AEM
Journals.ASM.org

Multigene Editing in the *Escherichia coli* Genome via the CRISPR-Cas9 System

Yu Jiang,[a,b] Biao Chen,[a,b] Chunlan Duan,[a] Bingbing Sun,[a,b] Junjie Yang,[a,b] Sheng Yang[a,b,c]

Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China[a]; Shanghai Research Center of Industrial Biotechnology, Shanghai, China[b]; Shanghai Collaborative Innovation Center for Biomanufacturing Technology, Shanghai, China[c]

An efficient genome-scale editing tool is required for construction of industrially useful microbes. We describe a targeted, continual multigene editing strategy that was applied to the *Escherichia coli* genome by using the *Streptococcus pyogenes* type II CRISPR-Cas9 system to realize a variety of precise genome modifications, including gene deletion and insertion, with a highest efficiency of 100%, which was able to achieve simultaneous multigene editing of up to three targets. The system also demonstrated successful targeted chromosomal deletions in *Tatumella citrea*, another species of the *Enterobacteriaceae*, with highest efficiency of 100%.

Metabolic engineering is widely applied to modify *Escherichia coli* to produce industrially relevant biofuels or biochemicals, including ethanol (1), higher alcohols (2), fatty acids (3), amino acids (4), shikimate precursors (5), terpenoids (6), polyketides (7), and polymeric precursors of 1,4-butanediol (8). An important example of a successful metabolic engineering project is the modification of *E. coli* to produce 1,3-propanediol, which was developed by Genencor and DuPont (9) and led to a commercial process. This industrially optimized strain required up to 26 genomic modifications, including insertions, deletions, and regulatory modifications. Such large numbers of genome editing targets require efficient tools to perform time-saving sequential manipulations or multiplex manipulations.

A wide variety of tools for targeted gene editing, which can be classified into homologous recombination and group II intron retrohoming, are available for *E. coli* (10, 11). The efficiency of introduction of mutations mediated by homologous recombination can be improved (i) by using counterselection markers, such as the typical *sacB*-based method (12), and (ii) by improving the frequency of homologous recombination by using phage-derived recombinases (RecET and λ-Red) (13–15), applying double-stranded (16, 17) or single-stranded donor DNAs (18), or inducing double-stranded breaks (DSBs) in a chromosomal target using I-SceI (12, 19, 20). The λ-Red recombinase method (13) and group II intron retrotransposition (21) leave scars in the genome that limit their application in allelic exchange. Of all the methods mentioned above, only single-stranded-DNA (ssDNA)-based gene modification mediated by λ-Red was further developed as a multiplex genome editing tool, known as multiplex automated genome engineering (MAGE) (22, 23), which greatly facilitates genome-scale engineering. However, the short ssDNA oligonucleotide-mediated MAGE has advantages in allelic exchange-based genome mutation but has challenges regarding targeted multiple gene insertions over a certain length (22).

The clustered regularly interspaced short palindromic repeats–CRISPR-associated system (CRISPR-Cas system) was used recently as efficient genome engineering technology in several prokaryotes and eukaryotes, including (but not limited to) *E. coli* (24), *Saccharomyces cerevisiae* (25), *Streptomyces* spp. (26), higher plants (27), *Bombyx mori* (28), *Drosophila* (29), and human cell lines (30–32). The type II CRISPR-Cas system from *Streptococcus pyogenes* uses a maturation CRISPR RNA (crRNA) and *trans*-activating crRNA (tracrRNA) guiding the nuclease Cas protein 9 (Cas9) to the target of any DNA sequence, known as a protospacer, with a protospacer-adjacent motif (PAM) present at the 3′ end (NGG in the case of *S. pyogenes*, where N represents any nucleotide) (33). In genome editing cases, the 20-bp complementary region (N_{20}) with the requisite NGG PAM matching genomic loci of interest was programmed directly into a heterologously expressed CRISPR array, and fused crRNA and tracrRNA as a single synthetic guide RNA (sgRNA) transcript obviated the need for processing the transcribed CRISPR array (pre-crRNA) into individual crRNA components (31).

In *E. coli*, the CRISPR-Cas9 system has been demonstrated to apply allelic exchange with efficiency as high as 65% ± 14% (24) and to control gene expression via a nuclease-deficient Cas9 protein (34, 35). No detailed method for applying the CRISPR-Cas9 system in precise genome editing, including gene insertions and knockouts, has been published. Therefore, we developed a CRISPR-Cas9 system-based continual genome editing strategy, including gene insertions and knockouts of both single and multiple (up to three) targets, and expanded the system to include *Tatumella citrea*, another species of the *Enterobacteriaceae*, for continual gene deletions.

MATERIALS AND METHODS

Strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are given in Table 1. *E. coli* DH5α was used as a cloning host, and *E. coli* MG1655 or *T. citrea* DSM 13699 was used in the

Received 10 December 2014 **Accepted** 17 January 2015

Accepted manuscript posted online 30 January 2015

Citation Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. 2015. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. Appl Environ Microbiol 81:2506–2514. doi:10.1128/AEM.04023-14.

Editor: R. M. Kelly

Address correspondence to Sheng Yang, syang@sibs.ac.cn.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.04023-14.

doi:10.1128/AEM.04023-14

TABLE 1 Strains and plasmids used in this study[a]

