Iterative *in vivo* cut'n'paste of chromosomal loci in *Escherichia coli* K-12 using synthetic DNA

Authors: Sebastian Gude[1,*], Andreas B. Bertelsen[1], and Morten H. H. Nørholm[1,2]

[1] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark

[2] current address: Novo Nordisk A/S, Hørsholm, Denmark

***Corresponding author:** Sebastian Gude, The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark, segu@biosustain.dtu.dk

Keywords:

genome engineering, genome refactoring, genetic engineering, CRISPR Cas9, λ Red recombineering, homologous recombination, *in vivo*, *Escherichia coli*, bacteria, scarless

Abstract

We have developed a simple and highly efficient *in vivo* method to iteratively relocate functional chromosomal loci onto an episome in *Escherichia coli* by utilizing synthetic DNA fragments. In this *in vivo* cut'n'paste procedure, "cutting" is executed by

26 the RNA-guided DNA endonuclease Cas9 and a set of guides, while “pasting” is
27 facilitated by the phage λ Red recombinase, which are all synthesized on easily
28 curable plasmids. To demonstrate the utility of this approach, we commercially
29 obtained synthetic DNA fragments containing locus-specific homology regions,
30 antibiotic marker cassettes, and standardized Cas9 target sequences. By using these
31 locus-specific synthetic DNA fragments, we relocated 7 functional chromosomal loci.
32 Scarless relocation mutants of the *trg*, *aer*, *tsr*, *malHM*, *malQT*, *macB-nadA*, and *folA*
33 chromosomal loci were obtained as antibiotic-resistant isolates by combining Cas9
34 counterselection with the restoration of an antibiotic marker cassette. The additional
35 antibiotic marker cassettes and standardized Cas9 target sequences present in the
36 synthetic DNA fragments are inherently eliminated upon completion of the procedure,
37 enabling iterative processing of loci. This procedure should be widely useful, especially
38 in genome (re-)engineering of *E. coli* and other bacteria because the procedure can
39 be done on wild-type cells.
40
41 **Introduction**
42
43 The ability to manipulate genomes is key for uncovering and leveraging the
44 potential of (micro-)biology. A milestone in our ability to edit bacterial genomes was the
45 development of the one-step inactivation of chromosomal genes procedure[1]. Among
46 many other scientific and engineering endeavours, it facilitated the creation of a
47 comprehensive single gene knockout library and the charting of essential genes in *E.*
48 *coli*[2], the study of bacterial motility[3,4], stringent response[5], the interplay of growth and
49 gene expression[6], two-component systems[7], evolution of antibiotic resistsance[8], and
50 chromosomal organisation[9] as well as the development of strains with reduced

genomes[10] and highly titratable induction systems[11] while also playing a crucial role in metabolic engineering applications[12,13]. Though the one-step inactivation procedure[1] was preceded by numerous other bacterial genome editing techniques[14–21], its ease of use and broad applicability ensured its still continuing success. The required short homology extensions to create genomic edits are easily attached to an antibiotic marker cassette via PCR, the recombineering machinery is expressed from an easily curable plasmid, and the procedure does not require any prior modification to the bacterial genome. Furthermore, the antibiotic marker cassette can subsequently be eliminated, resulting in an unmarked, yet not fully scarless, mutant that is compatible with another round of genome editing.

With the rise of CRISPR-Cas technologies[22], λ Red recombineering was successfully supplemented with sequence-specific RNA-guided counterselection[23–25]. Combining λ Red recombineering and CRISPR-Cas with intricate multi-layered selection, counterselection, and colour screening enables large-scale genome editing such as genome recoding[26,27], global genome rearrangement[28], and interspecies gene transfer[29]. While these advanced genome editing techniques accomplish feats vastly beyond the capabilities of the one-step inactivation of chromosomal genes procedure[1], they all fall short in at least one respect. Namely, they are relatively difficult to implement. Requiring prior assembly of DNA fragments in yeast[27], elaborate *in vitro* DNA manipulations of helper plasmids[26,28] or episomes[29], or the presence of particular chromosomal mutations enabling counter-selectability[26–28], and a lack of efficient methods to eliminate helper plasmids[26–28], not only limit their adaptation to a few specialized laboratories but furthermore impede their capacity to iteratively edit bacterial genomes.

75 Here, we present *in vivo* cut'n'paste, an easy-to-use, scarless, large-scale
76 genome editing procedure that can be performed in wild-type cells. Leveraging
77 CRISPR-Cas cutting and λ Red recombineering with standardized synthetic DNA
78 fragments and an antibiotic marker recycling scheme completely removes the need
79 for *in vitro* DNA manipulations and, in conjunction with easily curable helper plasmids,
80 inherently enables iterative genome editing. We successfully relocated seven
81 functional genomic loci, including metabolic and essential genes, in *E. coli* of up to 140
82 kb in size.
83
84 **Results**
85
86 *In vivo* cut'n'paste
87
88 *In vivo* cut'n'paste enables to iteratively relocate genomic loci from the native
89 chromosome onto an episome (Supp. Movie 1). In each round, a genomic locus,
90 defined by two pairs of tandem homology regions (HR^A-HR^B and HR^C-HR^D) is excised
91 from the native chromosome and subsequently integrated in between the two
92 homology regions HR^{OL} and HR^{TET} in an episome (Fig. 1A). The entire relocation
93 process occurs inside the cell without requiring any *in vitro* DNA manipulations.
94 Each round of *in vivo* cut'n'paste consists of two phases: (1) pre-loading and
95 (2) relocation. During pre-loading, two synthetic DNA fragments, PRE^{epi} and PRE^{chr}
96 are sequentially integrated via homologous recombination into the episome and the
97 native chromosome, respectively, to equip cells with the required modification for the
98 relocation of the genomic locus (Fig. 1B; for details see Supplementary Materials,
99 Supp. Fig. 1-3). PRE^{epi} is flanked by the homology regions HR^{OL} and HR^{TET}. Directly

100 downstream of HR^{OL} PRE^{epi} contains the inside homology region HR^{B} of the genomic
101 locus. The two outside homology regions HR^{A} and HR^{D}, flanked by the constant cut
102 sites CS^{1} and CS^{2}, are located next to HR^{B} followed by a chloramphenicol marker
103 cassette and another constant cut site, CS^{3}. Integration of PRE^{epi} at HR^{OL} and HR^{TET}
104 into the episome truncates the tetracycline maker present in the episome and, thus,
105 renders the cells sensitive to tetracycline yet resistant to chloramphenicol. PRE^{chr} is
106 flanked by the homology region HR^{C} and HR^{D}, enabling its integration directly adjacent
107 to the genomic locus of interest in the native chromosome. In PRE^{chr} HR^{C} is directly
108 followed by 5'TET^{R}, encoding a short, non-functional fragment of the tetracycline
109 marker cassette, and HR^{TET}, a kanamycin marker cassette, flanked by the constant
110 cut sites CS^{1} and CS^{2} completes PRE^{chr}.

