



- (51) **International Patent Classification:** Not classified
- (21) **International Application Number:**
PCT/US2012/069672
- (22) **International Filing Date:**
14 December 2012 (14.12.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/576,588 16 December 2011 (16.12.2011) US
61/700,650 13 September 2012 (13.09.2012) US
- (71) **Applicant:** UNIVERSITY OF DELAWARE [US/US]; 1 Innovation Way, Delaware Technolgy Park, Building 1, Suite 500, Newark, DE 19711 (US).
- (72) **Inventors:** PAPOUTSAKIS, Eleftherios; 7 Bowman Way, Newark, DE 19711 (US). AL-HINAL, Mohab, Ali; 814 Wharton Drive, Newark, DE 19711 (US). FAST, Alan, G.; 13 Sandalwood Drivem Apt. 9, Newark, DE 19713 (US).
- (74) **Agents:** PRESTIA, Paul, F. et al.; Ratnerprestia, P.O. Box 980, Valley Forge, PA 19482 (US).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

[Continued on next page]

(54) **Title:** RECOMBINANT CLOSTRIDIUM ORGANISM AND METHOD FOR ISOLATION OF DOUBLE-CROSSOVER ALLELIC EXCHANGE MUTANTS

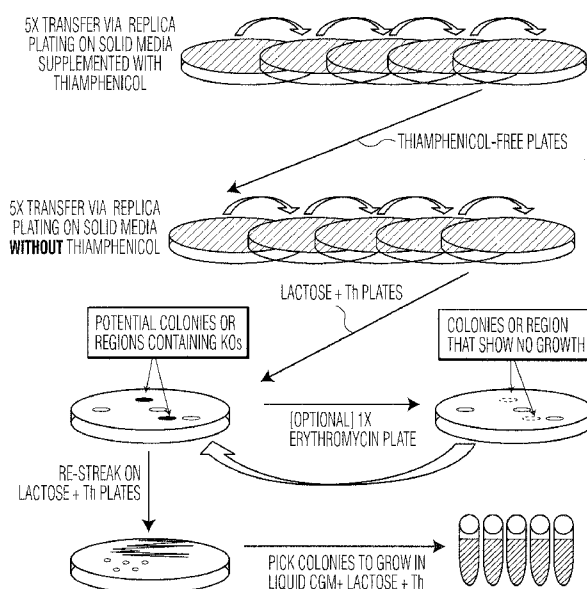


FIG. 3

(57) **Abstract:** The present invention provides recombinant *Clostridium* organisms having of double-crossover allelic exchange based on counter-selection and methods for isolation thereof.



— with sequence listing part of description (Rule 5.2(a))

~ 1 ~

**RECOMBINANT CLOSTRIDIUM ORGANISM AND METHOD FOR ISOLATION OF
DOUBLE-CROSSOVER ALLELIC EXCHANGE MUTANTS**

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 61/576,588,
5 filed December 16, 2011, and U.S. Provisional Application No. 61/700,650, filed
September 13, 2012, the contents of each of which are incorporated herein in their
entireties for all purposes.

GOVERNMENT INTERESTS

This invention was made with government support under Grant No. DE-SC0007092
10 awarded by the United States Department of Energy. The government has certain rights
in this invention.

FIELD OF THE INVENTION

The invention pertains to the field of biotechnology. In particular, the invention
provides recombinant *Clostridium* organisms having double-crossover allelic exchange and
15 methods for isolation thereof.

BACKGROUND OF THE INVENTION

The genus *Clostridium* is composed of Gram⁺, spore-forming bacteria that belong
to the phylum Firmicutes. Members of this genus are obligate anaerobes and contain both
medically and commercially important species. *C. botulinum*, *C. difficile*, *C. perfringens*
20 and *C. tetanii* are medically-important human pathogens. *C. sporogenes* and *C. novyi*
have the potential for use as anti-cancer drug delivery systems that specifically target
hypoxic and necrotic areas of tumors. Other species of the genus are industrially
important for producing commodity chemicals and biofuels from a variety of carbon
sources. These include strains that produce solvents, such as *C. acetobutylicum*, *C.*
25 *pasteurianum* and *C. beijerinckii*, the acetogens that grow on CO₂/CO/H₂, such as *C.*
ljungdahlii and *C. carboxidivorans*, and the cellulolytic clostridia, such as *C. thermocellum*,
C. cellulolyticum, and *C. phytofermentans* among others. Industrial and medical
applications of these organisms are hindered by the limited availability of effective genetic
tools, and the inherent difficulty of genetically modifying these organisms has been a
30 roadblock in the development of desirable strains. Gene deletions (knockouts, KOs) and
chromosomal integrations (knockins, KIs) for both fundamental investigations and
practical applications remain a difficult, slow and inefficient task.

Until recently, most of the relatively few homologous recombination mutants
isolated in *Clostridium* spp. have been single-crossover mutants brought about by a
35 Campbell-like integration of the entire plasmid backbone. To date, only a handful of gene
replacement mutants via allelic exchange have been isolated.

~ 2 ~

To bypass the difficulty of isolating allelic exchange mutants in *Clostridium*, three main methods are currently employed for gene inactivation and DNA chromosomal integration. The first is based on the mobile group II intron from the *ltrB* gene of *Lactococcus lactis* (commercially known as the TargeTron technology). These are large
 5 'retrohoming' ribozymes capable of self-splicing from pre-mRNA and can be designed to target a specific locus in the host genome thus disrupting the gene of interest. This method can be effective but is not without limitations. One possible drawback is that target sequences smaller than 400 bp may not contain optimal target sites for the mobile group II intron recognition sequence as discussed in the TargeTron gene knockout system
 10 user guide (Sigma Aldrich). Moreover, since the mutants generated by the mobile group II intron system are disruption mutants and not true deletion mutants, some variation in phenotype maybe observed depending on the intron targeting sequence, and hence its insertion into the gene of interest. This phenomenon, where different mutant phenotypes are observed has been reported when attempting to inactivate *ca_c0437* in *C.*
 15 *acetobutylicum*. Additionally, integration efficiency drastically decreases when attempting to deliver foreign "cargo" DNA >1 kb, and thus this method cannot be effectively used to integrate DNA into the chromosome.

A second method relies upon the creation of auxotrophic mutants, and specifically auxotrophic mutants for uracil. The *pyrE*, *pyrF* or *upp* genes are targeted for initial
 20 disruption leading to an auxotrophic mutant that requires uracil-supplemented media for growth. In addition, these mutants exhibit resistance to 5-Fluoroorotic Acid (5-FOA) or 5-Fluorouracil (5-FU), both toxic antimetabolites. Subsequently, these mutants serve as a genetic background for isolating further mutations in target gene(s) by utilizing a counter-selection method. A plasmid expressing the *pyrF* gene, for example, and containing an
 25 antibiotic resistance marker flanked by two regions of homology to the target gene is transformed into the host cells. To isolate double-crossover mutants, cells are plated onto media containing 5-FOA for negative-selection. Mutants that have undergone the recombination event survive while those that still maintain the plasmid, and hence express the *pyrF* gene, perish. This method has been used to inactivate the *pta* gene in *C.*
 30 *thermocellum*. A similar approach was utilized to isolate unmarked chromosomal mutants in *C. perfringens* albeit the initial mutant created (called the HN13 strain with a disrupted *galKT* operon) was not an auxotroph but rather was unable to metabolize galactose. The HN13 strain was used as genetic background to isolate unmarked chromosomal mutants. GalK is responsible for phosphorylating galactose, which in HN13
 35 is not further metabolized causing toxicity to the cells. Thus the integration vector expresses *galK* to enable negative or counter-selection once cells are plated on galactose supplemented media. This approach was utilized to disrupt *virRS* operon as well as six other genes in *C. perfringens* HN13.

