. doi:10.1128/AEM.01484-10

516   Wang, Y., Liu, Y., Liu, J., Guo, Y., Fan, L., Ni, X., . . . Ma, Y. (2018). MACBETH:

517   Multiplex automated *Corynebacterium glutamicum* base editing method. *Metab*

518   *Eng, 47*, 200-210. doi:10.1016/j.ymben.2018.02.016

519   Wang, Y., Wang, S., Chen, W., Song, L., Zhang, Y., Shen, Z., . . . Ji, Q. (2018). Precise

520   and efficient genome editing in *Klebsiella pneumoniae* using CRISPR-Cas9 and

521   CRISPR-assisted cytidine deaminase. *Appl Environ Microbiol*.

522   doi:10.1128/aem.01834-18

523   Xiao, H., Li, Z., Jiang, Y., Yang, Y., Jiang, W., Gu, Y., & Yang, S. (2012). Metabolic

524   engineering of D-xylose pathway in Clostridium beijerinckii to optimize solvent

525   production from xylose mother liquid. *Metab Eng*, 14(5), 569-578. doi:

526   10.1016/j.ymben.2012.05.003

527   Zhang, N., Shao, L., Jiang, Y., Gu, Y., Li, Q., Liu, J., . . . Yang, S. (2015). I-SceI-

528   mediated scarless gene modification via allelic exchange in *Clostridium*. *J*

529   *Microbiol Methods, 108*, 49-60. doi:10.1016/j.mimet.2014.11.004

530   Zhang, Y., Qin, W., Lu, X., Xu, J., Huang, H., Bai, H., . . . Lin, S. (2017). Programmable

531   base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat*

532   *commun, 8*. doi:10.1038/s41467-017-00175-6

533   Zheng, K., Wang, Y., Li, N., Jiang, F.-F., Wu, C.-X., Liu, F., . . . Liu, Z.-F. (2018).

534   Highly efficient base editing in bacteria using a Cas9-cytidine deaminase fusion.

535   *Communications biology, 1*, 32-32. doi:10.1038/s42003-018-0035-5

536   Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., . . . Gao, C. (2017). Precise

537   base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion.

538   *Nat Biotechnol, 35*(5), 438-+. doi:10.1038/nbt.3811

539

540

541

542

543

544   **Figure legends**

545   **Figure 1: Base editing in *C. beijerinckii* NCIMB 8052.** (A) Design and strategy for

546   CRISPR-Cas9^{D10A} nickase-mediated C·G to T·A conversion in *C. beijerinckii*. Guide

547   RNA, Apobec1, Cas9^{D10A} nickase and UGI are all expressed on one plasmid pCBEclos.

548   (B) The procedure for confirming the mutant strain edited by pCBEclos in *C.*

549   *beijerinckii.* After assembly of pBEclos, it is electroporated into *C. beijerinckii* NCIMB

550   8052 and cells are plated on CGM plates with erythromycin to select for pBEclos. In

551   the case of *pyrE* mutagenesis, an additional selection step is carried out by patching

552   single colonies on CGM plates with erythromycin and 5-FOA. Colonies from both the

553   transformation plate and the mutant selection plate are finally screened by PCR

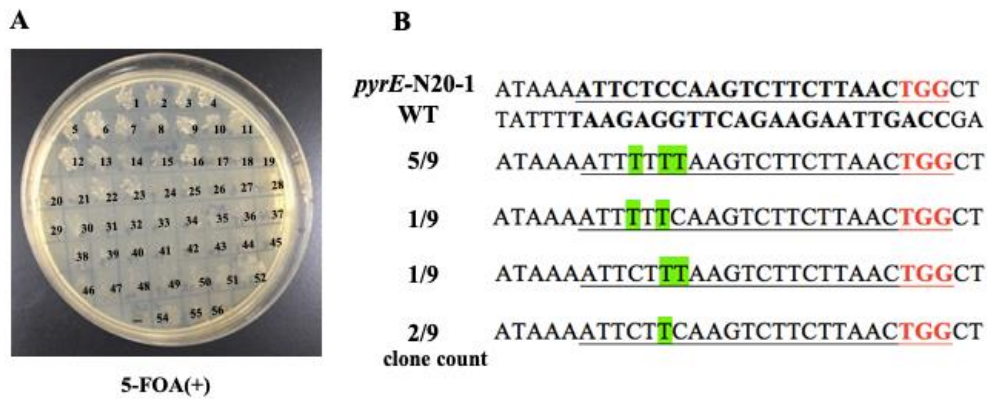554   amplification and subsequent Sanger sequencing. When revealed by the sequencing

results, mixed colonies are re-streaked on CGM plates with appropriate antibiotic for

