Research Article

Use of transposase and ends of IS608 enables precise and scarless genome modification for modulating gene expression and metabolic engineering applications in *Escherichia coli*

Chandresh Thakker[1,2]

Kevin Lin[1]

Heidi Martini-Stoica[1,3]

George N. Bennett[1]

[1]Department of BioSciences, Rice University, Houston, TX, USA

[2]TOTAL New Energies USA, Inc., 5858 Horton Street, Emeryville, CA, USA

[3]Interdepartmental Program in Translational Biology and Molecular Medicine, Baylor College of Medicine, Houston, TX, USA

Correspondence: Dr. George N. Bennett, Department of BioSciences, Rice University, 6100 Main St., 77005 Houston, TX, USA.

E-mail: gbennett@rice.edu

Keywords: counterselection, genome modification, IS*608,* Red recombinase, transposase

Abbreviations: LE, left End of IS608; **RE,** Right End of IS608; 5FOA. 5-fluorooroate; FLP, Flippase recombination enzyme; FRT, Flippase Recognition Target; RBS, Ribosome Binding Site

Abstract

Background:

Various methods have been developed for gene disruption in bacteria; however, extra in vitro manipulation steps or the residual presence of a scar in the host chromosome limits the use of such methods.

Methods and results:

By utilizing the unique properties of ISHp608, we have developed a simple and precise method for genome manipulation in Escherichia coli that alters the gene sequence without leaving foreign DNA in the chromosome. This strategy involves PCR amplification of a DNA cassette containing ISHp608-LE (left end)-antibiotic resistance gene-counterselection marker- ISHp608-RE (right end) by using primers containing extensions homologous to the adjacent regions of the target gene on the chromosome. The λ Red mediated recombination of the PCR product and antibiotic resistance screening results in transformants with a modified gene target. The ISHp608-LE-antibiotic resistance gene-counterselection marker-ISHp608-RE cassette can then be excised using a temperature sensitive plasmid expressing the tnpA transposase, which precisely cleaves ISHp608-LE and ISHp608-RE without leaving a scar sequence.

Conclusions:

We demonstrated lacZ gene point mutation repair, two precise disruptions of the lacZ gene and constructed a library of lacZ variants having variable β-galactosidase activity by changing its ribosome binding site sequences using the ISHp608 system. This technique can be used in E. coli genome modification and could be extended for use in other bacteria.

1 Introduction

Development of new molecular tools and advancement in genome engineering has been useful for various system biology and biotechnological applications of microbial systems. An important molecular tool in functional genomics studies is the targeted inactivation of any gene of interest. While a number of methods have been developed that allow disruption of specific genes in *E. coli* [1-9] and other organisms [10-14], many of these methods that utilize recombination introduce a scar in disrupted genes which could lead to unforeseen genetic consequences such as intrachromosomal rearrangements[8].

Transposable elements, a diverse group of mobile genetic elements present in most genomes, do not leave behind a scar sequence. Transposition requires cleavage by a transposase enzyme at the transposon ends for transfer into a target site and a variety of structurally and mechanistically distinct transposase enzymes are known [15]. The enzyme *tnpA* has been reported for the transposition of ISHp*608* (IS*608*) in *Helicobacter pylori* and is active in *E. coli* [16]. The IS*608* TnpA 155-amino acid transposase has unusual properties, uses a single tyrosine (Y) residue as a nucleophile to catalyze DNA strand breakage and transfers. IS*608* family does not possess terminal internal repeats (IRs), but does have left (LE) and right ends (RE) that include secondary structures located at some distance from the cleavage sites [16, 17]. The cleavage site at LE is 19 nucleotides 5' to the foot of a hairpin preceded by a flanking tetranucleotide sequence TTAC, whereas, the cleavage site (TCAA) at RE is 10 nucleotides 3' to the hairpin [18].

IS*608* insertion and excision occurs 3' to a conserved and essential tetranucleotide TTAC directly adjacent to LE [16]. IS*608* transposition is strand specific and thought to involve precise excision of insertion sequences in the top strand, resulting in circular transposon intermediates with abutted RE and LE copies (transposon joint or RE-LE junction). The circular transposon, a transposition intermediate, then undergoes insertion. Following excision of the transposon

cassette, the adjacent DNA strands are re-ligated, preserving the tetranucleotide TTAC and without any loss or gain of DNA sequences [17-19]. The unique capability of IS*608* to excise the transposable element without loss or gain of DNA sequences is of great interest with respect to targeted genomic modifications. Precise gene manipulations can be carried out without leaving any residual scar sequence in the host chromosome by employing IS*608* LE and RE together with IS*608* encoded transposase gene *tnpA*. With this objective, we developed a precise and scarless genome manipulation technique using *E. coli*. To introduce IS*608* LE and RE into the gene of interest, we utilize λ-Red recombination[7], which does not require *in vitro* manipulations such as restriction digestion, ligation or construction of a suicide vector. Here we describe a procedure based on the λ Red system and IS*608* to successfully demonstrate gene disruptions, point mutation repair and the construction of a gene library of mutations in the ribosome binding site (RBS).

2 Materials and methods

2.1 Media, culture conditions and reagents

E. coli strains were routinely cultivated in Luria-Bertani (LB) medium at 37°C. Strains harboring temperature sensitive plasmid were cultivated at 30°C. For curing of temperature sensitive plasmid, host strains were cultivated nonselectively at 42°C. An appropriate concentration of ampicillin (50-100 μg/ml), kanamycin (Km) (20-50 μg/ml), chloramphenicol (Cm) (15-25 μg/ml) or spectinomycin (Sp) (50-100 μg/ml) was added to the medium for selection. LB agar containing 10% sucrose was used to identify the sucrose resistant or sensitive phenotype strain. Minimal medium AB [20] supplemented with 0.3% sodium succinate, 0.3% sodium fumarate, 1% glucose, 10 μg/ml thiamine, 18 μg/ml 5-fluoroorotate (5FOA) and 1.5% agar was used, henceforth referred as ABO medium, while using *oroP* as a counterselection marker [21]. 5FOA was from Gold Biotechnology, Inc. (St. Louis, MO) and dissolved in DMSO. Isopropyl- *β*-D-galactopyranoside (IPTG)

and 5-bromo-4-chloro-3-indolyl-*β*-D-galactopyranoside (X-Gal) were used at 40 μg/ml for blue-white screening of colonies.

2.2 Bacterial strains

Bacterial strains used in this study are listed in Table 1. *E. coli* MG1655 (Wild type, $F^-\lambda^-$) obtained from CGSC, The Coli Genetic Stock Center (New Haven, CT, USA), and was used for propagating non-conditional replicative plasmids. AB2572 [Hfr(PO1), *lacZ13*(Oc), *relA1*, *thi-1*] and KL334 [Hfr(PO2A), lacZ118(Oc), lacI22, and λ^-, lysA23] are derivatives of the *E. coli* K-12 and were also obtained from CGSC. CT2572 [Hfr(PO1), *relA1*, *thi-1*], a derivative of AB2572, was constructed during this study by correcting *lacZ13*(Oc) mutation. Conditional replicative *oriRγ* plasmids were maintained in the *pir*$^+$ host PIR 1 [F^-Δ*lac169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA(*Δ*MluI)::pir-116*], which was obtained from Invitrogen (CA, USA). *E. coli* ABLE-C (Stratagene), which was *E. coli* C *lac*(LacZ$^-$) [Kan^R McrA$^-$ McrCB$^-$ McrF$^-$ Mrr$^-$ HsdR (r_k^- m_k^-)] [F′ *proAB lacI*^{q}ZΔM15 Tn*10*(Tet^R)], harboring pCS1966 was obtained from J. Martinussen (Technical University of Denmark, Denmark) [21].