~ 3 ~

A recently reported method for integrating DNA into the chromosome is based upon coupling the expression of a promoterless antibiotic resistance marker to that of the constitutively expressed thiolase (*thl*) promoter to screen for integrants. Integration mutants that have undergone the desired recombination event to fuse the silent antibiotic resistance marker immediately downstream of the *thl* gene (thus leading to the expression of the antibiotic marker) can be easily selected on solid media supplemented with the appropriate antibiotic. Current data on the strength of clostridial promoters suggest that there is a limited number of constitutively promoters that are suitable for this promoter "hijacking" strategy. This constitutes a limitation as to where the integration event can be targeted to in the chromosome.

Most recently, Cartman *et al.* described the use of the *codA* gene that codes for cytosine deaminase from *E. coli* as a counter-selection marker in *C. difficile* for isolating allelic exchange mutants of the *tcdC* gene. CodA converts cytosine to uracil. However, its non-stringent substrate specificity allows it to act upon the pyrimidine analog 5-fluorocytosine (5-FC) and convert it to 5-FU, a toxic antimetabolite. The *upp* gene, that codes for uracil phosphoribosyltransferase, is responsible for conferring toxicity in the presence of 5-FU. Thus, a precondition for the successful use of *codA* as a heterologous counter-selection marker is the presence of the *upp* gene in the host organism as well as the absence of the *codA* gene. *Clostridium* genome sequences indicate that several *Clostridium* organisms contain *codA* homologues. These include *C. perfringens*, *C. botulinum*, *C. lentocellum* and *C. ljungdahlii*.

Thus, in view of the limitations of currently available genetic tools, there remains a need for effective and efficient genetic tools to generate and isolate mutants in *Clostridium* having desirable gene modifications.

25 SUMMARY OF THE INVENTION

The present invention relates to recombinant *Clostridium* organisms having double-crossover allelic exchange and methods for isolation thereof.

A recombinant *Clostridium* organism comprising a recombinant DNA is provided. The recombinant DNA comprises (a) a first nucleic acid sequence encoding a first counter-selection marker, (b) a first promoter operably linked to the first nucleic acid sequence, and (c) a second nucleic acid sequence flanked by a first region at least 90% identical to an upstream region of a first target genomic locus in the recombinant *Clostridium* organism and a second region at least 90% identical to a downstream region of the first target genomic locus.

35 The first counter-selection marker may be a heterologous toxin.

The first promoter may be a heterologous inducible promoter.

The *Clostridium* organism may be selected from the group consisting of *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetanii*, *C. sporogenes*, *C. novyi*, *C.*

~ 4 ~

acetobutylicum, *C. pasteurianum*, *C. beijerinckii*, *C. ljungdahlii*, *C. carboxidivorans*, *C. thermocellum*, *C. cellulolyticum*, *C. phytofermentans*, *C. tyrobutyricum*, *C. butyricum*, *C. kluyveri*, *C. sticklandii*, *Moorella thermoacetica*, *C. autoethanogenum*, *C. ragsdalei*, *Moorella thermoautotrophica*, *C. propionicum*, *C. thermobutyricum*, *C. thermoaceticum*, *C. formicoaceticum*, and *Butyrivacterium methylotrophicum*. Preferably, the *Clostridium* organism is *C. acetobutylicum* ATCC 824.

The second nucleic acid sequence may be a heterologous nucleic acid sequence. The second nucleic acid sequence may comprise a gene, a series of genes, a gene transcription promoter, an antisense RNA, or a combination thereof. The second nucleic acid sequence may encode a positive selection marker.

A method for generating a mutant of the recombinant *Clostridium* organism of the present invention is also provided. The method comprises growing the recombinant *Clostridium* organism, whereby a mutant having double-crossover allelic exchange of the first target genomic locus with the second nucleic acid sequence is generated. The method may further comprise isolating the mutant.

In the present method, the recombinant *Clostridium* organism comprises a recombinant DNA, which comprises (a) a first nucleic acid sequence encoding a first counter-selection marker, (b) a first promoter operably linked to the first nucleic acid sequence, and (c) a second nucleic acid sequence flanked by a first region at least 90% identical to an upstream region of a first target genomic locus in the recombinant *Clostridium* organism and a second region at least 90% identical to a downstream region of the first target genomic locus. The first counter-selection marker may be a heterologous toxin. The first promoter may be a heterologous inducible promoter. The *Clostridium* organism may be selected from the group consisting of *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetanii*, *C. sporogenes*, *C. novyi*, *C. acetobutylicum*, *C. pasteurianum*, *C. beijerinckii*, *C. ljungdahlii*, *C. carboxidivorans*, *C. thermocellum*, *C. cellulolyticum*, *C. phytofermentans*, *C. tyrobutyricum*, *C. butyricum*, *C. kluyveri*, *C. sticklandii*, *Moorella thermoacetica*, *C. autoethanogenum*, *C. ragsdalei*, *Moorella thermoautotrophica*, *C. propionicum*, *C. thermobutyricum*, *C. thermoaceticum*, *C. formicoaceticum*, and *Butyrivacterium methylotrophicum*. Preferably, the *Clostridium* organism is *C. acetobutylicum* ATCC 824.

In the present method, where the first target genomic locus comprises a first target gene, the first target gene may be disrupted or deleted in the mutant.

In the present method, where the second nucleic acid sequence is a heterologous nucleic acid sequence, the heterologous nucleic acid sequence may be integrated into the first target genomic locus in the mutant. The second nucleic acid sequence may comprise a gene, a series of genes, a gene transcription promoter, an antisense RNA, or a combination thereof. Where the second nucleic acid sequence encodes a positive

~ 5 ~

selection marker, the positive selection marker may be integrated into the mutant. The method may further comprise removing the positive selection marker from the mutant.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a mode of action of the MazF mRNA interferase. MazF is part of the toxin-antitoxin system of *E. coli* coded by the *mazEF* operon. Under normal conditions, the antitoxin MazE binds to and inhibits the MazF protein. Once the cells are subjected to stress, the MazE is degraded by cellular proteases and the MazF protein is relieved from inhibition. MazF then targets mRNA at ACA sequences, thus leading to growth arrest followed by cell death.

Figure 2 shows an optimized *mazF* DNA sequence (SEQ ID NO: 1).

Figure 3 illustrates subcultivation via replica plating for isolating desired double-cross recombination mutants. Upon growth in Th-supplemented liquid media overnight (not shown), 200 μ l of culture were spread onto 2X YTG plates supplemented with 5 μ g/ml Th. Plates were then replica plated onto fresh 2X YTG supplement with Th every 24 hours for a total of 5 days. Subsequently, the plates were transferred onto 2X YTG lacking Th to enhance plasmid loss. After 5 days of replica plating, a replica of the plates was made onto Th plus lactose supplemented media and the cells were allowed to grow for 24 hours. Areas that showed growth on the lactose-containing media were streaked onto fresh plates immediately to isolate single colonies for confirmation.

Figures 4A-D show *bgaR*- P_{bgaL} promoter reporter activity in response to 30 min of induction by lactose and toxicity of MazF in *C. acetobutylicum*. (A) A schematic showing the arrangement of the lactose inducible promoter (P_{bgaL}) and divergent regulator (*BgaR*) upstream of the β glucuronidase (*gusA*) gene. (B) Specific activity of the reporter in response to lactose addition. Average activities ($n=2$) were 0.6 U/mg, 8.0 U/mg and 13.0 U/mg when 0 mM, 1 mM and 10 mM of lactose were added, respectively. (C) The pKRAH1_*mazF* expression plasmid, whereby the codon-optimized *mazF* gene was placed under the control of the lactose inducible promoter and cloned into *C. acetobutylicum*. Ori⁺, Gram⁺ origin of replication; Ori⁻, Gram⁻ origin of replication. (D) Plates showing the effects of *mazF* expression in *C. acetobutylicum*. Cells harboring the pKRAH1_*mazF* plasmid were grown in the absence of lactose (-lactose) and presence of lactose (+lactose). Without lactose in the media, cells grow abundantly on the plate, but no growth is observed when lactose is added to the media and thus *mazF* is expressed. One of two biological replicates is shown.

