1   *Classification*: Biological Sciences/ Applied Biological Sciences

2   **CRISPR-Cas9-assisted native end-joining editing offers a simple strategy for**

3   **efficient genetic engineering in *Escherichia coli***

4   Chaoyong Huang[a], Tingting Ding[a], Jingge Wang[a], Xueqin Wang[a], Jialei Wang[a], Lin

5   Zhu[a], Changhao Bi[b], Xueli Zhang[b], Xiaoyan Ma[a], Yi-Xin Huo[a,c,1]

6   [a]Key Laboratory of Molecular Medicine and Biotherapy, School of Life Sciences,

7   Beijing Institute of Technology, No. 5 South Zhongguancun Street, Beijing, PR China,

8   100081

9   [b]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin,

10   PR China, 300308

11   [c]UCLA Institute of Advancement (Suzhou), 10 Yueliangwan Road, Suzhou Industrial

12   Park, Suzhou, PR China, 215123

13   [1]To whom correspondence should be addressed. Email: huoyixin@bit.edu.cn. Phone

14   number: +86 18910254522

15   Emails of coauthors:

16   Chaoyong Huang: 2120171376@bit.edu.cn

17   Tingting Ding: 3120185701@bit.edu.cn

18   Jingge Wang: 3120181362@bit.edu.cn

19   Xueqin Wang: wxqin@bit.edu.cn

20   Jialei Wang: 3120181361@bit.edu.cn

21   Lin Zhu: 3120181373@bit.edu.cn

22   Changhao Bi: bi_ch@tib.cas.cn

23   Xueli Zhang: zhang_xl@tib.cas.cn

24   Xiaoyan Ma: xyma@bit.edu.cn

25

Abstract

Unlike eukaryotes, bacteria are less proficient in homologous recombination (HR) and non-homologous end joining (NHEJ). All existing genomic editing methods for *Escherichia coli* rely on exogenous HR or NHEJ systems to repair DNA double-strand breaks (DSBs). Although an *E. coli* native end-joining (ENEJ) system has been reported, its potential in chromosomal engineering has not yet been explored. Here, we present a CRISPR-Cas9-assisted native end-joining editing and show that ENEJ-dependent DNA repair can be used to conduct rapid and efficient knocking-out of *E. coli* genomic sequence of up to 83 kb. Moreover, the positive rate and editing efficiency is independent of high-efficiency competent cells. The method requires neither exogenous DNA repair systems nor introduced editing template. The Cas9 complex is the only foreign element in this method. This study is the first successful engineering effort to utilize ENEJ mechanism in genomic editing and provides an effective strategy for genetic engineering in bacteria that are inefficient in HR and NHEJ.

Keywords: *Escherichia coli*, genetic engineering, CRISPR/Cas9, end-joining, Large fragment deletion

Significance

The application in prokaryotes is difficult because of the weak homologous recombination and non-homologous end joining systems. *E. coli*, as the most-used prokaryote in metabolic engineering, has no NHEJ system. All existing genomic editing methods for *E. coli* rely on exogenous HR or NHEJ systems to repair double-strand breaks introduced by CRISPR/Cas9. In this report, we firstly demonstrate that the weak and previously ignored end-joining mechanism in *E. coli* can be used for efficient large-scale genetic engineering assisted by CRISPR/Cas9. Our efforts greatly simplify the genomic editing procedure of *E. coli* and provide an effective strategy for genetic engineering in bacteria that are inefficient in HR and NHEJ.

Introduction

53 The ability to easily and efficiently edit the bacterial genome is highly desirable in

54 metabolic engineering to produce valuable chemicals (1-4). Creating double strand

55 breaks (DSBs) is a key step in driving genomic engineering (5). Novel applications

56 of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-

57 associated protein 9 (Cas9) can generate DSBs at desired genomic loci (6,7). Under

58 the guidance of a trans-activating crRNA (tracrRNA):crRNA duplex or an

59 engineered single-guide RNA (sgRNA), the Cas9 protein recognizes and cleaves a

60 target DNA sequence with a required protospacer adjacent motif (PAM) (6,8). To

61 survive, it is necessary for bacteria to repair the DSBs via homologous recombination

62 (HR) or non-homologous end joining (NHEJ). Genomic editing events occur during

63 DSBs repair.

64   NHEJ is the prevalent DSB repair mechanism. It joins unrelated DNA ends and

65 creates new genomic combinations between sequences that do not share homology

66 (9-11). Unlike eukaryotes, bacteria are less proficient in or incapable of NHEJ (12).

67 Some bacteria, including phylogenetically distant *Mycobacteria* and *Bacillus subtilis*,

68 possess a rudimentary NHEJ machinery consisting of two proteins, namely the

69 multifunctional ATP-dependent Ligase-D and Ku protein that collectively protect,

70 process, and ligate DNA ends (13,14). Other bacteria, such as the prevalent gut

71 commensal *Escherichia coli* where Ligase-D-like and Ku-like proteins have not been

72 found are generally accepted to be NHEJ-free (12) and rely only on HR to repair

73 DSBs. *E. coli* is frequently used as a negative control for NHEJ experiments in other

74 bacteria (15).

75   Generally, the artificial DSBs in *E. coli* must be repaired by RecA-mediated HR

76 with a homologous sequence as the editing template (6). However, the native HR

77 pathway is commonly recognized to be of low efficiency in repairing CRISPR-Cas9-

78 introduced DSBs. The introduction of bacteriophage-derived λ-Red recombinases

79 made a breakthrough in the efficiency of making desired mutations in the genome

3

80   (6,16,17). Recently, *Su et al.* and *Zheng et al.* introduced a *Mycobacteria*-derived

81   NHEJ pathway into *E. coli* to repair the DSBs generated by CRISPR-Cas9 at desired

82   genomic loci and demonstrated the rapid and efficient inactivation of bacterial genes

83   in a HR-independent manner (1,18). According to the views of *Su et al.* and *Zheng*

84   *et al.*, *E. coli* acquired end-joining activity after overexpressing *Mycobacteria*-

85   derived Ligase-D and Ku protein. Taken together, all existing genomic editing

86   methods of *E. coli* rely on HR pathways, generally assisted by λ-Red recombinases,

87   or exogenous NHEJ pathways to repair DSBs generated by CRISPR-Cas9.

88   A Ligase-D- and Ku-independent *E. coli* native end-joining (ENEJ) mechanism

89   differs from classic NHEJ mechanism has been reported in *E. coli* and was named as

90   alternative end-joining (A-EJ) (11). In that study, DSBs generated by restriction

91   endonuclease I-SceI were repaired by A-EJ with an efficiency of 10^{-5}. However, we

92   have not seen any report on the application of ENEJ in genetic engineering. To fill

93   in the gap and add to the toolbox of genetic engineers, we developed C̲RISPR-Cas9-

94   assisted n̲ative e̲nd-joining e̲diting (CNEE), in which the ENEJ was utilized to repair

95   DSBs generated by CRISPR-Cas9. Using a systematically optimized protocol, we

96   demonstrated that ENEJ could be used to enable the rapid and efficient deletion of *E.*

97   *coli* genomic sequence up to 83 kb in length. The method requires neither editing

98   template and λ-Red recombinases nor exogenous NHEJ associated proteins, thus

99   realizing the *E. coli* genomic editing with the lowest number of foreign elements.

100   The method has a similar, if not higher, editing efficiency to previous methods that

101   overexpress exogenous DNA repair proteins. Moreover, the method does not require

102   high-competency host strains, greatly decreasing the difficulty of genomic editing of

103   wild-type strains.

104   We hypothesized that the strategies used in our method might be useful for other

105   CRISPR-Cas9-mediated genomic editing methods in which DSBs are repaired by

106   RecA-mediated HR, λ-Red-mediated HR or NHEJ. To verify that, we applied the

107   similar strategies in these methods. After modification, the efficiencies of these

108   methods significantly increased.

109   Taken together, this work is the first to utilize ENEJ in genomic editing in *E. coli*

110   and demonstrates that genomic editing of prokaryotes can be independent of exogenous

111   DNA repair systems.

112

113 **Results**

114 **All existing *E. coli* genomic editing methods are dependent on the exogenous DNA**

115 **repair system**

116 In *E. coli*, CRISPR-Cas9-introduced DSBs lead to cell death unless repaired in a timely

117 manner by DNA repair pathways (19) such as native RecA-mediated HR or A-EJ (11).

118 CRISPR-Cas9-based genomic editing methods achieve genetic modifications by

119 introducing DSBs in the genome and repairing it via DNA repair pathways. In all

120 existing methods, bacteriophage-derived λ-Red system, a HR system, or

121 *Mycobacteria*-derived NHEJ system is introduced into *E. coli* to increase the

122 efficiency of HR or end-joining.