2.3 Plasmids

Plasmids used and constructed in this study are listed in Table 1. Plasmid pMARMAR 4 [22] was obtained from M. Goulian (University of Pennsylvania, Philadelphia), *tnpA* expressing pBS107 [17] from M. Chandler (Laboratoire de Microbiologie et Génétique Moléculaires, France), pLOI2227 [23] from L. O. Ingram (University of Florida, Gainesville, FL), the Red recombinant plasmid pKM208 [6] from K. Murphy (University of Massachusetts Medical School, MA) and pAH68 [24] from B. Wanner (Purdue University, IN). To construct plasmid pCT104, the 2.5 Kb fragment containing *cat* (Cm^R) and *sacB* genes was amplified by PCR from pMARMAR4 using primers LE-NdeI and RE-NdeI followed by cloning into PCR 2.1-TOPO vector (Invitrogen). The *Xba*I-*Sac*I fragment of PCR 2.1-TOPO was then sub cloned into pAH68 vector. The plasmid pHMS102 was constructed by amplifying 2.5 Kb fragments containing *cat* and *sacB* genes from pMARMAR4 by PCR using primers

LE-PstI and RE-BamHI and then cloning the fragment into pAH68 vector using *Pst*I and *BamH*I. To construct plasmid pKJL12, we first made pCT200 by PCR amplification of the P_{syn}-o*roP* region from pCS1966 using primers oroP-BamF and oroP-PstR and cloned into pACYC177 using *BamH*I and *Pst*I. pCT200 was then used as template to amplify a 3 Kb region containing kanamycin resistance (*kan*) and P_{syn}-*oroP* by PCR using primers KpnI-LE-KanF and EcoRI-RE-KanR followed by cloning into pAH68 vector using *Kpn*I and *EcoR*I sites. Primers LE-PstI and KpnI-LE-KanF contain the 64 bp LE, whereas, primers RE-BamHI and EcoRI-RE-KanR contain the 42 bp RE similar flanking regions. Plasmids, pCT104, pHMS102 and pKJL12, were verified by nucleotide sequencing for sequence accuracy using primers pAH68F and pAH68R. Plasmids pCT104 and pHMS102 differ only with respect to the LE arm of IS*608*. The pCT104 harbors 42 bp of LE3, whereas, pHMS102 harbors 64 bp of LE1 as described by [17]. The *tnpA* expressing temperature sensitive plasmid pCTS21 was constructed by PCR amplifying an ~3.7 Kb region, containing the P*lac* promoter, *tnpA*, spectinomycin resistance gene and $lacI^q$, from pBS107 using primers pAPT-plac-F and pAPT-lacIq-R and cloning into pLOI2227 vector using *EcoR*I site. Linear representations of the template plasmids pCT104, pHMS102 and pKJL12 are shown in Fig. 1A and *tnpA* expressing temperature sensitive plasmid pCTS21 map is shown in Fig. 1B.

2.4 DNA manipulations and oligonucleotides

Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA) and used as per recommended conditions. Synthesized oligonucleotides were from Sigma-Aldrich (St. Louis, MO) and are shown in Table 2. Phusion High-Fidelity DNA polymerase (Finnzyme, Finland) or FailSafe PCR enzyme mix (Epicenter Biotechnologies, Madison, WI) was used for cloning of DNA fragments and gene disruption. *Taq* polymerase (5 Prime, Gaithersburg, MD) was used for the PCR verification of cloned fragments in plasmid or genome. PCR conditions were as recommended by the manufacturer. Nucleotide sequencing was carried out to verify desired DNA modifications. Plasmid DNA was prepared using either the GenElute HP plasmid

miniprep kit (Sigma-Aldrich) or the HiSpeed plasmid midi kit (Qiagen, Germany). DNA fragments and PCR products were recovered using a QIAquick gel extraction kit and a QIAquick PCR purification kit (Qiagen, Germany), respectively.

2.5 Chromosomal gene disruption

Host strains harboring a Red recombinant plasmid pKM208 were prepared according to the method described by [6]. Plasmid pKM208 expresses the λ Red system under control of an IPTG inducible promoter. PCR products, containing either a 2.5 Kb LE-Cm^R-*sacB*-RE or 3 Kb LE-Km^R-*oroP*-RE region, were obtained using primers P1 (5'-CAAAACTAACGCCTTAAAGCCC-3') and P2 (5'-TTGACATACTCCCCATAGCTAAAGC-3') with 50-75 bp homology extension of the target gene sequence. PCR products were purified by gel elution followed by *Dpn*I digestion and repurification. A 50 µl sample of cells was mixed with 3-5 µl (300-500 ng) of purified PCR products in pre-chilled 1 mm cuvette. Electroporation was performed by using Bio-Rad Gene Pulser (Bio-Rad, CA) with a setting of 1.8 kV, 25 µFD and 200 ohms. Cells were suspended immediately by adding 0.3 ml LB followed by dilution to 2.7 ml LB in 15 ml falcon plastic tube and incubated at 37°C, 250 rpm for 1-2 hr. Cells were plated on LB agar plates containing either 15 µg/ml chloramphenicol or 20 µg/ml kanamycin and incubated at 37°C for 12-15 hr. Occasionally, the remainder of the cell broth was kept overnight at room temperature and spread on LB agar plates to select Cm^R or Km^R candidates. The Cm^R or Km^R candidates were maintained on grid plates at 37°C and tested for ampicillin sensitivity at 30°C to check for the loss of temperature sensitive Red recombinase plasmid pKM208. If the plasmid was not lost, then Ap^R candidates were colony-purified once or twice at 42°C and confirmed for the loss of plasmid.

2.6 Verification by counterselection marker and PCR

Disruption of the target gene was confirmed by the presence of the counterselection marker *sacB* or *oroP.* The Cm^R transformants were streaked on LB agar + 10% sucrose grid plate and sucrose sensitive transformants were selected, while Km^R transformants were streaked on ABO

medium and 5FOA sensitive transformants were selected. Colony PCR was also used to confirm the sucrose and 5FOA sensitive transformants. Cells were scrapped using sterile tip from the grid plate and suspended in 25 µl sterile water in 1.5 mL Eppendorf tube followed by incubation at 95°C in heating block for 10-15 min. Immediately after incubation, tubes were transferred to ice bucket and were kept for 2-3 min followed by centrifugation at 10,000 × *g* for 2 min to remove cell debris. Supernatant (2-3 µl) was used as crude DNA in a 20 µl PCR reaction using primers P1 and P2 to confirm the insertion of the new fragment (LE-Cm^R-*sacB*-RE or LE-Km^R-*oroP*-RE) in the host chromosome. Alternatively, upstream and downstream primers of the target gene from the disruption site were also used to confirm the gene disruption.

2.7 In vivo excision of LE-antibiotic gene-counterselection marker-RE

Colony PCR verified Cm^R or Km^R candidates were transformed with pCTS21 and colonies were obtained at 30°C on LB agar containing 25 µg/ml kanamycin or spectinomycin. pCTS21 is a low copy temperature sensitive plasmid expressing *tnpA* under control of inducible P*lac* promoter. Several single colonies from the transformation plate were inoculated individually in 5 ml LB with 1 mM IPTG and incubated overnight at 30°C, 250 rpm. Overnight cultures were diluted appropriately and spread on LB agar + 10% sucrose or ABO plates and incubated at 37°C. Several sucrose or 5FOA resistant colonies were restreaked at 42°C and then tested for loss of all antibiotic resistances. Colony PCR and sequencing was performed to confirm the excision of foreign DNA fragment and scarless gene modification.