Figures 5A-D show gene replacement via allelic exchange at the *sigK* and *ca_p0167* loci. (A) Selection of unmarked *sigK* double-crossover deletion mutants. The boxed region of the *ca_c1688* and *pilT* show the approximate regions of homology (~ 1 kb each) incorporated in the KO vector. Upon plating cells on plates containing Th and lactose to negatively select against cells bearing the plasmid, putative deletion mutants are selected

~ 6 ~

and screened by PCR with the indicated primers. (B) Isolation of *ca_p0167* deletion mutants. Boxed areas indicate the approximate location of the two regions of homology (~ 1 kb each). Colonies growing on Th plus lactose plates were isolated and mutants confirmed with indicated primers. (C) PCR confirmation of double-crossover *sigK* deletion mutants in (A) with primers that anneal to the chromosome as indicated. Three (3) of the 6 colonies isolated and screened contained the desired double-crossover recombination. One of three is shown. Additionally, the Th^R marker was excised by the FLP recombinase as described in the text thus creating an unmarked *sigK* deletion ($\Delta sigK_{um}$). (D) PCR confirmation of deletion mutants in (B) with the indicated primers. 8 out of 10 colonies screened contained the expected double-crossover integration. One of eight is shown. M, 2-log DNA ladder (NEB).

Figures 6A-B show isolation of *sigF* and *ca_p0157* double-crossover disruption mutants with small internal regions of homology. Five mutants for each gene KO were examined for integration of the ThR gene. These colonies were re-streaked from the Th and lactose plates regions that showed no growth on the Em supplement plates as described in the text. (A) The pKOGEX_mazF plasmid, whereby the ThR gene is flanked by two regions of homology of the target gene. Upon plating on solid media containing both lactose and Th, cells which retained the plasmid perished while those that lose it and undergo homologous recombination with the target sequence survived. (B) Confirmation of gene knockouts using PCR for *sigF* and *ca_p0157* with primers indicated in (A). Genomic DNA was extracted from *C. acetobutylicum* cells that were lactose tolerant and Th resistant. For each gene, target-specific primers that are complementary only to chromosomal, not plasmid, sequences were used as indicated. M, 2-log DNA ladder (NEB). Optionally, an additional replica of the Th plus lactose plate was made on Em containing media to confirm plasmid loss, since EmR is encoded of the plasmid backbone.

Figure 7 illustrates a generic schematic showing the knock-out integration mechanism into the genome of *Cac*. Cells that undergo a double cross-over even while losing the plasmid backbone will be screened on plates containing both, the appropriate antibiotic and lactose (for induction of the mazF counter-selection marker).

Figures 8A-D show isolation of unmarked chromosomal integration of the *fdh* gene in the *sigK* locus. RH1 and RH2 indicate the first and second regions of homology, respectively. (A) Schematic showing the WT, marked and unmarked *fdh* integration in the M5 *sigK* locus as well as the expected amplicon sizes with the indicated primers. (B) PCR screening and confirmation of integration mutants with primers indicated in (A). Three integrants were isolated immediately following transformation of the pKOSIGK::FDH_mazF plasmid on Th and lactose plates, and data for one is shown. The unmarked (*sigK::fdh_um*) strain and WT strain are also shown. MW, λ HindIII digest ladder (NEB); M, 2-log ladder (NEB). (C) sq-RT-PCR confirming the expression of the

~ 7 ~

integrated *fdh* gene in the M5 *sigK* locus. Primers were designed to amplify 700 bp in the middle of the *fdh* transcript. Two biological replicates are shown for both the integration strain as well as the control M5 strain. M, 2-log ladder (NEB). (D) Effectiveness of asRNA-based induction of segregational plasmid instability: loss of the FLP-recombinase expressing plasmid following induction with lactose to express the asRNA against the *repL* origin of replication. After six serial transfers in liquid CGM, with and without lactose, equal volumes were plated onto solid media supplemented with Em (encoded on the plasmid backbone). Following incubation, the plates were visually inspected for growth. As anticipated, the culture that was induced with lactose did not show any growth while the uninduced culture showed growth.

Figures 9A-B show gene replacement via homologous recombination at the *sigF* locus without *recU* expression. (A) A schematic showing the KO vector design and isolation of *sigF* deletion mutants. Boxed areas indicate the approximate location of the two regions of homology (~ 1 kb each). (B) Colonies appearing on Th and lactose plates immediately following transformation were isolated and mutants confirmed with primers indicated in (A). Of the 14 colonies screened, 8 showed the correct PCR product sizes indicating a double-crossover as shown in (A). Data for one is shown. M, 2-log ladder (NEB).

Figure 10 illustrates serial transfers procedure in liquid culture to enrich for double-crossover mutants. A single colony bearing the pKOSIGK2_mazF plasmid was used to inoculate a 10 ml CGM supplemented with 5 µg/ml Th and allowed to grow overnight. A 5% inoculum from this culture was then used to inoculate a fresh CGM tube supplemented with 5 µg/ml Th and allowed to grow for 8-12 hours. This serial transfer procedure was repeated a total of 5 times. Finally, 500 µl from the last culture was transferred to a new CGM tube containing 5 µg/ml Th and 40 mM lactose and allowed to grow for 6 hrs. Serial dilutions were then plated onto solid media containing 5 µg/ml Th and 40 mM lactose and the plates were incubated at 37 °C until colonies were observed. Putative mutants were then isolated and confirmed. All 5 mutants isolated contained the desired double-crossover *sigK* in-frame deletion.

Figure 11 is a flow chart showing the chronology of the design and implementation of a counter-selection method for isolating double-crossover gene deletions or chromosomal integrations.

Figure 12 is a schematic showing a plasmid design for in-frame markerless deletions of genes.

Figure 13 is a schematic showing a plasmid design for gene knock-in the chromosome at a targeted DNA sequence.

~ 8 ~

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based on the discovery a novel system for efficient isolation of double-crossover allelic exchange mutants in *Clostridium* that enables markerless chromosomal gene deletions and DNA integration. The novel system may be obtained by the use of a heterologous counter-selection marker that facilitates identification of generally rare allelic exchange events, and may be used to generate quickly and efficiently markerless DNA deletions and integrations at any genomic locus without the need for auxotrophic mutants or the use of the mobile group II introns. This novel system may be applied in all prokaryotes where double-cross over events are rare or very infrequent (for the toxin gene) and/or where removal of plasmids after genetic recombination is difficult for lack of temperature sensitive plasmids.

The present invention provides a recombinant DNA comprising a first nucleic acid sequence encoding a counter-selection marker and a promoter operably linked to the first nucleic acid. The recombinant DNA further comprises a second nucleic acid sequence flanked by two regions, one homologous to an upstream region of a target genomic locus in a *Clostridium* organism and the other one homologous to a downstream region of the same target genomic locus.

The *Clostridium* organism may be any member of the genus of *Clostridium*, including medically or industrially important species. Medically important species may be used as drug delivery systems that target specific tumors. Industrially important species may be used for producing commodity chemicals and biofuels from a variety of carbon sources, and include acetogens that grow on CO₂/CO/H₂ and cellulolytic clostridia. Examples of the *Clostridium* organism include *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetanii*, *C. sporogenes*, *C. novyi*, *C. acetobutylicum*, *C. pasteurianum*, *C. beijerinckii*, *C. ljungdahlii*, *C. carboxidivorans*, *C. thermocellum*, *C. cellulolyticum*, *C. phytofermentans*, *C. tyrobutyricum*, *C. butyricum*, *C. kluyveri*, *C. sticklandii*, *Moorella thermoacetica*, *C. autoethanogenum*, *C. ragsdalei*, *Moorella thermoautotrophica*, *C. propionicum*, *C. thermobutyricum*, *C. thermoaceticum*, *C. formicoaceticum*, and *Butyribacterium methylotrophicum*. Preferably, the *Clostridium* organism is *C. acetobutylicum* ATCC 824.