Strain or plasmid	Characteristics	Source or reference
Strains		
E. coli DH5α	F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG* ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 hsdR17* (r_K^- m_K^+) λ$^-$	TaKaRa
E. coli MG1655	K-12; F$^-$ λ$^-$ *rph-1*	CGSC 6300
Streptococcus pyogenes MGAS5005	M1 serotype, wild type	ATCC BAA-947
Tatumella citrea		
DSM 13699		DSMZ
MGly1	MG1655 Δ*cadA*::*cat*	This study
MGly2	MGly1 Δ*cat*::*cadA*^{P}	This study
1655ΔcadA	MG1655 Δ*cadA*	This study
1655ΔcadAΔmaeAΔmaeB	MG1655 Δ*cadA* Δ*maeA* Δ*maeB*	This study
1655ΔmaeAΔmaeB	MG1655 Δ*maeA* Δ*maeB*	This study
1655ΔyjcS::ybaS	MG1655 Δ*yjcS*::*ybaS*	This study
1655ΔyjcS::evgAS	MG1655 Δ*yjcS*::*evgAS*	This study
1655ΔmaeB::gltPΔmaeA	MG1655 Δ*maeB*::*gltP* Δ*maeA*	This study
13699ΔtkrA	DSM 13699 Δ*tkrA*	This study
13699ΔtkrAΔglk	DSM 13699 Δ*tkrA* Δ*glk*	This study
General plasmids		
pKD46	*repA101*(Ts) *bla araC* P_{araB}-*Red*	13
pSU2718	p15A *cat*	52
pTrc99A	*pMB1 bla lacI*^{q}	53
pIJ778	*bla aadA FRT*	54
pKD46K	*repA101*(Ts) *kan araC* P_{araB}-*Red*	20
pTrc99A-spec	*pMB1 aadA lacI*^{q}	This study
Crispr-Cas9 system plasmids		
pCas series		
pCB001	p15A *cm* P_{cas}-*cas9*	This study
pCas	*repA101*(Ts) *kan* P_{cas}-*cas9* P_{araB}-*Red lacI*^{q} P_{trc}-sgRNA-*pMB1*	This study
pCasΔ*cas9*	*repA101*(Ts) *kan* P_{araB}-*Red lacI*^{q} P_{trc}-sgRNA-*pMB1*	This study
pTarget series harboring sgRNAs, with or without donor DNAs		
pTargetF-*pMB1*	*pMB1 aadA* sgRNA-*pMB1*	This study
pTargetF-*cadA*	*pMB1 aadA* sgRNA-*cadA*	This study
pTargetF-*kefB*	*pMB1 aadA* sgRNA-*kefB*	This study
pTargetF-*yjcS*	*pMB1 aadA* sgRNA-*yjcS*	This study
pTargetF-*cat*	*pMB1 aadA* sgRNA-*cat*	This study
pTargetF-*kefB-yjcS*	*pMB1 aadA* sgRNA-*kefB*, sgRNA-*yjcS*	This study
pTargetF-*cadA*^{P}	*pMB1 aadA* sgRNA-*cadA*^{P}	This study
pTargetTΔR-Δ*cadA*^{P}	*pMB1 aadA* Δ*cadA*^{P} (430 bp)	This study
pTargetT-Δ*cadA*^{P}	*pMB1 aadA* sgRNA-*cadA*^{P} Δ*cadA*^{P} (430 bp)	This study
pTargetT-Δ*cadA*	*pMB1 aadA* sgRNA-*cadA* Δ*cadA* (804 bp)	This study
pTargetT-Δ*maeA*	*pMB1 aadA*, sgRNA-*maeA* Δ*maeA* (708 bp)	This study
pTargetT-Δ*maeB*	*pMB1 aadA* sgRNA-*maeB* Δ*maeB* (829 bp)	This study
pTargetT-Δ*maeA*Δ*maeB*	*pMB1 aadA* sgRNA-*maeA*, Δ*maeA* (708 bp) sgRNA-*maeB* Δ*maeB* (829 bp)	This study
pTargetT-ΔcadAΔmaeAΔmaeB	*pMB1 aadA* sgRNA-*cadA* Δ*cadA* (804 bp) sgRNA-*maeA* Δ*maeA* (708 bp) sgRNA-*maeB* Δ*maeB* (829 bp)	This study
pTargetT-Δ*yjcS*::*ybaS*	*pMB1 aadA* sgRNA-*yjcS* Δ*yjcS* (733 bp):: *ybaS* (1.3 kb)	This study
pTargetT-Δ*yjcS*::*evgAS*	*pMB1 aadA* sgRNA-*yjcS* Δ*yjcS* (733bp):: *evgAS* (4.5 kb)	This study
pTargetT-Δ*maeB*::*gltP*	*pMB1 aadA* sgRNA-*maeB* Δ*maeB* (829 bp)::*gltP* (1.7 kb)	This study
pTargetT-Δ*maeB*::*gltP* Δ*maeA*	*pMB1 aadA* sgRNA-*maeB* Δ*maeB* (829 bp)::*gltP* (1.7 kb) sgRNA-*maeA* Δ*maeA* (708 bp)	This study
pTargetT-Δ*tkrA*	*pMB1 aadA* sgRNA-*tkrA* Δ*tkrA* (978 bp)	This study
pTargetT-Δ*glk*	*pMB1 aadA* sgRNA-*glk* Δ*glk* (963 bp)	This study

[a] *bla*, ampicillin resistance gene; *kan*, kanamycin resistance gene; *aadA*, spectinomycin resistance gene; *cat*, chloramphenicol resistance gene; P_{cas}-*cas9*, the *cas9* gene with its native promoter; P_{araB}-*Red*, the *Red* recombination genes with an arabinose-inducible promoter; P_{trc}-sgRNA-*pMB1*, sgRNA with an N_{20} sequence for targeting the *pMB1* region with a *trc* promoter; sgRNA-*cadA*, sgRNA with an N_{20} sequence for targeting the *cadA* locus; sgRNA-*cadA*^{P}, sgRNA with an N_{20} sequence for targeting the partial *cadA* fragment inserted inside the heterologous *cat* loci of strain MGly2. Δ*cadA* (804 bp), editing template with an 804-bp region homologous to the *cadA* locus; Δ*cadA*^{P} (430 bp), editing template with a 430-bp region homologous to the partial *cadA* fragment inserted inside the heterologous *cat* loci of strain MGly2; Δ*maeA* (708 bp), editing template with an 804-bp region homologous to the *maeA* locus; Δ*yjcS* (733 bp)::*ybaS* (1.3 kb), editing template with an 804-bp region homologous to the *yjcS* locus with a 1.3-kb *ybaS* insertion; Δ*yjcS* (733 bp):: *evgAS* (4.5 kb), editing template with a 733-bp region homologous to the *yjcS* locus with a 4.5-kb *evgAS* insertion; Δ*maeB* (829 bp)::*gltP* (1.7 kb), editing template with an 829-bp region homologous to the *maeB* locus with a 1.7-kb *gltP* insertion.

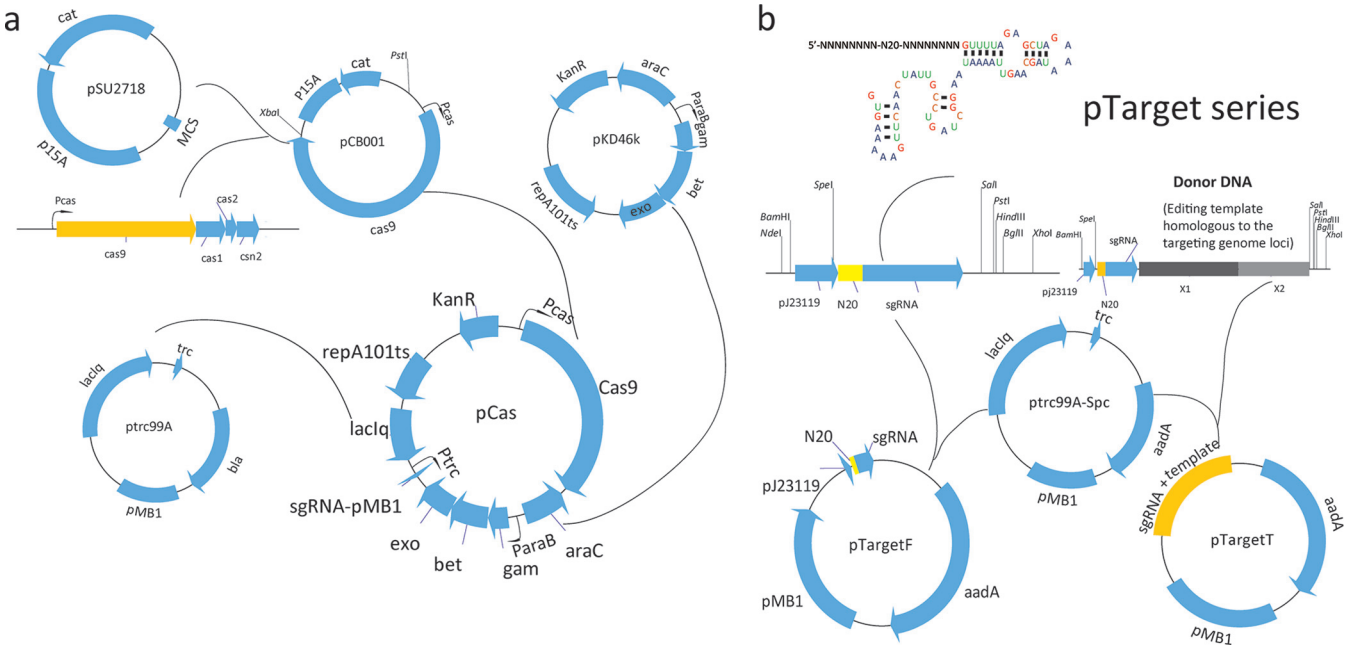

FIG 1 Construction of the CRISPR-Cas two-plasmid system. The *cas9* gene and the sgRNA directing it to the targeted region were separated in pCas and pTarget series. (a) pCas contains the *cas9* gene with a native promoter, an arabinose-inducible sgRNA guiding Cas9 to the *pMB1* replicon of pTarget, the λ-Red recombination system to improve the editing efficiency, and the temperature-sensitive replication *repA101*(Ts) for self-curing. sgRNA is displayed with its secondary structure (51). (b) pTarget was constructed to express the targeting sgRNA, with (pTargetT series) or without (pTargetF series) donor DNA as editing templates. Cas9, Cas9 endonuclease; pJ23119, synthetic promoter (38); N20, 20-bp region complementary to the targeting region (38); *araC*, arabinose-inducible transcription factor; pKD46K, a form of pKD46 in which the *bla* gene is replaced with the *aadA* gene that confers kanamycin resistance (21); pTrc99A-spec, a form of pTrc99A, in which *bla* was replaced by *aadA*, which confers spectinomycin resistance.