2.8 β-Galactosidase assay

β-Galactosidase activity was measured as originally described by [25] with the following modifications. Cultures were grown overnight in 5 ml LB at 37°C, 250 rpm. Cells were diluted 1:100 in fresh medium and were grown for 3 h at 37°C, 250 rpm. Two-milliliters of culture broth was centrifuged at 10,000 × *g* for 2 min at 4°C and cells were washed once with chilled Z buffer. Cells were suspended in 2 ml Z buffer and OD_{600} was measured. Cells were diluted 1:50 in 1 ml Z buffer

and mixed with 100 µl chloroform and 50 µl 0.1% SDS followed by brief vortexing. After 10 min of incubation at 30°C, 0.2 ml of ONPG (4 mg/ml) was added, and the suspension was vortexed for 10 sec and incubated further at 30°C for 1 h. Reaction was stopped by the addition of 0.5 ml of 1 M Na_2CO_3. After centrifugation, OD_{420} and OD_{550} values were measured using the supernatant. β-Galactosidase activity in Miller units was calculated as $(522 \cdot OD_{420})/(t \cdot v \cdot OD_{600})$, where t is time in minutes, v is the volume of culture used in the assay in milliliters.

3 Results

3.1 Precise excision of LE-antibiotic gene-counterselection marker-RE is mediated by IS608 TnpA

To demonstrate the precise excision, we carried out a preliminary experiment using a plasmid containing the LE-antibiotic-resistance gene-counterselection marker-RE cassette. After selection of cells containing the IS608 cassette based on antibiotic resistance, the LE-antibiotic resistance gene-counterselection marker-RE cassette can be eliminated using a temperature sensitive plasmid expressing the TnpA transposase, which precisely cleaves LE and RE without leaving behind a scar sequence and a counterselection can be used to identify those cells that have been relieved of the cassette. The essentials of the method are shown in figures illustrating the plasmids used (Figure 1) and how they are used in the chromosomal modifications (Figure 2).

The *in vivo* excision assay was carried out using pCT104 or pHMS102 as the target plasmid for the excision of the LE-Cm^R-*sacB*-RE fragment and pBS107 as the *tnpA* expressing plasmid. *E. coli* PIR1 carrying pCT104 or pHMS102 was transformed with pBS107 followed by overnight IPTG induction of the liquid culture at 37°C. Plasmid preparation from the overnight culture was carried out and transformed into the PIR1 strain to select for Ap^R transformants. These transformants were tested for chloramphenicol sensitivity (Cm^S) and sucrose resistance. Out of 47 Ap^R candidates, 6 were Cm^S and sucrose resistant indicating the loss of LE-Cm^R-*sacB*-RE fragment from pCT104 or

pHMS102. Nucleotide sequencing of these 6 Cm^S candidates confirmed the excision of the LE-Cm^R-*sacB*-RE fragment. The results of this experiment are consistent with the *in vivo* excision assay reported by [17] using pBS107. To obtain a host strain free of any plasmid after *in vivo* excision, we constructed a temperature sensitive plasmid pCTS21, as described in Materials and Methods, which carries *tnpA* gene under P*lac* promoter.

3.2 Precise scarless LacZ gene disruptions can be made by the method

To demonstrate precise, scarless gene disruption we targeted the *lacZ* gene using the following strategy. A DNA cassette containing LE-antibiotic resistance gene-counterselection marker-RE was PCR amplified by using primers containing 50-75 bp extensions homologous to the target region of the chromosome. The PCR product was then introduced into the host strain harboring a temperature sensitive plasmid expressing λ Red recombinase, which facilitates the incorporation of the cassette into the chromosome, disrupting the target gene. Similar to above, cells are selected for the excision cassette followed by elimination of the cassette using the TnpA transposase and counterselection to identify cells that no longer have the cassette. A schematic diagram depicting *lacZ* disruption in MG1655 is shown in Fig. 2. We made two *lacZ* gene disruptions by precisely deleting either a 950 bp or a 2957 bp region of the *lacZ* gene. Primers lacZ-F1 and lacZ-R1 were used to amplify a 2.5 Kb LE-Cm^R-*sacB*-RE fragment by PCR from pHMS102 for knocking out a 950 bp region of *lacZ* of CT2572 *lacZ*$^+$ and MG1655. To knockout and remove 2957 bp of *lacZ* of CT2572 *lacZ*$^+$, primers lacZ-F1 and lacZ-R2 were used to amplify a 3 Kb LE-Km^R-*oroP*-RE of pKJL12 by PCR. Blue-white screening was used to identify Cm^R or Km^R transformants with a nonfunctional *lacZ* gene. Precise excision and analysis was performed as above. Both *lacZ* gene disruptions were verified by sequencing.

3.3 A lacZ point mutation can be repaired using the method

We used this developed technique to repair point mutation in *E. coli* AB2572 (CGSC 2572) strain, which has LacZ13 (Oc) mutation (CAA → TAA) at position 70 in the *lacZ* gene [26]. LacZ-LEH and

LacZ13-RE primers were used to amplify a 2.5 Kb fragment LE-Cm^R-*sacB*-RE of pHMS102 for *lac*Z point mutation repair. Primers LacZ-LEH and LacZ13-RE contain 50 bp of *lac*Z flanking sequence to repair the point mutation by homologous recombination. The tetranucleotide sequence TTAC, located at position 65 to 68 bp within the *lac*Z gene, was used as an insertion site to introduce the 2.6 Kb PCR product within the *lacZ* gene. Excision of the 2.5 Kb LE-Cm^R-*sacB*-RE cassette from *lacZ* gene of the modified AB2572 was carried out by pCTS21. Selection of the *lacZ* positive derivative of AB2572 was performed by blue-white screening. The resulting strain with functional *lacZ* gene was referred as CT2572 *lacZ*$^+$. PCR confirmation and nucleotide sequencing using LacZ-up-F and LacZ-bet-R primers revealed that the *lacZ* point mutation was repaired (TAA → CAA) without leaving behind a scar sequence.

3.4 Construction of lacZ ribosome binding site (RBS) library allows a set of modified sequences to be introduced and give a spectrum of cells with altered lacZ expression levels

To further demonstrate the wide applicability of this technique, we constructed a *lacZ* library with variable β-galactosidase activity by precisely modifying the ribosome binding site and spacer region in CT2572 *lacZ*$^+$ and KL334 (CGSC-4345) strains. Strain KL334 has a nonfunctional *lacZ* gene as a result of an ochre mutation UAA [lacZ118(Oc)] at position 52 of the *lacZ* gene. Primers Lac-RBS-F and Lac-RBS-R, containing 80 bp flanking sequence for homologous recombination, were used to construct a *lacZ* library. TCAC sequence is located 13 bp upstream of the *lacZ* start codon in the *E. coli* chromosome. In order to create the tetranucleotide sequence TTAC for precise excision of the PCR cassette fragment by *tnpA,* we introduced a single base substitution T<u>C</u>AC → T<u>T</u>AC (corresponds to position 78 of the flanking region in forward primer Lac-RBS-F). Reverse primer Lac-RBS-R contains 4 mixed bases (Table 2) in the flanking sequence to introduce mutations in RBS and in the spacer region between RBS and the start codon. Using these primers and pHMS102 as template, 2.5 Kb PCR products were obtained. Cm^R transformants were verified by colony PCR using LacZ-up-F and LacZ-bet-R primers and further by nucleotide sequencing of purified PCR

products for the random mutations. Upon excision of the cassette using the pCTS21 plasmid, blue white screening was used to select blue mutant candidates with a functional *lacZ* gene. Blue colonies were further verified by nucleotide sequencing for the introduced mutations and scarless excision of the 2.5 Kb PCR fragment. β-galactosidase activity of *lacZ* variants was measured. Table 3 shows the RBS and spacer region sequence of wild type *lacZ* and 12 variants obtained in this study. β-galactosidase activity was measured in Miller units and expressed as percent normalized to wild type strain. All 12 *lacZ* variants resulted in decreased β-galactosidase activity as compared to the wild type strain.