111 Once pre-loading is completed, cells contain all features required for relocation.
112 Firstly, the genomic locus in the native chromosome now contains HR^{B} on one side
113 and HR^{TET} on the other side. Secondly, the episome also contains both HR^{B} and
114 HR^{TET}. Thirdly, a "repair band aid", consisting of HR^{A} and HR^{D}, which can fix the
115 double-stranded break that will be left behind by the excision of the genomic locus
116 from the native chromosome, is present in the episome. Optionally, additional features
117 may be introduced during pre-loading, for example, a terminator in between HR^{C} and
118 5'TET^{R}, which would, once relocation is completed, partially insulate the relocated
119 locus from its new local environment on the episome.

120 Relocation is initiated by transforming the helper plasmid pCut. pCut expresses
121 guides targeting the constant cut sites CS^{1}, CS^{2}, and CS^{3} from a CRISPR array, as
122 well as a locus-specific single guide targeting the junction between HR^{A} and HR^{B}. The
123 sequence encoding the locus-specific single guide can effectively be inserted into pCut
124 via I-CreI cutting counterselection-aided plasmid recombineering, thus completely

125 eliminating the need for *in vitro* DNA manipulations (Supp. Fig. 4; see Methods for
126 details). RNA-guide Cas9 endonuclease activity introduces double-stranded breaks
127 into six targeted locations across the episome and the native chromosome. These
128 double-stranded breaks release the "repair band aid" from the episome and the
129 genomic locus from the native chromosome while simultaneously exposing the
130 required homology regions for integration of the "repair band aid" into the native
131 chromosome and of the genomic locus into the episome. Importantly, integration of
132 the genomic locus into the episome not only fixes its double-stranded break but,
133 additionally, leads to the restoration of the tetracycline marker cassette, enabling
134 positive selection of successful relocation events in the presence of tetracycline.
135 Notably, the chloramphenicol and kanamycin marker cassettes, along with the
136 constant cut sites, which were introduced during pre-loading, are innately lost during
137 the relocation process. Since the helper plasmid pCut is easily curable from the cells
138 by inducing guides targeting its origin of replication and its antibiotic marker, relocation
139 mutants are inherently capable of undergoing additional rounds of *in vivo* cut'n'paste
140 (Fig. 1A).
141
142 <u>Relocating a genomic locus</u>
143
144 In a first attempt, we employed *in vivo* cut'n'paste to relocate the *trg* locus from
145 its native genomic location in CNP^{empty} (*E. coli* K-12 MG1655 WT + EPI^{empty}) onto an
146 empty episome (EPI^{empty}). Candidates of successful relocations, CNP^{trg} (*E. coli* K-12
147 MG1655 Δ*trg* + EPI^{trg}), were obtained under tetracycline selection. After curing the
148 helper plasmids, we sent a CNP^{trg} isolate for short-read whole genome re-sequencing.
149 Mapping reads of the CNP^{trg} isolate to the parental CNP^{empty} reference depicted a

150 depth of coverage of about 400x across the native chromosome and the episome (Fig.
151 2A, middle row). No mutations relative to the parental strain CNP^{empty} were observed.
152 Yet, notably, pairs of adjacent bases at three genomic locations were not spanned by
153 any reads. Namely, both edges of the *trg* locus on the native chromosome (HR^{A}-HR^{B}
154 and HR^{C}-HR^{D}) and one edge of HR^{OL} on the episome lacked spanning reads (Fig. 2A,
155 middle row). No lack of spanning reads at these locations was observed when reads
156 of the parental CNP^{empty} strain were mapped to the CNP^{empty} reference (Fig. 2A, top
157 row). The lack of spanning reads in the CNP^{trg} isolates suggested that though the *trg*
158 locus was present, it was not located in its native genomic location, and, additionally,
159 that something had likely been inserted into the episome next to HR^{OL}. We updated
160 the reference map to include the desired relocation of the *trg* locus based on our *in*
161 *silico* design of the *in vivo* cut'n'paste process. When mapping the reads of the CNP^{trg}
162 isolate to this updated reference, the lack of spanning reads disappeared (Fig. 2A,
163 bottom row), suggesting that the *trg* locus had indeed been relocated from the native
164 chromosome onto the episome. Long-read sequencing of an extracted episome
165 further supported the successful relocation of the *trg* locus. We observed eleven reads
166 that spanned from the backbone of the episome across the entire, mutation-free *trg*
167 locus.

168 The *trg* locus (≈1.8 kb) contains a single gene encoding the transmembrane
169 Trg chemoreceptor. Trg mediates chemotaxis toward D-ribose[30]. Expansion dynamics
170 of trg-relocated CNP^{trg} isolates in glycerol minimal medium soft agar plates in the
171 presence (dark colours) and absence (light colours) of D-ribose were comparable to
172 the behaviour of CNP^{empty} which encodes *trg* in its native genomic context (Fig. 2B,
173 black and blue). When the episome EPI^{trg} was removed, the expansion dynamics of
174 the CNP^{trg} isolates in presence of D-ribose were reduced to a slower aerotactic

175 behaviour[31] as observed in the absence of D-ribose, indicating that the ability to sense
176 D-ribose was lost along with the episome (Fig. 2B, red). As an additional test, we
177 separately transferred the episomes EPI^{trg} and EPI^{empty} into a strain background in
178 which the *trg* gene had been disrupted by a kanamycin marker cassette (*E. coli* K-12
179 MG1655 Δtrg::KanR). As expected, faster expansion dynamics in presence of D-
180 ribose were restored only when EPI^{trg}, containing the relocated *trg* locus, was present
181 (Fig. 2C).
182
183 Iteratively relocating genomic loci
184
185 Upon completion of the *trg* relocation the antibiotic marker cassettes and
186 constant cut sites acquired during pre-loading were inherently lost and the tetracycline
187 marker cassette on the episome was restored (Fig. 1A). Thus, CNP^{trg} was capable of
188 undergoing additional rounds of *in vivo* cut'n'paste. We iteratively applied *in vivo*
189 cut'n'paste to relocate two additional genomic loci using CNP^{trg} as a starting point (Fig.
190 3A). First, the *aer* locus (≈1.7 kb) encoding the aerotaxis receptor Aer was relocated,
191 resulting in $CNP^{trg,aer}$. Then, the *tsr* locus (≈1.8 kb) encoding the Serine and aerotaxis
192 receptor Tsr was relocated, resulting in $CNP^{trg,aer,tsr}$. The relocated loci were mutation-
193 free. The triple-relocated $CNP^{trg,aer,tsr}$ qualitatively retained the colony expansion
194 behaviour of its ancestor CNP^{empty} (Fig. 3 B,C). Furthermore, separately transferring
195 the episomes EPI^{empty} (Fig. 3D), EPI^{trg} (Fig. 3 E), and $EPI^{trg,aer}$ (Fig. 3F) into a Δ*trg* Δ*aer*
196 Δ*tsr* chromosomal background partially restored the colony expansion behaviour as
197 expected.
198 In an independent attempt, using CNP^{empty} as a starting point, we successfully
199 iteratively relocated two genomic loci (≈8.1 kb and ≈7.9 kb) encoding 10 genes