556 isolation of pure colonies and re-sequencing.
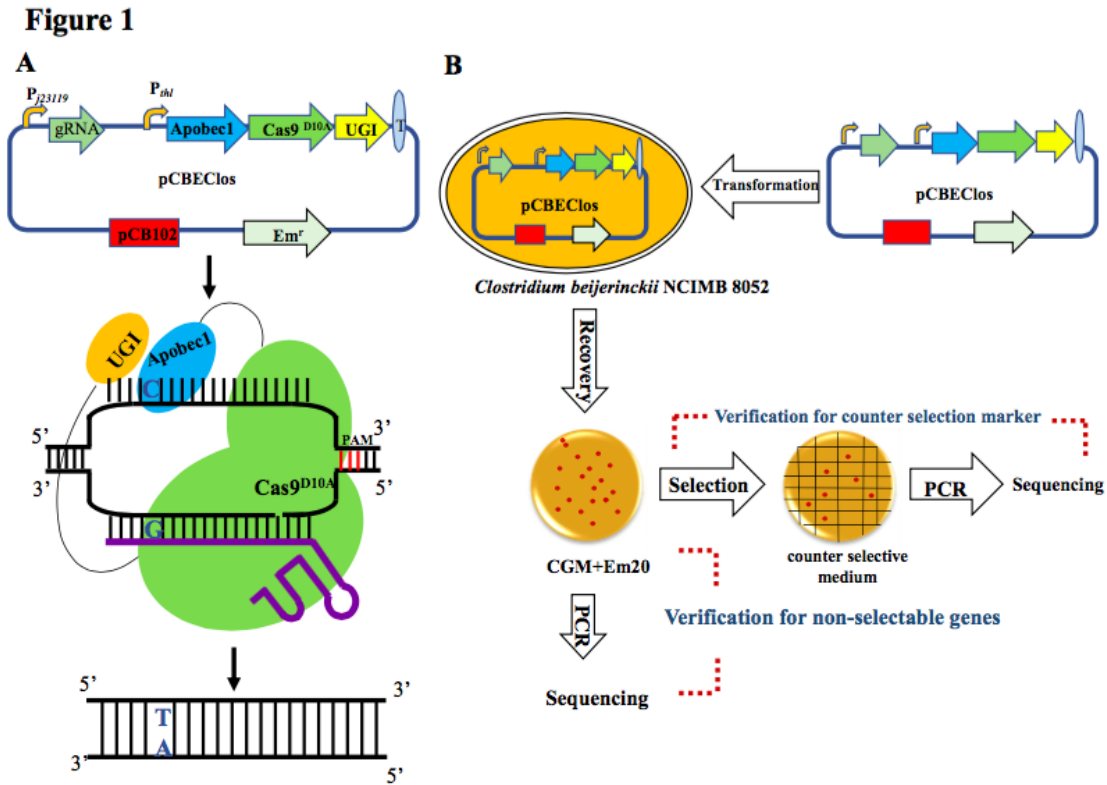
Figure 1

558 **Figure 2: Mutagenesis of *pyrE* gene in *C. beijerinckii* NCIMB 8052 via pCBEclos**

559 **system.** (A) *C. beijerinckii* were spread on CGM plates containing 5-FOA after

560 transformation with plasmid pCBEclos-*cbei1006*-g1, "-" represents the negative

561 control; (B) Sequence alignment of the *pyrE* mutants edited by pCBEclos system after

562 selection on 5-FOA plates. The bolded and underlined sequence is the targeted N20 site,