4 Discussion

4.1 Precision of excision is demonstrated which allows scarless sequence modifications to be made in the chromosome

In this study, we have described a new method for precise gene modification in *E. coli* that leaves no foreign DNA within the host genome. Using this method, we have demonstrated point mutation repair in the *lacZ* gene, two disruptions of the *lacZ* gene, and constructed a library of 12 *lacZ* variants by modifying the ribosome binding site and spacer region. This method has the great advantage of not leaving behind a scar sequence in the host genome which may be problematic under certain conditions, in particular situations where multiple genomic manipulations are necessary.

Datsenko and Wanner [8] developed a procedure which has been used to construct 3,985 single gene-knockout *E. coli* mutants [27] and has also been applied to other bacteria such as *Salmonella* and *Yersinia* [28, 29]. However, this method introduces Flippase Recognition Target (FRT) sites into host DNA upon recombination, and a scar sequence containing an FRT site (82- to 85-nt in size) is left following elimination of the resistance gene using FLP (Flippase recombination enzyme). As the Red system requires a short homologous region for gene disruption, a new PCR

fragment can recombine at the scar region of previously performed gene disruption in the chromosome. Moreover, the scar sequence could result in chromosomal rearrangement from FLP-promoted recombination between FRT sites at different loci in the host chromosome where multiple changes are made[8].

In addition, the FLP-FRT system is analogous to the Cre-*lox* [30] recombination system in which Cre recombinase catalyzes cofactor-independent recombination between two *loxP* recognition sites. This system leaves one copy of the *loxP* site in the genome after excision which can be recognized by Cre upon consecutive rounds of genome manipulation and would also result in intra-chromosomal rearrangements. To minimize genetic instability, *lox* sites (*lox66* and *lox71*) containing mutations within the inverted repeats were also developed [31].

We have overcome this problem by developing a new technique using ISHp*608* from *Helicobacter pylori* [16] and taking advantage of the special properties of the TnpA transposase [17]. This particular transposase family does not possess terminal internal repeats, but does have left (LE) and right ends (RE) that contain palindromic repeats that form secondary structure hairpin loops recognized by TnpA. Insertion and/or excision of IS*608* are mediated by TnpA precisely at 3' of a tetranucleotide sequence TTAC without a gain or loss in DNA sequences at either end [17]. Since a TTAC sequence is needed somewhere near the area to be modified the naturally occurring sequence would likely have this sequence within 100bp or so in one of the strands, and if not the TTAC sequence could be introduced during the recombination by using a primer sequence that alters a similar sequence in the native chromosome to a TTAC in one of the strands. We constructed special template plasmids pCT104, pHMS102 and PKJL12 that contain the left and right ends of ISHp*608* together with antibiotic resistance and counterselection markers and have a conditional replication origin *oriRγ* to reduce false positive transformants from carryover in the PCR step. Template plasmids pCT104 and pHMS102 were constructed with Cm^R and *sacB* gene of *Bacillus subtilis*, which encodes a levansucrase that transfers fructosyl residues from sucrose to

different cellular acceptors. Expression of *sacB* is toxic in a wide range of Gram negative bacteria [2, 32]. This marker has been widely used in gene replacement techniques [9, 33-35]. However, while making *lacZ* disruptions we found that this marker is not as effective as a counterselection tool as desired since several PCR verified and Cm^R colonies turned out to be sucrose resistant. This may be due to spontaneous point mutations arising in *sacB* [4]. The use of a counterselection or selectable phenotype during the process is useful since the frequency of excision by the ISH608 transposase is less efficient than the FLP or Cre recombinases such as are used in the Datsenko and Wanner method. The frequency of excision from the chromosome has been examined in detail [36] and depends on orientation of the transposon and if the strand recognized by the transposase is on the lagging strand, excision occurs at a higher frequency. As the size of the sequence bounded by the LE and RE of ISH608 becomes longer the frequency of excision is lower. Depending on these factors excision is low enough, 10^{-3} or so in some cases, to make a counter selection marker a needed component of the system. The frequency could be potentially improved by engineering higher expression of the transposase, using mutant hosts or modifying culture conditions.

4.2 Counterselections for excision

In order to have an alternative counterselection marker that could be used in wild type strains, we examined the *oroP* gene of *L. lactis* as a counterselection tool to construct template plasmid pKJL12. The *oroP* gene, which encodes a dedicated orotate transporter, has been used as selection/counterselection tool for *L. lactis* strains, *E. coli and B. subtilis* [21, 37] suggesting that it can be utilized as selection/counter marker in wide range of bacteria [38]. Cells harboring the *oroP* gene can be selected either by their ability to utilize orotate as a pyrimidine source or their sensitivity to 5-fluoroorotate. However, in order to use *oroP* as counterselection marker the background strain must be unable to transport 5-fluoroorotate. In *E. coli*, *Salmonella typhimurium* and *Sinorhizobium meliloti* the C_4-dicarboxylate transport protein DctA has been identified to mediate uptake of orotate [39, 40]. This limitation was overcome by supplementing 0.3% succinate

and 0.3% fumarate in AB medium. As a result, the DctA transport protein would mostly remain active in transporting the supplemented C_4-dicarboxylic acids and does not interfere with *oroP* function.

Both, *sacB* and *oroP* counterselection markers were proved useful in this study for screening of desired transformants. There are a variety of other counterselection markers that could be introduced into the cassette and used depending on the host being manipulated. Such alternative counterselection markers could be placed in the cassette instead of those we demonstrated. In our experiments we successfully used *sacB* and *oroP* markers in generating two *lacZ* gene disruptions (950 bp and 2957 bp) and *lacZ* point mutation repair. Furthermore, we obtained 12 variants of *lacZ* gene with variable expression by modifying RBS and/or the spacer region before the start codon (Table 3). These modifications were introduced using nonsense primer containing mixed bases. The modifications were based on the study of [41] describing 16 single point mutations near the beginning of the *lacZ* gene and their effect on *lacZ* expression. However, due to limited screening, we obtained five variants 8D, 27A, 27B, 22A and 22B with single point mutations either in the spacer region or RBS. The rest of the variants had two or more mutations resulting in decreased *lacZ* expression (Table 3). It is important to mention here that *lacZ* variants 1A, 8D, 207A and 210B were obtained from parent strain KL334 which has a nonfunctional *lacZ* gene due to the *lacZ*118 (Oc) mutation. This suggests the potential of this method to simultaneously repair the existing point mutation and introduce base substitutions in the RBS and or spacer region of *lacZ* gene. By undertaking extensive screening, a large library of variants having variable gene expression can be constructed for any gene of interest, which could provide significant insight into gene function and precise balancing of chromosomal genes expressing optimal protein and enzyme levels for metabolic engineering purposes. Similarly, by utilizing this method in combination with the Ribosome Binding Site (RBS) calculator [42],

synthetic RBS sequence variants of bacteria can be achieved for targeted translation and its application in metabolic engineering studies.