200 involved in maltose utilisation (Fig. 4A). Again, the relocated loci were mutation-free.
201 The relocation mutants CNP^{malHM} and $CNP^{malHM,malQT}$ retained their ability to uptake
202 and ferment maltose (Fig. 4B).
203
204 Relocating a large genomic locus and an essential gene
205
206 To illustrate the utility of *in vivo* cut'n'paste to process very large genomic loci,
207 we attempted to relocate the *macB-nadA* locus (≈140 kb) containing 135 genes from
208 its native genomic location onto the episome. The relocation was successful and
209 mutation-free.
210 As opposed to traditional sequential deletion and complementation procedures,
211 all genes, including those located inside the genomic locus that is edited, remain
212 present throughout the entire *in vivo* cut'n'paste process. Hence, *in vivo* cut'n'paste
213 should be well suited for manipulating hard-to-edit genomic features such as essential
214 genes. Indeed, we successfully relocated the essential gene *folA* onto an episome,
215 resulting in CNP^{folA}. An episome containing an essential gene that is simultaneously
216 absent from the native chromosome, as is the case in CNP^{folA}, is expected to be stably
217 maintained even in the absence of antibiotic selection. As a direct test, we replaced
218 the tetracycline antibiotic marker cassette of the episome in CNP^{folA} by a green
219 fluorescent reporter (GFP), resulting in $CNP^{folA,TetR::gfp}$. During serial culturing of
220 $CNP^{folA,TetR::gfp}$ in LB medium without any antibiotic supplementation 100% of the
221 bacterial population depicted green fluorescence (Fig. 5, Supp. Fig. 5), clearly
222 confirming the stable maintenance of the antibiotic marker-free episome as an extra-
223 chromosomal element (see Methods for details).
224

225 Unintended side effects

226

227 All relocated genomic loci were mutation-free. Yet, some of the relocation
228 mutants acquired background mutations in their native chromosome while undergoing
229 *in vivo* cut'n'paste (Supp. Table 7). In general, it remains unclear whether these
230 background mutations have any functional relevance or are merely coincidental
231 genetic drift.

232

233

234 **Discussion**

235

236 We developed *in vivo* cut'n'paste - an easy-to-use procedure for large-scale
237 genome editing in *E. coli* that does not rely on any *in vitro* DNA manipulations. We
238 successfully relocated multiple genomic loci encoding metabolic and essential genes,
239 generated a genomic alteration >100 kb, and performed up to three consecutive edits
240 in a single strain. All relocated loci, including the 140 kb *macB-nadA* relocation, were
241 mutation-free. Due to its robustness, as each construction step is accompanied by
242 positive antibiotic selection, *in vivo* cut'n'paste is well-suited for implementation via lab
243 automation.

244 *In vivo* cut'n'paste can be adapted to precisely characterise gene function. One
245 round of *in vivo* cut'n'paste effectively creates a complementation mutant in which the
246 gene of interest is located on a single-copy extra-chromosomal replicon. Thus, *in vivo*
247 cut'n'paste inherently overcomes the limitations of traditional sequential "delete and
248 complement" procedures hampering the investigation of essential genes while also
249 circumventing any convoluting copy number effects. By iterating *in vivo* cut'n'paste

250 arbitrary modules of genes, like those obtained here (Fig. 3 & 4), can be created. For
251 example, genes that are co-modulated or functionally connected but natively
252 dispersed across the chromosome can easily be clustered onto a single extra-
253 chromosomal replicon[32]. Utilising the origin of transfer (oriT) encoded on the episome,
254 such modules can be transferred to other strains or species via conjugation (oriT)
255 without the size-limitations imposed by transformation[33], and, hence, potential future
256 adaptation of *in vivo* cut'n'paste to other bacterial species will only broaden its utility
257 (see Supplementary Text). Moreover, relocations of essential genes, such as
258 demonstrated here in CNP^{folA}, are highly desirable in the context of large-scale
259 fermentation applications in green biotechnology as they allow to stably maintain
260 transferable extra-chromosomal replicons in production strains without incurring the
261 metabolic costs and legal requirements associated with continual antibiotic selection[34].
262 *In vivo* cut'n'paste empowers experimenters to refactor genomes, akin to the
263 seven genomic edits demonstrated here, without requiring expert knowledge in
264 advanced bacterial genome engineering methodologies. Consequently, *in vivo*
265 cut'n'paste can contribute to diverse areas of research, from studying bacterial
266 motility[3,4], to examining stringent response[5], to dissecting the interplay of growth and
267 gene expression[6], to investigating two-component systems[7], to elucidating the
268 evolution of antibiotic resistance[8], to characterizing chromosomal organisation[9], and
269 beyond. We envision that *in vivo* cut'n'paste can also enable novel experimental
270 evolution approaches by enabling the "freezing" of parts of the bacterial genome
271 during adaptation. By relocating sets of genes onto an episome that is periodically
272 "reset", i.e., replaced by its unevolved ancestor, evolutionary space can thus be
273 restricted as chosen by the experimenter. The ability to relocate genomic loci makes
274 *in vivo* cut'n'paste particularly well-suited for studying how genome architecture

275 impacts cell functioning, a topic of which we are only beginning to scratch the
276 surface[35,36]. Ultimately, *in vivo* cut'n'paste opens up novel avenues to obtain minimal
277 genomes via bottom-up approaches to complement traditional, ineffectual top-down
278 genome reduction attempts.
279
280 **Materials and Methods**
281
282 See "Methods" document.
283
284 **Data availability**
285
286 All data is publicly available (see Supp. Table 8 for details). All custom-written
287 scripts are supplied in the Supplementary Materials. Sequencing data was deposited
288 at NCBI (PRJNA1271273; see Supp. Table 1, 2, and 3 for assession numbers).
289 Sequence-verified, annotated maps of episomes and helper plasmids are supplied in
290 the Supplementary Materials. Strains are available for non-commercial use through
291 the Belgian Coordinated Collections of Microorganisms (https://bccm.belspo.be; see
292 Supp. Table 1 for assession numbers). Helper plasmids are available for non-
293 commercial use through Addgene (https://addgene.org; see Supp. Table 3 for
294 assession numbers).
295
296 **Authors' contributions**
297
298 SG, ABB, and MHHN developed the initial concept of *in vivo* cut'n'paste. SG
299 implemented, tested, and refined *in vivo* cut'n'paste iteratively. SG performed all

experiments, analysed the data, and wrote the manuscript. All authors read the final version of the manuscript.

Conflicts of interest

The authors do not have any conflicts of interest to disclose.

Acknowledgement and funding information

The authors would like to acknowledge Valeria Stavila (The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark) for technical assistance. S.G. would like to acknowledge Alex T. Nielsen (The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark) for hosting the final period of research activity in his laboratory. *E. coli* K-12 MG1655 WT was a kind gift of Tom Shimizu (AMOLF, Amsterdam, The Netherlands). The title and abstract of this manuscript pay homage to the article “One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products”. This work was supported by The Novo Nordisk Foundation.