genome engineering procedures. The genomic DNA of *S. pyogenes* strain MGAS5005, kindly provided by Xuesong Sun of Jinan University (Guangdong, China), was used to amplify the *cas9* gene. *E. coli* or *T. citrea* (36, 37) was grown in LB medium (1% [wt/vol] tryptone, 0.5% [wt/vol] yeast extract, 1% [wt/vol] NaCl) at 37°C or 30°C. Ampicillin (100 mg/liter), kanamycin (50 mg/liter), spectinomycin (50 mg/liter), or chloramphenicol (25 mg/liter) was added as needed.

Plasmid construction. All constructs used in this study are given in Table 1, and the sgRNA, primer, and N_{20} sequences followed by the PAM used in this study are given in Tables S1 and S2 in the supplemental material. Plasmids and genomic DNA were extracted using the AxyPrep kit (Corning) according to the manufacturer's instructions. PCR used the polymerases *Taq* (Thermo Scientific) and KOD-plus-neo (Toyobo). Restriction endonucleases and T4 DNA ligase were purchased from Thermo Scientific.

The two-plasmid system, in which the *cas9* gene and the sgRNA directing it to the targeted region were separated in the pCas and pTarget series, was used for genome editing as shown in Fig. 1. pCas in the two-plasmid system consisted of *cas9*, λ-Red, a temperature-sensitive replicon, and the sgRNA with a *lacI*^{q}-P_{trc} promoter guiding the pMB1 replication of pTarget. pCB001 was constructed by amplifying the *cas9* sequence and the native promoter from *S. pyogenes* MGAS5005 with primers pA001 and pA002, followed by ligation to pSU2718, which was digested with PstI/XbaI. The *kanR-repA101*(Ts) fragment containing the kanamycin-resistant gene *kanR* and the temperature-sensitive replicon *repA101*(Ts) were amplified from pKD46K (21) by primers pA006/pA007, the *lacI*^{q} gene and the P_{trc} promoter (*lacI*^{q}-P_{trc} fragment) were amplified from pTrc99A by pA008/pA009, and the sgRNA-pMB1 sequence was amplified from pTarget, the construction of which is described below, by pA010/pA011. The λ-Red recombinase gene was amplified from pKD46 by pA012/pA013, and digested by XbaI. pCas was constructed by ligating the *cas9* cassette digested from pCB001 by PstI/XbaI, with the PstI/BglII-digested overlap PCR product of *kanR-repA101*(Ts), the *lacI*^{q}-P_{trc} fragment, and the XbaI/BglII-digested λ-Red gene.

The pTarget series had two versions, pTargetT and pTargetF, which had donor DNA for recombination supplied in the plasmid pTarget and not supplied, respectively (Fig. 1b). pTargetF consists of the sgRNA sequence, the N_{20} sequence, and the multiple restriction sites, with the donor DNA supplied as fragments. pTrc99A-spec was constructed by ligating the MluI/XhoI-digested pTrc99A framework, including the *pMB1* replicon amplified by pA003/pA056 from pTrc99A, with the spectinomycin-resistant gene *aadA* amplified by pA054/pA055 from pIJ778. The sgRNA sequence with promoter pJ23119 and the multiple restriction sites was synthesized *de novo* as described previously (35) (GenScript) and was inserted into NdeI/XhoI-digested pTrc99A-spec (Fig. 1). The pTargetF series, used in target single-gene modification with a targeting N_{20} sequence of gene loci of interest, was obtained by inverse PCR with the modified N_{20} sequence hanging at the 5′ ends of primers and followed by self-ligation (38). pTargetF-*kefB-yjcS* consisting of double sgRNAs was achieved by BioBrick cloning with BamHI and BglII (39). sgRNA-*yjcS* with its promoter was digested from pTargetF-*yjcS* with BamHI/BglII and inserted into the BglII-digested pTargetF-*kefB*. The pTargetT series consisted of the sgRNA sequence, N_{20}, the multiple restriction sites, and the donor DNA used as the genome editing template. The editing templates had a 250- to 550-bp sequence homologous to each side (upstream or downstream) of the targeted region in the genome. pTargetT-Δ*cadA*, pTargetT-Δ*maeB*, and pTargetT-Δ*maeA* were constructed by inserting the editing template through overlap PCR of the three fragments amplified by primers pB014/pB015, pB027/pB030, and pB016/pB017 to form upstream editing templates and pB029/pB028, pB058/pB059, and pB060/pB061 to form downstream editing templates from the MG1655 genome. The sgRNA fragment amplified by primers pB019/pB018, pB025/pB018, or pB053/pB018 from pTargetF was inserted into the SpeI/SalI-digested pTargetF. pTargetT-Δ*maeA*Δ*maeB* was constructed by inserting the