563 the red underlined is the PAM sequence, and the mutated nucleotides are highlighted in
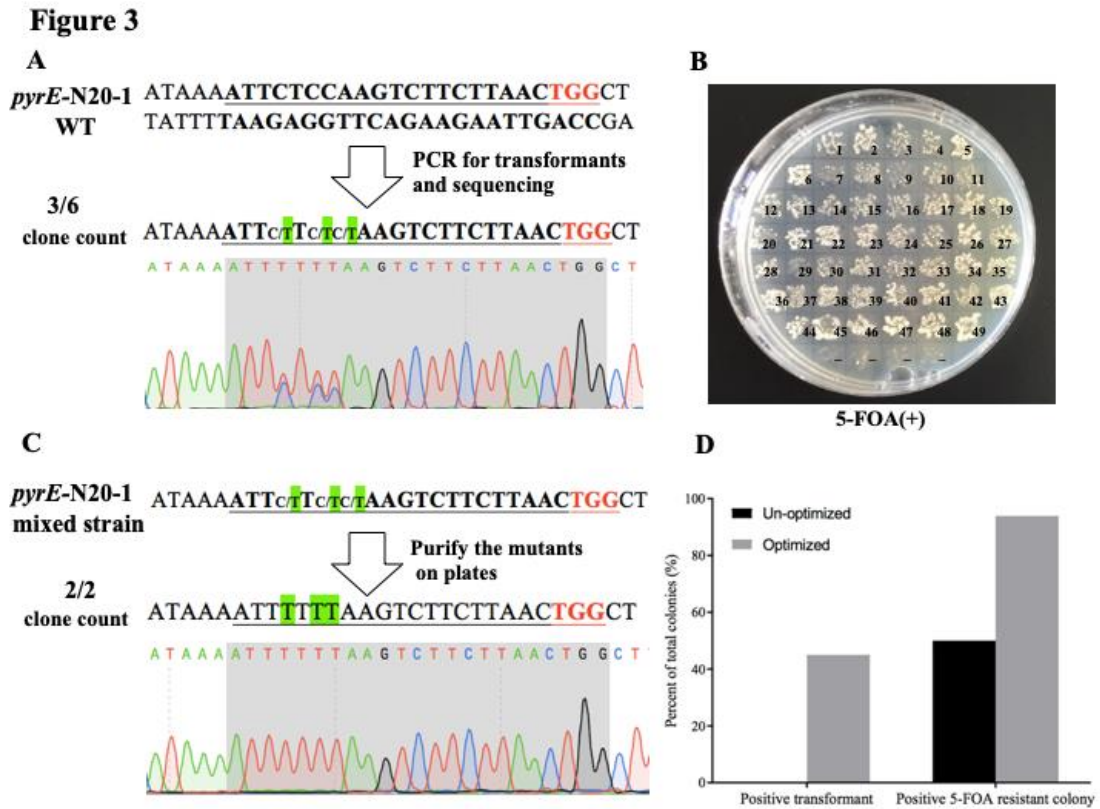
564 green.

Figure 2

A

5-FOA(+)

B

pyrE-N20-1	ATAAAATTCTCCAAGTCTTCTTAACTGGCT
WT	TATTTTAAGAGGTTCAGAAGAATTGACCGA
5/9	ATAAAATTTTTTAAGTCTTCTTAACTGGCT
1/9	ATAAAATTTTTCAAGTCTTCTTAACTGGCT
1/9	ATAAAATTCTTTAAGTCTTCTTAACTGGCT
2/9	ATAAAATTCTTCAAGTCTTCTTAACTGGCT
clone count	

Figure 3: Base editing efficiency was improved by pCBEclos-opt carrying the codon optimized Apobec1 and UGI. (A) Sequence alignment of the *pyrE* mutants obtained by transformation with pCBEclos-*cbei1006*-g1-opt; (B) *C. beijerinckii* carrying the plasmid pCBEclos-*cbei1006*-g1-opt were patched on plates containing 5-FOA, "-" represents the negative control; (C) Sequence alignment of the *pyrE* mutants isolated after re-streaking mixed colonies on CGM+ erythromycin plate; (D) The ratio of positive transformants and 5-FOA resistant strains was compared between the pCBEclos and pCBEclos-opt. The bolded and underlined sequence is the targeted N20 site, the red underlined is the PAM sequence, and the mutated nucleotides are highlighted in green.
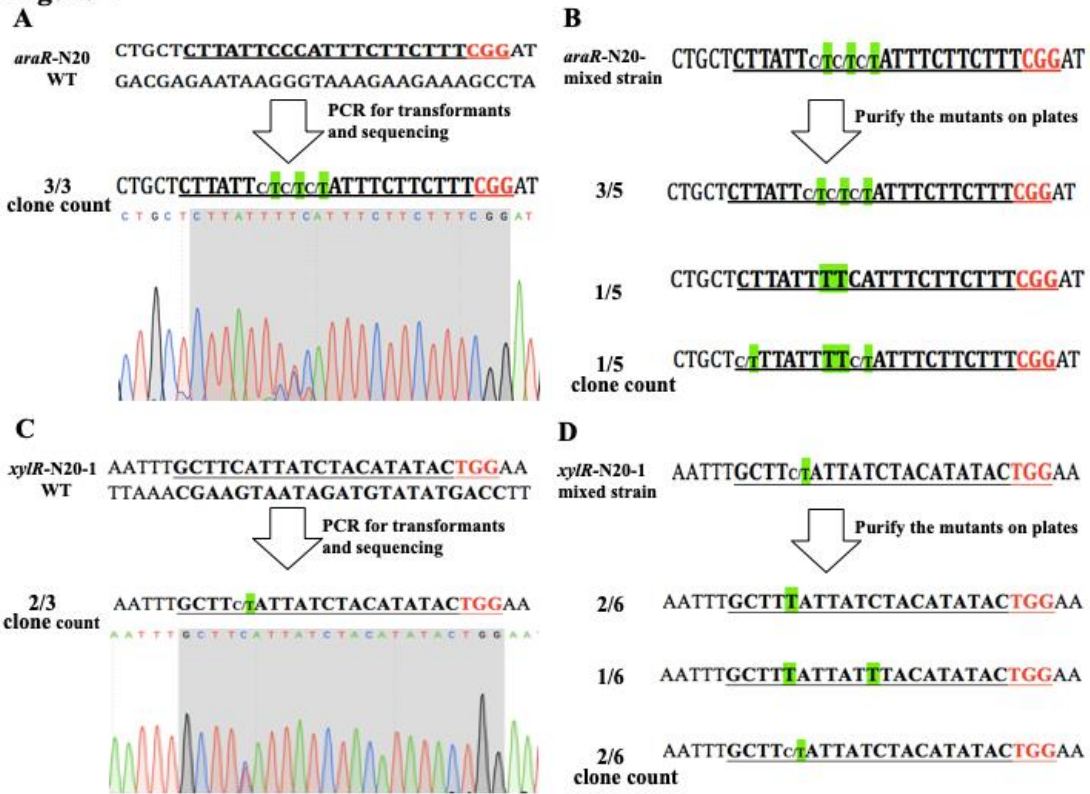
Figure 3