References

1. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).

2. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
3. Chapman, M. R. *et al.* Role of *Escherichia coli* Curli Operons in Directing Amyloid Fiber Formation. *Science* **295**, 851–855 (2002).
4. Smith, K. D. *et al.* Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nat. Immunol.* **4**, 1247–1253 (2003).
5. Traxler, M. F. *et al.* The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol. Microbiol.* **68**, 1128–1148 (2008).
6. Klumpp, S., Zhang, Z. & Hwa, T. Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell* **139**, 1366–1375 (2009).
7. Zhou, L., Lei, X.-H., Bochner, B. R. & Wanner, B. L. Phenotype MicroArray Analysis of *Escherichia coli* K-12 Mutants with Deletions of All Two-Component Systems. *J. Bacteriol.* **185**, 4956–4972 (2003).
8. Liu, J., Gefen, O., Ronin, I., Bar-Meir, M. & Balaban, N. Q. Effect of tolerance on the evolution of antibiotic resistance under drug combinations. *Science* **367**, 200–204 (2020).
9. Lau, I. F. *et al.* Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* **49**, 731–743 (2003).
10. Kolisnychenko, V. *et al.* Engineering a Reduced *Escherichia coli* Genome. *Genome Res.* **12**, 640–647 (2002).
11. Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J. & Voigt, C. A. Escherichia coli “Marionette” strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.* **15**, 196–204 (2019).

12. Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U. & Lee, S. Y. Systems metabolic engineering of *Escherichia coli* for L -threonine production. *Mol. Syst. Biol.* **3**, 149 (2007).

13. Atsumi, S. *et al.* Metabolic engineering of Escherichia coli for 1-butanol production. *Metab. Eng.* **10**, 305–311 (2008).

14. Winans, S. C., Elledge, S. J., Krueger, J. H. & Walker, G. C. Site-directed insertion and deletion mutagenesis with cloned fragments in Escherichia coli. *J. Bacteriol.* **161**, 1219–1221 (1985).

15. Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. New method for generating deletions and gene replacements in Escherichia coli. *J. Bacteriol.* **171**, 4617–4622 (1989).

16. Russell, C. B., Thaler, D. S. & Dahlquist, F. W. Chromosomal Transformation of Escherichia coli recD Strains with Linearized Plasmidst. *J BACTERIOL* **171**, (1989).

17. Metcalf, W. W. *et al.* Conditionally Replicative and Conjugative Plasmids Carrying lacZa for Cloning, Mutagenesis, and Allele Replacement in Bacteria. *Plasmid* **35**, 1–13 (1996).

18. Link, A. J., Phillips, D. & Church, G. M. Methods for generating precise deletions and insertions in the genome of wild-type Escherichia coli: application to open reading frame characterization. *J. Bacteriol.* **179**, 6228–6237 (1997).

19. Kato, C., Ohmiya, R. & Mizuno, T. A Rapid Method for Disrupting Genes in the Escherichia coli Genome. *Biosci. Biotechnol. Biochem.* **62**, 1826–1829 (1998).

20. Zhang, Y., Buchholz, F., Muyrers, J. P. P. & Stewart, A. F. A new logic for DNA engineering using recombination in Escherichia coli. *Nat. Genet.* **20**, 123–128 (1998).

21. Posfai, G., Kolisnychenko, V., Bereczki, Z. & Blattner, F. R. Markerless gene replacement in Escherichia coli stimulated by a double-strand break in the chromosome. *Nucleic Acids Res.* **27**, 4409–4415 (1999).

22. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

23. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).

24. Pyne, M. E., Moo-Young, M., Chung, D. A. & Chou, C. P. Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in Escherichia coli. *Appl. Environ. Microbiol.* **81**, 5103–5114 (2015).

25. Jiang, Y. *et al.* Multigene Editing in the Escherichia coli Genome via the CRISPR-Cas9 System. *Appl. Environ. Microbiol.* **81**, 2506–2514 (2015).

26. Wang, K. *et al.* Defining synonymous codon compression schemes by genome recoding. *Nature* **539**, 59–64 (2016).

27. Zürcher, J. F. *et al.* Continuous synthesis of E. coli genome sections and Mb-scale human DNA assembly. *Nature* **619**, 555–562 (2023).

28. Wang, K., De La Torre, D., Robertson, W. E. & Chin, J. W. Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly. *Science* **365**, 922–926 (2019).

29. Teufel, M., Klein, C. A., Mager, M. & Sobetzko, P. A multifunctional system for genome editing and large-scale interspecies gene transfer. *Nat. Commun.* **13**, 3430 (2022).

30. Kondoh, H., Ball, C. B. & Adler, J. Identification of a methyl-accepting chemotaxis protein for the ribose and galactose chemoreceptors of Escherichia coli. *Proc. Natl. Acad. Sci.* **76**, 260–264 (1979).

31. Zhulin, I. B., Rowsell, E. H., Johnson, M. S. & Taylor, B. L. Glycerol elicits energy taxis of Escherichia coli and Salmonella typhimurium. *J. Bacteriol.* **179**, 3196–3201 (1997).

32. Tan, J. *et al.* Independent component analysis of E. coli's transcriptome reveals the cellular processes that respond to heterologous gene expression. *Metab. Eng.* **61**, 360–368 (2020).

33. Hanahan, D. Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* **166**, 557–580 (1983).

34. Sevillano, L., Díaz, M. & Santamaría, R. I. Development of an antibiotic marker-free platform for heterologous protein production in Streptomyces. *Microb. Cell Factories* **16**, 164 (2017).

35. El Houdaigui, B. *et al.* Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling. *Nucleic Acids Res.* **47**, 5648–5657 (2019).

36. Dame, R. T., Rashid, F.-Z. M. & Grainger, D. C. Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.* **21**, 227–242 (2020).

Figures and figure legends

See "Figures" document.

Supplementary Materials

See “SupplementaryMaterials” file.

Iterative *in vivo* cut'n'paste of chromosomal loci in *Escherichia coli* K-12 using synthetic DNA