4.3 Potential and relation to other new methods

Recently the use of CRISPRi for metabolic engineering in bacterial cells has been reported [43] and this technique offers a new and widely applicable means of regulating gene expression of multiple targets using guide RNAs. The CRISPRi method is different in that it acts by controlling transcription by means of dCas9 localization on the chromosome to affect RNA polymerase action in the article above. The use of this CRISPRi method is just being introduced into engineering of bacterial cells and will play a large role in the future as advanced techniques are developed. The method reported here builds on the widely used lambda rec system and allows specific mutations to be made in the chromosome so that long term gene modification or expression occurs without the continued need for added proteins to be made in the cell for regulation.

In conclusion, this method has advantages over existing methods for scarless genome engineering in *E. coli.* In this study, we have demonstrated the applicability of this method for repairing a point mutation, precisely disrupting a chromosomal gene and constructing a gene library to study expression level in *E. coli*. This method could also be used to introduce a point mutation, to delete a large genome region and to integrate foreign genes into host chromosome and may be explored for the genome engineering of other organisms.

Acknowledgement

This work and publication was supported by grants from the National Science Foundation CBET-0828516 and DBI-1262296. We thank Mary Harrison for technical support and helpful discussions.

Conflict of interest

The authors declare no financial or commercial conflict of interest.

5 References

[1] Yang, J., Sun, B., Huang, H., Jiang, Y.*, et al.*, High-Efficiency Scarless Genetic Modification in Escherichia coli by Using Lambda Red Recombination and I-SceI Cleavage. *Appl Environ Microbiol* 2014, *80*, 3826-3834.

[2] Zhou, Q. M., Dong Jie Fan, Jiang Bi Xie, Chuan Peng Liu, Jun Mei Zhou, A method for generating precise gene deletions and insertions in *Escherichia coli*. *World J Microbiol Biotechnol* 2010, *26*, 1323-1329.

[3] Sharan, S. K., Thomason, L.C., Kuznetsov, S.G., Court, D.L., Recombineering: a homologous recombination-based method of genetic engineering. *Nature Protocol* 2009, *4*, 206-223.

[4] Heermann, R., Zeppenfeld, T., Jung, K., Simple generation of site-directed point mutations in the Escherichia coli chromosome using Red(R)/ET(R) Recombination. *Microb Cell Fact* 2008, *7*, 14.

[5] Datta, S., Costantino, N., Zhou, X., Court, D.L., Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc Natl Acad Sci U S A* 2008, *105*, 1626-1631.

[6] Murphy, K. C., Campellone, K. G., Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic E. coli. *BMC Mol Biol* 2003, *4*, 11.

[7] Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A.*, et al.*, An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* 2000, *97*, 5978-5983.

[8] Datsenko, K. A., Wanner, B. L., One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000, *97*, 6640-6645.

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

[10] Park, Y. N., Masison, D., Eisenberg, E., Greene, L. E., Application of the FLP/FRT system for conditional gene deletion in yeast *Saccharomyces cerevisiae*. *Yeast* 2011, *28*, 673-681.

[11] Yu, H., Yan, X., Shen, W., Shen, Y.*, et al.*, Efficient and precise construction of markerless manipulations in the *Bacillus subtilis* genome. *J Microbiol Biotechnol* 2010, *20*, 45-53.

[12] Liang, R., Liu, J., Scarless and sequential gene modification in Pseudomonas using PCR product flanked by short homology regions. *BMC Microbiol* 2010, *10*, 209.

[13] Yamamoto, S., Izumiya, H., Morita, M., Arakawa, E., Watanabe, H., Application of lambda Red recombination system to Vibrio cholerae genetics: simple methods for inactivation and modification of chromosomal genes. *Gene* 2009, *438*, 57-64.

[14] Gust, B., Chandra, G., Jakimowicz, D., Yuqing, T.*, et al.*, Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. *Adv Appl Microbiol* 2004, *54*, 107-128.

[15] Curcio, M. J., Derbyshire, K. M., The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol* 2003, *4*, 865-877.

[16] Kersulyte, D., Velapatino, B., Dailide, G., Mukhopadhyay, A. K.*, et al.*, Transposable element ISHp608 of Helicobacter pylori: nonrandom geographic distribution, functional organization, and insertion specificity. *J Bacteriol* 2002, *184*, 992-1002.

[17] Ton-Hoang, B., Guynet, C., Ronning, D. R., Cointin-Marty, B.*, et al.*, Transposition of ISHp608, member of an unusual family of bacterial insertion sequences. *Embo J* 2005, *24*, 3325-3338.

[18] Guynet, C., Achard, A., Hoang, B. T., Barabas, O.*, et al.*, Resetting the site: redirecting integration of an insertion sequence in a predictable way. *Mol Cell* 2009, *34*, 612-619.

[19] Guynet, C., Hickman, A. B., Barabas, O., Dyda, F.*, et al.*, In vitro reconstitution of a single-stranded transposition mechanism of IS608. *Mol Cell* 2008, *29*, 302-312.

[20] Clark, D., Maaloe O., DNA replication and the division cycle in *Escherichia coli*. *J Mol Biol* 1967, *23*, 99-112.

[21] Solem, C., Defoor, E., Jensen, P. R., Martinussen, J., Plasmid pCS1966, a new selection/counterselection tool for lactic acid bacterium strain construction based on the oroP gene, encoding an orotate transporter from Lactococcus lactis. *Appl Environ Microbiol* 2008, *74*, 4772-4775.

[22] Goulian, M., van der Woude, M., A simple system for converting lacZ to gfp reporter fusions in diverse bacteria. *Gene* 2006, *372*, 219-226.

[23] Martinez-Morales, F., Borges, A. C., Martinez, A., Shanmugam, K. T., Ingram, L. O., Chromosomal integration of heterologous DNA in Escherichia coli with precise removal of markers and replicons used during construction. *J Bacteriol* 1999, *181*, 7143-7148.

[24] Haldimann, A., Wanner, B. L., Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J Bacteriol* 2001, *183*, 6384-6393.

[25] Miller, J. H., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 1972.

[26] Schwartz, N. M., Supression of a lacO-O mutation in Escherichia coli. *J Bacteriol* 1964, *88*, 996-1001.

[27] Baba, T., Ara, T., Hasegawa, M., Takai, Y.*, et al.*, Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2006, *2*, 2006 0008.

[28] Sun, W., Wang, S., Curtiss, R., 3rd, Highly efficient method for introducing successive multiple scarless gene deletions and markerless gene insertions into the Yersinia pestis chromosome. *Appl Environ Microbiol* 2008, *74*, 4241-4245.

[29] Uzzau, S., Figueroa-Bossi, N., Rubino, S., Bossi, L., Epitope tagging of chromosomal genes in Salmonella. *Proc Natl Acad Sci U S A* 2001, *98*, 15264-15269.

[30] Abremski, K., Hoess, R., Sternberg, N., Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell* 1983, *32*, 1301-1311.

[31] Albert, H., Dale, E.C., Lee, E., Ow, D.W., Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. *The Plant Journal* 1995, *7*, 649-659.

[32] Quandt, J., Hynes, M. F., Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* 1993, *127*, 15-21.

[33] Blomfield, I. C., Vaughn, V., Rest, R. F., Eisenstein, B. I., Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature-sensitive pSC101 replicon. *Mol Microbiol* 1991, *5*, 1447-1457.