A

pyrE-N20-1
WT
ATAAAA**TTCTCCAAGTCTTCTTAAC****TGG**CT
TATTTT**AAGAGGTTCAGAAGAATTG**ACCGA

⇩ PCR for transformants and sequencing

3/6
clone count

ATAAAA**TTcTTcTcTAAGTCTTCTTAAC****TGG**CT

A T A A A A T T T T T A A G T C T T C T T A A C T G G C T

B

5-FOA(+)

C

pyrE-N20-1
mixed strain
ATAAAA**TTcTTcTcTAAGTCTTCTTAAC****TGG**CT

⇩ Purify the mutants on plates

2/2
clone count

ATAAAA**TTTTTAAGTCTTCTTAAC****TGG**CT

A T A A A A T T T T T A A G T C T T C T T A A C T G G C T

D

Percent of total colonies (%)

■ Un-optimized
▨ Optimized

Positive transformant   Positive 5-FOA resistant colony

577

Figure 4: Mutagenesis of gene *araR* and *xylR* in *C. beijerinckii* NCIMB 8052 via pCBEclos-opt system. (A) Sequence alignment of the *araR* mutants obtained by transformation with pCBEclos-*cbei3835*-g1-opt; (B) The pure *araR* mutant was obtained after single-round streaking the mixed colony on plate. (C) Sequence alignment of the *xylR* mutants obtained by transformation with pCBEclos-*cbei2385*-g1-opt. (D) The pure *xylR* mutant was isolated by single-round streaking a mixed colony on plate. The bolded and underlined sequence is the targeted N20 site, the red underlined is the PAM sequence, and the mutated nucleotides are highlighted in green.

30

Figure 4

Figure 5: Pure mutants were obtained directly from the transformants of *C. beijerinckii* NCIMB 8052 via plasmid pCBEclos-opt. (A) Sequence alignment of the *xylR* pure mutant obt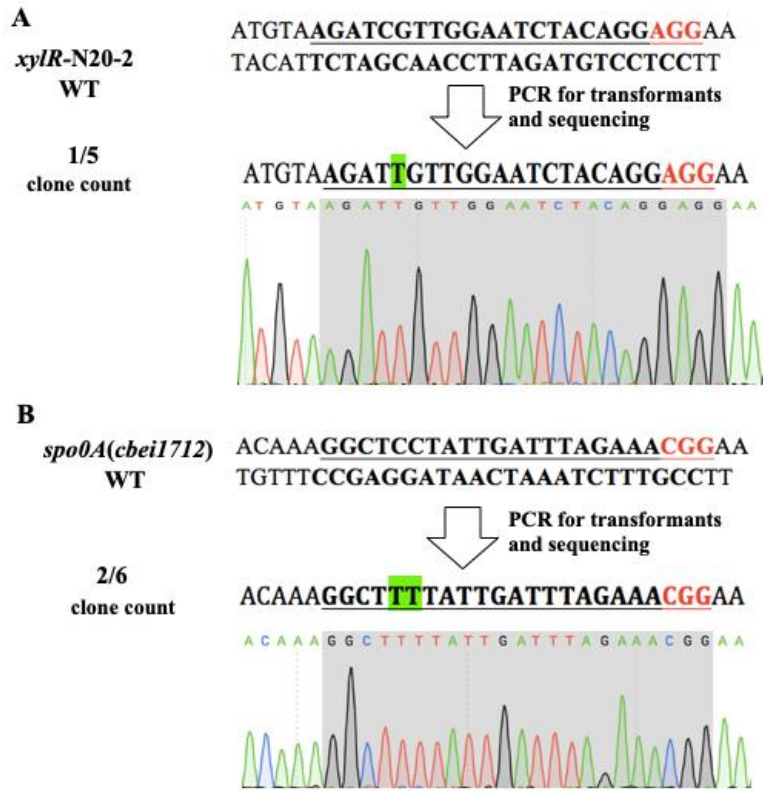ained by transformation with pCBEclos-*cbei2385*-g2-opt; (B) Sequence alignment of the *spo0A* pure mutant obtained by transformation with pCBEclos-*spo0A*-g2-opt. The bolded and underlined sequence is the targeted N20 site, the red underlined is the PAM sequence, and the mutated nucleotides are highlighted in green.