Figures and figure legends

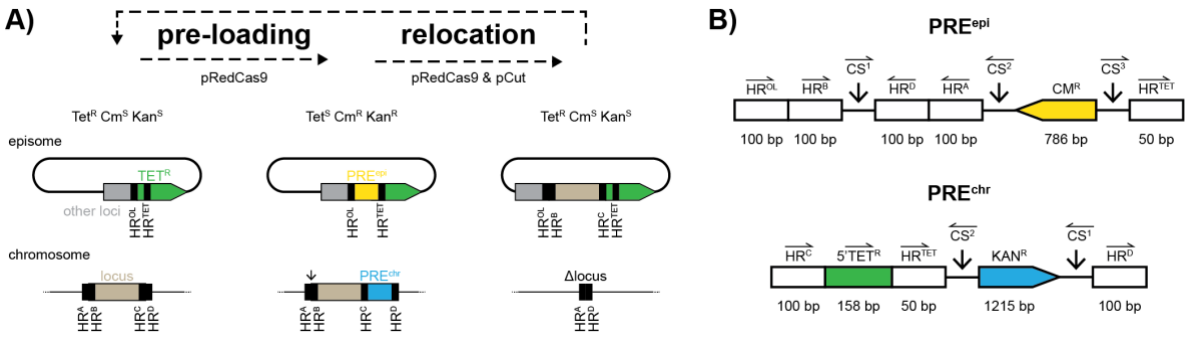


Figure 1: Iterative *in vivo* cut'n'paste

a) Overview of the iterative *in vivo* cut'n'paste workflow. **b)** Illustration of the synthetic pre-loading DNA fragments and their features. A more detailed description of the design choices is provided in the Supplementary Material. Half-arrows indicate the relative orientations of features. Solid, downwards arrows mark cut sites. pRedCas9 and pCut: easily curable helper plasmids, HR: homology region, CS: cut site, TET^R: tetracycline marker cassette, CM^R: chloramphenicol marker cassette, KAN^R: kanamycin marker cassette, 5'TET^R: non-functional fragment of TET^R, Tet: tetracycline, Cm: chloramphenicol, Kan: kanamycin, R: resistant, S: susceptible, and bp: base pairs. Not to scale.

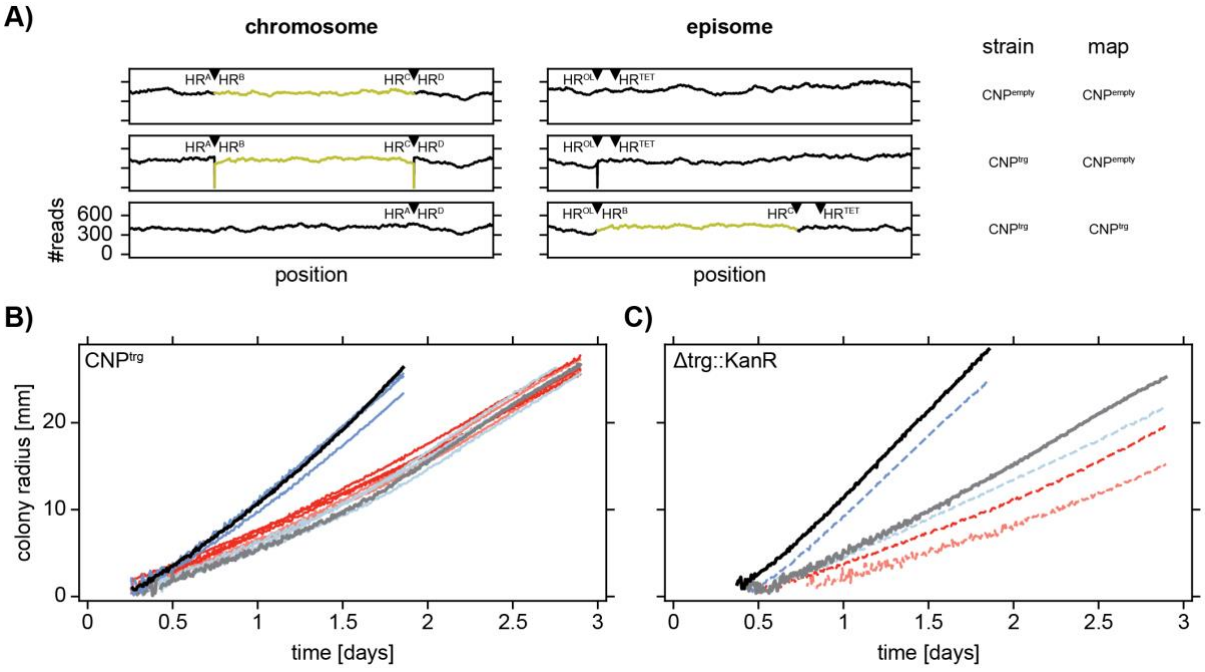


Figure 2: *In vivo* cut'n'paste of the trg locus

a) Number of mapped reads in 3.2 kb segments of the native chromosome and the episome. Only reads mapping to a position and to both of its direct neighbors are counted. Strain identities and reference maps as indicated within the panel. The trg locus is highlighted in yellow. Triangles mark locations of homology regions. **b-c)** Colony expansion as function of time in glycerol minimal medium with (dark) or without (light) 100 μM D-ribose. **b)** CNP^{trg} (blue, solid), CNP^{trg} without EPI^{trg} (red, solid), and CNP^{empty} (black, solid). Three independent CNP^{trg} isolates and a single replicate of CNP^{empty}. **c)** *E. coli* K-12 MG1655 Δtrg::KanR with EPI^{trg} (blue, dashed), *E. coli* K-12 MG1655 Δtrg::KanR with EPI^{empty} (red, dashed), and CNP^{empty} (black, solid). A single replicate each.

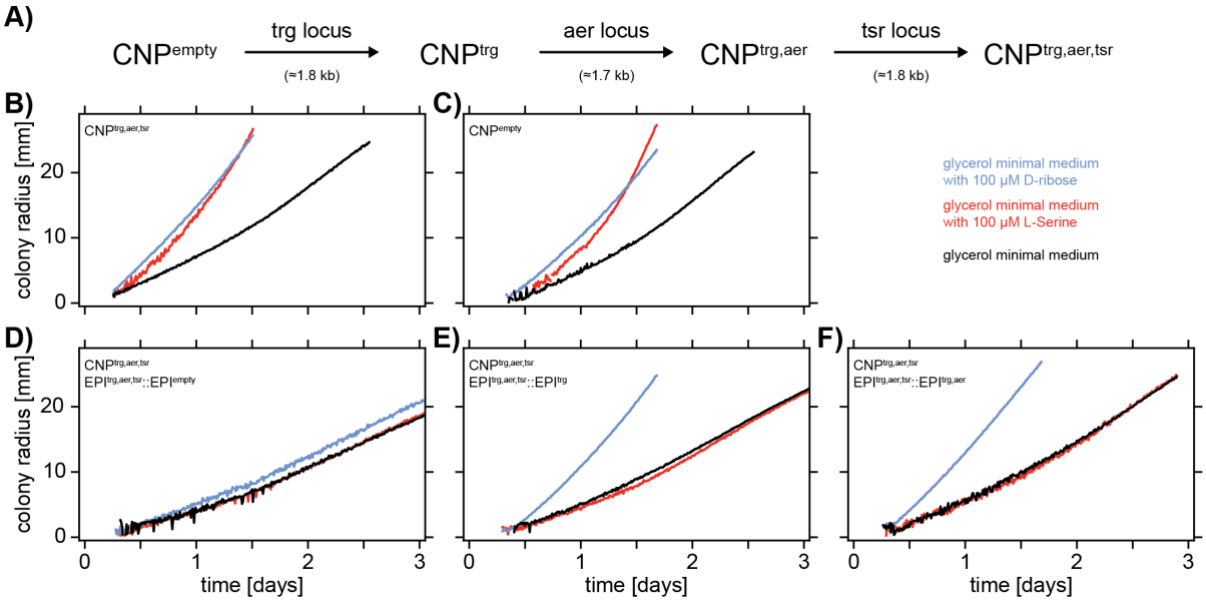


Figure 3: Iterative *in vivo* cut'n'paste of three chemoreceptor loci

a) Illustration of iterative *in vivo* cut'n'paste of the trg, aer, and tsr loci. **b-f)** Colony expansion as function of time in glycerol minimal medium with 100 µM D-ribose (blue), with 100 µM L-Serine (red), or without any additions (black). Strain identities as defined within the individual panels.