123 To compare the performances of existing *E. coli* genomic editing methods, we

124 tested all reported methods for *lacZ* inactivation in the wild-type MG1655 [MG1655

125 (WT)] strain (Table 1). These methods rely on either λ-Red mediated-HR or NHEJ to

126 repair CRISPR-Cas9-induced DSBs. When using HR-dependent methods, a 20-bp-

127 designed DNA sequence was inserted into the open reading frame (ORF) of the *lacZ*

128 gene to generate a frameshift mutation. When using NHEJ-dependent protocols, the

129 stochastic length of the DNA sequence of the ORF was deleted, thus destroying the

130 *lacZ* gene (Fig. 1a). Edited colonies were verified via blue-white screening, colony PCR,

131 and sequencing. The edited colonies were white in the Luria-Bertani (LB) plate

132 containing IPTG and X-gal, while the WT colonies were blue.

133 Six λ-Red-dependent methods (methods 1−6) were tested (Table 1). The

134 procedures of these six methods are summarized in Fig. S1−S6, respectively. The

135 results showed that five out of six methods achieved a high positive rate of > 65% (Fig.

136 1c), and the editing efficiency ranged from 4.6×10^{-6} to 2.3×10^{-4} (Table 1, column 2).

137 These findings indicated that CRISPR-Cas9-induced DSBs were efficiently repaired by

138 the λ-Red mediated HR. Then, the same six methods were tested without

139 overexpressing λ-Red recombinases (Table 1, column 3). In the absence of λ-Red

6

140　recombinases, RecA was the only recombinase involved in HR that repaired CRISPR-

141　Cas9-induced DSBs. As shown in Fig. 1d, all methods achieved a low positive rate of

142　$< 6\%$, and methods 5 and 6 even showed no positive colonies (white colonies). For

143　methods 1–4, the editing efficiency correspondingly decreased to 6.0×10^{-8}–3.6×10^{-7}

144　^{7} (Table 1, column 3). The six methods were also tested in the absence of both λ-Red

145　recombinases and DNA template (Table 1, column 5). The results showed that only

146　methods 2 and 4 generated positive colonies, with rates of 2.8% and 1.5%, respectively

147　(Fig. 1e), and the editing efficiency were 5.0×10^{-8} and 2.0×10^{-8}, respectively (Table

148　1, column 4).

149　　　　Two NHEJ-dependent methods (methods 7 and 8) were also tested (Table 1), and

150　the procedures of these two methods are summarized in Fig. S7 and S8, respectively.

151　Both methods achieved a high positive rate of $> 50\%$ (Fig. 1b). However, the editing

152　efficiencies of the two methods were only 2.7×10^{-7} and 9.9×10^{-7}, respectively (Table

153　1, column 4). Methods 7 and 8 were then tested in the absence of overexpressed Ligase-

154　D and Ku proteins. The results showed that no positive colonies (white colonies) were

155　obtained (Table 1, column 5).

156　**Establishment of CNEE**

157　No previous reports have reported genomic editing in the absence of exogenous DNA

158　repair system and editing template. However, by testing all previous methods for

159　genomic editing, for the first time, we show that ENEJ may be potentially used in

160　genomic editing (Table 1, column 5). To fully utilize ENEJ-mediated DSB repair and

161　increase the efficiency of CRISPR-Cas9, we established CNEE and conducted

162　systematic optimizations (Table S1).

163　　　　We constructed a two-plasmid system in which the Cas9 and sgRNA were

164　expressed by different plasmids, namely p-P$_{BAD}$-*cas9* and p-P$_{BAD}$-sgRNA-X (X refers

165　to the target DNA) (Fig. 2). The two-plasmid system was more convenient in plasmid

166　construction than the one-plasmid system (16,20). The Cas9 used here was xCas9-3.7,

167   an evolved SpCas9 variant with broad PAM compatibility and high DNA specificity

168   (7). Furthermore, sgRNA other than tracrRNA:crRNA duplex was applied to increase

169   the performance and stability of the CRISPR-Cas9 system (1,8). The combination of

170   xCas9-3.7 and sgRNA increased the efficiency of the CRISPR-Cas9 system, thus

171   reducing the survival rate of WT cells. The expressions of both Cas9 and sgRNA were

172   controlled by an inducible promoter, thus the CRISPR-Cas9 system functioned only

173   when an inducer was added. The L-arabinose-induced P_{BAD} promoter and two IPTG-

174   induced promoters, namely P_{T5} and $P_{L}lacO_{1}$, were tested individually. We found that

175   the P_{BAD} promoter was stricter than both P_{T5} and $P_{L}lacO_{1}$, so we used the P_{BAD} promoter

176   for Cas9 and sgRNA expressions (Table S1). Then, we optimized the working

177   concentration of L-arabinose. Different concentrations of L-arabinose, ranging from

178   1−30 mM, were tested to induce the CRISPR-Cas9 system, and L-arabinose at a

179   concentration of 20 mM led to the highest positive rate and editing efficiency (Table

180   S1). We found that L-arabinose at low concentrations (*i.e.,* 1 mM) were inefficient to

181   induce the CRISPR-Cas9 system; and that high-concentration L-arabinose (*i.e.,* 30 mM)

182   resulted in higher rate of cell death, which might be due to the toxicity effect of high

183   amounts of the Cas9 protein (21). Culture temperature was also an important factor in

184   genomic editing. Low culture temperatures resulted in a reduction in cell growth rate,

185   whereas high culture temperatures reduced the stability of plasmid p-P_{BAD}-sgRNA-X,

186   which is temperature sensitive (22). Different culture temperatures, ranging from 25°C

187   to 37°C, were tested for genomic editing, and 30°C showed the highest positive rate

188   and editing efficiency (Table S1). In addition, we also optimized the culture time and

189   medium in order to further improve the positive rate and editing efficiency (Table S1).

190   The schematic procedures of the optimized CNEE are presented in Fig. 2. Detailed

191   procedures are provided in Fig. 9. First, the plasmid p-P_{BAD}-*cas9* expressing Cas9 and

192   the plasmid p-P_{BAD}-sgRNA expressing sgRNA were transformed into the target strain.

193   Then, after one of the transformants was cultured for a period of time, L-arabinose was

194   applied to induce the expression of Cas9 and sgRNA, resulting in site-specific

195 generations of DSBs and deletion mutations. Briefly, the Cas9−sgRNA complex binds

196 and cleaves the target DNA strands and generates DSBs at desired genomic loci. The

197 ENEJ pathway repairs the DSBs, and the cells survive the CRISPR-Cas9 mediated

198 cleavage but carry stochastic deletion in the target genomic loci (11). Finally, a small

199 fraction of the culture was plated in a selection plate, and the edited strains were verified

200 by colony PCR and sequencing. The plasmid p-P$_{BAD}$-sgRNA-X was removed by

201 increasing the cultivation temperature to 37−42℃ (Fig. S10a), and a new p-P$_{BAD}$-

202 sgRNA-X plasmid targeting a different genomic locus was introduced for the next

203 round of genomic editing. Finally, the plasmid p-P$_{BAD}$-Cas9 was cured by culturing in

204 LB antibiotic-free medium (Fig. S10b).

CNEE is convenient for gene inactivation

206 In principle, the ENEJ system repairs CRISPR-Cas9-induced DSBs and results in

207 stochastic genomic sequence deletion, thus disrupting the target ORF (Fig. 3a). We first

208 inactivated the *lacZ* gene in the MG1655 (WT) strain (Fig. 3b). The plasmid p-P$_{BAD}$-

209 sgRNA-*lacZ* containing the spacer that target the *lacZ* gene was transformed into

210 MG1655 (WT) harboring plasmid p-P$_{BAD}$-*cas9*; the mutations of the *lacZ* gene in the

211 resulting strains was examined by blue-white screening, colony PCR, and sequencing.

212 Almost all colonies in the experimental group turned white, indicating that most

213 cells surviving CRISPR-Cas9-mediated cleavage were edited, whereas all colonies in

214 the control group were blue in color, indicating that genomic editing did not occur in

215 the absence of CRISPR-Cas9 (Fig. 3c). The positive rate was 96% and the editing

216 efficiency was 3×10^{-5} (Fig. 3f). One hundred white colonies were randomly picked

217 for PCR screening using primer pairs F1/R1 and F2/R2. Theoretically, an agarose gel

218 electrophoresis band could be obtained if the sequence deletion was within the

219 complementary regions of primer pair. The results showed that using primer pair F1/R1,

220 52% of the white colonies generated one band, indicating that 52% of total positive

221 colonies carried a sequence deletion within the complementary regions of F1/R1 (Fig.

222 3f). Part of the PCR results are shown in Fig. 3d. The control group showed a band of

9

223 1,226 bp in size, the experimental group, *i.e*. Nos. 1 and 5, exhibited a band that are

224 obviously less than 1,226 bp in size. For colonies that did not exhibit a band, *i.e*. Nos.