31

Figure 5

A

xylR-N20-2
WT

ATGTA**AGATCGTTGGAATCTACAGG****AGG**AA
TACATTCTAGCAACCTTAGATGTCCTCCTT

⬇ PCR for transformants and sequencing

1/5
clone count

ATGTA**AGATTGTTGGAATCTACAGG****AGG**AA

B

spo0A(*cbei1712*)
WT

ACAAA**GGCTCCTATTGATTTAGAAA****CGG**AA
TGTTTCCGAGGATAACTAAATCTTTGCCTT

⬇ PCR for transformants and sequencing

2/6
clone count

ACAAA**GGCTTTTATTGATTTAGAAA****CGG**AA

594

595

596

597

598

599

600

601

602

603

604

605

32

Supporting information, Table S1: Plasmids and strains used in this study.

Strains or plasmids	Description	Source or reference
Strains		
C. beijerinckii NCIMB 8052	Wild type	NCIMB
E. coli DH5α	Commercial transformation host	GIBCO BRL, Life Technologies
Plasmids		
pXY1	pCB102, MLS^R, P$_{thl}$ promotor,ColE1 origin , Amp^R, *E.coli-Clostridium* shuttle vector	This study
pCBEclos-*cbei1006*-g1	Derived from pXY1-Cas9n, pJ23119-sgRNA1-*cbei1006*, P$_{thl}$-rAPOBEC1-XTEN Cas9n-UGI	This study
pCBEclos-*cbei1006*-g2	Derived from pCBEclos-*cbei1006*-g1, pJ23119-sgRNA2-*cbei1006*	This study
pCBEclos-*cbei1006*-g3	Derived from pCBEclos-*cbei1006*-g1, pJ23119-sgRNA3-*cbei1006*	This study
pCBEclos-*cbei1006*-g1-opt	Derived from pCBEclos-*cbei1006*-g1, P$_{thl}$-rAPOBEC1(optimized)-XTEN Cas9n-UGI(optimized)	This study
pCBEclos-*cbei4456*-g1-opt	Derived from pCBEclos-*cbei1006*-g1-opt, P$_{thl}$-rAPOBEC1(optimized)-XTEN Cas9n-UGI(optimized)	This study
pCBEclos-*cbei2385*-g1-opt	Derived from pCBEclos-*cbei1006*-g1-opt, P$_{thl}$-rAPOBEC1(optimized)-XTEN Cas9n-UGI(optimized)	This study
pCBEclos-*cbei2385*-g2-opt	Derived from pCBEclos-*cbei1006*-g1, P$_{thl}$-rAPOBEC1(optimized)-XTEN Cas9n-UGI(optimized)	This study
pCBEclos-*cbei1712*-opt	Derived from pCBEclos-*cbei1006*-g1-opt, pJ23119-sgRNA-*cbei1712*	This study

607

608

609

Supporting information, Table S2: Oligonucleotides used in this study.

Oligos	Sequence (5'→3')
cbei1006-gRNA1-up1	aggtataatactagt**attctccaagtcttcttaac**gttttagagctagaaatagcaagt
cbei1006-gRNA1-up2	aatgtgctgcattgacagctagctcagtcctaggtataatactagtattctccaagtct
BE-gRNA-dn	ctattattattttatcaatatatttgttaaaaaggtaccaaaaaagcaccgactcgg
BE-P$_{thl}$-up	agtcggtgcttttttggtacctttttaacaaaatatattgataaaaataataatagtgg
BE-P$_{thl}$-dn(apo-hm)	gccactgggccagtctctgagctcatggatcctctaactaacctcctaaattttgatac
Apobec1-hm-up	gttaccccgtatcaaaatttaggaggttagttagaggatccatgagctcagagactggc
Apobec1-hm-dn	cgctatttgtgccgatagctaagcctattgagtatttcttatcactttcgggtgtggcg
Cas9nclos-up	cccgggacctcagagtccgccacacccgaaagtgataagaaatactcaataggcttagc
Cas9nclos-dn	tccttttcaataatatctgacagattagtagaaccaccagagtcacctcctagctgact
UGI-hm-up	tgaaacacgcattgatttgagtcagctaggaggtgactctggtggttctactaatctgt
UGI-hm-dn	gtcacgacgttgtaaaacgacggccagtgaattcccgggttaagaaccaccagagagca

