BIOLOGY Methods & Protocols

Biology Methods and Protocols, 2016, 1–8

doi: 10.1093/biomethods/bpw004
Methods Manuscript

METHODS MANUSCRIPT

Leveraging modern DNA assembly techniques for rapid, markerless genome modification

Ilya B. Tikh[1] and James C. Samuelson[1,*]

[1]Protein Expression and Modification Division, New England BioLabs, Inc., Ipswich, MA, 01938-2723, USA

*Correspondence address. 240 County Road, Ipswich, MA, USA, 01938-2723. Tel: +1 978 380 7288; Fax: +1 978 921 1350; Email: samuelson@neb.com

Abstract

The ability to alter the genomic material of a prokaryotic cell is necessary for experiments designed to define the biology of the organism. In addition, the production of biomolecules may be significantly improved by application of engineered prokaryotic host cells. Furthermore, in the age of synthetic biology, speed and efficiency are key factors when choosing a method for genome alteration. To address these needs, we have developed a method for modification of the *Escherichia coli* genome named FAST-GE for Fast Assembly-mediated Scarless Targeted Genome Editing. Traditional cloning steps such as plasmid transformation, propagation and isolation were eliminated. Instead, we developed a DNA assembly-based approach for generating scarless strain modifications, which may include point mutations, deletions and gene replacements, within 48 h after the receipt of polymerase chain reaction primers. The protocol uses established, but optimized, genome modification components such as I-SceI endonuclease to improve recombination efficiency and SacB as a counter-selection mechanism. All DNA-encoded components are assembled into a single allele-exchange vector named pDEL. We were able to rapidly modify the genomes of both *E. coli* B and K-12 strains with high efficiency. In principle, the method may be applied to other prokaryotic organisms capable of circular dsDNA uptake and homologous recombination.

Keywords: Genome modification; DNA assembly; I-SceI; synthetic biology

Introduction

The ability to manipulate the genetic material of a cell has proven to be invaluable for understanding the organism's biology. The basis of our understanding of many essential biological processes has been enabled by the ability to perform gene deletion and complementation studies [1–4]. Furthermore, countless strains have been specifically engineered to facilitate the production of valuable proteins requiring modified cellular environments or helper factors [5–9].

A number of the genome modification methods used today were developed decades ago and have had only minor updates due to new techniques becoming available [10–16]. Most of the current methods for modifying common lab organisms such as *Escherichia coli* fall into two categories. The first set of methods is driven by homologous recombination and is reliant on proteins such as endogenous RecA [15, 16]. A primary advantage of methods that utilize the cell's own recombination machinery is that helper plasmids are not necessary. Homologous recombination results in a single crossover event between the incoming DNA and the chromosome resulting in a tandem arrangement of the wild-type gene and mutant gene. Then, a second crossover event potentially leaves the mutation of interest in the chromosome. These protocols (labelled as 'Classic allelic-exchange' in Fig. 1) generally do not leave a scar in the genome upon removal of the selection marker. However, these methods are often cumbersome, taking as much as a week for

Received: 17 June 2016; **Revised:** 6 September 2016. **Accepted:** 26 September 2016

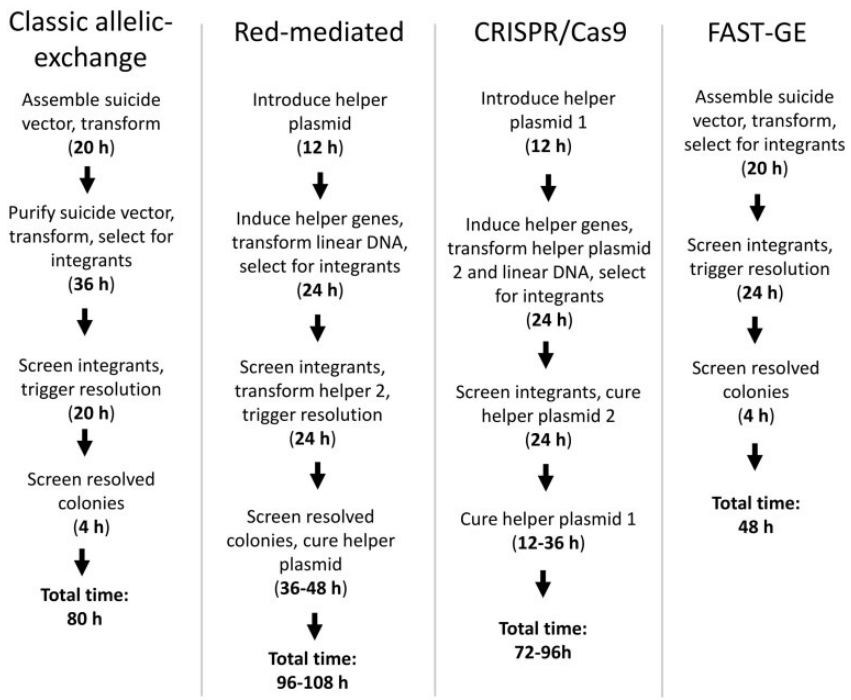

Figure 1: Timeline comparison of the most common methods for genome modification in *E. coli* (h = hours). During Red-mediated protocols, helper plasmid 1 contains the red exo, beta and gamma genes, which facilitate the incorporation of linear DNA into the chromosome. Helper plasmid two traditionally contains Flp recombinase to help remove the antibiotic resistance marker from the chromosome. When utilizing the CRISPR/Cas9, the first helper plasmid also contains the red exo, beta and gamma genes in addition to Cas9. The second helper plasmid contains the guide RNA, used to eliminate any WT cells which did not undergo desired recombination. Regardless of the method of construction, all modifications should be sequenced in order to verify the presence of the desired modification.

a single modification. A second set of commonly used methods is derived from the phage lambda recombination system, most famously described by Datsenko and Wanner, and utilizes the expression of exogenous phage proteins to mediate the recombination steps [10, 11, 17, 18]. The lambda Red system allows for faster integration and modification of a desired locus using a linear PCR product that contains short segments of homology, but this method requires the presence of a helper plasmid driving the expression of the phage genes and is less efficient for the replacement of large pieces of DNA. Though more rapid than classic methods, optimized lambda Red techniques still require approximately 4 days from start to finish, counting the time it takes to initially transform and subsequently cure the helper plasmids [10]. The lambda Red system has also been applied for DNA oligonucleotide-mediated genome modification; however, this method will not be discussed further as it is primarily limited to the generation of point mutations or codon substitutions [19].