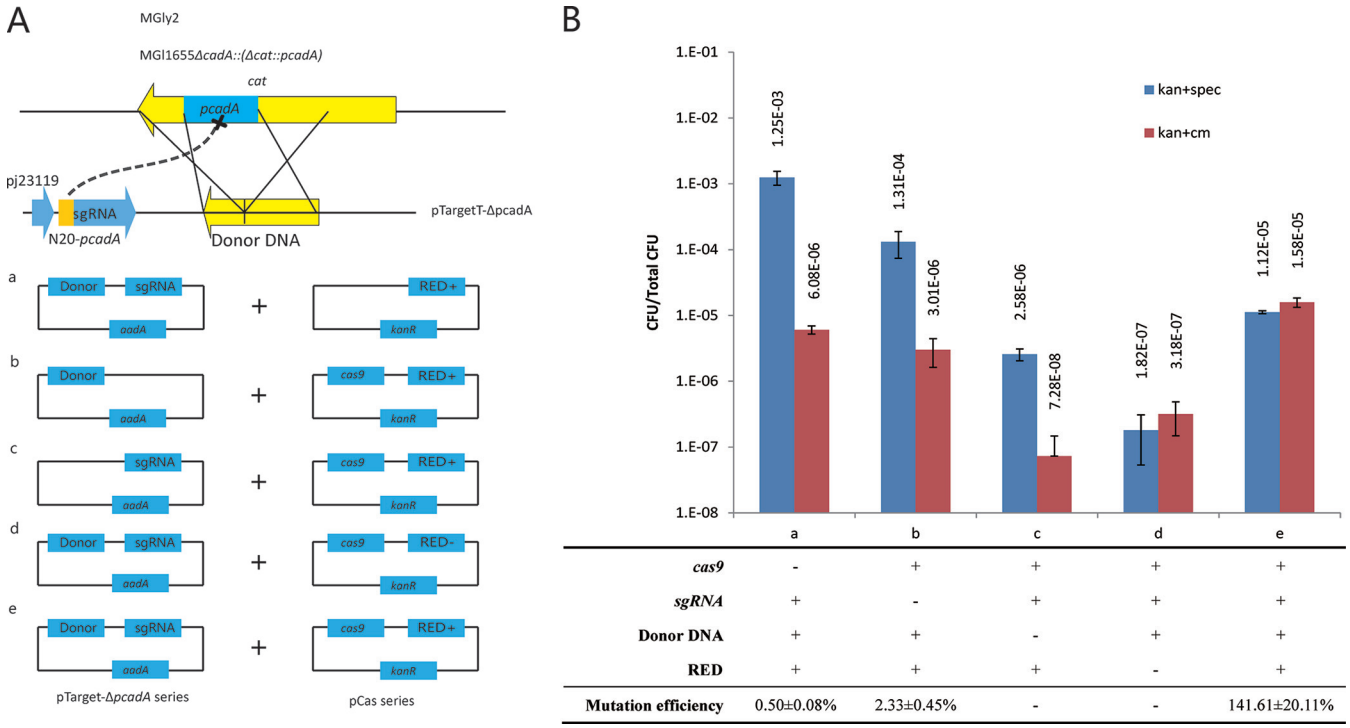

FIG 2 Effects of *cas9*, targeting sgRNA, donor DNA, and λ-Red in the CRISPR-Cas two-plasmid system. (A) Diagram of the experimental conditions. (a) *cas9* was deficient in pCas; (b and c) targeting sgRNA (b) or donor DNA (c) was deficient in pTargeting series; (d and e) λ-Red with (RED+) (e) or without (RED-) (d) induction. (B) Mutation efficiency. The fraction of spectinomycin-resistant (spec) and kanamycin-resistant (kan) or chloramphenicol-resistant (cm) and kanamycin-resistant (kan) CFU calculated from total CFU was determined under the experimental conditions shown under the histogram and depicted in panel A. Data are means ± standard deviations from three independent experiments.

fragment amplified from pTargetT-Δ*maeA* by pB062/pB063 into the SalI/BglII-digested pTargetT-Δ*maeB*. pTargetT-Δ*cadA*Δ*maeA*Δ*maeB* was constructed by inserting the fragment amplified from pTargetT-Δ*cadA* by primers pB064/pB065 into the HindIII-digested pTargetT-Δ*maeA*Δ*maeB*. pTargetT-Δ*yjcS*::*ybaS* or pTargetT-Δ*yjcS*::*evgAS* was constructed by inserting the fragment joined by overlap extension PCR amplified using primers pB037/pB041 or pB037/pB045, pB040/pB042 (to form the *ybaS* fragment) or pB044/pB046 (to form the *evgAS* fragment), and pB043/pB036 or pB047/pB036 into the PstI/HindIII-digested pTargetF-*yjcS*. pTargetT-Δ*maeB*::*gltP* was constructed by inserting the fragment overlapped by PCR and amplified by primers pB073/pB018, pB074/pB075, pB076/pB077 (to form the *gltP* fragment), and pB078/pB050 into SpeI/SalI-digested pTargetF-*yjcS*. pTargetT-Δ*maeB*::*gltP*Δ*maeA* was constructed by ligating the sgRNA and editing template fragment digested from pTargetT-Δ*maeA* by BamHI/SalI to BglII/XhoI-digested pTargetT-Δ*maeB*::*gltP*.

For the control experiment, strain MGlyl was designed by inserting the *cat* gene amplified from pSU2718 by pB068/pB069 into the *cadA* loci of MG1655. Strain MGly2 was constructed by inserting a 275-bp fragment of *cadA* (*cadA*^{P}) amplified from MG1655 by pB070/pB071 in the *cat* loci of MGlyl to inactivate the chloramphenicol resistance activity by standard CRISPR-Cas system protocol (described below) using pCas and pTargetF-*cat*. pCasΔ*cas9* was constructed by digestion of a 1,435-bp fragment of *cas9* from pCas by NdeI followed by self-ligation. pTargetF-*cadA*^{P} was constructed routinely as described above by inverse PCR with primers pB079/pB033. pTargetT-Δ*cadA*^{P} was constructed by inserting the pB066/pB067-amplified fragment (donor DNA) into the BglII/XhoI-digested pTargetF-*cadA*^{P} and the pTargetTΔR-Δ*cadA*^{P} missing the targeting sgRNA, which was constructed by inserting the pB066/pB067-amplified fragment into the BamHI/XhoI-digested pTargetF-*cadA*^{P}.

pCas and pTargetF were deposited in Addgene under the numbers 62225 and 62226.

Genome editing. MG1655 and DSM 13699 competent cells harboring pCas were prepared as described previously (16, 36, 37). Arabinose (10 mM final concentration) was added to the culture for λ-Red induction according to the protocol. For electroporation, 50 μl of cells was mixed with 100 ng of pTargetT series DNA; electroporation was done in a 2-mm Gene Pulser cuvette (Bio-Rad) at 2.5 kV, and the product was suspended immediately in 1 ml of ice-cold LB medium. When the donor DNA was supplied in a PCR fragment, 100 ng of pTargetF series DNA and 400 ng of donor DNA were coelectroporated. Cells were recovered at 30°C for 1 h before being spread onto LB agar containing kanamycin (50 mg/liter) and spectinomycin (50 mg/liter) and incubated overnight at 30°C. Transformants were identified by colony PCR and DNA sequencing.

For control experiments, the strain MGly2 modified from MG1655 was used as the host. pCas and pCasΔ*cas9* with a *cas9* deletion were cotransformed with pTargetT-Δ*cadA*^{P} and pTargetTΔR-Δ*cadA*^{P} with targeting sgRNA deletion, respectively; pCas was also cotransformed with pTargetF-*cadA*^{P} without a *cat* homologous fragment (Fig. 2A). Both λ-Red induction and noninduction were done by adding arabinose (10 mM final concentration) or not, according to the previous protocol. Cells were recovered at 30°C for 1 h before being spread onto LB agar containing kanamycin (50 mg/liter) and spectinomycin (50 mg/liter) or kanamycin (50 mg/liter) and chloramphenicol (25 mg/liter) and incubated at 30°C overnight.

Plasmid curing. For the curing of pTarget series, the edited colony harboring both pCas and pTarget series was inoculated into 2 ml of LB medium containing kanamycin (50 mg/liter) and IPTG (isopropyl-β-D-thiogalactopyranoside; 0.5 mM). The culture was incubated for 8 to 16 h, diluted, and spread onto LB plates containing kanamycin (50 mg/liter). The colonies were confirmed as cured by determining their sensitivity to spectinomycin (50 mg/liter). The colonies cured of pTarget series were used in a second round of genome editing. pCas was cured by growing the colonies overnight at 37°C nonselectively (13).

TABLE 2 Mutation efficiency of the CRISPR-Cas two-plasmid system[a]