The counter-selection marker may be a heterologous toxin. The toxin may be a protein or polypeptide that inhibits the growth of the *Clostridium* organism by, for example, at least 50%, 60%, 70%, 80%, 90%, 95%, 99% or 100%, preferably at least 80%, more preferably at least 90%. Examples of the toxin include mRNA interferase, CcdB, Doc, ParE, Kid, PasB, HigB, RelE, MvpT, Tex, and ζ. The mRNA interferase may be MazF, ChpBK or RelE. The nucleic acid sequence encoding the heterologous counter-selection marker may be optimized to improve expression and translation of the marker in a *Clostridium* organism. For example, an optimized *mazF* DNA sequence is shown in Figure 2 (SEQ ID NO: 1).

~ 9 ~

The promoter may be a homologous or heterologous promoter to a *Clostridium* organism. The promoter is preferably not constitutive activated in a *Clostridium* organism. It may be an inducible promoter or a temporally expressed promoter. The inducible promoter may be selected from the group consisting of a lactose inducible promoter, an
5 IPTG inducible promoter, a metabolite inducible promoter, a light inducible promoter, a catabolite inducible promoter, a promoter induced by a synthetic or non-metabolizable analog of any of the inducers for the foregoing inducible promoters, and a temperature inducible promoter.

10 The term "operably linked" used herein refers to the arrangement of a promoter and a nucleic acid sequence encoding a protein (e.g., a counter-selection marker) such that activation of the promoter leads to the expression of the protein.

A target genomic locus may be any locus in the chromosome of a *Clostridium* organism. The locus may comprise one or more regulatory elements involved in regulating the expression or translation of a target gene. A regulatory element may be a
15 promoter or enhancer of the target gene. The locus may also comprise one or more target genes. A target gene may encode a protein involved in a biological activity or function of the *Clostridium* organism. The biological activity or function may relate to the medical or industrial utility of the *Clostridium* organism. For example, the target genomic locus may comprise the *adc* gene encoding an acetoacetate decarboxylase protein
20 involved in producing acetone by the *Clostridium* organism.

The second nucleic acid sequence may be of any length, and is flanked by two regions of homology. Each flanking region may be of any length, preferably at least about 100, 500 or 2000 nucleotides, more preferably at least about 500 nucleotides. One flanking region is homologous to an upstream of the target genomic locus while the other
25 flanking region is homologous to an downstream of the same target genomic locus.

The term "homology" or "identical" as used herein refers to similarity between two nucleic acid sequences. Preferably, a flanking region may be at least about 60%, 70%, 80%, 85%, 90%, 95%, 99% or 100%, preferably at least about 85%, more preferably at least about 90%, most preferably at least about 95%, identical to the upstream or
30 downstream region of the target genomic locus.

For example, the second nucleic acid sequence may be flanked by a first region at least about 85% identical to an upstream region of the target genomic locus and a second region at least about 85% identical to a downstream region of the target genomic locus. Preferably, the first region is at least about 90% identical to the upstream region while the
35 second region is at least about 90% identical to the downstream region.

The second nucleic acid sequence may be a heterologous nucleic acid sequence. The heterologous nucleic acid sequence may comprise a gene, a series of genes, a gene transcription promoter, an antisense RNA or a combination thereof.

~ 10 ~

The second nucleic acid sequence may encode a protein. The protein may be a marker, preferably, a positive selection marker suitable for the *Clostridium* organism. The positive selection marker may be an antibiotic resistance marker for thiamphenicol (Th), erythromycin (EM), tetracycline (TET), ampicillin (AP), spectinomycin (SP), chloramphenicol (CM), kanamycin (KM), neomycin (NM), gentamicin (GM), clindamycin (CL) or streptomycin (SM). The protein may be essential to an auxotrophic host bacterium.

The second nucleic acid sequence may comprise a gene desirable to be integrated into the chromosome of the *Clostridium* organism. For example, the desirable gene may encode a secondary alcohol dehydrogenase, which improves the biological activity or function of the *Clostridium* organism and reduces acetone to isopropanol.

According to another aspect of the present invention, a recombinant *Clostridium* organism comprising any of the recombinant DNAs of the present invention is provided. The recombinant DNA comprises a first nucleic acid sequence encoding a counter-selection marker, a promoter operably linked to the first nucleic acid, and a second nucleic acid sequence flanked by two regions. While one flanking region is at least about 60%, 70%, 80%, 85%, 90%, 95%, 99% or 100%, preferably at least about 85%, more preferably at least about 90%, most preferably at least about 95%, identical to the upstream of a target genomic locus while another flanking region is at least about 60%, 70%, 80%, 85%, 90%, 95%, 99% or 100%, preferably at least about 85%, more preferably at least about 90%, most preferably at least about 95%, identical to the downstream of the same target genomic locus.

In the recombinant *Clostridium* organism of the present invention, the counter-selection marker may be a heterologous toxin. The toxin may be a protein or polypeptide that inhibits the growth of the *Clostridium* organism by, for example, at least about 50%, 60%, 70%, 80%, 90%, 95%, 99% or 100%, preferably at least about 80%, more preferably at least about 90%. Examples of the toxin include mRNA interferase, CcdB, Doc, ParE, Kid, PasB, HigB, RelE, MvpT, Tex, and ζ . The mRNA interferase may be MazF, ChpBK or RelE. The nucleic acid sequence encoding the heterologous counter-selection marker may be optimized to improve expression and translation of the marker in a *Clostridium* organism. For example, an optimized *mazF* DNA sequence is shown in Figure 2 (SEQ ID NO: 1).

In the recombinant *Clostridium* organism of the present invention, the promoter may be homologous or heterologous. The promoter is preferably not constitutive activated in a *Clostridium* organism. It may be an inducible promoter or a temporally expressed promoter. The inducible promoter may be selected from the group consisting of a lactose inducible promoter, an IPTG inducible promoter, a metabolite inducible promoter, a light inducible promoter, a catabolite inducible promoter, a promoter induced

~ 11 ~

by a synthetic or non-metabolizable analog of any of the inducers for the forgoing inducible promoters, and a temperature inducible promoter.

The term "operably linked" used herein refers to the arrangement of a promoter and a nucleic acid sequence encoding a protein (e.g., a counter-selection marker) such that activation of the promoter leads to the expression of the protein.

For example, where the promoter in the recombinant *Clostridium* organism is an inducible promoter, the growth of the recombinant *Clostridium* organism is inhibited by, for example, at least about 50%, 60%, 70%, 80%, 90%, 95%, 99% or 100%, preferably at least about 80%, more preferably at least about 90%, when the recombinant *Clostridium* organism is grown in a culture medium comprising an inducer of the inducible promoter.

For the recombinant *Clostridium* organism of the present invention, the *Clostridium* organism may be any member of the genus *Clostridium*, including *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetani*, *C. sporogenes*, *C. novyi*, *C. acetobutylicum*, *C. pasteurianum*, *C. beijerinckii*, *C. ljungdahlii*, *C. carboxidivorans*, *C. thermocellum*, *C. cellulolyticum*, *C. phytofermentans*, *C. tyrobutyricum*, *C. butyricum*, *C. kluyveri*, *C. sticklandii*, *Moorella thermoacetica*, *C. autoethanogenum*, *C. ragsdalei*, *Moorella thermoautotrophica*, *C. propionicum*, *C. thermobutyricum*, *C. thermoaceticum*, *C. formicoaceticum*, and *Butyrivacterium methylotrophicum*. Preferably, the *Clostridium* organism is *C. acetobutylicum* ATCC 824.