cbei1006-gRNA2-up(A-up)	attgacagctagctcagtcctaggtataatactagt**aacttccgccattgtaacta**gttttagagctagaaatagcaag
cbei1006-gRNA3-up(A-up)	attgacagctagctcagtcctaggtataatactagt**ttgtgccatagttacaatgg**gttttagagctagaaatagcaag
pBEclos-A-dn	ttgactacttcttcacttgga
pBEclos-B-up	gttctgataaaaatcgtggtaaa
pBEclos-B-dn	atcctttgatcttttctacgg
pBEclos-C-up	taacgtgagttttcgttcca
pBEclos-C-dn	actagtattatacctaggactgag
Apobec1-opt-up	ttaccccgtatcaaaatttaggaggttagttagaggatccatgtcaagtgaaacaggac
Apobec1-opt-dn	tatttgtgccgatagctaagcctattgagtatttcttatcagattcaggagttgcagat
Cas9nclos-up(for opt)	ccaggaacatcagaatctgcaactcctgaatctgataagaaatactcaataggcttagc
Cas9nclos-dn(for opt)	tcttttctattatatctgaaagatttgttgatcctccactgtcacctcctagctgact
UGI-opt-up	tttatgaaacacgcattgatttgagtcagctaggaggtgacagtggaggatcaacaaat
UGI-opt-dn	cgacgttgtaaaacgacggccagtgaattcccgggttatgatcctccagataacatttt
cbe44565-gRNA1-up(A-up)	attgacagctagctcagtcctaggtataatactagt**cttattcccatttcttcttt**gttttagagctagaaatagcaag
cbei2385-gRNA1-up(A-up)	attgacagctagctcagtcctaggtataatactagt**gcttcattatctacatatac**gttttagagctagaaatagcaag
cbei2385-gRNA2-up(A-up)	attgacagctagctcagtcctaggtataatactagt**agatcgttggaatctacagg**gttttagagctagaaatagcaag
cbei2385-verf-up	ttgatagaagtaaatcacagtaaaataaaag
cbei2385-verf-dn	gaagcatacacatctatgaattctc
cbei1712-gRNA-up(A-up)	attgacagctagctcagtcctaggtataatactagt**ggctcctattgatttagaaa**gttttagagctagaaatagcaag
cbei4456-verf-up	gggttacataaaggccct
cbei4456-verf-dn	ttaaactctagaacaagaatctctaaca
cbei1006-verf-up	acgagattataggaataatataaattgatc
cbei1006-verf-dn	tcacagtcctgagaaacatatat
cbei1712-verf-up	atacaatgcaattggaaaaggt
cbei1712-verf-dn	atttgttggcttacctttatcat

611   The bolded and underlined sequence represent the target sites used in base editing.
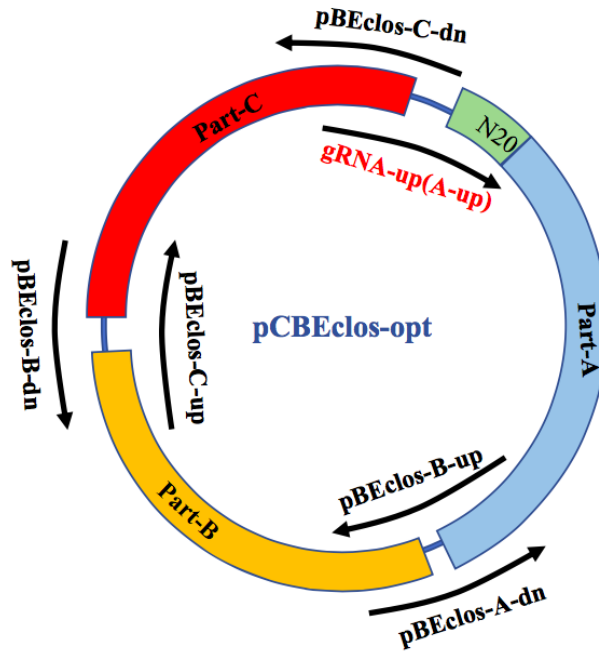
612

613

614

615   **Supporting information, Figure. S1: Schematic for construction of pCBEclos**

616   **series plasmids.**

34

617

618   Primers pBEclos-B-up/pBEclos-B-dn, pBEclos-C-up/pBEclos-C-dn are used to

619   amplify the universal fragments BEclos-B-opt(part B), BEclos-C-opt(part C). Only the

620   part A is changed for the construction of new plasmid by the primer gRNA-up(A-up)

621   and the universal primer pBEclos-A-dn. Then, part A, part B and part C are assembled

622   to generated the new base editing plasmid.