225 2 and 6, deletions beyond the complementary regions of F1/R1 were observed. Part of

226 the sequencing results are shown in Fig. 3e. Deletions of different lengths that were

227 within the complementary regions of F1/R1 were observed. Similarly, using primer pair

228 F2/R2, 74% of the white colonies exhibited one band of less than 3,075 bp, indicating

229 that 74% of total positive colonies carried a sequence deletion within the ORF of *lacZ*

230 gene (Fig. 3f). For colonies that did not exhibit a band, deletions beyond the ORF of

231 *lacZ* gene were observed. Generally, these colonies were excluded because sequence

232 deletions that exceed the length of the ORF might impair the function of neighbouring

233 genes. We still found blue colonies in the experimental group, so we examined 10 of

234 these by colony PCR and sequencing. All tested blue colonies were determined to have

235 an intact *lacZ* gene, generally, these colonies are called 'escapers'. Although 'escapers'

236 still existed, these only accounted for 4% of all the colonies, making colony screening

237 much less tedious.

238   Using the same approach, we then iteratively knocked out *dinB*, a gene with an

239 ORF that was smaller than *lacZ* (Fig. 3b). One hundred colonies in the experimental

240 group were randomly picked for PCR screening and sequencing, with F3/R3 as primers.

241 The positive rate was 97%, and the editing efficiency was 2.2×10^{-5}. All colonies that

242 were not positive (3%) were proven to have an intact *dinB* gene. Among the positive

243 colonies, 60% carried a sequence deletion within the ORF of *dinB* (Fig. 3f).

CNEE is efficient in knocking out large fragments

245 Then, CNEE was applied for large-fragment knocking-out in MG1655 (WT), that is,

246 an 81-kb large fragment (3,609,883−3,690,267) and an 83-kb large fragment

247 (2,348,515−2,431,673). To delete the 81-kb large fragment from the bacterial

248 chromosome, we designed the p-P$_{BAD}$-sgRNA-81kb plasmid which contains two

249 sgRNA expression chimeras (Fig. 4a). The plasmid was transformed into MG1655 (WT)

250 that harbored plasmid p-P$_{BAD}$-*cas9*. Theoretically, under the guidance of two sgRNAs,

10

251 the Cas9 protein cuts off the fragment from the genome by simultaneously introducing

252 a pair of DSBs in the genome, followed by the re-joining of the distant ends (Fig. 4b).

253 There is no essential gene in the 81-kb fragment, thus the knocking-out of this fragment

254 does not impair the growth of the cell. Flanking the fragment are two essential genes

255 *yhhQ* and *bcsB*, and the deletion of either gene is lethal to the cells (23). Therefore, the

256 survived cells should have intact *yhhQ* and *bcsB*.

257 To check the mutants by colony PCR, three primer pairs were designed to target

258 the regions within the *yhhQ* (F4/R4), the 81-kb fragment (F5/R5), and the *bcsB* (F6/R6)

259 (Fig. 4b). One hundred colonies were randomly selected for PCR screening, and part

260 of agarose gel electrophoresis results are shown in Fig. 4c. The desired mutants had

261 only the first and the third bands, whereas the control [MG1655 (WT)] had all three

262 bands. The positive rate was 99%, and the editing efficiency was 4.4×10^{-6} (Fig. 4d).

263 Using the same approach, we then iteratively knocked out an 83-kb large fragment

264 (Fig. 4b). There is no essential gene in the 83-kb fragment, thus the knocking-out of

265 this fragment does not impair the growth of the cells. Franking the fragment are two

266 essential genes *nrdB* and *folC*, and the deletion of each gene is lethal to the cells (23).

267 Therefore, the survived cells should have intact *nrdB* and *folC*. Primer pairs F7/R7,

268 F8/R8 and F9/R9 were designed to check the positive mutants, and part of agarose gel

269 electrophoresis results are shown in Fig. 4c. The positive rate was 85%, and the editing

270 efficiency was 2.6×10^{-6} (Fig. 4d). We also obtained 15% of false positive colonies,

271 *i.e.* the No. 4 in the experimental group (Fig. 4c). Further investigation showed that

272 these false-positive colonies did not lose the 83-kb fragment, but contained stochastic

273 length of deletion in the region of both target sites. The results indicated that the

274 fragment being cut off could, with a certain probability, be re-integrated back into the

275 chromosome through ENEJ.

276 **CNEE does not require high-efficiency competent cells**

277 Generally, CRISPR-Cas9-based genomic editing methods depend on high-efficiency

278 competent cells to achieve relatively high editing efficiency, and electroporation is

279 commonly used to obtain high transformation efficiency (1,6,16,18,20,24-26).

280 However, highly competent cells, especially chemically competent cells, are not always

281 easy to prepare. Some strains, such as the MG1655 (WT) strain are very difficult to

282 prepare highly competent cells due to their genetic properties (27,28). The CNEE

283 allows transformants to reproduce before inducing genome editing by adding L-

284 arabinose, and thus does not depend on high-competency host strains. Theoretically, as

285 long as one transformant is obtained, genome editing can proceed normally (Fig. 6a).

286 To prove this advantage, we tested the CNEE for gene inactivation using MG1655 (WT)

287 competent cells of different transformation efficiencies. The *lacZ* gene was selected as

288 the target gene. The CaCl$_2$ method, Inoue method (27), and electroporation method

289 were used to obtain competent cells of different competencies, ranging from 10^4 (cfu/μg

290 pUC19) to 10^8 (cfu/μg pUC19). As expected, there was no apparent difference in either

291 the positive rate or the editing efficiency (Table 2).

292 In previous methods, the minimal transformation efficiency of competent cells

293 was 10^8–10^9 (cfu/μg pUC19) to obtain a relatively high editing efficiency (Table 3).

294 However, in CNEE, a competency of 10^4 (cfu/μg pUC19) is sufficient to obtain a

295 comparable editing efficiency (Tables 2 and 3).

296 **The DSB repair in CNEE is independent of RecA and LigB, but is dependent on**

297 **RecBCD**

298 The CNEE system has been proved to be convenient and efficient for gene inactivation

299 and large-fragment knocking-out. However, the detailed mechanism of ENEJ involved

300 in CNEE was unknown. Chayot *et al.* proposed a model for the mechanism of A-EJ,

301 which is the first ENEJ pathway found in *E. coli* (Fig. 5a) (11). In that model, A-EJ

302 relies on the degradation of DNA ends by the RecBCD complex and the ligation by the

303 LigA. In addition, the regions of microhomology (1−9 nt) are essential to promote the

304 ligation of two DNA ends.

305   To explore whether A-EJ involves in the CNEE, the *lacZ* was inactivated in the

306   MG1655 (WT), MG1655 (*ΔrecA*), and MG1655 (*ΔrecBCD*) strains. Both the MG1655

307   (*ΔrecA*) and MG1655 (*ΔrecBCD*) strains were obtained by CNEE. In the MG1655

308   (*ΔrecA*) and MG1655 (*ΔligB*) strains, the positive rate were 92% and 95%, respevtively,

309   and the editing efficiency were 3.2×10^{-5} and 3.15×10^{-5}, respectively, close to that of

310   the MG1655 (WT) strain (96%, 3.0×10^{-5}) (Figs. 5b and 5c), indicating that DSBs

311   repair is independent of the RecA recombinase and LigB. However, no DSB repair

312   event was found in the MG1655 (*ΔrecBCD*) strain (Figs. 5b and 5c), demonstrating that

313   DSBs repair relied on the RecBCD complex. Moreover, microhomology (1−3 nt) in the

314   joining site were observed in every DSB repair event (Fig. 5d). These results agreed

315   with the A-EJ model. We were unable to verify the role of LigA in DSB repair via gene

316   deletion, as the *ligA* is an essential gene for *E. coli*. Taken together, the DSB repair is

317   independent of RecA and LigB, but is dependent on RecBCD complex. Therefore, the

318   A-EJ was likely the ENEJ involved in the CNEE.

319   **Strategies used in CNEE are also applicable for other CRISPR-Cas9-mediated**

320   **genome editing methods**

321   Accumulating the offsprings of the transformants harboring the editing systems and

322   recovering the targeted cells are the keys to ensure high efficient CNEE. This strategy

323   was applied to other CRISPR-Cas9-mediated methods depending on RecA mediated

324   HR, λ-Red mediated HR, or NHEJ. The *Mycobacterium tuberculosis* H37Rv-derived

325   NHEJ pathway (Mt-NHEJ) is efficient in joining non-homologous DNA ends (1). The

326   adjacent genes *ligD* and *mku* are functional in Mt-NHEJ, and these were added to

327   plasmid p-P$_{BAD}$-*cas9* with the IPTG-induced P$_{T5}$ promoter to generate plasmid p-P$_{BAD}$-

328   *cas9*/P$_{T5}$-*ligD-mku* (Fig. 6b). The newly constructed plasmid was then transformed into

329   MG1655 (WT) strain along with plasmid p-P$_{BAD}$-sgRNA-*lacZ*. IPTG was used to

330   induce the expression of Ligase-D and Ku protein. Other than this, all procedures were

331   the same as CNEE. We obtained a 100% positive rate (Fig. 6c) and a 2.4×10^{-4} editing

332   efficiency (Fig. 6d). The editing efficiency was eight-fold as high as that of CNEE,

13

333   indicating that the Mt-NHEJ was more efficient than ENEJ in repairing CRISPR-Cas9-

334   induced DSBs.