Recently, the CRISPR/Cas9 system has been adapted for modifying genomes in multiple prokaryotic organisms [18, 20–23]. Unfortunately, to take full advantage of this system, the organism must be highly recombinogenic, which *E. coli* is not [22]. In general, the application of the CRISPR/Cas9 system in *E. coli* requires the use of the lambda Red system for the initial allele exchange step and then the Cas9 nuclease is employed to eliminate the WT allele. This method is effective but the overall combined use of lambda proteins and Cas9 is subject to the same drawbacks of other Red-mediated methods [18, 20–22]. The exception is a recent report that employed a Cas9 nickase to initiate the generation of large genomic deletions [23]. The Cas9 nickase method is dependent upon naturally occurring (or pre-engineered) repetitive DNA elements in order for intervening DNA to be eliminated by homologous recombination. Due to its ability to target multiple loci utilizing different guide RNAs, the CRISPR/Cas9 system holds promise as a way to multiplex genome editing; however, currently the ability to simultaneously introduce changes into multiple locations of the *E. coli* chromosome is limited. For an overview of current genome engineering methods, see Fig. 1.

Given the various drawbacks of current methods, we sought to push the limits of speed, efficiency and versatility by leveraging recent advances in molecular biology. By utilizing a state-of-the-art DNA assembly approach to create sealed, circular DNA *in vitro*, we are able to completely eliminate the need to propagate plasmids in order to perform a genomic modification in *E. coli*. Our work utilized a modified version of Gibson assembly, however, any DNA assembly technique able to efficiently generate high levels of fully circularized DNA should be applicable in our protocol. Consequently, the described protocol allowed us to perform scarless modifications of the *E. coli* genome in approximately 48 h from the reception of PCR primers, significantly reducing the time relative to any other published method. This rapid turn-around allows for fast creation of complex engineered *E. coli* hosts tailored for specific applications, potentially saving weeks of effort if multiple modifications are required. Finally, given the simplicity of the protocol, we believe it should be possible to apply this method to a wide range of prokaryotic organisms, so long as the host is capable of efficient circular DNA uptake, homologous recombination, and the promoter driving the genes of the pDEL vector are functional in the target organism. The method described herein is named FAST-GE, an acronym chosen to indicate Fast Assembly-mediated Scarless Targeted Genome Editing.

Materials and methods

Bacterial strains, plasmids and enzymes

E. coli strain NEB10-beta (New England BioLabs, Ipswich, MA) was used for all cloning steps to create the pDEL vector. Plasmids containing the R6K origin were replicated in BW23473 [24]. *E. coli* strains NEB Express, T7 Express and ER2744 (New England BioLabs, Ipswich, MA) [25] were utilized in genome modification experiments. All enzymes were from New England BioLabs (Ipswich, MA). Chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Primers and gBlock® synthetic DNAs were synthesized by IDT (Coralville, IA). The DasherGFP gene was obtained from DNA2.0 (Menlo Park, CA). Kanamycin and ampicillin were used at 40 μg/ml and 100 μg/ml, respectively, unless otherwise stated. The *rhaBAD* and *lac* promoters were induced with 0.2% rhamnose and 0.5 mM IPTG, respectively. Counter-selection plates contained 5% (w/v) sucrose on plates containing 5 g/l of yeast extract, 10 g/l of tryptone and 7.5 g/l of agar, in addition to inducers as described above. Routine cell growth was performed at 37 °C in lysogeny broth medium supplemented with 0.1% glucose. Colony PCR screens were performed using Quick-Load Taq 2X Master Mix, while Q5 Hot Start High Fidelity 2X Master Mix (New England BioLabs, Ipswich, MA) was used to PCR-amplify genomic DNA for sequence verification.

Construction of pDEL

The open reading frame for the *sacB* gene was amplified from pRE112 [26] and inserted into pMAL-c5X using NEBuilder® HiFi DNA assembly to create pMAL-sacB. Subsequently, the *sacB* gene was amplified with a *lac* promoter. Low-level basal expression of *sacB* is not harmful in the absence of sucrose, as such the lac repressor is not required. If the strain encodes *lacI*, then IPTG is necessary during the resolution step. Plasmid pKD4 was used as the source of the kanamycin resistance gene and associated promoter [11]. The R6K origin of replication was amplified from pCD13PKS [27]. The region containing the I-SceI endonuclease site as well as the *I-SceI* gene under the control of *rha*BAD promoter was ordered as a gBlock®; the sequence is shown in Supplementary Table S1. The rhamnose promoter was chosen as it should be tightly repressed in the absence of the inducer, but will also generate sufficient levels of *I-SceI* transcripts in the majority of *E. coli* strains. To generate the pDEL vector, the following PCR fragments were assembled using an NEBuilder® HiFi DNA assembly reaction: the R6K origin, kanamycin resistance gene, *I-SceI* gBlock® and the P*lac-sacB* fragment. The final circular vector is shown in Supplementary Fig. S1 and the entire pDEL vector sequence is included in the Supplementary Material.