623

624   **Supporting information, Figure. S2: Target sites on *pyrE* gene were changed to**

625   **test the pCBEclos system.**
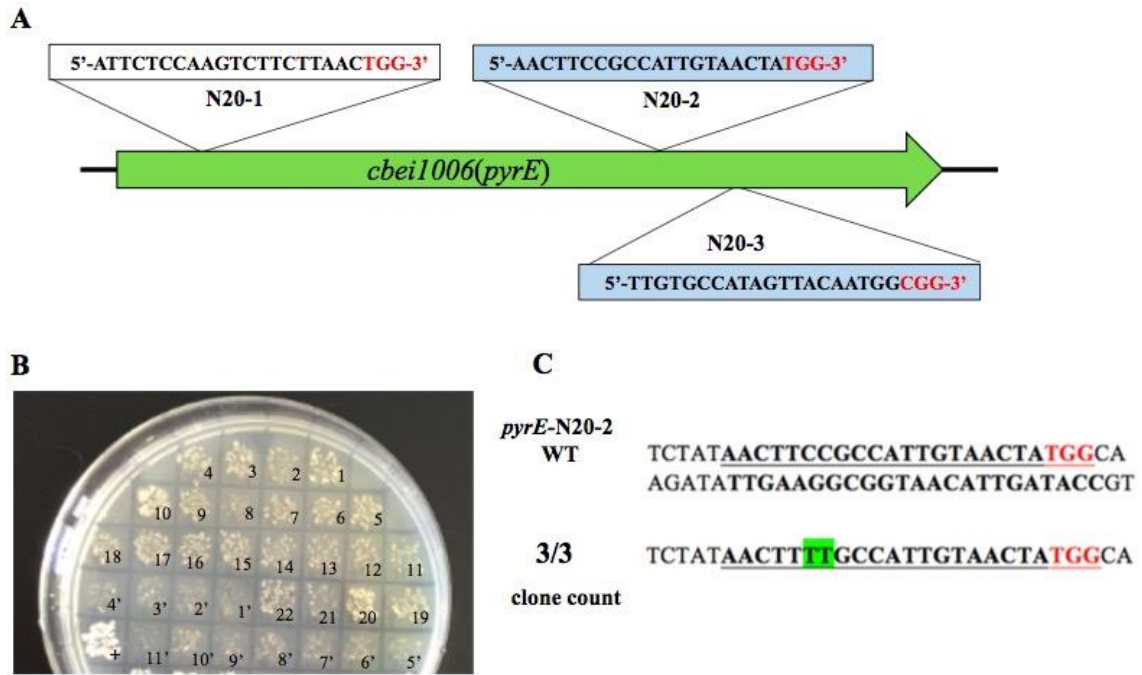
626

627 (**A**) The sequence of target sites in the *pyrE* gene; (B) *C. beijerinckii* were spread on

628 plates containing 5-FOA after transformation with plasmid pCBEclos-*cbei1006*-g2 and

629 pCBEclos-*cbei1006*-g3; Strains 1-22 are the transformants of plasmid pCBEclos-

630 *cbei1006*-g2, while strains 1'-11' are the transformants of plasmid pCBEclos-*cbei1006*-

631 g3; "+" represents the positive control; (C) Sequence alignment of the *pyrE* mutants

632 edited by pCBEclos system after selection on 5-FOA plates. The bolded and underlined

633 sequence is the targeted N20 site, the red underlined is the PAM sequence, and the

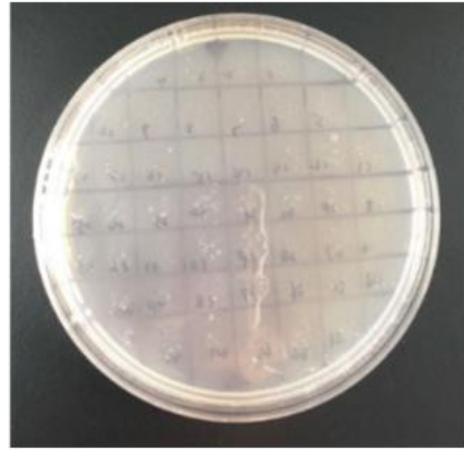634 mutated nucleotides are highlighted in green.

635

636 **Supporting information, Figure. S3: Clearance of plasmid pCBEclos-*cbei1006*-g1-**