Expt no.	Host cell	Plasmid pTarget	Targeting genome locus of sgRNA	Donor DNA supplied in pTarget or in PCR fragment (F)	Length (bp) of homologous extensions (upstream, downstream)	Mutation efficiency (%)[b]	Plasmid pTarget curing efficiency (%)
1	MG1655	pTargetT-Δ*cadA*	*cadA*	pTargetT-Δ*cadA*	523, 281	86 ± 4	100
2	MG1655	pTargetT-Δ*maeA*Δ*maeB*	*maeA*, *maeB*	pTargetT-Δ*maeA*Δ*maeB*	250, 550	97 ± 4	ND
3	MG1655	pTargetT-Δ*cadA*Δ*maeA*Δ*maeB*	*cadA*, *maeA*, *maeB*	pTargetT-Δ*cadA*Δ*maeA* Δ*maeB*	250, 550	47 ± 8	ND
4	MG1655	pTargetT-Δ*yjcS*::*ybas*	*yjcS*	pTargetT-Δ*yjcS*::*ybaS*	373, 360	92 ± 0	ND
5	MG1655	pTargetT-Δ*yjcS*:: *evgAS*	*yjcS*	pTargetT-Δ*yjcS*::*evgAS*	373, 360	75 ± 18	ND
6	MG1655	pTargetT-Δ*maeB*::*gltP* Δ*maeA*	*maeB*, *maeA*	pTargetT-Δ*maeB*::*gltP* Δ*maeA*	250, 550	78 ± 26	ND
7	1655ΔcadA[c]	pTargetT-Δ*yjcS*::*evgAS*	*yjcS*	pTargetT-Δ*yjcS*::*evgAS*	373, 360	92 ± 7	100
8	MG1655	pTargetF-*cadA*	*cadA*	Δ*cadA* (F)	523, 281	69 ± 4	ND
9	MG1655	pTargetF-*yjcS*	*yjcS*	Δ*yjcS*::*evgAS* (F)	40, 40	6 ± 4	ND
10	MG1655	pTargetF-*yjcS*	*yjcS*	Δ*yjcS*::*evgAS* (F)	373, 360	28 ± 10	ND
11	MG1655	pTargetF-*kefB*-*yjcS*	*kefB*, *yjcS*	Δ*kefB*, Δ*yjcS*::*evgAS* (F)	250, 550	0	ND
12	DSM 13699	pTargetT-Δ*tkrA*	*tkrA*	pTargetT-Δ*tkrA*	483, 513	100 ± 0	100
13	DSM 13699 Δ*tkrA*[c]	pTargetT-Δ*glk*	*glk*	pTargetT-Δ*glk*	500, 500	94 ± 8	100

[a] The genome editing was performed with the CRISPR-Cas two-plasmid system with pCAS and pTarget, as shown. ND, not determined.
[b] Determined from triple electroporation experiments by colony PCR from 12 transformants for each mutation (agarose electrophoresis gels of colony PCR and relative sequencing results are shown in Fig. S1 and S2 in the supplemental material).
[c] Second round of genome editing.

RESULTS

Establishment of a two-plasmid-based CRISPR-Cas9 system. Two-plasmid systems were designed initially to use the CRISPR-Cas9 system, as reported for *E. coli* (24, 35), which separated *cas9* and the sgRNA in pCas and pTarget series, respectively (Fig. 1). pCas was constructed by introducing the Cas9 protein from *S. pyogenes* MGAS5005 with its native promoter, the temperature-sensitive replicon *repA101*(Ts) from plasmid pKD46 for self-curing (13), the λ-Red gene under the control of the P_{araB} promoter, which is induced by l-arabinose (40), and an sgRNA containing an N_{20} sequence targeting the pTarget *pMB1* replicon (sgRNA-*pMB1*) under the control of an IPTG-inducible promoter, P_{trc}. The sgRNA targeting the genome loci of interest located in the pTarget series was expressed from a minimal constitutive promoter with a *pMB1* origin of replication (Fig. 1b).

This CRISPR-Cas9 system was first tested for the effect of a deficiency of any of the four motifs *cas9*, sgRNA targeting the genome loci, donor DNA, and λ-Red gene. MGly2 was designed specifically as a control host modified from MG1655 to have a heterologous chloramphenicol *cat* resistance gene insertion at the *cadA* locus, which was inactivated by a DNA fragment [*cadA*^{P}] inserted inside the *cat* gene locus. Modified MGly2 colonies harboring pCas series (kanamycin resistant) with the *cadA*^{P} deletion were expected to retrieve chloramphenicol resistance activity and thus survival on agar containing chloramphenicol and kanamycin. The total CFU were calculated by growth on agar containing kanamycin and agar containing spectinomycin resulting from pCas and pTarget-Δ*cadA*^{P} (spectinomycin resistant) cotransformation into MGly2 (Fig. 2A). A deficiency of *cas9* in pCas or targeting sgRNA in pTarget resulted in a low level of recombination efficiency (<5%) through λ-Red recombination and a low survival rate on chloramphenicol selection medium, as expected. A deficiency of *cat* homologous fragments (donor DNA) or the λ-Red gene without induction resulted in a very low survival rate, even in the absence of chloramphenicol selection compared to that without *cas9*, since most of the strains were killed by *cas9* through introduction of dsDNA breaks into the chromosome. The CRISPR-Cas9 system using pCas and pTargetT-Δ*cadA*^{P} with *cas9*, targeting sgRNA, *cat* homologous fragments, and the λ-Red gene resulted in a 100% mutation rate and a relatively high survival rate (1.12E−05), which indicated that expression of the λ-Red protein increased the target site mutation rate by CRISPR-Cas9 significantly (24). In addition, using the CRISPR-Cas9 system by introducing dsDNA breaks into the chromosome increased the rate of recombination of the damaged DNA, as reported elsewhere (27). We counted 2.61-fold more colonies (1.58E−05/6.08E−06) after cotransformation with pCas and pTargetT-Δ*cadA*^{P} compared to a deficiency of the *cas9* construct (Fig. 2B). Without the donor DNA, the few colonies observed on chloramphenicol selection medium likely resulted from an escape from the death effect of DSBs by alternative end joining (41).

The two-plasmid-based CRISPR-Cas9 system makes multiplex gene modifications continuously. This CRISPR-Cas9 system was tested for (i) single, double, and multiple gene deletions and (ii) single and double gene insertions. For a single gene deletion, as shown for *cadA*, 86% ± 4% of the transformants showed the expected genotype and 100% of the cells lost pTargetT-Δ*cadA* (Table 2, experiment 1). We then doubled and tripled the number of editing targets. When MG1655 harboring pCas was transformed with pTargetT-Δ*maeA*Δ*maeB* or pTargetT-Δ*cadA* Δ*maeA*Δ*maeB*, which were expected to perform a *maeA-maeB* double deletion or a *cadA-maeA-maeB* triple deletion, the mutation efficiencies were as high as 97% ± 4% and 47% ± 8%, respectively (Table 2, experiment 2 and 3).

We used pCas to perform a single insertion and a mixed gene insertion and deletion. High mutation rates (92% ± 0% and

75% ± 18%, respectively) were obtained when *ybaS* (1.3 kb) and *evgAS* (4.5 kb) were inserted into *yjcS* (Table 2, experiment 4 and 5). For mixed gene insertion and deletion, 78% ± 26% of the colonies showed the expected genotype for the deletion of *maeA* and the insertion of *gltP* (1.7 kb) into the *maeB* locus (Table 2, experiment 6).

Continual gene editing was tested (Table 2, experiment 7). When 1655ΔcadA, cured of pTargetT-Δ*cadA*, was transformed with pTargetT-Δ*yjcS*::*evgAS*, the mutation efficiency for the insertion of *evgAS* into the *yjcS* locus was relatively high, 92% ± 7%. pCas was finally cured by the end of the procedure by culture at 37°C overnight. In all our experiments, >90% of colonies regained kanamycin sensitivity, indicating successful clearance of the temperature-sensitive plasmid pCas, in accordance with published data (13). Agarose gel electrophoresis of colony PCR and the sequencing results are supplied in Fig. S1 and S2 in the supplemental material. This demonstrated the feasibility of performing multiple rounds of genome editing to engineer novel bacterial strains.

We did not attempt multiple gene deletions or insertions of more than three genes because the cloning procedure for pTargetT was complicated and time-consuming when multiple donor DNAs were included. The method will not have the level of efficiency needed for metabolic engineering of an industrially relevant strain. Thus, although the problem of low efficiency of gene insertion was solved and double or multiple gene deletions or insertions were achieved, a simpler procedure for genome editing is needed.

Simplified genome editing by a CRISPR-Cas9 system with donor DNA supplied as a fragment. To simplify the cloning procedure for the pTarget series, the donor DNA was designed to be supplied in fragments. For single-gene editing, pTarget could thus be cloned simply by changing the N_{20} sequence of the sgRNA when different genomic loci are being targeted, which could be done by inverse PCR with mutations incorporated into the primers (38), resulting in the pTargetF version (Fig. 1b). Double- or multiple-gene editing of the pTargetF series with double or multiple sgRNAs could be done easily by the BioBrick method (38).