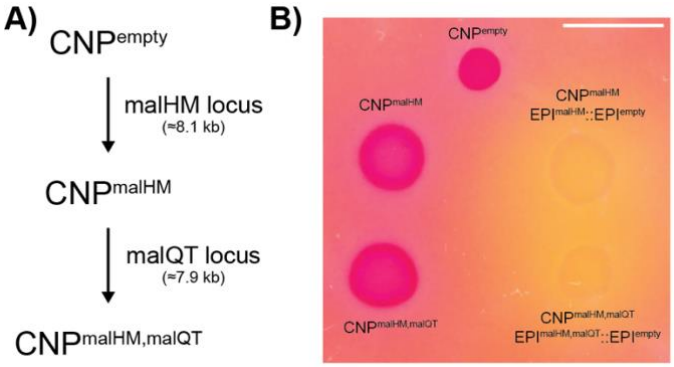


Figure 4: Iterative *in vivo* cut'n'paste of two maltose-utilization loci

a) Illustration of iterative *in vivo* cut'n'paste of the malHM (i.e., *malEFGH* and *malKlamBmalM*) and malQT (i.e., *malPQ* and *malT*) loci. **b)** Maltose utilization phenotypes on maltose-MacConkey+Tet/agar at 37 °C after 9 hours. Dark red: cells can utilize maltose, pale yellow: cells cannot utilize maltose. Strain identities as defined within the panel. Scale bar: 10 mm.

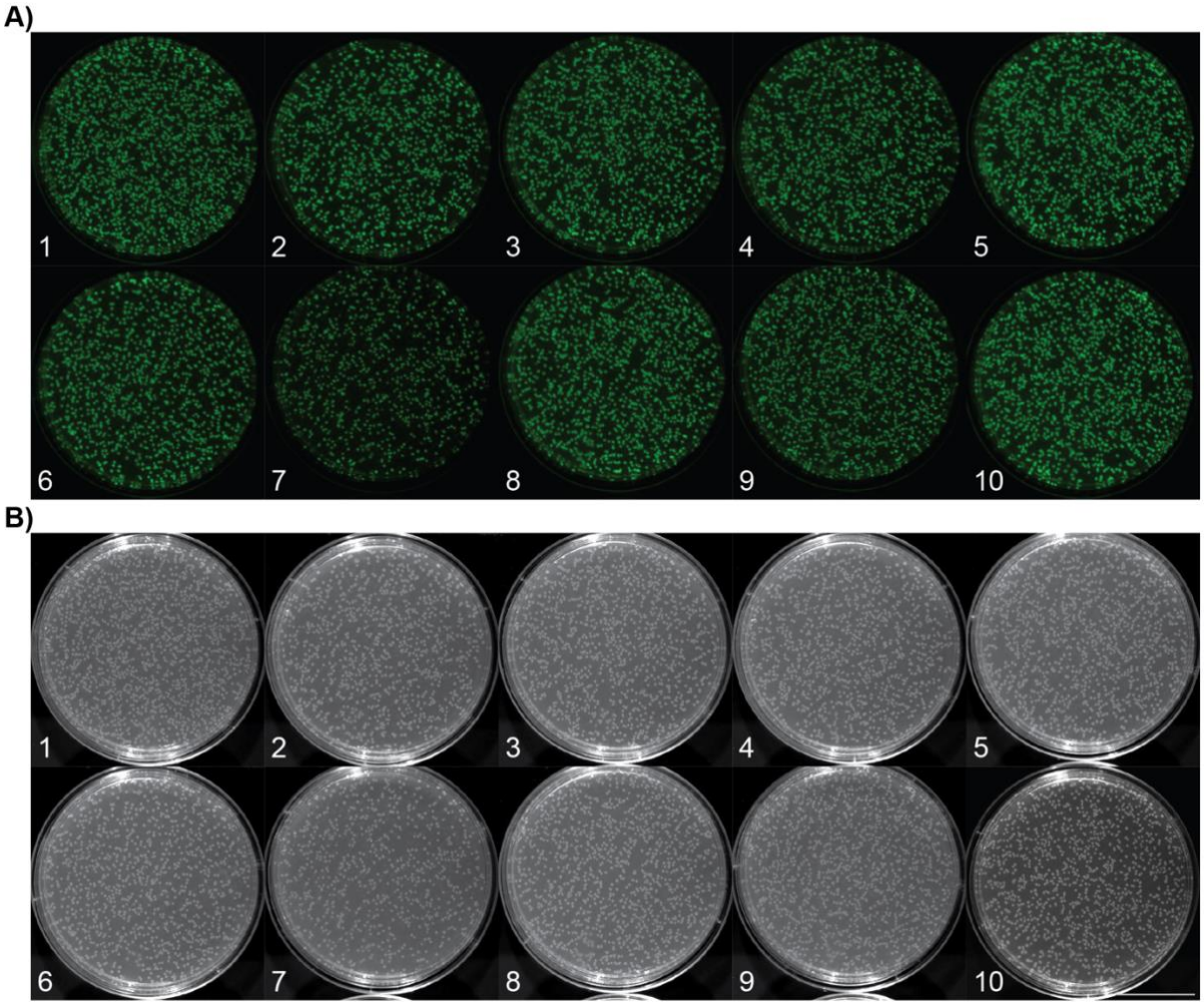
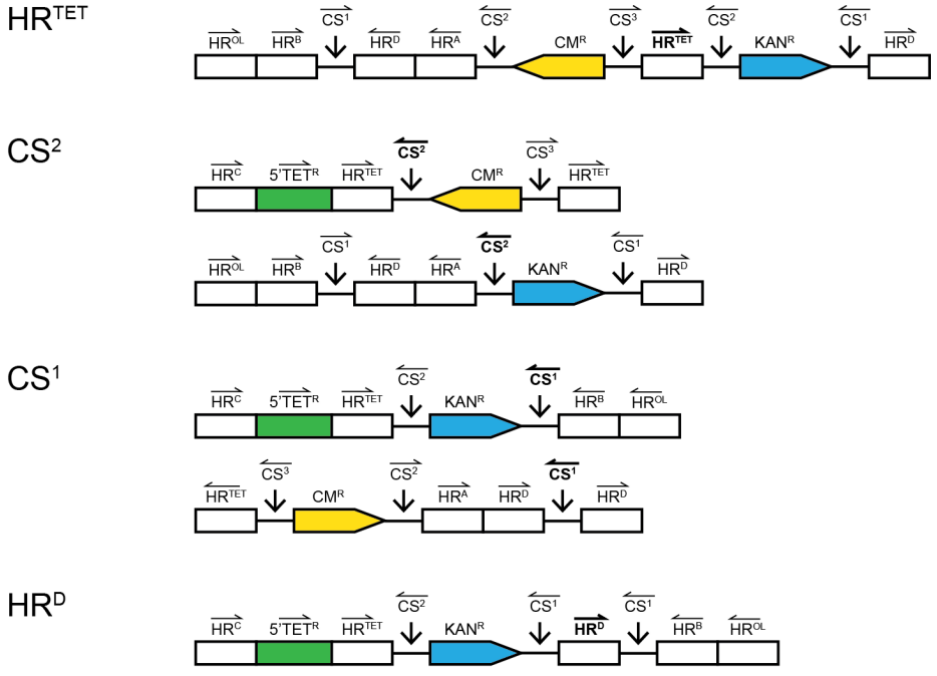