335   Most reported methods used HR to repair CRISPR-Cas9-induced DSBs. However,

336   according to previous reports, the RecA-mediated HR pathway in *E. coli* showed low

337   efficiency in repairing DSBs, thus, λ-Red system must be introduced to enhance HR

338   activity in these study (6,20). Now that the ENEJ, a weak DSB repair mechanism,

339   worked well in CNEE, it was interesting to know how RecA-mediated HR performs

340   using the same strategies of CNEE. We added a donor sequence containing two

341   homologous arms of 500 bp to plasmid p-P$_{BAD}$-sgRNA-*lacZ* to generate plasmid p-

342   P$_{BAD}$-sgRNA-Donor-*lacZ* (Fig. 6b). The plasmid was transformed into MG1655 (WT)

343   strain along with plasmid p-P$_{BAD}$-*cas9*. All procedures were the same as CNEE. As a

344   result, a 50% positive rate and a 4×10^{-5} editing efficiency were obtained (Fig. 6c). The

345   editing efficiency was slightly higher than CNEE, and the positive rate was lower than

346   CNEE. We found that nearly all the remaining colonies (50%) that were not positive

347   went through DSB repair by ENEJ mechanism, indicating that RecA-mediated HR was

348   equivalent to ENEJ in repairing CRISPR-Cas9-induced DSBs. To enhance HR activity,

349   we added λ-Red system to plasmid p-P$_{BAD}$-*cas9* to generate plasmid p-P$_{BAD}$-*cas9*/P$_{T5}$-

350   Redγβα (Fig. 6b). The plasmid was transformed into MG1655 (WT) strain, along with

351   plasmid p-P$_{BAD}$-sgRNA-Donor-*lacZ*. The positive rate was 100% (Fig. 6c), indicating

352   that all colonies grown in the plate were edited. The editing efficiency was 2×10^{-2},

353   about 667-fold of CNEE, indicating that 1 out of 50 cells was edited.

354

14

Discussion

In this study, we developed the CNEE for genomic editing. The method utilizes ENEJ to repair DSBs, thus is independent of exogenous DNA repair systems such as NHEJ. To our knowledge, this is the first successful engineering effort to utilize ENEJ mechanism in genomic editing. Accumulating the offsprings of the transformants harboring the editing systems and recovering the targeted cells are the keys to ensure the high efficiency of the CNEE. The method has been applied for gene inactivation and large-fragment knocking-out. For gene inactivation, the positive rates were above 96%, and the editing efficiencies were above 2.2×10^{-5}. For large-fragment knocking-out, the positive rates were above 85%, and the editing efficiencies were above 2.6×10^{-6}. Our method has three advantages compared to the existing methods: (i) it requires neither a DNA template nor an exogenous DSBs repair system such as NHEJ, (ii) it is simple and efficient in gene inactivation and large-fragment knocking-out, (iii) it is independent of high-efficiency competent cells.

Classic recombineering in *E. coli* relies on λ-Red recombinases to achieve genomic editing with an introduced homologous DNA as the editing template (Table 3). Either double-stranded DNA (dsDNA) or single-stranded DNA (ssDNA) is functional as the editing template. Generally, dsDNA mediated recombineering is less efficient, and it requires a selection marker such as an antibiotic resistance gene (29-31). On the contrary, ssDNA mediated recombineering does not require selection marker due to its higher efficiency; however, it is only applicable to small-scale genomic editing such as codon displacement (32,33). Recombineering is unable to knock out large fragments from the chromosome, which limits its application in large-scale genetic engineering. The introduction of CRISPR-Cas9 to *E. coli* provides an efficient way for complicated chromosome modifications such as knocking out large fragments (Table 3). Many CRISPR-Cas9-dependent methods have been reported. In these methods, the CRISPR-Cas9 system was combined with an exogenous DSB repair system, λ-Red system, or NHEJ system to achieve site-specific genomic editing.

15

383 However, our testing showed that these methods are strictly dependent on the

384 exogenous DSB repair system. In CNEE, ENEJ is employed to repair CRISPR-Cas9-

385 induced DSBs in the absence of an editing template. The ENEJ mechanism naturally

386 exists in *E. coli*. As a much weaker DSB repair mechanism than λ-Red system or NHEJ

387 system, it still shows a good performance in genomic editing. By gene disruption, we

388 demonstrated that the ENEJ is independent of RecA recombinase and LigB, but

389 dependent on RecBCD, a protein complex that has 5' to 3' exonuclease, 3' to 5'

390 exonuclease, and helicase activities. Theoretically, ligase is involved in the end joining

391 of DNA ends. In *E. coli*, LigA and LigB are the only ligases. Now that the LigB is

392 useless in DSB repair, the LigA is likely the ligase involved in the ENEJ mechanism.

393 We were unable to verify the role of LigA in end joining because the *ligA* is an essential

394 gene in *E. coli*. The sequencing results of over thirty samples of DSB repair indicated

395 that microhomology of 1–3 nt was essential for end joining. The A-EJ model proposed

396 by Chayot et al. explains the end-joining mechanism well, but the exact biochemical

397 mechanism remains to be explored.

398   CRISPR-Cas9 system has been conveniently employed to introduce site-specific

399 DSBs, and the cleavage of dsDNA kills WT cells, thereby facilitating the selection of

400 edited cells. In eukaryotes such as *Saccharomyces cerevisiae*, the HR and NHEJ

401 activities are high enough to efficiently repair CRISPR-Cas9-induced DSBs. Therefore,

402 the application of CRISPR-Cas9 in eukaryotes is simple and convenient. However, in

403 bacteria, the situation is very different. A small fraction of bacteria has relatively high

404 HR activity such as *Streptococcus pneumoniae* or relatively high NHEJ activity such

405 as *Mycobacterium tuberculosis*. In many bacteria such as *E. coli*, both HR and end-

406 joining activities are weak. The application of CRISPR-Cas9 in these bacteria is, thus,

407 difficult. However, in this study, we show that the *E. coli* native HR pathway or end-

408 joining pathway is applicable for efficient genomic editing, which makes genomic

409 editing in *E. coli* as simple as that in *S. cerevisiae*. This observation implies that all

410 bacteria have the potential to be edited, rely on their native DSB repair mechanism.

16

411   Existing genomic editing methods for *E. coli* are strictly dependent on high-
412   efficiency competent cells, and electroporation or expensive commercial kits are
413   generally used to obtain high-efficiency competent cells. However, it is difficult to
414   make highly competent cells for some *E. coli* strains even using special and complicated
415   methods. Therefore, genomic editing in these strains is highly limited. However, in this
416   study, we showed that CNEE is independent of high-efficiency competent cells, and
417   the required minimal transformation efficiency is 10^4 cfu/µg DNA, which is very easy
418   to obtain using traditional, chemical methods. Moreover, the same editing efficiency is
419   achieved as long as the transformation efficiency is over 10^4 cfu/µg DNA. We believe
420   that CNEE is more advantageous in other bacteria that have a lower competency than
421   *E. coli*.

422   Gene inactivation and large-fragment knocking-out are frequently employed in
423   synthetic biology and metabolic engineering. And researchers always choose to utilize
424   an efficient and simple method. As opposed to the indirect selection provided by
425   recombineering, the use of the CRISPR-Cas9 system makes it possible to directly select
426   for the desired mutation. CRISPR-Cas9-induced DNA cleavage kills WT cells, and thus
427   edited cells can be selected easily. However, the major limitation is the presence of a
428   background of cells that escape CRISPR-Cas9-induced cell death and lack the desired
429   mutation, and the limitation is more prominent in strains that have a weak DSB repair
430   mechanism. Improving the effieicncy of CRISPR/Cas9 system is a key way to solve
431   this problem. In this study, the Cas9 derivative xCas9-3.7 and the engineered sgRNA
432   were utilized to successfully increase the efficiency of introducing DSBs (7,8).
433   Moreover, the engineering of an inducible system helps circumvent this limitation. Off-
434   target is another limitation for the application of CRISPR/Cas9 technology. In this study,
435   the use of xCas9-3.7 helps to reduce the off-target event. In addition, designing an
436   appropriate sgRNA sequence is important to avoid off-target. Therefore, a reliable
437   sgRNA designing software or website is needed.