637 **opt.**
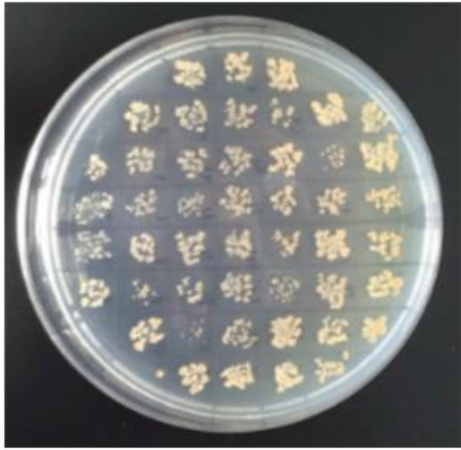
CGM+uracil   **CGM+uracil+Em20**
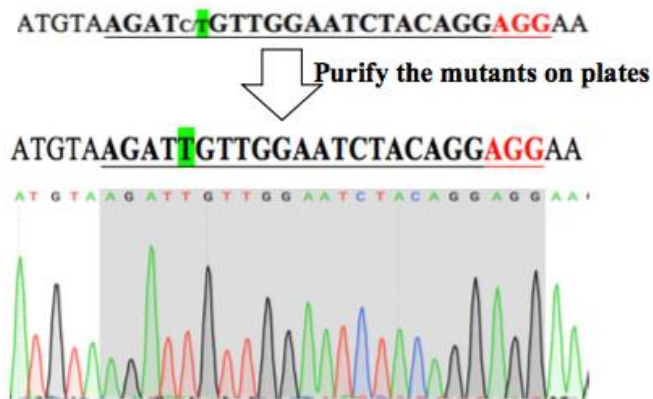
638

639

640

641

642 **Supporting information, Figure. S4: Purification of the mixed *xylR* and *spo0A***

643 **mutants on plates.**
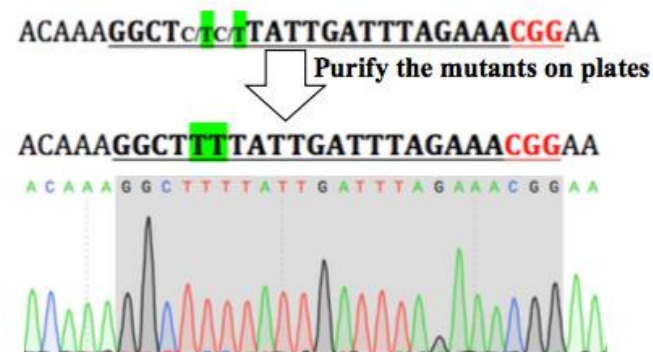
A
xylR-N20-2
mixed strain

ATGTA**AGAT**c/**T**GTTGGAATCTACAGG**AGG**AA

Purify the mutants on plates

3/3
clone count

ATGTA**AGATTGTTGGAATCTACAGGAGG**AA

B
spo0A(cbei1712)
mixed strain

ACAAA**GGCT**c/**TC/T**TATTGATTTAGAAA**CGG**AA

Purify the mutants on plates

2/2
clone count

ACAAA**GGCTTT**TATTGATTTAGAAA**CGG**AA

644

645 (A) Sequence alignment of the pure *xylR* mutants after streaking the mixed strain on

646 plate. (B) Sequence alignment of the pure *spo0A* mutants after streaking the mixed
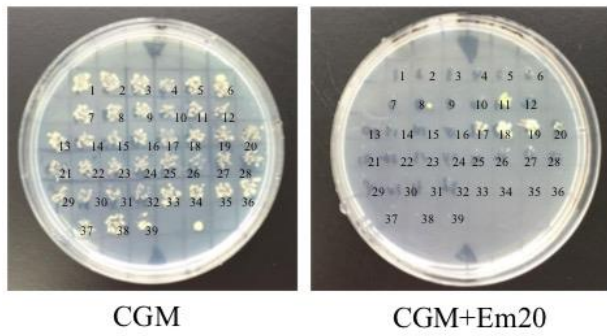
647 strain on plate.

648

649

650 **Supporting information, Figure. S5: Clearance of plasmid pCBEclos-*cbei2385*-g2-**
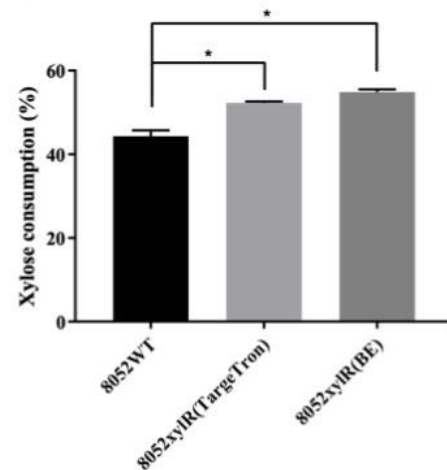
651 **opt and the xylose consumption of 8052WT, 8052*xylR*(TargeTron) and**

652 **8052*xylR*(BE).**

653

654   (A) Clearance of plasmid pCBEclos-*cbei2385*-g2-opt. (B) Xylose consumption of

655   strains 8052WT, 8052*xylR*(TargeTron) (*xylR* was disrupted by TargeTron technology)

656   and 8052*xylR*(BE) (*xylR* was disrupted by pCBE-opt) in XHP2 medium containing 60

657   g/l D-xylose. Samples were taken after 72 h of fermentation. Fermentations were

658   performed in triplicate.

659

660

661

662

663

664

665