In the recombinant *Clostridium* organism of the present invention, the second nucleic acid sequence may be a heterologous nucleic acid sequence. The heterologous nucleic acid sequence may comprise a gene, a series of genes, a gene transcription promoter, an antisense RNA or a combination thereof.

The second nucleic acid sequence may encode a protein. The protein may be a marker, preferably, a positive selection marker suitable for the *Clostridium* organism. The positive selection marker may be an antibiotic resistance marker for thiamphenicol (Th), erythromycin (EM), tetracycline (TET), ampicillin (AP), spectinomycin (SP), chloramphenicol (CM), kanamycin (KM), neomycin (NM), gentamicin (GM), clindamycin (CL) or streptomycin (SM). The protein may be essential to an auxotrophic host bacterium.

According to yet another aspect of the present invention, a method for generating a mutant of the recombinant *Clostridium* organism of the present invention is provided. The method comprises growing the recombinant *Clostridium* organism under conditions permitting double-crossover allelic exchange of the target genomic locus in the recombinant *Clostridium* organism, thereby generating a mutant having the double-crossover allelic exchange of the first target genomic locus with the second nucleic acid sequence.

~ 12 ~

The method may further comprise isolating the mutant. The term "isolating" used herein refers to separating the mutant from the *Clostridium* organism.

In the method of the present invention, the target genomic locus may comprise a target gene. Upon the double-crossover allelic exchange of the target genomic locus, the target gene may be deleted or disrupted in the mutant. The deletion may be in-frame. The mutant may have a knockout (KO) of the target gene.

In the method of the present invention, the second nucleic acid sequence may be a heterologous nucleic acid sequence, which may be integrated into the first target genomic locus in the mutant. The heterologous nucleic acid sequence may comprise a gene, a series of genes, a gene transcription promoter, an antisense RNA or a combination thereof. Where the gene encodes a positive selection marker, the resulting mutant carries the positive selection marker upon double-crossover allelic exchange. The method may further comprises removing the positive selection marker from the mutant. The marker may be removed from the mutant by various conventional techniques, for example, asRNA-based induction of segregational plasmid instability.

The method of the present invention may further comprise multiple sequential knockouts and/or knockins of multiple target genes. After each knockout or knockin, a positive selection marker used to confirm the knockout or knockin may be introduced into the mutant, and may be removed to generate a markerless mutant.

The term "about" as used herein when referring to a measurable value such as an amount, a percentage, and the like, is meant to encompass variations of $\pm 20\%$ or $\pm 10\%$, more preferably $\pm 5\%$, even more preferably $\pm 1\%$, and still more preferably $\pm 0.1\%$ from the specified value, as such variations are appropriate.

Example 1. Isolation of double-crossover allelic exchange mutants

A method to enable gene replacement KOs as well as KIs in *Clostridium* has been developed based on counter-selection to overcome the limitations methods currently employed to isolate allelic exchange mutants in *Clostridium*. The method includes two components: an inducible promoter and a toxic gene based on a general mechanism that would be applicable to *Clostridium* and many, if not all, other organisms. A recent study identified a lactose inducible promoter and its divergent regulator (*bgaR*; CPE0770) in *C. perfringens* strain 13. The *Escherichia coli* gene *mazF*, an mRNA interferase, was chosen as the toxic gene. The toxin, MazF exhibits a general mechanism for toxicity, and has been successfully used as a counter-selection marker in *B. subtilis*. The MazE-MazF complex is part of the toxin-antitoxin system in *E. coli* (Fig. 1). MazF is stable while the anti-toxin MazE is labile. The *mazF* gene was synthetically constructed for optimized translation in *C. acetobutylicum*. Once expressed and translated, MazF cleaves mRNA at ACA sequences therefore arresting cell growth (Fig. 1). The combination of the lactose-

~ 13 ~

inducible promoter with this *mazF* gene led to the development of an effective counter-selection method for generating both KOs and KIs in *C. acetobutylicum*.

MATERIALS AND METHODS

Bacterial strains and plasmids. Relevant characteristics of bacterial strains and

5 plasmids used in this example are listed in Table 1.

Table 1. Bacterial strains and plasmids

Strains or plasmids	Relevant Characteristics	Sources or Reference
• <i>Strains</i>		
<i>E. coli</i>		
ER2275	<i>hsdR mcrA recA1 endA1</i>	NEB
Top 10	<i>hsdR mcrA recA1 endA1</i>	Life Technologies
DB3.1	<i>ccdb</i> survival cells	Life Technologies
Turbo	<i>lacI^r endA1</i>	NEB
<i>C. acetobutylicum</i>		
ATCC 824	Wild-Type strain	ATCC
M5	ATCC 824 pSOL1	ATCC
824 Δ <i>sigK</i>	Mostly deleted <i>sigK</i> , Th ^R	Example 1
824 Δ <i>sigK</i> <i>um</i>	Unmarked <i>sigK</i> deletion	Example 1
824 Δ <i>ca_p0167</i>	Mostly deleted <i>ca_p0167</i> , Th ^R	Example 1
824 <i>sigF::Th^R</i>	Disrupted <i>sigF</i> , Th ^R	Example 1
824 <i>ca_p0157::Th^R</i>	Disrupted <i>ca_p0157</i> , Th ^R	Example 1
824 Δ <i>sigF</i>	Mostly deleted <i>sigF</i> , Th ^R	Example 1
M5 <i>sigK::fdh</i>	<i>fdh</i> integrated in <i>sigK</i> locus, Th ^R	Example 1
M5 <i>sigK::fdh</i> <i>um</i>	Unmarked <i>fdh</i> integrated in <i>sigK</i> locus	Example 1
• <i>Plasmids</i>		
pAN3	Km ^R ; Φ 3T I gene	Example 1
pKORECU	Th ^R ; MLS ^R ; <i>recU</i>	Tracy 2011
pCR8/GW/TOPO TA	Sp ^R ; topoisomerase	Life Technologies
pCR8/ <i>sigF</i> /Th ^R	pCR8/GW/ <i>sigF::Th^R</i>	Jones
pKORD_ <i>mazF</i>	Th ^R ; MLS ^R ; <i>ccdb</i> , <i>recU</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> ; DEST Cassette	Example 1
pKOD_ <i>mazF</i>	Th ^R ; MLS ^R ; <i>ccdb</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> ; DEST Cassette	Example 1
pAH2	Th ^R ; <i>bgaR</i> and P _{bgaL} upstream of <i>cpe-gusA</i>	Hartman
pKRAH1	Th ^R ; <i>bgaR</i> and P _{bgaL} upstream of MCS	Hartman
pKRAH1_ <i>mazF</i>	Th ^R ; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i>	Example 1
P94FLP_ <i>asRepL</i>	Em ^R ; flippase; <i>ptb</i> promoter; <i>bgaR</i> and P _{bgaL} upstream of <i>asRNA</i> to the <i>repL</i> ori	Example 1
pCR8/ <i>cap0157</i> /Th ^R	pCR8/GW/ <i>ca_p0157::Th^R</i>	Example 1
pCR8/ <i>cap0167</i> /Th ^R	pCR8/GW/Th ^R -FRT flanked by <i>ca_p0167</i> ~1000 bp up- and downstream region of homology	Example 1
pCR8/ <i>sigK</i> /Th ^R	pCR8/GW/Th ^R -FRT flanked by <i>sigK</i> ~1000 bp of up- and downstream region of homology	Example 1
pCR8/ <i>sigF2</i> /Th ^R	pCR8/GW/Th ^R -FRT flanked by <i>sigF</i> ~1000 bp of up- and downstream region of homology	Example 1
pKOSIGF_ <i>mazF</i>	Th ^R -FRT; MLS ^R ; <i>recU</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> , <i>sigF::Th^R</i>	Example 1
pKOCAP0157_ <i>mazF</i>	Th ^R ; MLS ^R ; <i>recU</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> , <i>ca_p0157::Th^R</i>	Example 1
pKOSIGK_ <i>mazF</i>	Th ^R -FRT; MLS ^R ; <i>recU</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL}	Example 1