Genome modification experiments

For genome modification experiments, electrocompetent *E. coli* were prepared as follows: cells were grown in 3 ml cultures of LB until late exponential phase (OD600 ~ 0.8). Cells were harvested by centrifugation and washed 3 times with 500 μl of ice cold 10% glycerol. Finally, cells were resuspended in 50 μl of 10% glycerol.

Genome modification constructs were made by assembling upstream and downstream homology regions (Fig. 2, boxes A and C, respectively), created using site-specific primers containing regions of microhomology to each other as well as to a PCR product of pDEL amplified with primers located between the R6K origin of replication and the *I-SceI* gene. In cases when the desired modification was gene replacement, the PCR product of the new gene fragment was assembled between the upstream and downstream regions of homology. Typical assembly reactions contained 250–300 ng of total DNA, at manufacturer's recommended ratios in a total volume of 10 μl. Sequencing of individual clones generated from each of the HiFi DNA assemblies confirmed a very low error rate. Results are shown in Supplementary Table S2.

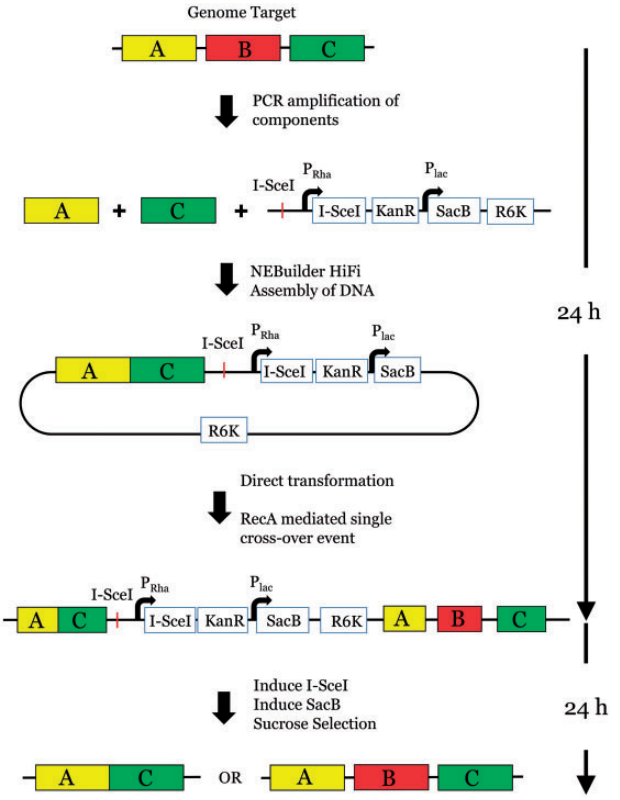

Figure 2: Schematic overview of the genome modification process. In order to modify *E. coli* chromosomal region B, flanking regions with at least 500 bp of homology on each side of B (boxes A and C) are generated. PCR is the easiest method to generate this linear DNA. To enable DNA assembly, the flanking regions should also contain small homology regions of 18–24 nucleotides to each other as well as to the linearized deletion cassette (amplified from pDEL). If region B is being mutated or replaced, as opposed to deleted, an additional fragment containing that modification must be included and will be assembled between A and C. Linear DNA fragments are subsequently assembled into a circular construct using NEBuilder® HiFi DNA assembly, transformed into electrocompetent cells, and the initial integration event is selected for by kanamycin resistance. The integration location is verified by PCR and colonies containing the desired insertion are subcultured into media containing the inducers for I-SceI and SacB, and subsequently on counter-selection plates with the same inducers. Expression of I-SceI promotes homologous recombination in response to a double-strand break in the genome and SacB is used to select against cells that failed to remove the integration cassette. The final colonies are screened by PCR for the presence of the desired genome modification.

For transformation, 2 μl of assembled DNA was mixed with electrocompetent cells and electroporated in a 1 mm cuvette according to manufacturer's instructions (BioRad). Electroporated cells were resuspended in 950 μl of SOC and allowed to recover at 37 °C for 1.5–2 h. Following the recovery step, 100 and 900 μl aliquots were plated on fresh LB-Kan plates and incubated at 37 °C overnight. Colonies from the LB-Kan plate were picked, transferred into a liquid LB-Kan culture and simultaneously screened by colony PCR to confirm the location of the initial recombination event. Cultures containing the desired integration were allowed to reach early log phase (approx. $OD_{600} = 0.2$), at

which point they were diluted 1:500 into fresh LB containing 100 μM IPTG and 0.2% w/v rhamnose but lacking kanamycin. Post induction, cultures were allowed to grow for an additional 3 h at 37 °C, after which they were plated on counter-selection plates containing 5% sucrose, 100 μM IPTG and 0.2% rhamnose. Following an overnight incubation at 37 °C, colonies from the counter-selection plates were picked and grown in LB to be analysed by PCR with locus-specific primers for removal of the integration cassette. In order to confirm the desired changes, the PCR products were column purified and sequenced using the Sanger method.

Results

Composition of the insertion cassette

When designing the pDEL vector, several critical factors were considered. First, the vector size was maintained as compact as possible in order to optimize amplification yield and fidelity. This led us to critically evaluate the components currently used with other genome modification approaches. We determined that at the minimum an antibiotic selection marker and a counter-selection marker were required. The frequency of homologous recombination in wild-type *E. coli* is low; however, recombination can be increased locally by the presence of a double-strand break. Thus, an I-SceI restriction site and the corresponding endonuclease gene were added to the construct to encourage homologous recombination by inducing a double-strand break. Additionally, we customized the promoters of the *sacB* and *I-SceI* genes, to improve the overall efficiency of the protocol. The final composition of the deletion cassette is shown in Fig. 2.