Figure 5: Maintenance of an antibiotic marker-free episome during serial culturing **a)** False-colored fluorescence (GFP; excitation: 455-485 nm, emission: 508-557 nm, exposure time: 15 msec) and **b)** brightfield images of $CNP^{folA,TetR::gfp}$ plated onto LB agar at transfers during serial culturing in LB without any antibiotic supplementation. Numerals within the panels indicate transfer numbers. Cultures reached comparable densities of $OD_{600} \approx 4$ at each transfer. Scale bar: 30 mm.

59



60

61

62 **Supplementary Figure 1:** Merging of synthetic DNA fragments during simultaneous

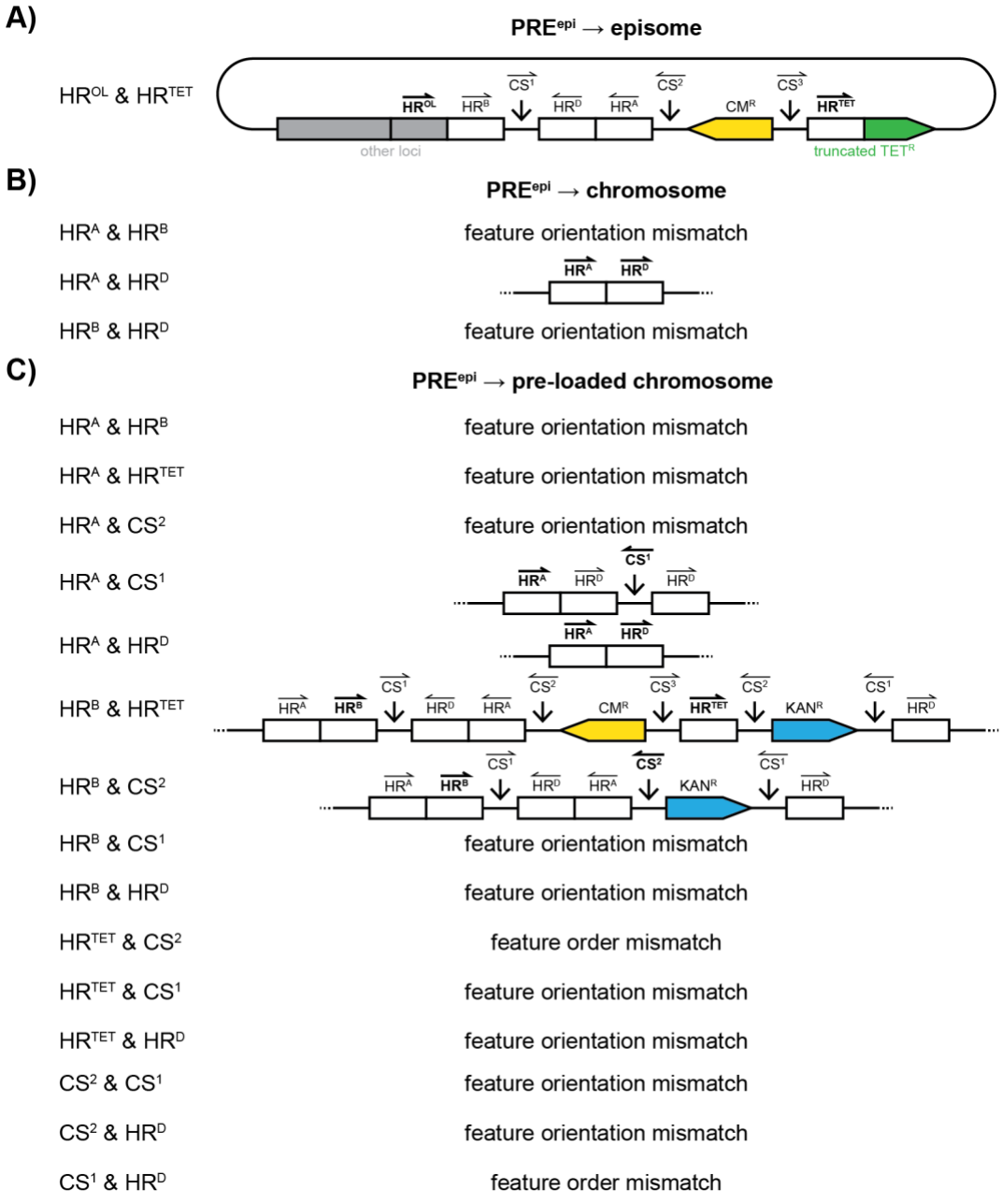
63 pre-loading

64 Illustration of merged dimers of the synthetic DNA fragments PRE^{epi} and PRE^{chr} due

65 to sequence homology at shared features. Merging of PRE^{epi} and PRE^{chr} causes

66 feature loss. Locations of merging are highlighted in bold. Definitions as defined in Fig.

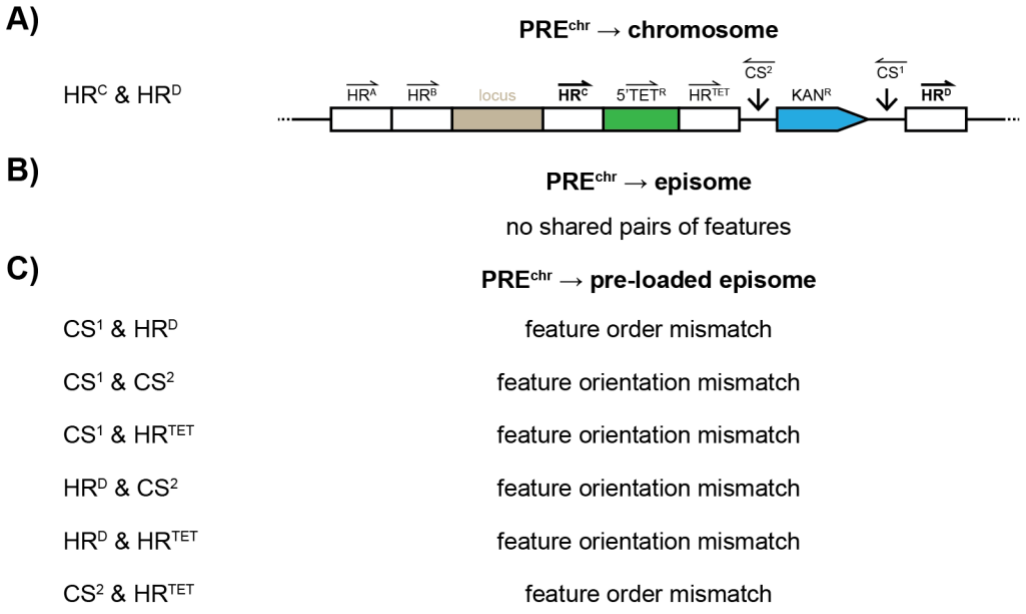
67 1. Not to scale.