By using the pTargetF series with donor DNA supplied as fragments, we obtained single-gene *cadA* deletion efficiency as high as 69% ± 4% when pTargetF-*cadA* and the fragments homologous to the upstream and downstream regions of the *cadA* locus (obtained by overlap PCR) were cotransformed into MG1655 harboring pCas (Table 2, experiment 8). For gene insertions, because λ-Red recombination can be obtained efficiently with homologous regions of ≥40 bp (13), we reduced the homologous length from 300 to 500 bp to 40 bp, which could be incorporated directly into the PCR primers for the donor DNA fragment. However, a very low mutation efficiency of 6% ± 4% was obtained when we inserted *evgAS* into the *yjcS* locus (Table 2, experiment 9). We extended the homologous length in the donor DNA to ~400 bp for the same targeting site, and a higher insertion rate of 28% ± 10% was obtained (Table 2, experiment 10).

We attempted to perform double-gene editing (Table 2, experiment 11) by combined deletion of locus *kefB* with the insertion of *evgAS* into locus *yjcS*, but we obtained no double mutation.

Results of agarose gel electrophoresis of colony PCR and sequencing are supplied in Fig. S1 and S2 in the supplemental material.

Application of two-plasmid-based CRISPR-Cas9 system in *Tatumella citrea* for continuous gene deletion. To evaluate the possibility of a broader applicability of the system described above, *T. citrea* DSM 13699, another member of the *Enterobacteriaceae*, was selected. Two genes, encoding a subunit of glyoxylate reductase (*tkrA*) and glucokinase (*glk*) were chosen as individual targets. The system fit DSM 13699 well without any modification, with 100% ± 0% *tkrA* deletion efficiency and 94% ± 8% second-gene *glk* deletion efficiency (Table 2, experiments 12 and 13). pTargetT-Δ*cadA*, pTargetT-Δ*yjcS*::*evgAS*, pTargetT-Δ*tkrA*, and pTargetT-Δ*glk* were 100% cured (Table 2). The observed efficient genome editing of *T. citrea* without strain-specific backbone modification of the two-plasmid-based CRISPR-Cas9 system suggests a possible broader applicability of this system in various *Enterobacteriaceae* species.

DISCUSSION

In this study, we expanded the application of the CRISPR system from the published allelic exchange procedure (24) to targeted single or multiple gene deletions and insertions in *E. coli* and another *Enterobacteriaceae* species, *T. citrea*. Compared to published scarless genome modification methods, such as those involving *sacB* (12), I-SceI (12, 19, 20), and MAGE (22, 23), the CRISPR-based targeted genome modification method can perform multiple gene insertions or deletions, whereas *sacB* or I-SceI could be used to modify only single targets each time. ssDNA oligonucleotide-mediated MAGE was used successfully for multiple allelic exchange, but small-fragment (30 bp) insertion decreased mutation efficiency dramatically (12, 22). In addition, the CRISPR-based gene modification system offers unprecedented convenience and efficiency in design and manipulation. Targeting any site of interest requires the insertion of only a short spacer into a targeting sgRNA construct, pTargetF in this study, which can be achieved by inverse PCR and self-ligation within 2 days, with donor DNA supplied as PCR fragments (Fig. 1b). The manipulation time for the procedure was reduced to 2 days for each round of modification, and up to three gene targets can be modified simultaneously; an additional 2 days are required for the entire procedure (Fig. 3). For the metabolic engineering case that required 26 genomic modifications, as mentioned in the introduction, the total manipulation time can be 20 to 54 days. Metabolic engineering is based on the cell system network in which simple gene engineering might result in unexpected phenotypes, and with the rapid development of genome sequencing technology (42), more sequenced genotypes need to be illustrated biologically. This CRISPR-based time-saving genome modification method will be a powerful tool in the metabolic engineering field and will facilitate the output of genetically modified strains, thus increasing the likelihood of engineering complex strains. *T. citrea* is an important host for production of the industrially relevant vitamin C precursor 2-keto-D-gluconic acid (43). The application of the CRISPR-based gene modification system in *T. citrea* will greatly facilitate metabolic engineering of this strain compared to the only traditional homologous recombination-based gene knockout system as described previously (43, 44). The successful expansion of this system without any specific modification to *T. citrea* indicated its wide adaptability and flexibility in other *Enterobacteriaceae* species.

The challenge facing DSB-based, multiplex genome modification techniques might be the toxicity of simultaneous multiple chromosomal breaks and the high rate of nonhomologous end

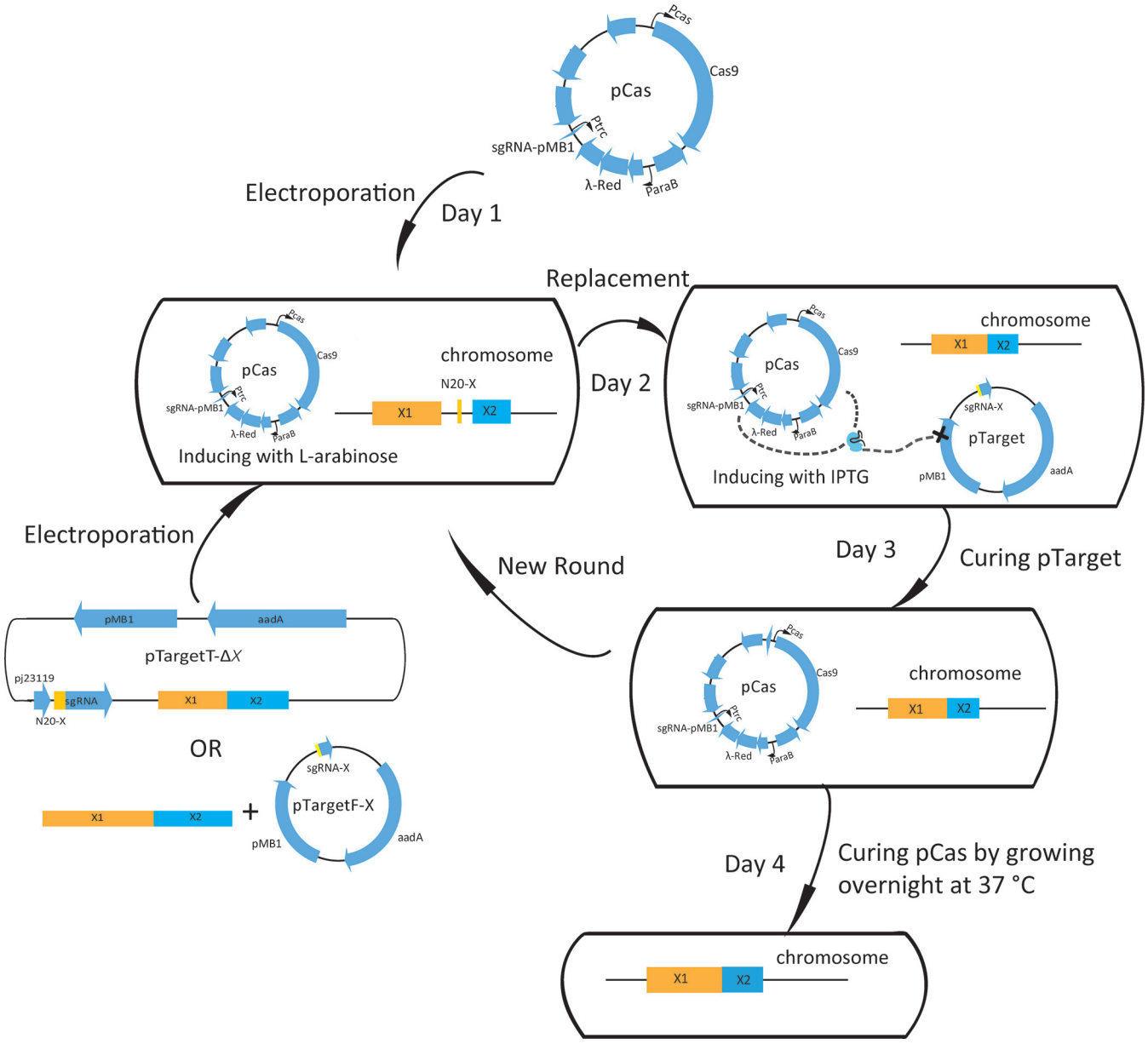

FIG 3 Detailed diagram of continual genome editing with the two-plasmid system.