Kanamycin resistance was chosen as the primary selection marker as growth on kanamycin plates has proven to be a consistent indicator of single-copy genome integration [11, 16]. For counter-selection SacB was chosen as its activity and toxicity in *E. coli* has been well documented over the past several decades [28, 29]. The *sacB* gene originates from *B. subtilis* and encodes a periplasmic levansucrase. The exact mechanism of toxicity by SacB is still not well understood, but it is thought that periplasmic SacB creates large levan polymers when cells are grown in the presence of sucrose [29, 30]. In order to optimize *sacB* expression and periplasmic localization, we replaced the native *Bacillus* promoter with a *lac* promoter. Counter-selection on sucrose proved to be very robust as a large majority of surviving colonies were free of the *sacB* gene (Supplementary Table S3).

Inclusion of the I-SceI homing endonuclease as well as the cognate recognition site was based on several recent papers demonstrating the increased recombination frequency during genome alteration by the generation of a unique double-strand break in the *E. coli* chromosome by I-SceI. In order to augment I-SceI cleavage efficiency, two terminators were introduced immediately upstream of the I-SceI cut site as active transcription through the recognition sequence was recently reported to lower I-SceI cleavage efficiency [10].

In addition to the components actively involved in the recombination process, we chose to include the R6K origin of replication on the pDEL plasmid. The R6K origin is unique in that it requires the presence of the *pir* gene product for the plasmid to replicate; common *E. coli* strains lack this gene. The conditional origin was included as an alternative strategy, in cases where plasmid isolation is a more practical first step. In our experience, the DNA assembly products were successfully transformed and integrated directly in all strains tested.

Rapid, targeted genome modification

Three changes were made to non-essential chromosomal regions: deletion of the *lac* operon (Fig. 3A), a point mutation in the *lacZ* gene (Supplementary Fig. S1A), and insertion of the T7 RNA polymerase gene (Fig. 4). Changes were introduced into both *E. coli* B and K-12 backgrounds. These modifications were chosen to demonstrate the capability of the FAST-GE method to easily generate a large deletion, a point mutation and an insertion, respectively. The detailed illustration of the construct used to generate the deletion of the *lac* operon in ER2523 (an *E. coli* B derivative) is shown in Fig. 3A. Electrocompetent ER2523 cells were transformed with 150 ng of DNA from the assembly reaction. This protocol was sufficient to result in several correctly integrated transformants (Fig. 3B). Two different primer pairs are used to identify strains with the pDEL construct integrated correctly at the desired locus. All eight colonies analysed contained a pDEL integration at the *lac* locus when combining the results of the F1-R1 and F2-R2 PCR analyses. Each primer pair has two potential products depending on whether recombination occurred via the 5′ flank homology region or the 3′ flank homology region. In each diagnostic PCR reaction, the extension time is set to amplify the shorter of the two possible PCR products. We found that this experimental approach was acceptable as shown in Fig. 3B.

Recombinant number one was chosen for the resolution step. After reaching noticeable turbidity grown in the presence of kanamycin (OD_{600} = approx. 0.2), the culture from recombinant number one was diluted 1:500 into fresh LB medium lacking kanamycin but containing rhamnose and IPTG, for induction of *I-SceI* and *sacB*, respectively. After 2 h of incubation in antibiotic-free medium, dilutions were plated on sucrose agar plates. Following counter-selection on sucrose, final resolution and removal of the integration cassette was highly effective. Table 1 presents a summary of the process for deletion of the *lac* operon. The removal of the counter-selection cassette relies upon homologous recombination and depending on the recombination site the process can lead to one of two outcomes. First, the entire integration cassette may be removed such that the original genome sequence is restored (henceforth referred to as WT sequence). Alternatively, the integration cassette may be removed resulting in the desired genomic modification (Fig. 3A). Blue colony colour (X-gal conversion) was used as an indicator of *lac* operon function and Fig. 3C shows that 9 of 16 resolved clones had regained β-galactosidase activity, suggesting resolution to WT sequence. PCR analysis of the same clones showed that six of the recombinants had the desired deletion of the *lac* operon. Several of the resulting 1.3 kb F1-R2 PCR products were sequenced to find that no unintended mutations were introduced.

The described method of genome engineering is not biased according to the modification type. To demonstrate the ability of the FAST-GE method to introduce large insertions in the *E. coli* genome, the T7 RNA polymerase (*T7 gene 1*) was inserted into the chromosome of MC1061, a common *E. coli* K-12 strain. *T7 gene 1* may be inserted into the *lac* locus to facilitate IPTG-inducible expression of the T7 RNA polymerase for the purpose of recombinant protein expression. Accordingly, we chose to recreate the *lacZ-T7gene1* operon fusion (approx. 3 kb), as found in T7 Express cells, between the *yahF* and *mhpT* genes (Fig. 4A). The integration and counter-selection steps were followed by colony PCR analysis as described above (Fig. 4). As with other experiments, we were able to obtain desired clones in 48 h from the beginning of the process. Using a similar procedure, we were also able to replace a single nucleotide in order to create an E462A active site substitution within LacZ (Supplementary Fig. S1).