Supplementary Figure 2: Integration of synthetic DNA fragment PRE^{epi} during sequential pre-loading

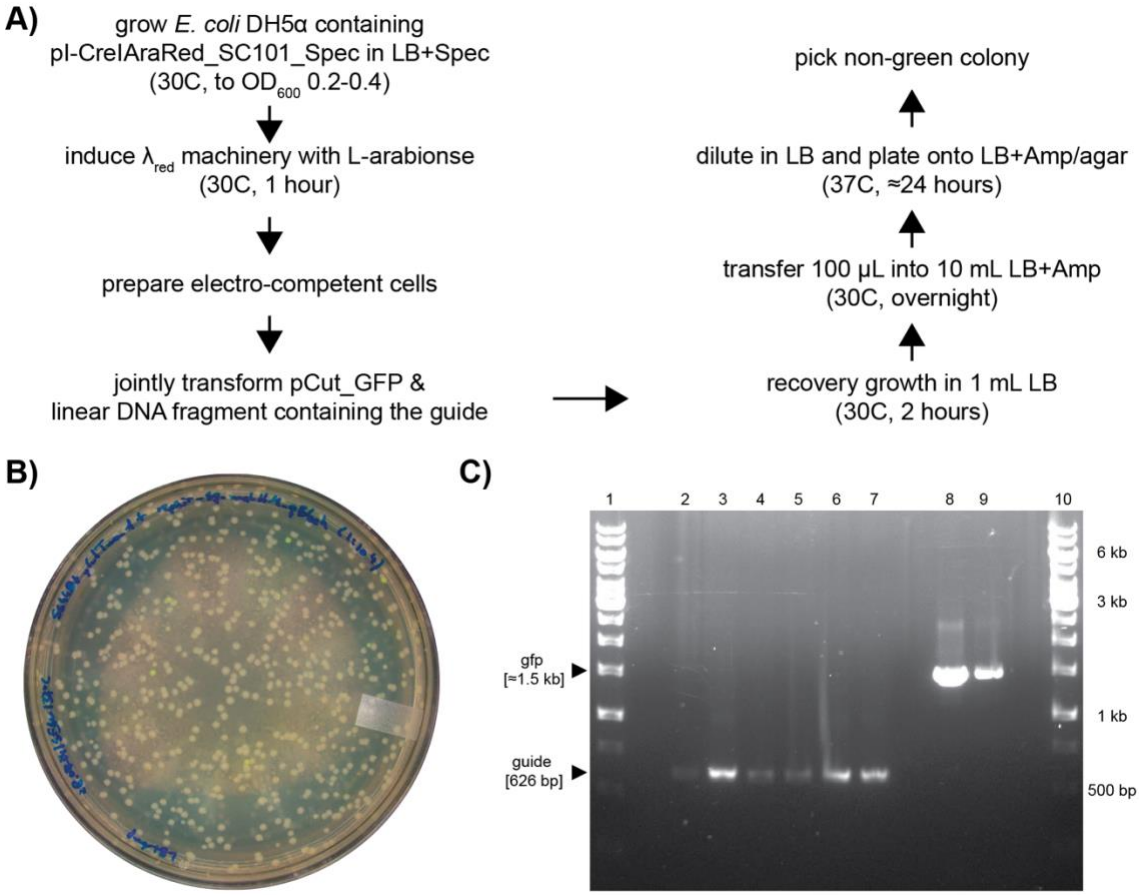
Illustration of integration events of the synthetic DNA fragment PRE^{epi} due to homology at pairs of shared features into A) the episome, B) the native chromosome, or C) a native chromosome containing PRE^{chr}. Integration events may cause feature loss. Many pairs of shared features cannot undergo successful integration due to their

relative orientations (feature orientation mismatch) or locations (feature order mismatch). Homologies used for integration are highlighted in bold. Definitions as defined in Fig. 1. Not to scale.



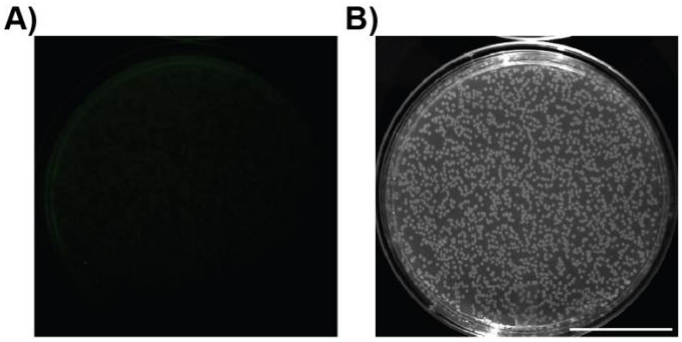
Supplementary Figure 3: Integration of synthetic DNA fragment PRE^{chr} during sequential pre-loading

Illustration of integration events of the synthetic DNA fragment PRE^{chr} due to homology at pairs of shared features into A) the native chromosome, B) the episome, or C) an episome containing PRE^{epi}. Many pairs of shared features cannot undergo successful integration due to their relative orientations (feature orientation mismatch) or locations (feature order mismatch). Homologies used for integration are highlighted in bold. Definitions as defined in Fig. 1. Not to scale.



Supplementary Figure 4: Insertion of locus-specific guide sequence into pCut_GFP via I-CreI cutting counterselection-aided plasmid recombineering

a) Overview of I-CreI cutting counterselection-aided plasmid recombineering. **b)** Image of transformants (1:10^4 dilution) plated onto LB-agar containing ampicillin [100 mg/L] (plate diameter: 90 mm). **c)** Agarose gel electrophoresis of diagnostic PCRs across the engineering location in pCut (primers: diag_guide_A, diag_guide_B). Non-green transformants (lanes 2-7), green transformants (lanes 8 and 9), and “gene ruler [1 kb]” DNA ladder (lanes 1 and 10). Expected band sizes are indicated by black triangles.



Supplementary Figure 5: *E. coli* K-12 MG1655 WT plated on LB agar

a) False-colored fluorescence (GFP; excitation: 455-485 nm, emission: 508-557 nm, exposure time: 15 msec) and **b)** brightfield images of *E. coli* K-12 MG1655 WT plated onto LB agar. Identical contrast settings as in Fig. 5. Scale bar: 30 mm.