joining (NHEJ), which could lead to unintended rearrangements (10). However, *E. coli* lacks the NHEJ mechanism, although a small possibility of an alternative end-joining mechanism exists (41), and is highly reliant on a homology-directed repair system to repair DSBs in the chromosome (45, 46). Thus, the success of multiplexing depends on the fine-tuning of Cas9 activity and the rescue efficiency of homology-directed repair. Double-stranded, λ-Red-mediated recombination successfully rescued the low efficiency of the *E. coli* native homology-directed repair system and, thus, succeeded in multigene editing even when Cas9 was expressed constitutively, while single DSB generated by constitutively expressed Cas9 could not be repaired without induction of λ-Red (Fig. 2B). Originally, we used an arabinose-inducible promoter to express Cas9, and thus, some cells escaped cleavage on the chromosome caused by the induction efficiency of the arabinose promoter (47); as a result, the curing efficiency of pTarget cleaved by *cas9* did not reach 100% (data not shown). We failed to clone the IPTG-inducible *trc*, which was expressed targeting sgRNA, and constitutively expressed *cas9* in one plasmid, which might cause by leakage of the *trc* promoter. We did not investigate the possibility of using other inducible promoters for both *cas9* and sgRNA or λ-Red to incorporate these into one plasmid, because the two-plasmid-based system has the advantage of a shorter total manipulation time than the inducible one-plasmid system if the number of targets is >2 (the two-plasmid system needs $2n+2$ days, and the one-plasmid system needs $3n$ days).

For the CRISPR-Cas9 system given in Table 2, we can perform up to three gene deletions and mixed gene deletions and insertions with acceptable levels of efficiency (47% ± 8% and 78% ± 26%, respectively). However, cloning of pTargetT, which contains mul-

tiple targeted sgRNAs and donor DNAs, was both time-consuming and labor-intensive. Thus, donor DNAs supplied in fragments that can be cotransformed into the cell with sgRNAs contained in pTargetF reduced the amount of time and labor needed for the constructions. However, the efficiency decreased dramatically when increasing the batch targets numbers (0%) (Table 2, experiment 11) or decreasing the length of homologous extensions from 300 to 400 bp to 40 bp (6% ± 4%) (Table 2, experiment 9). This was because the efficiency of double-stranded, λ-Red-mediated recombination was not sufficiently high, or the transformation efficiency of the dsDNA in *E. coli* was low. The recombination efficiency might be improved by using ssDNA as the donor DNA, as λ-Red-like proteins also facilitate the recombination of smaller ssDNA fragments, such as those used in MAGE. If CRISPR and MAGE are combined, the challenge might be that the multiple, repeated sgRNAs in pTarget will lead to rearrangements by self-homologous recombination, as well as its limitation in gene insertion manipulation.

Off-target effects of Cas9 in human and murine cells have been reported (48, 49), and some methods have been applied to mitigate these effects, including cooperative use of offset nicking and a *cas9* nickase mutant (50). To reduce the off-target effects of Cas9 in this study, an N_{20} sequence was selected to ensure the last 12 bp was highly specific for the targets (24).

ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China (2012CB721105, 2014CB745101), the National High Technology Research and Development Program of China (2012AA02A704), and Knowledge Innovation Program (KSZD-EW-Z-016-1, KSZD-EW-Z-019) and Science and Technology Service Network Initiative (KFJ-EW-STS-030) of the Chinese Academy of Sciences.

We thank Liuyang Diao from Shanghai Institutes for Biological Sciences for helpful discussion, and we thank Qiming Tian, Song Cui, and Yazhuo Sun from Global Bio-chem Technology Group for experimental support.

REFERENCES

1. **Ingram LO, Gomez PF, Lai X, Moniruzzaman M, Wood BE, Yomano LP, York SW.** 1998. Metabolic engineering of bacteria for ethanol production. Biotechnol Bioeng **58:**204–214. http://dx.doi.org/10.1002/(SICI)1097-0290(19980420)58:2/3<204::AID-BIT13>3.0.CO;2-C.
2. **Atsumi S, Hanai T, Liao JC.** 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature **451:**86–89. http://dx.doi.org/10.1038/nature06450.
3. **Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD.** 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature **463:**559–562. http://dx.doi.org/10.1038/nature08721.
4. **Leuchtenberger W, Huthmacher K, Drauz K.** 2005. Biotechnological production of amino acids and derivatives: current status and prospects. Appl Microbiol Biotechnol **69:**1–8. http://dx.doi.org/10.1007/s00253-005-0155-y.
5. **Bongaerts J, Kramer M, Muller U, Raeven L, Wubbolts M.** 2001. Metabolic engineering for microbial production of aromatic amino acids and derived compounds. Metab Eng **3:**289–300. http://dx.doi.org/10.1006/mben.2001.0196.
6. **Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD.** 2003. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat Biotechnol **21:**796–802. http://dx.doi.org/10.1038/nbt833.
7. **McDaniel R, Thamchaipenet A, Gustafsson C, Fu H, Betlach M, Ashley G.** 1999. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. Proc Natl Acad Sci U S A **96:**1846–1851. http://dx.doi.org/10.1073/pnas.96.5.1846.
8. **Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R, Estadilla J, Teisan S, Schreyer HB, Andrae S, Yang TH, Lee SY, Burk MJ, Van Dien S.** 2011. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat Chem Biol **7:**445–452. http://dx.doi.org/10.1038/nchembio.580.
9. **Nakamura CE, Whited GM.** 2003. Metabolic engineering for the microbial production of 1,3-propanediol. Curr Opin Biotechnol **14:**454–459. http://dx.doi.org/10.1016/j.copbio.2003.08.005.
10. **Esvelt KM, Wang HH.** 2013. Genome-scale engineering for systems and synthetic biology. Mol Syst Biol **9:**641. http://dx.doi.org/10.1038/msb.2012.66.
11. **Enyeart PJ, Chirieleison SM, Dao MN, Perutka J, Quandt EM, Yao J, Whitt JT, Keatinge-Clay AT, Lambowitz AM, Ellington AD.** 2013. Generalized bacterial genome editing using mobile group II introns and Cre-lox. Mol Syst Biol **9:**685. http://dx.doi.org/10.1038/msb.2013.41.
12. **Yu BJ, Kang KH, Lee JH, Sung BH, Kim MS, Kim SC.** 2008. Rapid and efficient construction of markerless deletions in the Escherichia coli genome. Nucleic Acids Res **36:**e84. http://dx.doi.org/10.1093/nar/gkn359.
13. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640–6645. http://dx.doi.org/10.1073/pnas.120163297.
14. **Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL.** 2000. An efficient recombination system for chromosome engineering in Escherichia coli. Proc Natl Acad Sci U S A **97:**5978–5983. http://dx.doi.org/10.1073/pnas.100127597.
15. **Zhang Y, Buchholz F, Muyrers JP, Stewart AF.** 1998. A new logic for DNA engineering using recombination in Escherichia coli. Nat Genet **20:**123–128. http://dx.doi.org/10.1038/2417.
16. **Sharan SK, Thomason LC, Kuznetsov SG, Court DL.** 2009. Recombineering: a homologous recombination-based method of genetic engineering. Nat Protoc **4:**206–223. http://dx.doi.org/10.1038/nprot.2008.227.
17. **Warner JR, Reeder PJ, Karimpour-Fard A, Woodruff LB, Gill RT.** 2010. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. Nat Biotechnol **28:**856–862. http://dx.doi.org/10.1038/nbt.1653.
18. **Costantino N, Court DL.** 2003. Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. Proc Natl Acad Sci U S A **100:**15748–15753. http://dx.doi.org/10.1073/pnas.2434959100.
19. **Posfai G, Kolisnychenko V, Bereczki Z, Blattner FR.** 1999. Markerless gene replacement in Escherichia coli stimulated by a double-strand break in the chromosome. Nucleic Acids Res **27:**4409–4415. http://dx.doi.org/10.1093/nar/27.22.4409.
20. **Yang J, Sun B, Huang H, Jiang Y, Diao L, Chen B, Xu C, Wang X, Liu J, Jiang W, Yang S.** 2014. High-efficiency scarless genetic modification in Escherichia coli using lambda-red recombination and I-SceI cleavage. Appl Environ Microbiol **80:**3826–3834. http://dx.doi.org/10.1128/AEM.00313-14.
21. **Karberg M, Guo H, Zhong J, Coon R, Perutka J, Lambowitz AM.** 2001. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. Nat Biotechnol **19:**1162–1167. http://dx.doi.org/10.1038/nbt1201-1162.
22. **Wang HH, Church GM.** 2011. Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. Methods Enzymol **498:**409–426. http://dx.doi.org/10.1016/B978-0-12-385120-8.00018-8.
23. **Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM.** 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature **460:**894–898. http://dx.doi.org/10.1038/nature08187.
24. **Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA.** 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol **31:**233–239. http://dx.doi.org/10.1038/nbt.2508.
25. **DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM.** 2013. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res **41:**4336–4343. http://dx.doi.org/10.1093/nar/gkt135.
26. **Cobb RE, Wang Y, Zhao H.** 8 December 2014. High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system. ACS Synth Biol. http://dx.doi.org/10.1021/sb500351f.
27. **Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C.** 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol **31:**686–688. http://dx.doi.org/10.1038/nbt.2650.