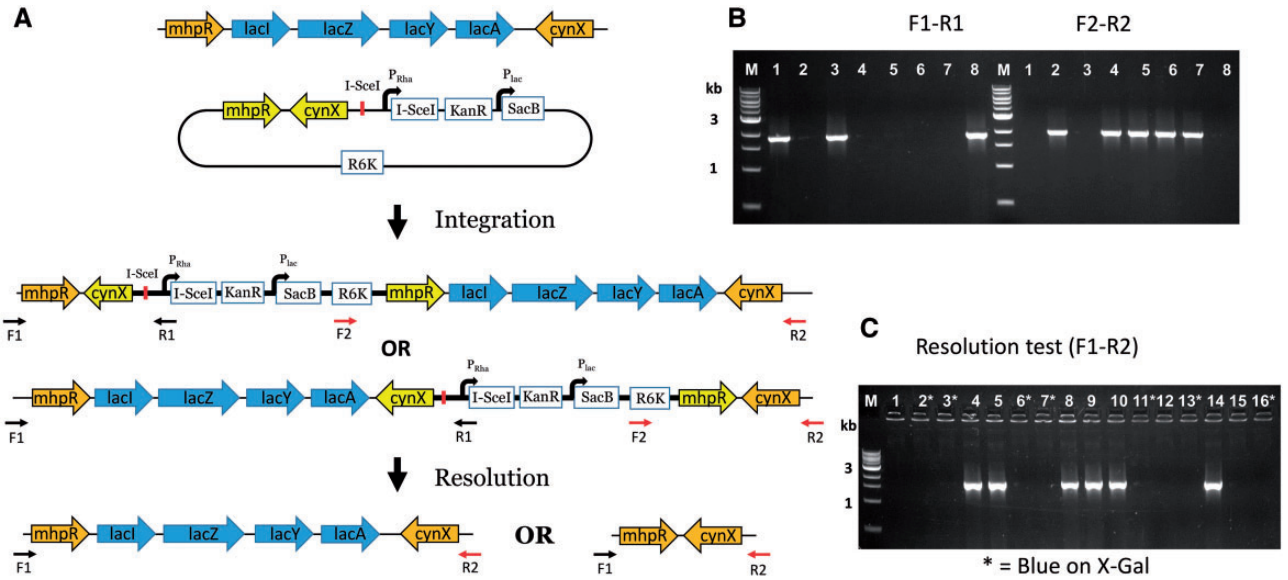

Figure 3: Deletion of the *lac* operon. **(A)** Overview of the genomic locus around the *lac* operon in strain ER2523 as well as the composition of the deletion cassette. The scheme demonstrates the two potential integration outcomes as well as the location of primer pairs F1-R1 and F2-R2 used to determine the location of the integration event. Finally, two potential outcomes of the resolution step are shown. **(B)** An agarose gel showing a representative subset of PCR products obtained with the F1-R1 and F2-R2 primer pairs. Expected sizes for the F1-R1 PCR products are either 1.6 kb if the crossover event occurred in the *mhpR* gene or 8.5 kb if the first crossover event occurred in the *cynX* gene. The expected sizes for the PCR products from the F2-R2 primer pair are either 7 kb if the crossover event occurred in the *mhpR* gene or 1.7 kb if the first crossover event occurred in the *cynX* gene. A DNA molecular ladder (Quick-load 1 kb ladder) is designated as M and 1 and 3 kb markers are labelled. The same set of individual colonies numbered 1-8 served as a template for both F1-R1 and F2-R2 PCR reactions. **(C)** An agarose gel of PCR results utilizing the primer pair F1 and R2 from individual colonies post counter-selection on sucrose. Expected PCR product size from a successful deletion of the lac operon is 1.3 kb.