438   Materials and methods

439   Strains and culture conditions

440   The *E. coli* strain DH5α was used as the host strain for molecular cloning and

441   manipulation of plasmids. The MG1655 (WT), MG1655 (*recA-*), or MG1655

442   (*ΔrecBCD*) were employed as host strains for genetic engineering. Strains MG1655

443   (*recA-*) and MG1655 (*ΔrecBCD*) were obtained by the CNEE (proposed in this study)

444   in the basis of MG1655 (WT). Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L

445   yeast extract, and 10 g/L NaCl) was used for cell growth in all cases unless otherwise

446   noted. Agar was added at 20 g/L for solid medium. SOC medium (20 g/L tryptone,

447   5 g/L yeast extract, 0.5 g/L NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, and

448   20 mM glucose) was used for cell recovery. Ampicillin (Amp), kanamycin (Kan) and

449   chloramphenicol (Cm) were used at concentrations of 0.1 g/L, 0.05 g/L and 0.025

450   g/L, respectively. When appropriate, IPTG, X-gal, L-arabinose, and glucose were

451   supplemented into media or cultures as follows: IPTG (1 mM), X-gal (0.1 g/L), L-

452   arabinose (20 mM), and glucose (10 g/L).

453   Plasmids construction

454   The primers and plasmids used in this study are listed in Tables S2 and S3,

455   respectively. Detailed construction procedures for plasmids p-P$_{BAD}$-sgRNA-X are

456   shown in Fig. S11. The complete sequences of the plasmids p-sgRNA-template A, p-

457   sgRNA-template B, p-sgRNA-template C, p-P$_{BAD}$-*cas9*, p-P$_{BAD}$-*cas9*/P$_{T5}$-*ligD-mku*,

458   and p-P$_{BAD}$-*cas9*/P$_{T5}$-Redγβα are presented in Notes S1–S6. The CRISPR target

459   sequences used in this study are listed in Table S4.

460   CNEE procedure

461   First, the Kan-resistant (Kan^R) plasmid p-P$_{BAD}$-*cas9* was transformed into the target

462   strain such as MG1655 to obtain the corresponding transformant, such as MG1655/p-

463   P$_{BAD}$-*cas9*. A series of temperature-sensitive Amp-resistant (Amp^R) plasmids

464   containing   specifically   designed   spacers   were   constructed   to   express   the

18

465   corresponding sgRNA and were generally named as p-P$_{BAD}$-sgRNA-X. Then,

466   specific plasmid p-P$_{BAD}$-sgRNA-X was transformed into the MG1655/p-P$_{BAD}$-*cas9*

467   strain, and the MG1655/p-P$_{BAD}$-*cas9*/p-P$_{BAD}$-sgRNA-X strain was screened in a LB

468   plate with Amp, Kan, and glucose at 30 ℃. A single colony was picked and

469   inoculated into 0.5 mL SOC medium and cultured at 30 ℃ for 2 h. Then, 4.5 mL LB

470   medium, 5 µL Amp, and 5 µL Kan were added to the cultures. After 1 h, 100 µL of

471   L-arabinose was added, and the cultures were cultured for another 3 h before plating.

472   A 10-µL aliquot of the cultures were plated onto a LB plate containing Amp, Kan,

473   and L-arabinose, and the plate was cultured overnight at 30 ℃. A flowchart of CNEE

474   is shown in Fig. 2 and Fig. S9. Details of the reagents and media used in the CNEE

475   are listed in Table S5.

476   **Plasmids curing and iterative editing**

477   Positive mutants were verified by colony PCR and sequencing and were cultured in

478   LB media individually in the presence of only Kan at 37 ℃ for 12 h to remove the

479   temperature-sensitive Amp^R plasmid p-P$_{BAD}$-sgRNA-X (Fig. S10a). Then, the

480   obtained edited strain containing only the plasmid p-P$_{BAD}$-*cas9* could was used as

481   the starting strain for the next round of genomic editing. The Kan^R plasmid p-P$_{BAD}$-

482   *cas9* is not stable in the host strain in the absence of Kan. When the last round of

483   genomic editing was completed, the edited strain was cultured in LB media without

484   Kan at 37 ℃ for 16 h to remove both Amp^R plasmids p-P$_{BAD}$-sgRNA-X and Kan^R

485   plasmid p-P$_{BAD}$-*cas9* (Fig. S10b). The overnight culture was diluted and plated on a

486   LB plate. Strains sensitive to both Amp and Kan are plasmid-free. Single plasmid-

487   free colonies were selected and verified using LB media with or without

488   corresponding antibiotics. The flowchart of the procedure is shown in Fig. S9b.

489   **Measurement of transformation efficiency**

490   Pure, supercoiled pUC19 was used to determine the transformation efficiency of

491   competent cells. First, 1 µl of pUC19 (1 ng/µl) was added to one tube of competent

492   cells. Next, the mixture was incubated for 30 minutes before conducting heat-shock

493   for 90 seconds in a 42 ℃ water bath. Then, the tube was placed on ice for two minutes

494   before adding 1 mL of SOC medium, and the tube was shaked at 200–230 rpm (37 ℃)

495   for 40 minutes. Last, 100 μl of the cultures was plated on a LB plate containing

496   ampicillin, and the plate was incubated overnight at 37 °. The transformation

497   efficiency is N $\times 10^{4}$ cfu/μg pUC19 ("N" refers to the number of colonies).

498   **Calculation of genomic editing positive rate**

499   One hundred colonies in the LB plate containing Amp, Kan, and L-arabinose were

500   tested by colony PCR to screen for positive mutants. Twenty of the positive mutants

501   were sequenced for further verification. The positive rate was calculated by the

502   proportion of positive colonies to the total number of colonies.

503   **Calculation of genomic editing efficiency**

504   In previous reports, the positive rate was frequently taken as the editing efficiency in

505   a genomic editing. In this study, we proposed a new definition that better reflected

506   editing efficiency. One control group was set along with the experimental group to

507   calculate the editing efficiency. In the control group, L-arabinose was not added to

508   the cultures, thus no Cas9 protein was expressed. In addition, all other operations

509   were the same as the experimental group. The editing efficiency was calculated by

510   the proportion of positive colonies in the experimental group to the total number of

511   colonies in the control group.

512

References

1. Su T, et al. (2016) A CRISPR-Cas9 assisted non-homologous end-joining strategy for one-step engineering of bacterial genome. *Sci Rep* 6:37895.

2. Gu P, et al. (2015) A rapid and reliable strategy for chromosomal integration of gene(s) with multiple copies. *Sci Rep* 5:9684.

3. Smanski MJ, et al. (2016) Synthetic biology to access and expand nature's chemical diversity. *Nat Rev Microbiol* 14:135–149.

4. Esvelt KM, Wang HH (2013) Genome-scale engineering for systems and synthetic biology. *Mol Syst Biol* 9:641.

5. Thomas G, Gersbach CA, Barbas CF (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31:397–405.

6. Wenyan J, David B, David C, Feng Z, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31:233–239.

7. Hu JH, et al. (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556:57–63.

8. Jinek M, et al. (2012) A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821.

9. Ren éD, et al. (2004) Evidence that stable retroviral transduction and cell survival following DNA integration depend on components of the nonhomologous end joining repair pathway. *J Virol* 78:8573–8581.

10. Claire W, Roland K (2006) DNA Double-strand break repair: all's well that ends well. *Annu Rev Genet* 40:363–383.

11. Romain C, Benjamin M, Didier M, Miria RJ (2010) An end-joining repair mechanism in *Escherichia coli*. *Proc Natl Acad Sci U.S.A.* 107:2141–2146.

12. Wilson TE, Topper LM, Palmbos PL (2003) Non-homologous end-joining: bacteria join the chromosome breakdance. *Trends Biochem Sci* 28:62–66.

13. Marina D, et al. (2004) *Mycobacterial* Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* 306:683–685.

14. Weller GR, et al. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297:1686–1689.

15. Malyarchuk S, et al. (2007) Expression of *Mycobacterium tuberculosis* Ku and Ligase D in *Escherichia coli* results in RecA and RecB-independent DNA end-joining at regions of microhomology. *DNA Repair* 6:1413–1424.

16. Li Y, et al. (2015) Metabolic engineering of *Escherichia coli* using CRISPR–Cas9 meditated genome editing. *Metab Eng* 31:13–21.

17. Pines G, et al. (2015) Codon compression algorithms for saturation mutagenesis. *ACS Synth Biol* 4:604–614.

18. Zheng X, Li SY, Zhao GP, Wang J (2017) An efficient system for deletion of large DNA fragments in *Escherichia coli* via introduction of both Cas9 and the non-homologous end joining system from *Mycobacterium smegmatis*. *Biochem Biophys Res* 485(4):768–774.