~ 14 ~

Strains or plasmids	Relevant Characteristics	Sources or Reference
	upstream of <i>mazF</i> , Th ^R -FRT flanked by <i>sigK</i> ~1000 bp of up- and downstream region of homology	
pKOCAP0167_ <i>mazF</i>	Th ^R -FRT; MLS ^R , <i>recU</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> , Th ^R -FRT flanked by <i>ca_p0167</i> ~1000 bp of up- and downstream region of homology	Example 1
pKISIGK::FDH_ <i>mazF</i>	Th ^R -FRT; MLS ^R , <i>recU</i> ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> , Th ^R and <i>fdh</i> flanked by <i>sigK</i> ~1000 bp of up- and downstream region of homology	Example 1
pKOSIGF2_ <i>mazF</i>	Th ^R -FRT; MLS ^R ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> , Th ^R -FRT flanked by <i>sigF</i> ~1000 bp of up- and downstream region of homology	Example 1
pKOSIGK2_ <i>mazF</i>	Th ^R -FRT; MLS ^R ; <i>repL</i> ; ori; <i>bgaR</i> and P _{bgaL} upstream of <i>mazF</i> , Th ^R -FRT flanked by <i>sigK</i> ~1000 bp of up- and downstream region of homology	Example 1
p94MCS	Amp ^R ; MLS ^R ; <i>ptb</i> promoter; MCS	Example 1
p94FDH	Amp ^R MLS ^R ; <i>ptb</i> promoter; <i>fdh</i>	Example 1

Codon optimization of *mazF* gene. The *mazF* gene was synthetically constructed (DNA 2.0, Melano Park, CA) for optimized codon usage in *C. acetobutylicum* (Fig. 2). Codons that had a usage frequency of less than 5 per one thousand were exchanged for those that exhibited a higher frequency. Additionally, the *lacI* operator sequence was placed upstream of the optimized *mazF* gene to enable the cloning of *mazF*-containing vectors in *E. coli* strains that express *lacI* or *lacI*^q.

Construction of plasmids. All primers are listed in Table 2. The pKOSIGK_*mazF* plasmid was constructed by PCR amplifying the two large regions of homology flanking the *sigK* gene utilizing splicing by overhang extension (SOE) PCR. The generated amplicon (~2 kb) was then cloned into the pCR8/GW/TOPOTA (Life Technologies [LT], Grand Island, NY), linearized with ScaI (New England Biolabs [NEB], Ipswich, MA), and finally ligated to the FRT-flanked thiamphenicol (Th^R-FRT) resistance gene. The resulting plasmid (pCR8/*sigK*/Th^R) was then recombined with either the pKORD_*mazF* or the pKOD_*mazF* (no *recU*) destination plasmids by utilizing the Gateway LR Clonase (LT) reaction yielding plasmids pKOSIGK_*mazF* and pKOSIGK2_*mazF*, respectively.

Table 2. List of primers used

Primer	Sequence (5' --> 3')	SEQ ID NO	Description
sigK F1 UP	GACAGCAGTGGATTGAAACAGG T	2	amplify sigK upstream region of homology
sigK R1 UP ScaI	AGTACTCATTACAAACACAGGTA ATCCTCC	3	amplify sigK upstream region of homology

~ 15 ~

Primer	Sequence (5' --> 3')	SEQ ID NO	Description
sigK F1 DN	GGCAACTAAACGGAACTGC	4	amplify sigK downstream region of homology
sigK R1 DN	GCCGCAGTAATAGTTTCCATATC TC	5	amplify sigK downstream region of homology
sigK PCA F2 DN	GGAGGATTACCTGTGTTTGTAAAT GAGTACTGGCAACTAAACGGAA ACTGC	6	SOE PCR primer for sigK region of homology
sigK PCA R2 UP	GCAGTTTCCGTTTAGTTGCCAGT ACTCATTACAAACACAGGTAATC CTCC	7	SOE PCR primer for sigK region of homology
sigK F conf	CTAATTGCAACTAATGGTGGGC	8	sigK mutant confirmation and sequencing primer primer
sigK R conf	TCATACCGTACTTGCTTCTCTGTC	9	sigK mutant confirmation and sequencing primer primer
CAP0167 F1 UP	CCAAGCTTCATGTACGGTAACAT C	10	amplify ca_p0167 region of homology
CAP0167 R1 DN	AACTGTTTCTCAGGGGAGAATG AC	11	amplify ca_p0167 region of homology
CAP0167 F conf	GAATGTGAGAGCTAGAAACAAT CTC	12	ca_p0167 mutant confirmation and sequencing primer
CAP0167 R conf	GAGCTGTTATTGTAAATGAAGG CGAC	13	ca_p0167 mutant confirmation and sequencing primer
sigF F1 UP	CCGGGAAATTCTAGATAAAATC CCGG	14	amplify sigF upstream region of homology
sigF R1 UP ScaI	AGTACTCTACTTATTGCTGATAC TAGCGGAAG	15	amplify sigF upstream region of homology
sigF R1 DN	CTCTCAGCCGCAGAAAAATGTG ATG	16	amplify sigF downstream region of homology
sigF F2 DN	CTCAGGTCGCAAATATGCTGGG	17	amplify sigF downstream region of homology
sigF PCA F2 DN	CTTCCGCTAGTATCAGCAATAAG TAGAGTACTCTCAGGTCGCAAA TATGCTGGG	18	SOE PCR primer for sigF region of homology
sigF PCA R2 UP	CCCAGCATATTTGCGACCTGAG AGTACTCTACTTATTGCTGATAC TAGCGGAAG	19	SOE PCR primer for sigF region of homology
Th inframe R	AACCGCCATTTCAGAGTTTAGGA CG	20	reverse primer that binds inside thiamphenicol resistance marker
Th inframe F	CGGAGCATTTGGCTTTCCTTCCA T	21	forward primer that binds inside thiamphenicol resistance marker
bgaR_mazF F2 ClaI	AAAAAAATCGATCCTCGAGATG AAAAGTATTAGGGC	22	amplify the bgaR-mazF fragment
bgaR_mazF R2 ClaI	AAAAAAATCGATCCCGATTGAC GAATACTATGAAAC	23	amplify the bgaR-mazF fragment
mazF F1	ACAATTTACAGGAGGGCTGAA ATGGTAAGCAGGTACGTA	24	amplify codon optimized mazF from DNA 2.0 plasmid