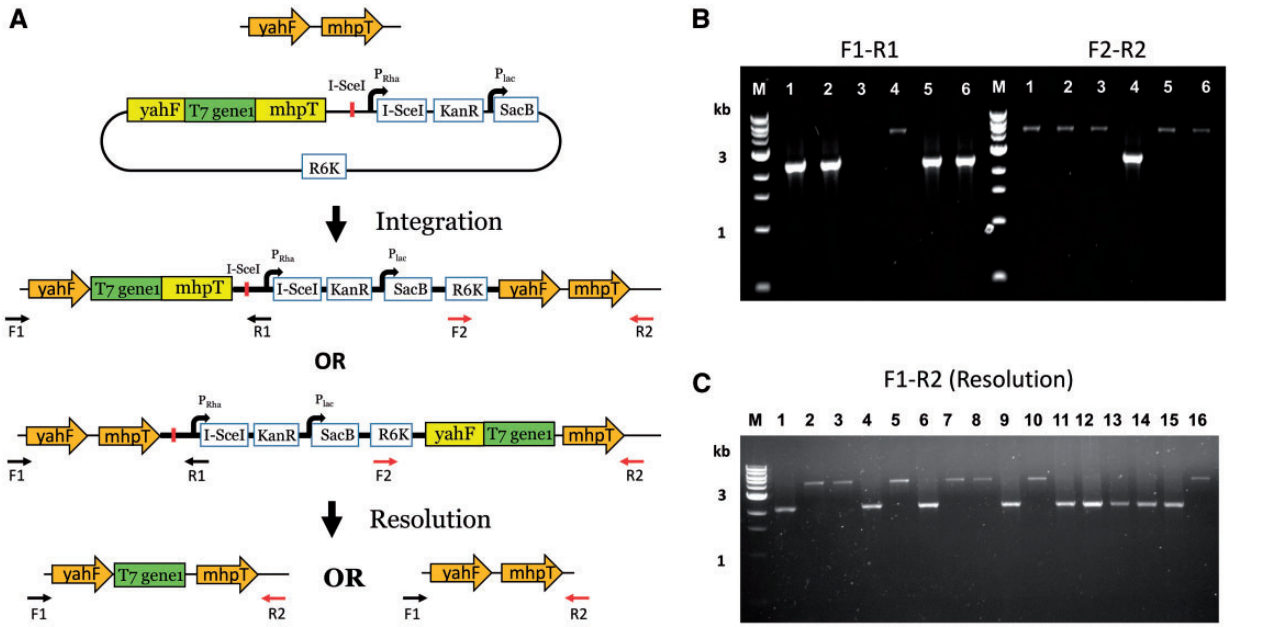

Figure 4: Insertion of the T7 RNA polymerase gene into K-12 strain MC1061. **(A)** Overview of the genomic locus around the *yahF-mhpT* locus as well as the composition of the T7 RNA polymerase gene insertion cassette. The scheme demonstrates the two potential integration outcomes after the first crossover event and the location of primer pairs F1-R1 and F2-R2 used to determine the location of the integration event. Finally, two potential outcomes of the resolution step are shown. **(B)** An agarose gel showing a representative subset of PCR products obtained with the F1-R1 and F2-R2 primer pairs. The expected sizes for the F1-R1 PCR products are either 4.5 kb if the crossover event occurred in the *yahF* gene or 2.5 kb if the first crossover event occurred in the *mhpT* gene. The expected sizes for the PCR products from the F2-R2 primer pair are either 2.6 kb if the crossover event occurred in the *yahF* gene or 4.7 kb if the first crossover event occurred in the *mhpT* gene. A DNA molecular ladder (Quick-load 1 kb ladder) is designated as M and 1 and 3 kb markers are labelled. The same set of individual colonies numbered 1-6 served as a template for F1-R1 and F2-R2 PCR reactions. **(C)** An agarose gel showing PCR results from 16 representative colonies post counterselection on sucrose (1–16) from primer pair F1 and R2. The expected PCR product size for successful T7 *gene* 1 insertion is 5.2 kb. Resolution to WT *yahF-mhpT* locus yields a PCR product of 2 kb. Lack of a PCR product would suggests that the integration cassette was not successfully removed from the genome.

Table 1: The efficiency of the genome modification protocol at various loci and modification types

Allele	Integration efficiency	Resolution efficiency	Desired modification
LacZ E462A	3/8	5/16	3/5
Δ*lac* (B strain)	8/8	13/16	6/13
T7 gene 1	5/6	16/16	7/16
rplI-Dasher	1/4	7/16	0/7

The ratio of isolated colonies reverting to WT sequence as opposed to the desired modification varied depending on the modification and the integration locus (Table 1).

Working with essential genes

The method described in this article is also beneficial for determining whether a gene is essential for growth of the bacterium or if a specific mutation destroys function. Since a fully functional copy of the target gene is maintained during the integration event, integration should always be possible as the present method separates the integration and resolution steps. If all of the resolved colonies in a sufficiently large sample have reverted to WT sequence, then one can reasonably assume that the desired modification is not tolerated by the cell (Fig. 5). In contrast, in lambda Red-derived methods, the absence of viable transformants does not necessarily indicate that a gene is essential or that the modification being attempted is not tolerated by the cell. An additional advantage of this method is that resolution can be carried out under different culture conditions, such as varied temperatures or growth medium composition. Thus, one can determine if a gene's activity is essential for survival under varied experimental conditions.

When we attempted to fuse a fluorescent protein (DasherGFP) to the C-terminus of the ribosome L9 subunit, corresponding to the *rplI* gene, colonies containing the initial crossover event were isolated (Fig. 5B). Out of the four colonies tested, only one returned the expected fragment size for the two diagnostic PCR reactions (Fig. 5B). The colony showing the expected PCR pattern was chosen for resolution. However, none of the colonies tested after sucrose selection contained the desired gene fusion (Table 1) (Fig. 5C). This led us to conclude that this fusion protein is not tolerated by *E. coli* due to disruption of ribosome function.

Best practices for genome modification success

When optimizing the FAST-GE protocol, we have noted several points where special care should be practiced to achieve best results and we offer several guidelines to improve the likelihood of success for first-time users. First, as with other homology-driven genome modification methods, the size of the flanking homology regions will affect the efficiency of the initial recombination event. Following standard practice for RecA-mediated homologous recombination, we recommend designing homology regions of at least 500 bases on both sides of the desired