[34] Lalioti, M., Heath, J., A new method for generating point mutations in bacterial artificial chromosomes by homologous recombination in Escherichia coli. *Nucleic Acids Res* 2001, *29*, E14.

[35] Oswald, W., Tonpitak, W., Ohrt, G., Gerlach, G., A single-step transconjugation system for the introduction of unmarked deletions into Actinobacillus pleuropneumoniae serotype 7 using a sucrose sensitivity marker. *FEMS Microbiol Lett* 1999, *179*, 153-160.

[36] Ton-Hoang, B., Pasternak, C., Siguier, P., Guynet, C.*, et al.*, Single-stranded DNA transposition is coupled to host replication. *Cell* 2010, *142*, 398-408.

[37] Defoor, E., Kryger, M. B., Martinussen, J., The orotate transporter encoded by oroP from Lactococcus lactis is required for orotate utilization and has utility as a food-grade selectable marker. *Microbiology* 2007, *153*, 3645-3659.

[38] Martinussen, J., Defoor, Novel recombinant marker gene encoding orotate transporter polypeptide, useful as selection marker, screening marker, counter-selection marker or bi-directional marker for identifying and/or selecting cells. Patent WO2005078106-A1, 2005.

[39] Baker, K. E., Ditullio, K. P., Neuhard, J., Kelln, R. A., Utilization of orotate as a pyrimidine source by Salmonella typhimurium and Escherichia coli requires the dicarboxylate transport protein encoded by dctA. *J Bacteriol* 1996, *178*, 7099-7105.

[40] Yurgel, S. N., Kahn, M. L., Sinorhizobium meliloti dctA mutants with partial ability to transport dicarboxylic acids. *J Bacteriol* 2005, *187*, 1161-1172.

[41] Munson, L. M., Stormo, G. D., Niece, R. L., Reznikoff, W. S., lacZ translation initiation mutations. *J Mol Biol* 1984, *177*, 663-683.

[42] Salis, H. M., Mirsky, E. A., Voigt, C. A., Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotech* 2009, *27*, 946-950.

[43] Lv, L., Ren, Y. L., Chen, J. C., Wu, Q., Chen, G. Q., Application of CRISPRi for prokaryotic metabolic engineering involving multiple genes, a case study: Controllable P(3HB-co-4HB) biosynthesis. *Metab Eng* 2015, *29*, 160-168.

Table 1. List of bacterial strains and plasmids used in this study

Strains or plasmids	**Description**	**Reference or source**

Strains		
MG1655	Wild type (*F- λ- rph-1*)	CGSC, Yale Univ., CT (CGSC 7740)
PIR 1	F-Δ*lac169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA(*Δ*MluI)::pir-116*	Invitrogen, CA
AB2572	Hfr(PO1), *lacZ13*(Oc), *relA1*, *thiE1*	CGSC, Yale Univ., CT (CGSC 2572)
KL334	Hfr(PO2A), *lacZ118*(Oc), *lacI22*, *λ-*, *lysA23*	CGSC, Yale Univ., CT (CGSC 4345)
ABLE-C	*E. coli* C *lac*(LacZ⁻) [Kan^R McrA⁻ McrCB⁻ McrF⁻ Mrr⁻ HsdR (r_k^- m_k^-)] [F′ *proAB lacI*^{q}*Z*ΔM15 Tn*10*(Tet^R)]	[38]
CT2572 *lacZ*$^+$	Hfr(PO1), *relA1*, *thiE1*	This study
Plasmids		
pMARMAR4	*oriR6K*, *sacB*, Cm^R	[22]
pAH68	*oriR*, tL3, *rgnB*, HK, Ap^R	[24]
pLOI2227	pSC101*ori ts*, T7, SP6, FRT, Kan^R	[23]
pKM208	pSC101*ori ts*, P*tac*, *red-gam*, *lacI*^{q}, Ap^R	[6]
pBS107	p15A*ori*, P*lac*, *tnpA*, *laqI*^{q}, Kan^R, Spect^R	[17]
pCS1966	ColE1*ori*, P_{syn}, *oroP*, Ery^R	[38]
pACYC177	p15A*ori*, Kan^R, Ap^R	New England Biolabs, MA, USA
pCTS21	pSC101*ori*, Rep^{ts}, T7, SP6, FRT, P*lac*, *tnpA*, *laqI*^{q}, Kan^R, Str^R/Spect^R	This study

pCT104	*oriR*, tL3, *rgnB*, HK, LE, *sacB*, RE, Cm^R, Ap^R	This study
pHMS102	*oriR*, tL3, *rgnB*, HK, LE, *sacB*, RE, Cm^R, Ap^R	This study
pCT200	p15A*ori*, P_{syn}, *oroP*, Kan^R	This study
pKJL12	*oriR*, tL3, *rgnB*, HK, LE, P_{syn}, *oroP*, RE, Kan^R, Ap^R	This study

Table 2. Oligonucleotides used in this study

Primer	Sequence[#] (5' to 3')
LacZ-up-F	caatacgcaaaccgcctctcc
LacZ-bet-R	ccgtgcatctgccagtttgag
LE-PstI	catagc*ctgcag*tattacCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATA CAAGGCGAAACGCCTTTAGGtctgtaagcggatgccgggagcagac
RE-BamHI	ctactc*ggatcc*tttgTTGACATACTCCCCATAGCTAAAGCTAGGGGATTCGGGCAAAgtgta tactttggcgtcacccctac
pAH68F	gctctcctgagtaggacaaatc
pAH68R	cagcctcgctttgtaacgg
Lac-RBS-F	**ctcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataaca atttt<u>t</u>ac**caaaactaacgccttaaagccc
Lac-RBS-R	**taacgccagggtttt<u>c</u>ccagtcacgacgttgtaaaacgacggccagtgaatccgtaatcatggtcatagcy<u>r</u> ty<u>t</u><u>m</u>ctgt**ttgacatactccccatagctaaagc
LacZ-LEH	**tcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttac**caaaactaacgccttaaagc
LacZ13-RE	**tattacgccagctggcgaaaggggga<u>t</u>gtgctgcaaggcgattaagttg<u>g</u>g**ttgacatactccccatagcta
LacZ-F1	**tcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttac**caaaactaacgccttaaagc
LacZ-R1	**gttaacgcctcgaatcagcaacggcttgccgttcagcagcagcagaccat**ttgacatactccccatagcta
LacZ-R2	**ttatttttgacaccagaccaactggtaatggtagcgaccggcgctcagct**ttgacatactccccatagcta
oroP-BamF	ctatca*ggatcc*ggaaggatcccccatagtt
oroP-PstR	gtacta*ctgcag*atacgactcactatagggc
pAPT-plac-F	ctatct*gaattc*ctttacactttatgcttccggc
pAPT-laqIq-R	cactac*gaattc*tcactgcccgctttcca
KpnI-LE-KanF	catagt*ggtacc*ttacCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATAC AAGGCGAAACGCCTTTAGGgaacttttgctttgccacggaacg

EcoRI-RE-KanR	ctactc*gaattc*TTGACATACTCCCCATAGCTAAAGCTAGGGGATTCGGGCAAAcagtgctg caatgataccgcgag
LE-NdeI	cggc*catatg*tattacCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGAtctgt aagcggatgccgggagcagac
RE-NdeI	gcgc*catatg*tttgTTGACATACTCCCCATAGCTAAAGCTAGGGGATTCGGGCAAAgtgtata ctttggcgtcacccttac