~ 16 ~

Primer	Sequence (5' --> 3')	SEQ ID NO	Description
mazF R1	CGATTGACGAATACTATGAAAC AATATT	25	amplify codon optimized mazF from DNA 2.0 plasmid
mazF F2 SalI	AAAAAAGTCGACGTGGAATTGT GAGCGGATAACAATTCACAGG AGGGCTGAAAT	26	amplify codon optimized mazF and incorporate SalI site
mazF R2 BamHI	AAAAAAGGATCCCGATTGACGA ATACTATGAAACAATATT	27	amplify codon optimized mazF and incorporate BamHI site
CAP0157 F	GCAGCACATATGCCGCTCAGGA ATTA	28	amplify internal fragment of ca_p0157
CAP0157 R	CGTTTCCTGTTCTACCCTCTATC C	29	amplify internal fragment of ca_p0157
CAP0157 F conf	TAGTCATGGGAAGGTTTGGC	30	ca_p0157 mutant confirmation primer
CAP0157 R conf	AACAGTATGGAGGCATAGGTG	31	ca_p0157 mutant confirmation primer
sigF F	TTTGAGAGATGACGGAATAAT	32	amplify sigF internal fragment
sigF R	CAGCATATTTGCGACCTGAG	33	amplify sigF internal fragment
sigF F conf	GGAAACAGCGGGGAAGCTAATG TT	34	sigF disruption mutant confirmation primer
sigF R conf	GTCCATGTAGTTCCTCCCTAGA TA	35	sigF disruption mutant confirmation primer
C08930_ncol_r	TATACCCCATGGTTGATAAAAGT TAATTTGCAGTTAC	36	Amplify the fdh gene
C08930_sacii_f	TTTACTCCGCGGGTTATTATGGA GAATAAAGTATTAA	37	Amplify the fdh gene
antiRepL_BamHI	TTACTGGGATCCATGAAAGAAA GATATGGAACAGTCTATAAAG	38	antisense repl
antiRepL_SacII	TACTTACCGCGGTTTGTTTACGG TATAACTTGTCTACCTCTAT	39	antisense repl
P94_sphi_f	TTAATAGCATGCTCGACTGTGG ATGGAGTTAA	40	amplify fdh gene from p94FDH with promoter and terminator
P94_sphi_r	TTAATCGCATGCTGACCATGATT ACGAATTCTATGA	41	amplify fdh gene from p94FDH with promoter and terminator
C08930_sqRT-PCR_f	AGCTGGATTGGCACTACTAACA GGAAATTAT	42	sq-RT-PCR primers for the fdh gene
C08930_sqRT-PCR_r	AACAGGAGTTCCTGGATGATCT TCTGT	43	sq-RT-PCR primers for the fdh gene
sigF conf F2	GTATTCCAGCCCCTTAAATTAT GAAC	44	sigF deletion mutant confirmation primer
sigF conf R2	CTCCCATACTATCAGGACTTGTT ATTCC	45	sigF deletion mutant confirmation primer
M13 F (-20)	GTAAAACGACGCCAG	46	sequencing primer for KO and KI fragments in sequencing plasmids
M13 R	CAGGAAACAGCTATGAC	47	sequencing primer for KO and KI fragments in sequencing plasmids

~ 17 ~

Primer	Sequence (5' --> 3')	SEQ ID NO	Description
bgaR_mazF R1	CCC GAT TGA CGA ATA CTA TGA AAC	48	Reverse primer to amplify the bgaR_mazF fragment from pKRAH1_mazF
bgaR_mazF F1	CCT CGA GAT GAA AAG TAT TAG GGC	49	Forward primer to amplify the bgaR_mazF fragment from pKRAH1_mazF

The pKOCAP0167_mazF plasmid was constructed by PCR amplifying the complete region of homology, including the *ca_p0167* ORF, and the resulting amplicon was cloned into pCR8/GW/TOPOOTA (LT). This plasmid was then double-digested with SspI/BsgI (NEB), resulting in the removal of 285 bp from the *ca_p0167* ORF, blunt ended with DNA polymerase I, Large (Klenow) fragment (NEB) and finally ligated the Th^R-FRT gene. The resulting plasmid (pCR8/cap0167/Th^R) was shuffled into the pKORD_mazF destination plasmid using the Gateway LR Clonase (LT) reaction to yield plasmid pKOCAP0167_mazF.

The *sigF* disruption plasmid was constructed by amplifying a 391 bp internal fragment and cloning it into pCR8/GW/TOPOOTA (LT). The resulting plasmid was then ScaI linearized (NEB) (ScaI has a single digest site in the middle of the amplified *sigF* region), ligated to the Th^R marker to yield plasmid pCR8/sigF/Th^R, which was then recombined with the pKORD_mazF destination plasmid to generate pKOSIGF_mazF.

The *ca_p0157* disruption plasmid was constructed by amplifying a 498 bp internal fragment, and cloning it into pCR8/GW/TOPOOTA (LT). The resulting plasmid was HindIII (NEB) linearized (HindIII has a single cut site approximately in the middle in the *ca_p0157* fragment), blunt ended with DNA polymerase I, Large (Klenow) fragment (NEB) and ligated to the Th^R gene yielding plasmid pCR8/cap0157/Th^R. This plasmid was then shuffled into the pKORD_mazF destination plasmid as before yielding plasmid pKOCAP0157_mazF.

The pKISIGK::FDH_mazF was constructed as follows. First, the formate dehydrogenase (*fdh*) gene (CLJU_c08930) from *C. ljungdahlii* (RS: NC_014328) was PCR amplified from *C. ljungdahlii* genomic DNA (gDNA) using primers incorporating the NcoI and SacII digest sites and then cloned into the NcoI/SacII (NEB) double-digested p94MCS plasmid yielding plasmid p94FDH. In this construct, the *fdh* gene is flanked upstream by a strong clostridial promoter (P_{ptb}) and downstream by the *adc* gene's rho-independent terminator. Next, the full length fragment, including the promoter and terminator, was PCR amplified from p94FDH and cloned into the NotI (NEB) linearized and blunt-ended pKOSIGK_mazF plasmid yielding plasmid pKISIGK::FDH_mazF.

The p94FLP-asRepL plasmid was constructed by treating the p94FLP plasmid with SbfI (NEB) to linearize it and subsequently clone the *bgaR*-asRepL fragment yielding plasmid p94FLP_asRepL. The lactose inducible promoter and its divergent regulator

~ 18 ~

(*bgaR*-P_{bgaL}) were utilized and fused immediately upstream of a 100 bp antisense-RNA (asRNA) that targets the first 100 bp of the *repL* transcript coding for the replication protein (RepL) creating plasmid p94FLP_asRepL. The asRNA is designed to be a reverse and complement RNA strand to that of the *repL* transcript. Once transcribed, the asRNA binds to the transcript and inhibits its translation.

The pKOSIGF2_mazF plasmid designed to generate a *sigF* deletion mutant was constructed by PCR amplifying the two large regions of homology flanking the *sigF* gene using SOE PCR. The resulting PCR product (~ 2 kb) was cloned into the pCR8/GW/TOPOTA (LT), linearized with ScaI (NEB) and ligated to the Th^R-FRT gene. The resulting plasmid (pCR8/sigF2/Th^R) was then recombined with the pKOD_mazF destination plasmid that is lacking the *recU* gene by utilizing the Gateway LR Clonase (LT) reaction yielding plasmid pKOSIGF2_mazF.

Culture conditions. *E. coli* strains were grown aerobically at 37° C and 220 rpm in liquid LB media or solid LB with agar (1.5%) supplemented with ampicillin (Amp) (50 µg/ml) or chloramphenicol (Cm) (35µg/ml) as needed. Single *C. acetobutylicum* colonies grown anaerobically on solid 2X YTG for at least 5 days (pH 5.8) were transferred to 10 ml of liquid CGM and heat shocked at 80° C for 10 min. Recombinant strains were grown as before in media supplemented with erythromycin (Em) (40 µg/ml for solid media and 100 µg/ml for liquid media) or thiamphenicol (Th) (5 µg/ml) as necessary. Where needed, 40 mM of β-lactose (Sigma Aldrich, St. Louis, MO) was added to the culture media to induce the expression from the lactose-inducible promoter.

Reporter enzyme assay. *C. acetobutylicum* cultures were grown to an A₆₀₀ of 1.0 in liquid 2X YTG supplemented with 5 µg/ml of Th where needed. Lactose was then added to the respective culture to a final concentration of 0 mM, 1 mM or 10 mM and cultures were incubated for 30 min at 37 °C in an anaerobic chamber. Cells were then collected by centrifugation at 5000 × *g* for 10 min at 4 °C and the pellets were stored at -20 °C until used. The β-glucuronidase enzyme assay was carried out as described in Hartman et al. All experiments were done with at least two biological replicates.