19. Cui L, Bikard D (2016) Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Re*s 44:4243–4251.

562   20.   Zhao D, et al. (2016) Development of a fast and easy method for *Escherichia coli* genome
563   editing with CRISPR/Cas9. *Microb Cell Fact* 15(1):205.

564   21.   Cho S, et al. (2018) High-level dCas9 expression induces abnormal cell morphology in
565   *Escherichia coli*. *ACS Synth Biol* 7:1085–1094.

566   22.   Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia*
567   *coli* K-12 using PCR products. *Proc Natl Acad Sci U.S.A.* 97:6640–6645.

568   23.   Baba T, et al. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout
569   mutants: the Keio collection. *Mol Syst Biol* 2:8.

570   24.   Yu J, et al. (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9
571   system. *Appl Environ Microbiol* 81(7):2506–2514.

572   25.   Zhang H, Cheng QX, Liu AM, Zhao GP, Wang J (2017) A novel and efficient method for
573   bacteria genome editing employing both CRISPR/Cas9 and an antibiotic resistance cassette.
574   *Front Microbiol* 8:812.

575   26.   Zhao D, et al. (2017) CRISPR/Cas9-assisted gRNA-free one-step genome editing with no
576   sequence limitations and improved targeting efficiency. *Sci Rep* 7(1):16624.

577   27.   Sambrook J, Russell DW (2006) The inoue method for preparation and transformation of
578   competent *E. Coli*: "ultra-competent" cells. *CSH Protoc* 2006:418–424.

579   28.   Li J, et al. (2018) T5 exonuclease-dependent assembly offers a low-cost method for efficient
580   cloning and site-directed mutagenesis. *Nucleic Acids Res* 47:3.

581   29.   Pines G, Freed EF, Winkler JD, Gill RT (2015) Bacterial recombineering: genome
582   engineering via phage-based homologous recombination. *ACS Synth Biol* 4:1176–1185.

583   30.   Santos CN, Regitsky DD, Yoshikuni YJ (2013) Implementation of stable and complex
584   biological systems through recombinase-assisted genome engineering. *Nat Commun* 4:2503.

585   31.   Hailong W, et al. (2014) Improved seamless mutagenesis by recombineering using *ccdB* for
586   counterselection. *Nucleic Acids Res* 42:e37.

587   32.   Wang HH, et al. (2009) Programming cells by multiplex genome engineering and accelerated
588   evolution. *Nature* 460:894–898.

589   33.   Ronda C, Pedersen LE, Sommer MO, Nielsen AT (2016) CRMAGE: CRISPR optimized
590   MAGE recombineering. *Sci Rep* 6:19452.

591   34.   Yu D, et al. (2000) An efficient recombination system for chromosome engineering in
592   *Escherichia coli*. *Proc Natl Acad Sci U.S.A.* 97(11):5978–5983.

593   35.   Ellis HM, Yu D, Ditizio T, Court DL (2001) High efficiency mutagenesis, repair, and
594   engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci*
595   *U.S.A.* 98:6742–6746.

596
597

598   **Figure legends**

599   **Fig. 1. Genomic editing tests of existing methods.** (**a**) Schematic of *lacZ* inactivation

600   via HR pathway or end-joining pathway. (**b**) Positive rates of methods 7 and 8 in the

601   presence of overexpressed Ligase-D and Ku protein. (**c**) Positive rates of methods 1–6

602   in the presence of editing template and overexpressed λ-Red recombinases. (**d**) Positive

603   rates of methods 1–6 without overexpressed λ-Red recombinases but with editing

604   template. (**e**) Positive rates of methods 1–6 in the absence of both editing template and

605   overexpressed λ-Red recombinases. Data are expressed as means ± s.d. from three

606   independent experiments.

607

608   **Fig. 2. Schematic of CNEE.** Step 1: Plasmids p-P$_{BAD}$-*cas9* and p-P$_{BAD}$-sgRNA-X were
609   transformed into a host strain. Step 2: a transformant was cultured for reproduction
610   before inducing genomic editing by adding L-arabinose. Step 3: Plasmids in the edited
611   strain were removed to obtain a plasmid-free strain.

612

613   **Fig. 3. CNEE based gene inactivation.** (**a**) Principle of gene inactivation. (**b**) Genes
614   that were inactivated. (**c**) Representative results of blue-white screening of *lacZ*
615   inactivation experiment. (**d**) Representative PCR results of *lacZ* inactivation
616   experiment. White colonies were randomly picked to conduct PCR screening with a
617   blue colony served as the control. F1/R1 were used as primers. (**e**) Representative
618   sequencing results of *lacZ* inactivation experiment. (**f**) Positive rates and editing
619   efficiencies of *lacZ* inactivation and *dinB* inactivation. "Total" represents all edited
620   colonies. "Fn/Rn (n=1, 2, 3)" represent colonies that contain sequence deletion within
621   the corresponding primer pairs. Data are expressed as means ± s.d. from three
622   independent experiments.

623

624   **Fig. 4. CNEE based large-fragment knocking-out.** (**a**) Structure of plasmid p-P$_{BAD}$-

625   sgRNA-X used for large-fragment knocking-out. The plasmid contains two sgRNA

626   expression chimeras. (**b**) Knocking-out of an 81 kb fragment and an 83 kb fragment. (**c**)

627   Representative PCR results of large-fragment knocking-out experiments. Colonies

628   were randomly picked for PCR screening, and a WT colony served as control. (**d**)

629   Positive rates and editing efficiencies of large-fragment knocking-out. Data are

630   expressed as means ± s.d. from three independent experiments.

631   24

632 **Fig. 5. Mechanism of ENEJ involved in CNEE.** (**a**) Model of A-EJ mechanism
633 proposed by Chayot *et al*. Step 1: the RecBCD complex binds and degrades DNA ends
634 and generates micro sticky ends in the microhomology regions. Step 2: the sticky ends
635 mediates the synapsis of DNA. Step 3: the Ligase-A repairs the gaps and generates
636 complete dsDNA. (**b**, **c**) Results of *lacZ* inactivation in three strains. *ΔrecA* represents
637 *recA* deficient. *ΔrecBCD* represents *recBCD* deficient. Strains MG1655 (*ΔrecA*) and
638 MG1655 (*ΔrecBCD*) were obtained by CNEE. (**d**) Representative DSB repair events
639 by CNEE. Data are expressed as means $\pm$ s.d. from three independent experiments.
640
641 **Fig. 6. Extended methods of CNEE.** (**a**) The core strategy of CNEE. Transformants
642 are cultured for reproduction to increase the cells number and cells activity before
643 triggering DNA cleavage and DSB repair. The white dots represent WT cells. The green
644 dots represent edited cells. The orange dots represent cells that die from the cleavage
645 of Cas9. "N" represents the number of living cells. "N_0" represents the number of
646 transformants. "n" represents the number of cell division. "1/a" represents the
647 efficiency of DSB repair. (**b**) Plasmids constructed for extended methods. Plasmid p-
648 P_{BAD}-*cas9*/P_{T5}-*ligD-mku* was obtained by adding the NHEJ associated genes (*ligD* and
649 *mku*) to plasmid p-P_{BAD}-*cas9*. Plasmid p-P_{BAD}-*cas9*/P_{T5}-Redγβα was obtained by
650 adding λ-Red system (Redγβα) to plasmid p-P_{BAD}-*cas9*. Plasmid p-P_{BAD}-sgRNA-
651 Donor-*lacZ* was obtained by adding a donor sequence to plasmid p-P_{BAD}-sgRNA-*lacZ*.
652 (**c-d**) Positive rates and editing efficiencies of CNEE and three extended methods. In
653 extended methods, NHEJ pathway, RecA-mediated HR pathway, or λ-Red-mediated
654 HR pathway served as DSB repair pathway. Data are expressed as means $\pm$ s.d. from
655 three independent experiments.