[#] The nucleotides that are homologous to regions adjacent to the *lacZ* gene are bold. The sequence of restriction endonuclease sites are in italics. The sequence of LE and RE are in capitals. The sequence marked in bold and underlined indicates introduced changes. Key to symbols: Y=C+T, R=G+A, M=C+A

Table 3. *lacZ* variants and their β-galactosidase activity

***lacZ* variant**	**Ribosome binding site (RBS)***	**Spacer region***	**β-galactosidase activity (% normalized to wild type)**
Wild type	A G G A	A C A G C T	100
207A	A G T A	A T A G C T	24.3
1A	A G T A	A T A G C T	19.9
22A, 22B	A G T A	A C A G C T	19.7
15A, 15B	A G T A	A T G G C T	4.2
8D, 27A, 27B	A G G A	A C G G C T	2.1
210B	A G T A	A T G G C T	1.9
32A, 32B	A G T A	A T G G C T	0.9

*The nucleotides that are underlined indicates the introduced changes in RBS and spacer region

Figure legends

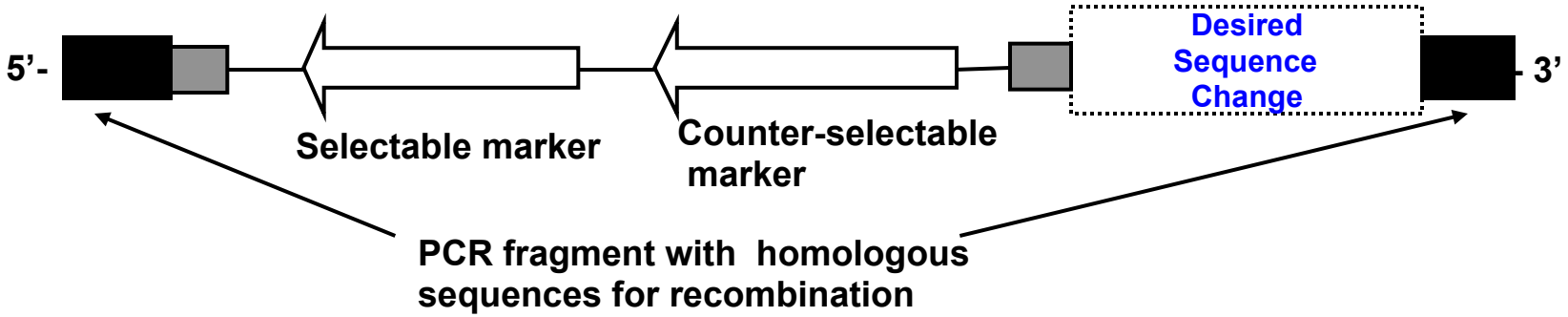
Figure 1. Diagram of plasmids containing selectable and excisable IS608 cassettes.

Panel A shows the linear view of plasmid pCT104 that contains the left (LE) and right (RE) ends of the IS608 transposon, the chloramphenicol resistance element Cm^R, the sacB locus for counter selection on sucrose, tL3 is the bacteriophage lambda transcription terminator, the HK phage attachment site from pAH68, oriRγ is the conditional replication origin of pAH68 derived from R6K, which requires *pir* for replication, ampicillin resistance Ap^R, and rgnB is a bacterial transcription terminator from pAH68. pHMS102 is similar to pCT104 but was cloned into pAH68 using different restriction enzymes. pKJL12 bears a kanamycin resistance element KmR and an oroP counterselection marker. The construction of the plasmids was described in the plasmids section of the methods. Panel B shows the circular map of the plasmid pCTS21 that contains a recogniton sequence (FRT) for FLP mediated excision of the cassette. This cassette contains the tnpA transponase gene from IS608 expressed under control of the lac promoter Plac, and the spectinomycin resistance element $Spect^R$ as well as the lac repressor gene with the more active promoter lacIq. The vector the cassette resides on is a low copy number vctor pSC101*ori*, with a temperature sensitive origin of replication $RepA101^{ts}$, and the plasmid also contains the promoters for T7 and SP6 RNA polymerase and is derived from pLOI2227, see methods.

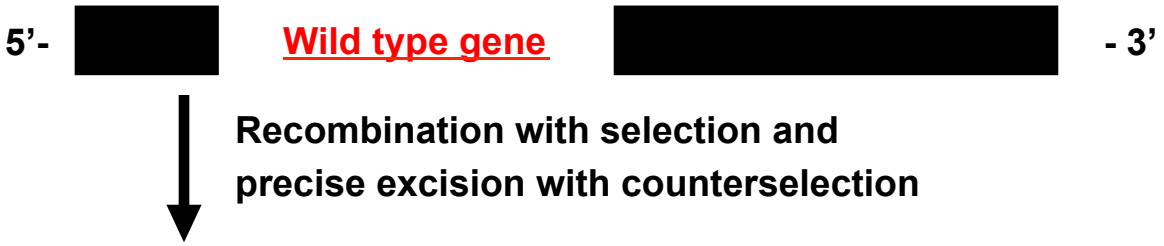
Figure 2. Diagram showing the genetic and DNA manipulations for correcting a point mutation and introducing a library of ribosome binding site mutations.

Panel A shows the schematic of the PCR fragment with the corrected point mutation (red colored C) in the downstream sequence that will recombine with the chromosome (black blocks). The position of the IS608 specific ends, TTAC at the left end (LE) and TCAA at the right (RE) are shown. The selectable markers Cm^R and *sacB* are shown within the cassette.

The corresponding sequence of the lac region of *E. coli* AB2572 bearing the point mutation red colored T to be corrected. The recombined chromosomal segment is shown on the third line and the corrected *lacZ* sequence is shown on the fourth line illustrating a corrected functional *lacZ*. In panel B the top line shows the cassette bounded by the altered sequence incorporating several sequence alterations in the region of the ribosome binding sequence just upstream from the ATG of the translation initiation codon of *lacZ*. Recombination into E. coli KL334 can produce a variety of sequence changes in the ribosome binding sequence as shown in the second line while at the same time correcting the early mutation in lacZ. The third line shows the recombined cassette fragment in the chromosome (insert) and the fourth line shows the chromosome after excision of the cassette by tnpA with the variety of modifications in the ribosome binding site and the functional *lacZ*.