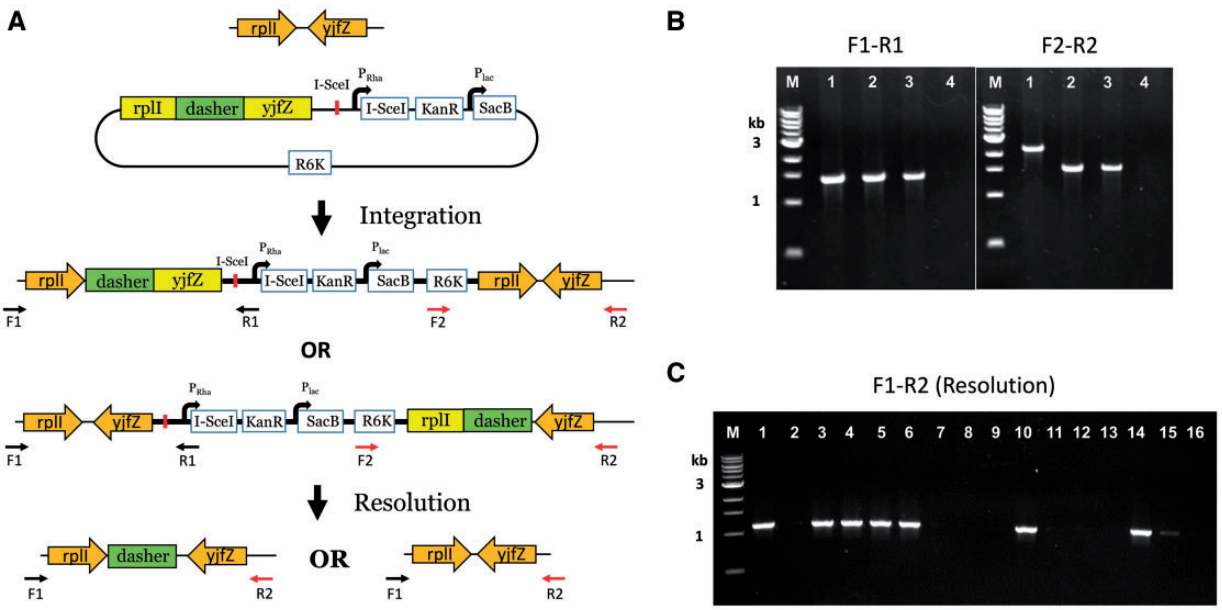

Figure 5: Working with an essential gene. **(A)** Overview of the *rplI* genomic locus as well as the composition of the *rplI*-dasher fusion cassette. The scheme demonstrates the two potential integration outcomes after the first crossover event and the location of primer pairs F1-R1 and F2-R2, used to determine the location of the integration event. Finally, two potential outcomes of the resolution step are shown. **(B)** An agarose gel showing a sample of a PCR product obtained with the F1-R1 and F2-R2 primer pairs. The expected sizes for the F1-R1 PCR products are either 2 kb if the crossover event occurred in the *rplI* gene or 1.3 kb if the first crossover event occurred in the *yjfZ* ORF. Similarly, the expected sizes for the PCR products produced by the F2-R2 primer pair are either 1.4 kb if the crossover event occurred in the *rplI* gene or 2.1 kb if the first crossover event occurred in the *yjfZ* ORF. A DNA molecular ladder (Quick-Load 1 kb ladder) is designated as M and 1 and 3 kb markers are labelled. The same set of individual colonies numbered 1–4 served as a template for both F1-R1 and F2-R2 PCR reactions. **(C)** An agarose gel of PCR results from 16 representative colonies post counter-selection on sucrose from primer pair F1 and R2. The expected PCR product size for successful *rplI*-dasher fusion is 1.6 kb. Resolution to WT *rplI* sequence yields a PCR product of 1 kb. Lack of a PCR product suggests that the integration cassette was not successfully removed from the genome.

modification [16]. Secondly, researchers may choose to transform the assembled genome modification construct into a *pir*$^+$ strain in parallel with transformation into the desired host. This optional, extra step ensures that a suicide plasmid is available to use, in case direct transformation of the assembly reaction does not yield colonies or if electroporation is not an option for transformation. Thirdly, if multiple changes to the genome are desired, we suggest that all of the necessary genome modification constructs be assembled in parallel in a *pir*$^+$ host and after the first modification, purified suicide plasmids can be used to perform subsequent modifications, simplifying and expediting generation of the desired strain.

Fourthly, when attempting direct transformations, we strongly recommend that a high concentration of DNA be used for the DNA assembly reaction, up to 250–300 ng of DNA can be assembled in a 10 µl reaction, with subsequent transformation of 2 µl of this reaction. If working with strains known to have low competence, larger assemblies on the scale of 500–1000 ng of DNA in a 50 µl reaction may be necessary. Then, concentrate the DNA before the transformation step.

Discussion

For decades, genome editing in bacteria has relied upon the use of plasmids, either as a template for recombination or as the source of helper proteins to facilitate the recombination event. The reliance on replicative plasmids, requires an initial transformation step to introduce them into the hosts, as well as a curing procedure to remove them after the genome modification protocol is completed. In some methods, plasmid curing is not straightforward. In allele exchange methods where the plasmid is carrying the WT gene after resolution, if plasmid curing fails, then the researcher is left wondering whether the WT gene might be essential for viability. Some allele exchange methods are particularly time-consuming and may require 2–3 weeks from construct design to verification of the final strain [16]. In this study, we demonstrate the ability to modify a bacterial genome without the need for a replicative plasmid, thus drastically reducing protocol time. This feat is accomplished by leveraging recent advances in DNA assembly technologies in order to transform the cells with a non-replicative, circularized piece of DNA. The assembled construct may be immediately transformed into the desired strain without the need for sequence verification due to high fidelity of the assembly process. Importantly, only a small number of resolved strains (typically fewer than eight) need to be analysed by focused sequencing reactions to verify the genome alteration as well as fidelity of the assembly process.

Eliminating the requirement for helper plasmid(s) enables the generation of markerless genomic modifications within 48 h from receipt of the necessary primers. Minimum requirements for this protocol are fragments of DNA sequence flanking the site of the desired modification and linear DNA encoding the pDEL deletion cassette, with all pieces containing small homology regions sufficient to allow assembly via the NEBuilder® HiFi DNA assembly method (or other high-efficiency DNA assembly methods). A high fidelity DNA polymerase should be employed to obtain the necessary DNA fragments and bacterial colonies may serve as the source of chromosomal template DNA to generate fragments for the DNA assembly reaction. As a result, time-consuming plasmid or genomic DNA purification procedures may be bypassed. To our knowledge, this protocol is at least 2 days faster, from start to finish, than any other method currently described. In addition, the FAST-GE method is accessible to anyone familiar with PCR and electroporation-based transformation protocols.

While we were able to generate all of our desired mutants in both K-12 and B derivatives using direct transformation of the assembly reaction, we do understand that some researchers may prefer to use a purified plasmid when working with difficult to transform strains. For those preferring to work with an isolated assembly clone we have included the conditional R6K origin of replication on the pDEL vector, which allows transformation of the assembly reaction into a *pir*$^+$ host for multi-copy replication.

In a drive to improve the overall speed of the protocol, versatility was not compromised as insertions, deletions and point mutations are all equally possible. The only limitation of this method (and many other methods) is that the strain to be modified must be capable of homologous recombination. For example, typical cloning strains (e.g. DH5α) containing a *recA* mutation are not suitable for modification by this method. Colonies containing the desired initial crossover event are easy to identify by colony PCR and require at most the screening of eight colonies, which is on par or better than any other protocols reported to date. Additionally, no difference in the frequency of integration was apparent regardless of whether the target gene was essential for *E. coli* growth. This combination of efficiency and speed is unique and is of especially high value to researchers interested in making multiple genome modifications.