25

656 **Tables**

657 **Table 1. Test results of reported genomic editing methods**

Foreign elements	DSB repair mechanisms							
	λ-Red-mediated HR		RecA-mediated HR		NHEJ		ENEJ	
Cas9 protein	+		+		+		+	
Editing template	+		+		−		−	
Redγβα	+		−		−		−	
Ku, Ligase-D	−		−		+		−	
Tested methods	PR (%)	EE	PR (%)	EE	PR (%)	EE	PR (%)	EE
Method 1 (6)	65 (±7.2)	$2.0\,(\pm0.32)\times10^{-5}$	0.71 (±0.21)	$1.0\,(\pm0.22)\times10^{-7}$	NA	NA	0	0
Method 2 (24)	86 (±6.7)	$1.5\,(\pm0.14)\times10^{-5}$	5.6 (±0.79)	$1.3\,(\pm0.18)\times10^{-7}$	NA	NA	2.8 (±0.99)	$5.0\,(\pm1.7)\times10^{-8}$
Method 3 (16)	95 (±3.8)	$2.3\,(\pm0.26)\times10^{-4}$	0.58 (±0.11)	$6.0\,(\pm0.48)\times10^{-8}$	NA	NA	0	0
Method 4 (20)	19 (±6.6)	$1.4\,(\pm0.12)\times10^{-5}$	2.1 (±0.34)	$3.6\,(\pm0.42)\times10^{-7}$	NA	NA	1.5 (±0.76)	$2.0\,(\pm0.90)\times10^{-8}$
Method 5 (25)	93 (±5.4)	$3.9\,(\pm0.38)\times10^{-5}$	0	0	NA	NA	0	0
Method 6 (26)	100 (±0)	$4.6\,(\pm0.44)\times10^{-6}$	0	0	NA	NA	0	0
Method 7 (1)	NA	NA	NA	NA	51 (±8.2)	$2.7\,(\pm0.30)\times10^{-7}$	0	0
Method 8 (18)	NA	NA	NA	NA	90 (±6.2)	$9.9\,(\pm1.6)\times10^{-7}$	0	0

658 PR: positive rate, EE: editing efficiency, NA: not available. Data are means ± s.d. from three independent experiments.

659

660 **Table 2. Test results of CNEE using cells with different competencies**

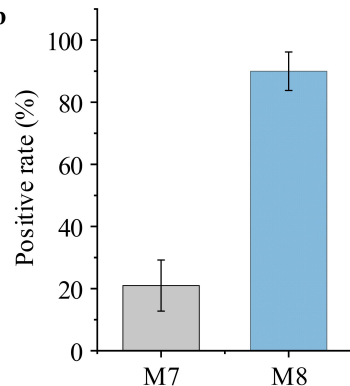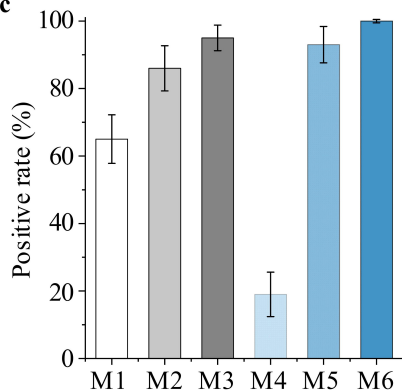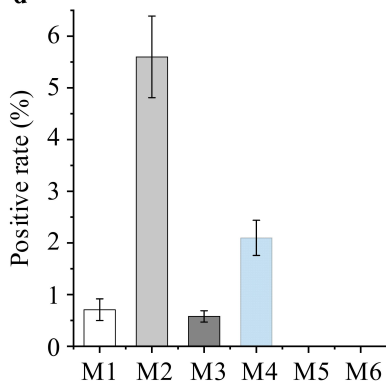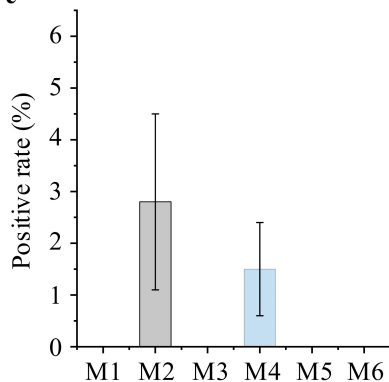| Competency (cfu/μg pUC19)	PR (%)	EE	Deletion size less than 1 kb (%)
10^4 (CaCl₂ method)	98 ± 1.4	$3.2\,(\pm0.21)\times10^{-5}$	55 ± 6.5
10^5 (CaCl₂ method)	96 ± 3.8	$3.0\,(\pm0.23)\times10^{-5}$	52 ± 7.6
10^6 (Inoue method)	97 ± 2.1	$3.5\,(\pm0.26)\times10^{-5}$	54 ± 7.1
10^7 (Inoue method)	95 ± 4.2	$3.3\,(\pm0.32)\times10^{-5}$	53 ± 4.3
10^8 (Electroporation method)	96 ± 3.4	$3.5\,(\pm0.28)\times10^{-5}$	58 ± 5.7