28. **Wang Y, Li Z, Xu J, Zeng B, Ling L, You L, Chen Y, Huang Y, Tan A.** 2013. The CRISPR/Cas System mediates efficient genome engineering in Bombyx mori. Cell Res **23:**1414–1416. http://dx.doi.org/10.1038/cr.2013.146.
29. **Yu Z, Ren M, Wang Z, Zhang B, Rong YS, Jiao R, Gao G.** 2013. Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Genetics **195:**289–291. http://dx.doi.org/10.1534/genetics.113.153825.
30. **Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F.** 2013. Multiplex genome engineering using CRISPR/Cas systems. Science **339:**819–823. http://dx.doi.org/10.1126/science.1231143.
31. **Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM.** 2013. RNA-guided human genome engineering via Cas9. Science **339:**823–826. http://dx.doi.org/10.1126/science.1232033.
32. **Zhang Q, Rho M, Tang H, Doak TG, Ye Y.** 2013. CRISPR-Cas systems target a diverse collection of invasive mobile genetic elements in human microbiomes. Genome Biol **14:**R40. http://dx.doi.org/10.1186/gb-2013-14-4-r40.
33. **Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E.** 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature **471:**602–607. http://dx.doi.org/10.1038/nature09886.
34. **Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E.** 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science **337:**816–821. http://dx.doi.org/10.1126/science.1225829.
35. **Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA.** 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell **152:**1173–1183. http://dx.doi.org/10.1016/j.cell.2013.02.022.
36. **Pujol CJ, Kado CI.** 2000. Genetic and biochemical characterization of the pathway in Pantoea citrea leading to pink disease of pineapple. J Bacteriol **182:**2230–2237. http://dx.doi.org/10.1128/JB.182.8.2230-2237.2000.
37. **Cha JS, Pujol C, Kado CI.** 1997. Identification and characterization of a Pantoea citrea gene encoding glucose dehydrogenase that is essential for causing pink disease of pineapple. Appl Environ Microbiol **63:**71–76.
38. **Ochman H, Gerber AS, Hartl DL.** 1988. Genetic applications of an inverse polymerase chain reaction. Genetics **120:**621–623.
39. **Shetty RP, Endy D, Knight TF, Jr.** 2008. Engineering BioBrick vectors from BioBrick parts. J Biol Eng **2:**5. http://dx.doi.org/10.1186/1754-1611-2-5.
40. **Guzman LM, Belin D, Carson MJ, Beckwith J.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol **177:**4121–4130.
41. **Chayot R, Montagne B, Mazel D, Ricchetti M.** 2010. An end-joining repair mechanism in Escherichia coli. Proc Natl Acad Sci U S A **107:**2141–2146. http://dx.doi.org/10.1073/pnas.0906355107.
42. **Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER.** 2013. The next-generation sequencing revolution and its impact on genomics. Cell **155:**27–38. http://dx.doi.org/10.1016/j.cell.2013.09.006.
43. **Dodge TC, Valle F, Rashid MH.** February 2005. Metabolically engineered bacterial strains having enhanced 2-keto-D-gluconate accumulation. US patent WO2005012486-A2.
44. **Banta S, Boston M, Jarnagin A, Anderson S.** 2002. Mathematical modeling of in vitro enzymatic production of 2-keto-L-gulonic acid using NAD(H) or NADP(H) as cofactors. Metab Eng **4:**273–284. http://dx.doi.org/10.1006/mben.2002.0231.
45. **Wilson TE, Topper LM, Palmbos PL.** 2003. Non-homologous end-joining: bacteria join the chromosome breakdance. Trends Biochem Sci **28:**62–66. http://dx.doi.org/10.1016/S0968-0004(03)00005-7.
46. **Malyarchuk S, Wright D, Castore R, Klepper E, Weiss B, Doherty AJ, Harrison L.** 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 (Amst) **6:**1413–1424. http://dx.doi.org/10.1016/j.dnarep.2007.04.004.
47. **Siegele DA, Hu JC.** 1997. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. Proc Natl Acad Sci U S A **94:**8168–8172. http://dx.doi.org/10.1073/pnas.94.15.8168.
48. **Cradick TJ, Fine EJ, Antico CJ, Bao G.** 2013. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res **41:**9584–9592. http://dx.doi.org/10.1093/nar/gkt714.
49. **Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD.** 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol **31:**822–826. http://dx.doi.org/10.1038/nbt.2623.
50. **Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F.** 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell **154:**1380–1389. http://dx.doi.org/10.1016/j.cell.2013.08.021.
51. **Anders C, Niewoehner O, Duerst A, Jinek M.** 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature **513:**569–573. http://dx.doi.org/10.1038/nature13579.
52. **Martinez E, Bartolome B, de la Cruz F.** 1988. pACYC184-derived cloning vectors containing the multiple cloning site and lacZ alpha reporter gene of pUC8/9 and pUC18/19 plasmids. Gene **68:**159–162. http://dx.doi.org/10.1016/0378-1119(88)90608-7.
53. **Amann E, Ochs B, Abel KJ.** 1988. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. Gene **69:**301–315. http://dx.doi.org/10.1016/0378-1119(88)90440-4.
54. **Gust B, Challis GL, Fowler K, Kieser T, Chater KF.** 2003. PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci U S A **100:**1541–1546. http://dx.doi.org/10.1073/pnas.0337542100.