Author Contributions

I.B.T. and J.C.S. designed research. I.B.T. performed experiments. I.B.T. and J.C.S. analysed data and wrote the paper.

Acknowledgements

We would like to thank Elisabeth A. Raleigh and William Jack for critical reading of the manuscript and Don Comb and Jim Ellard for creating and fostering a creative research environment.

Funding

Funding for open access charge: New England BioLabs, Inc.

Conflict of interest statement. The authors of this manuscript are employees of New England Biolabs, which produces molecular biology reagents, including NEBuilder HiFi DNA Assembly Master Mix.

Supplementary data

Supplementary data is available at *Biology Methods and Protocols* online.

References

1. Trinh CT, Li J, Blanch HW *et al*. Redesigning *Escherichia coli* metabolism for anaerobic production of isobutanol. *Appl Environ Microbiol* 2011;**77**:4894–904.
2. Baba T, Ara T, Hasegawa M *et al*. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2006;**2**:2006.0008. doi:10.1038/msb4100050.
3. Liu H, Yu C, Feng D *et al*. Production of extracellular fatty acid using engineered *Escherichia coli*. *Microb Cell Fact* 2016;**11**:41–54.

4. Snyder WB and Silhavy TJ. Enhanced export of beta-galactosidase fusion proteins in prlF mutants is Lon dependent. *J Bacteriol* 1992;**174**:5661–68.
5. Lobstein J, Emrich CA, Jeans C *et al*. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb Cell Fact* 2012;**11**:56.
6. Ajikumar PK, Xiao WH, Tyo KE *et al*. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* 2010;**330**:70–4.
7. Lee SJ, Lee DY, Kim TY *et al*. Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Appl Environ Microbiol* 2005;**71:**7880–87.
8. Kolisnychenko V, Plunkett G, Herring CD *et al*. Engineering a reduced *Escherichia coli* genome. *Genome Res* 2002;**12**:640–47.
9. Robichon C, Luo J, Causey TB *et al*. Engineering *Escherichia coli* BL21 (DE3) derivative strains to minimize *E. coli* protein contamination after purification by immobilized metal affinity chromatography. *Appl Environ Microbiol* 2011;**77:**4634–46.
10. Kim J, Webb AM, Kershner JP *et al*. A versatile and highly efficient method for scarless genome editing in *Escherichia coli* and *Salmonella enterica*. *BMC Biotechnol* 2014;**14:**84.
11. Datsenko KA and Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 2000;**97**:6640–45.
12. Yu BJ, Kang KH, Lee JH *et al*. Rapid and efficient construction of markerless deletions in the *Escherichia coli* genome. *Nucleic Acids Res* 2008;**36:**e84.
13. Herring CD and Blattner FR. Conditional lethal amber mutations in essential *Escherichia coli* genes. *J Bacteriol* 2004;**186**:2673–81.
14. Yang J, Sun B, Huang H *et al*. High-efficiency scarless genetic modification in *Escherichia coli* by using lambda Red recombination and I-SceI cleavage. *Appl Environ Microbiol* 2014;**80:**3826–34.
15. Link AJ, Phillips D and Church GM. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* 1997;**179:**6228–37.
16. Hamilton CM, Aldea M, Washburn B *et al*. New method for generating deletions and gene replacements in *Escherichia coli*. *J Bacteriol*1989;**171:**4617–22.
17. Pósfai G, Kolisnychenko V, Bereczki Z *et al*. Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Res* 1999;**27:**4409–15.
18. Pyne ME, Moo-Young M, Chung DA *et al*. Coupling the CRISPR/Cas9 system to lambda Red recombineering enables simplified chromosomal gene replacement in *Escherichia coli*. *Appl Environ Microbiol* 2015;**81**:5103–14.
19. Wang HH, Xu G, Vonner AJ *et al*. Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion. *Nucleic Acids Res* 2011;**39**:7336–47.
20. Touchon M and Rocha EP. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS One* 2010;**5:**e11126.
21. Jiang Y, Chen B, Duan C *et al*. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* 2015;**81:**2506–14.
22. Jiang W, Bikard D, Cox D *et al*. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 2013;**31:**233–39.
23. Standage-Beier KS, Zhang Q and Wang X. Targeted large-scale deletion of bacterial genomes using CRISPR-Nickases. *ACS Synth Biol* 2015;**4**:1217–25.
24. Haldimann A, Prahalad MK, Fisher SL *et al*. Altered recognition mutants of the response regulator PhoB: a new genetic strategy for studying protein-protein interactions. *Proc Natl Acad Sci USA* 1996;**93:**14361–366.
25. Samuelson JC, Zhu Z and Xu S. The isolation of strand-specific nicking endonucleases from a randomized SapI expression library. *Nucleic Acids Res* 2004;**32:**3661–71.
26. Edwards RA, Keller LH and Schifferli DM. Improved allelic exchange vectors and their use to analyze 987P fimbria gene expression. *Gene* 1998;**207:**149–57.
27. Platt R, Drescher C, Park SK and Phillips GJ. Genetic system for reversible integration of DNA constructs and lacZ gene fusions into the *Escherichia coli* chromosome. *Plasmid* 2000;**43:**12–23.
28. Gay P, Le Coq D, Steinmetz M *et al*. Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. *J Bacteriol* 1983;**153**:1424–31.
29. Blomfield I, Vaughn V, Rest R *et al*. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. *Mol Microbiol* 1991;**5:**1447–57
30. Reyrat JM, Pelicic V, Gicquel B *et al*. Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infect Immun* 1998;**66**:4011–17.