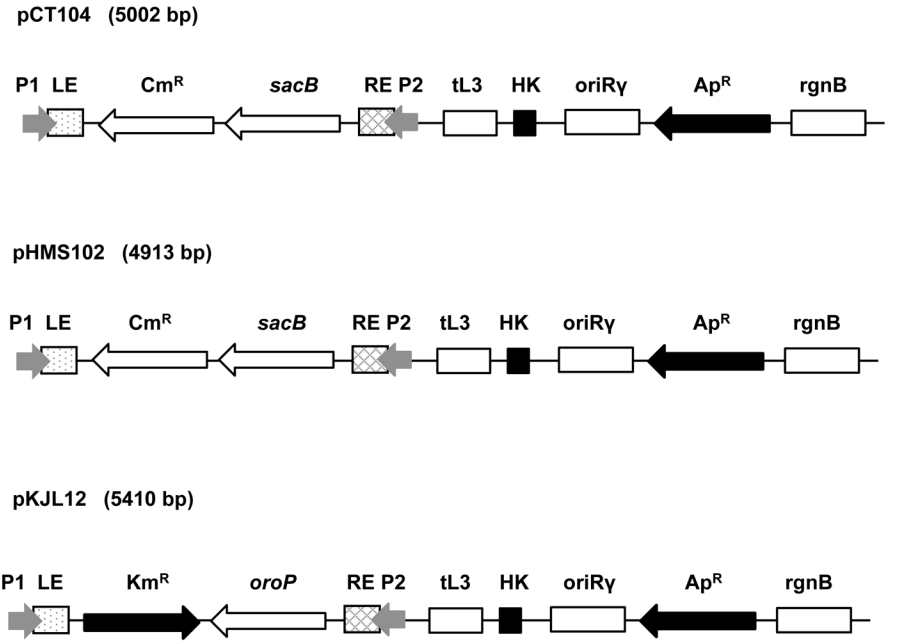


Host chromosome

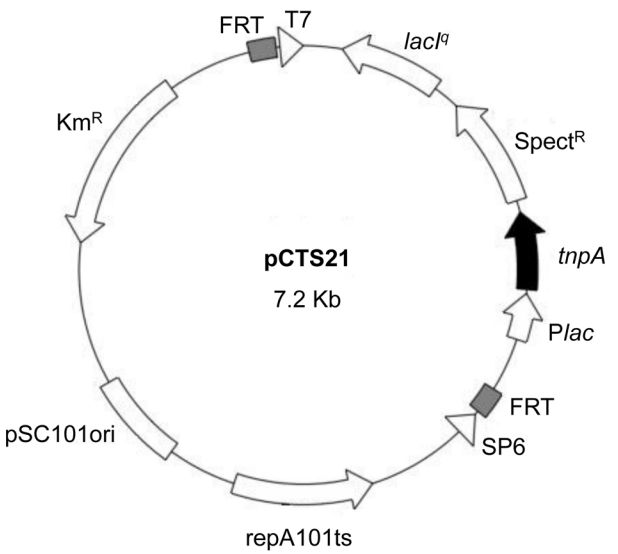


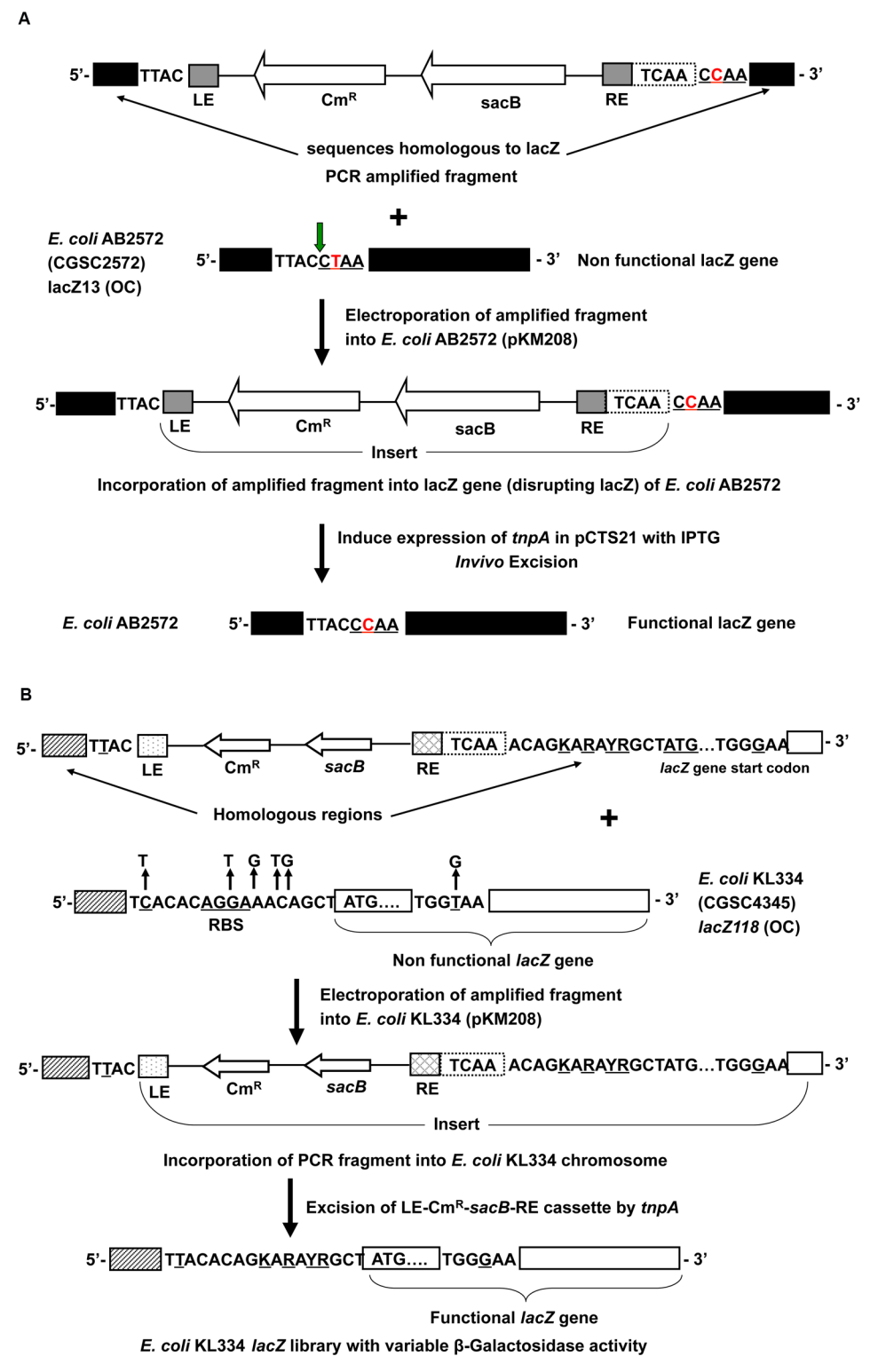
Precise modification 5'-

A


pCT104 (5002 bp)
P1 LE
Cm^R
sacB
RE P2
tL3
HK
oriRγ
Ap^R
rgnB
pHMS102 (4913 bp)
P1 LE
Cm^R
sacB
RE P2
tL3
HK
oriRγ
Ap^R
rgnB
pKJL12 (5410 bp)
P1 LE
Km^R
oroP
RE P2
tL3
HK
oriRγ
Ap^R
rgnB

B


FRT
T7
lacI^q
Km^R
Spect^R
pCTS21
7.2 Kb
tnpA
Plac
FRT
SP6
pSC101ori
repA101ts


A
5'-
TTAC
LE
Cm^R
sacB
RE
TCAA
CCAA
- 3'
sequences homologous to lacZ
PCR amplified fragment
+
E. coli AB2572
(CGSC2572)
lacZ13 (OC)
5'-
TTACCTAA
- 3'
Non functional lacZ gene
Electroporation of amplified fragment
into E. coli AB2572 (pKM208)
Insert
Incorporation of amplified fragment into lacZ gene (disrupting lacZ) of E. coli AB2572
Induce expression of tnpA in pCTS21 with IPTG
Invivo Excision
E. coli AB2572
TTACCCAA
Functional lacZ gene
B
TTAC
sacB
ACAGKARAYRGCTATG…TGGGAA
lacZ gene start codon
Homologous regions
T
T G TG
G
TCACACAGGAAACAGCT
ATG....
TGGTAA
RBS
E. coli KL334
(CGSC4345)
lacZ118 (OC)
Non functional lacZ gene
Electroporation of amplified fragment
into E. coli KL334 (pKM208)
Incorporation of PCR fragment into E. coli KL334 chromosome
Excision of LE-Cm^R-sacB-RE cassette by tnpA
TTACACAGKARAYRGCT
Functional lacZ gene
E. coli KL334 lacZ library with variable β-Galactosidase activity