661 PR: positive rate, EE: editing efficiency. Data are expressed as means ± s.d. from three independent experiments.

662

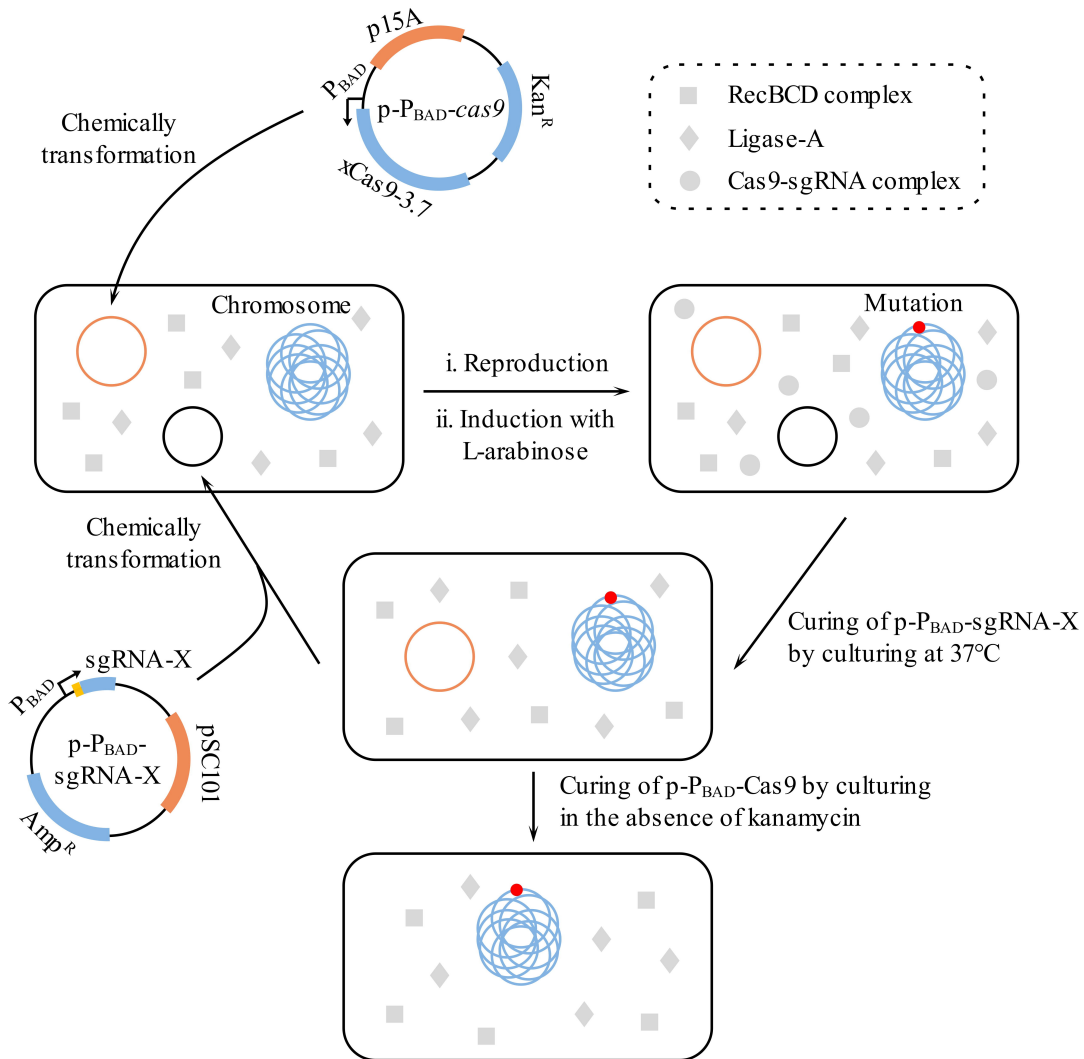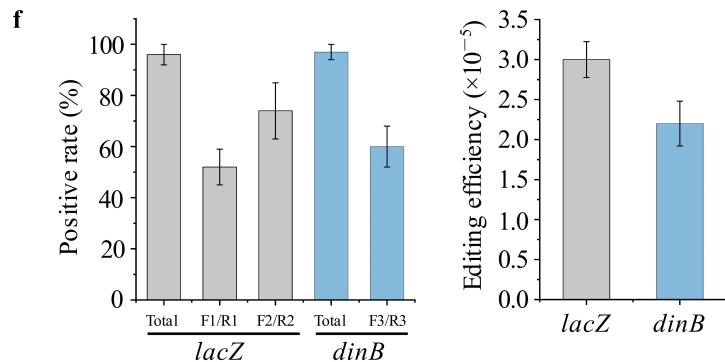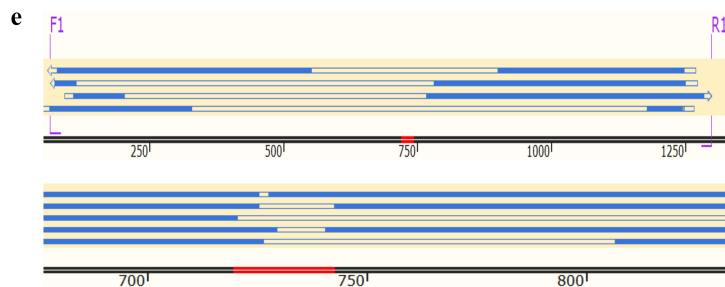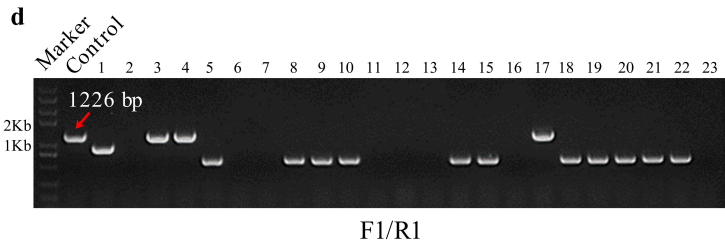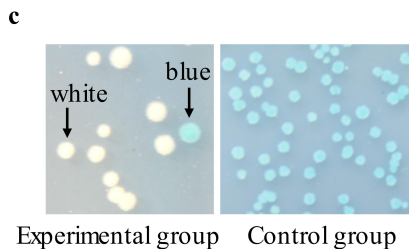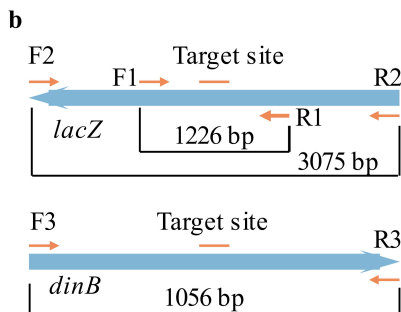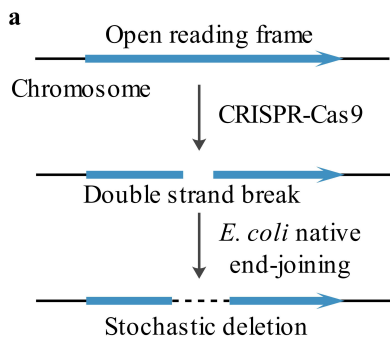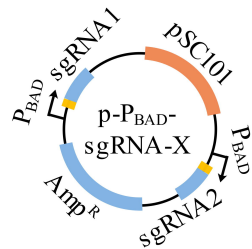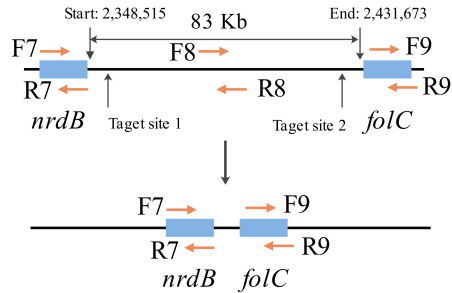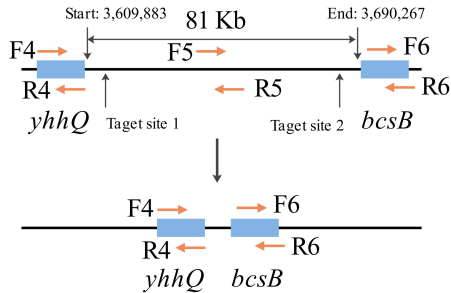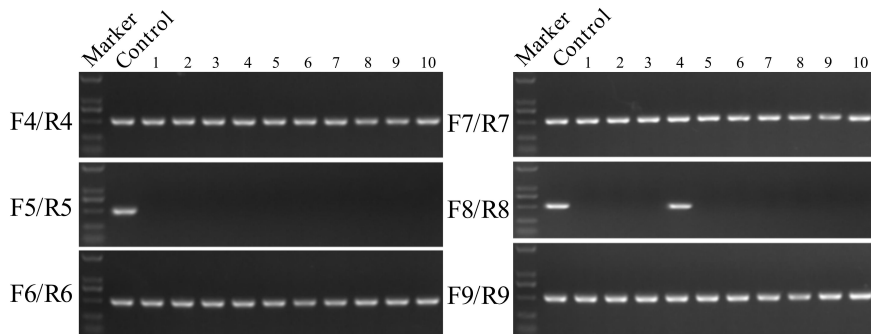663 **Table 3. Comparison of CNEE with existing genomic editing methods**
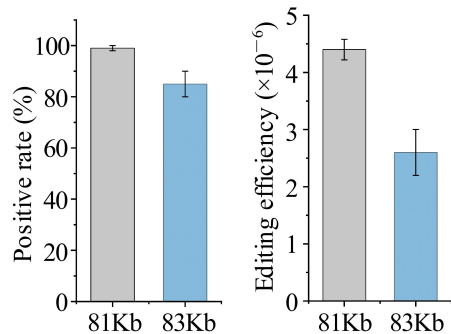
Principle	Method	Selection marker	Foreign proteins	Donor DNA	Efficiency	Large fragment knocking-out	Minimal competency (cfu/μg DNA)	Reference
ENEJ	CNEE	Not required	Cas9	Not required	Efficient	Available	10^4	This study
NHEJ	CA-NHEJ	Not required	Cas9, Ku, Ligase-D	Not required	Efficient	Available	10^8	(1,18)
HR	CA-HR	Not required	Cas9, Redγβα	Required	Efficient	Available	10^8	(16,26)
HR	dsDNA-HR	Required	Redγβα	Required	Difficult	Incapable	10^8	(22,34)
HR	ssDNA-HR	Not required	Redβ	Required	Efficient	Incapable	10^9	(32,35)

664

a Open reading frame — Chromosome → CRISPR-Cas9 → Double strand break → *E. coli* native end-joining → Stochastic deletion

b F2, F1, Target site, R2, R1, *lacZ*, 1226 bp, 3075 bp; F3, Target site, R3, *dinB*, 1056 bp

c Experimental group (white, blue); Control group

d Marker, Control, 1–23; 1226 bp; 2Kb, 1Kb; F1/R1

e F1, R1; 250, 500, 750, 1000, 1250; 700, 750, 800

f Positive rate (%): Total, F1/R1, F2/R2 (*lacZ*); Total, F3/R3 (*dinB*). Editing efficiency (×10⁻⁵): *lacZ*, *dinB*

a

b

Start: 3,609,883   81 Kb   End: 3,690,267

Target site 1   Target site 2

Start: 2,348,515   83 Kb   End: 2,431,673

Target site 1   Target site 2

c

d

a Double strand break

DNA resection

RecBCD   RecBCD

1-9 nt   1-9 nt

DNA-directed
end synapsis

Ligase-A

Ligase-A

Ligation

b

Positive rate (%)

MG1655 (*WT*)
MG1655 (Δ*recA*)
MG1655 (Δ*recBCD*)
MG1655 (Δ*ligB*)

c

Editing efficiency (×10⁻⁵)

MG1655 (*WT*)
MG1655 (Δ*recA*)
MG1655 (Δ*recBCD*)
MG1655 (Δ*ligB*)

d

Before editing		After editing
-----TCCGTCAGCG----CGGGCAAATAAT-----	→	-----TCCGTCAGCGGGCAAATAAT-----
-----GAAGCCGCCC-----CGGGGATACTG-----	→	-----GAAGCCGCCCGGGGATACTG-----
-----AGACCAGACC-----ACCACAGCGGAT-----	→	-----AGACCAGACCACAGCGGAT-----
-----GACGAAGCCG-----CGGGGATACTGA-----	→	-----GACGAAGCCGGGGATACTGA-----
-----GGCGTGCGGT-----TGGCACCATGC-----	→	-----GGCGTGCGGTGGCACCATGC-----

a

Reproduction → Induction of Cas9 and sgRNA →

Transformants

$N = N_0$ $N = N_0 \times 2^n$ $N = N_0 \times 2^n \times 1/a$

Edited cells, survive

Unedited cells, die

b

p15A — P_{T5} — ligD-mku

P_{BAD}

p-P_{BAD}-Cas9/ P_{T5}-ligD-mku

Kan^R

xCas9-3.7

P_{BAD} — sgRNA

Donor

p-P_{BAD}-sgRNA -Donor-lacZ

pSC101

Amp^R

p15A — P_{T5} — Redγβα

P_{BAD}

p-P_{BAD}-Cas9/ P_{T5}-Redγβα

Kan^R

xCas9-3.7

c

Positive rate (%)

ENEJ, NHEJ, RecA-HR, λ-Red-HR

d

Editing efficiency

ENEJ, NHEJ, RecA-HR, λ-Red-HR