Plasmid transformation. Newly constructed plasmids were transformed into *E. coli* Top10 (LT) cells and subsequently isolated and confirmed using restriction digests. Plasmid(s) destined to be electrotransformed into *C. acetobutylicum* were first transformed into electrocompetent *E. coli* ER2275 (pAN3) cells for DNA methylation. pAN3-bearing cells are a derivative of the ER2275 (pAN1) cells except that they have kanamycin resistance. The vectors were then isolated and confirmed from their respective *E. coli* ER2275 (pAN3) strains and subsequently transformed into *C. acetobutylicum* ATCC 824 or M5 by electroporation in an anaerobic chamber as described in Mermelstein et al. except that 5-10 µg of plasmid DNA were used to transform the cells.

~ 19 ~

Isolation and enrichment of gene replacement mutants via allelic exchange.

Upon transforming the appropriate plasmid(s) in *C. acetobutylicum* and following 4 hrs of outgrowth in liquid 2X YTG, cells were collected by centrifugation at $5000 \times g$ for 10 min and resuspended in 500 μ l of 2X YTG. 100 μ l was spread onto solid 2X YTG media supplemented with 5 μ g/ml Th and 40 mM lactose (4 plates) and Th only plates (1 plate). The plates were incubated for 48-72 hrs in an anaerobic chamber at 37 °C or until colonies were observed. Putative mutants appearing on the Th+lactose plates were isolated and screened by PCR to confirm the presence of the desired allelic exchange event.

If no colonies were seen on the Th and lactose plates, colonies from the Th-only plate were grown in liquid CGM and a series of subcultivation was undertaken to enrich for desired gene replacement mutants using a similar approach to the one described in Tracy et al. (2011) except that lactose containing plates were utilized for the final selection plate (Fig. 3). Transformants were vegetatively transferred every 24 hrs for 5 days on 2X YTG solid media supplemented with Th at 5 μ g/ml. After five days, cells were again vegetatively transferred for an additional 5 days under no antibiotic selection to facilitate plasmid curing. After five days of curing, cells were plated on Th+lactose plates, allowed to grow for 24 hrs and colonies were isolated and confirmed. Optionally, these plates were replica plated (21, 49) onto plates supplemented with Em at 40 μ g/ml (Em^R is encoded on the vector backbone) and allowed to grow for 24 hrs. Plates were then visually compared to the Th+lactose plates. Ideally, areas of growth on Th+lactose plates should show no growth on Em plates and this indicates chromosomal integration via allelic exchange. Areas on the Th+lactose plates that showed no growth on the Em supplemented plates were re-streaked on fresh plates to isolate single colonies for screening. Alternatively, a liquid transfer protocol may be undertaken. One colony from the Th-only plate was used to inoculate 10 ml liquid CGM supplemented with 5 μ g/ml Th and allowed to grow overnight at 37 °C. Subsequently, a 5% inoculum was transferred to a fresh 10 ml CGM tube supplemented with 5 μ g/ml Th and allowed to grow for 8-12 hrs. This transfer process was repeated a total of 5 times. Finally, 500 μ l from the final transfer were used to inoculate 10 ml of CGM supplemented with 5 μ g/ml Th and 40 mM of β -lactose and allowed to grow for 6 hours. Serial dilutions were plated onto solid 2X YTG media supplemented with 5 μ g/ml Th and 40 mM of β -lactose. Single colonies were then isolated and screened to confirm the presence of the desired allelic exchange event.

RNA isolation and semiquantitative reverse transcription PCR (sq-RT-PCR). RNA isolation, cDNA generation and semiquantitative real-time PCR were carried out as previously described except that 2 μ g of total RNA was reverse transcribed. Primers were designed to amplify a ~700 bp fragment from the reverse transcribed *fdh* transcript. All amplification reactions were carried out using Phusion® High-Fidelity DNA Polymerase (NEB).

~ 20 ~

Analytical techniques. Cell density was measured as A_{600} using a Beckman Coulter DU 730 spectrophotometer.

Confirmation of chromosomal or megaplasmid insertions by PCR and sequencing.

To confirm allelic exchange chromosomal integration mutants, gDNA was isolated from cells using the DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA) following the manufacturer's protocol. Appropriate primer combinations (Table 2) were used to confirm the putative integrants. Amplification reactions were carried out using Phusion® High-Fidelity DNA Polymerase (NEB). Mutants were also confirmed by Sanger sequencing.

RESULTS

10 Development of an inducible counter-selection marker. A lactose-inducible promoter that was developed for use in *C. perfringens* strain 13 was tested for functionality in *C. acetobutylicum*. Plasmid pAH2 with the inducible promoter placed upstream of the reporter gene β -glucuronidase (Fig. 4A), was introduced into *C. acetobutylicum*, giving rise to strain 824 (pAH2), and its response to induction with

15 lactose was tested. Previous work reported that lactose concentrations above 10 mM did not significantly increase expression levels from this promoter. The results showed that the promoter indeed did respond to induction with lactose in *C. acetobutylicum*. After 30 min of induction, the reporter's specific activity increased by 10 and 15 fold upon 1 mM and 10 mM of lactose addition, respectively (Fig. 4B). Baseline levels of the reporter's

20 activity (i.e., without addition of lactose) were close to WT levels. Without lactose induction the reporter's specific activity in strain 824 (pAH2) was 0.6 ± 0.2 U/mg compared to 0.2 ± 0.01 U/mg for the WT strain. These results suggest slight promoter leakiness but, it is hypothesized, that it would not be sufficient to induce the expression of the *mazF* gene (described below) to high enough levels that would promote cell death.

25 The synthetic *mazF* gene was cloned downstream of the lactose-inducible promoter and its divergent regulator (*bgaR*- P_{bgaL}) on plasmid pKRAH1. This generated plasmid pKRAH1_*mazF* (Fig. 4C). The plasmid was electrotransformed into *C. acetobutylicum*. Once the strain carrying the plasmid was confirmed, colonies were grown in liquid media supplemented with 5 μ g/ml of Th for 24 hrs. Subsequently, equal volumes of culture were

30 plated onto solid 2X YTG media supplemented with 5 μ g/ml of Th, with and without lactose. After 24 hrs of anaerobic incubation at 37° C, plates were visually inspected for growth. As anticipated, plates containing lactose did not have any colonies while those lacking lactose showed an abundance of growth (Fig. 4D). These data suggested that the toxicity of MazF might enable us to utilize it for counter-selection to screen for cells that

35 have targeted allelic exchange events.

Utilizing *mazF* as a counter-selectable marker to isolate *sigK* and *ca_p0167* gene deletions via allelic exchange using large regions of homology. To demonstrate the application and utility of the developed inducible counter-selection marker, the sporulation

~ 21 ~

- specific sigma-factor K (σ^K) and the gene encoding for CA_P0167, a putative sigma factor of unknown function encoded on the *C. acetobutylicum* pSOL1 megaplasmid were targeted for inactivation. The KO plasmids were designed to delete the majority of the ORF of each gene. Plasmid pKOSIGK_mazF included a Th^R -FRT that is further flanked by two large
- 5 regions of homology ~ 1000 bp both upstream and downstream of the *sigK* gene (with only 12 bp of the 5' *sigK* ORF and 29 bp of the 3' end included in the upstream and downstream regions of homology, respectively) (Fig. 5A). This design was to ensure that no DNA regulatory elements of adjacent genes would be affected. Consequently, the resulting *sigK* mutant would have 664 bp of the 705 bp of the *sigK* gene deleted.
- 10 Similarly, plasmid pKOCAP0167_mazF was designed with two regions of homology for the *ca_p0167* gene that ensured the removal of 285 bp of the 555 bp coding sequence (Fig. 5B). Allelic exchange mutants were isolated by utilizing the solid replica plating protocol described in the materials and methods section. Putative mutants appearing on the
- 15 Th +lactose plates were isolated and screened by PCR to confirm the presence of the desired gene replacement event using primers listed in Table 2. Of the 6 colonies screened for the *sigK* mutant, 3 contained the targeted gene replacement via allelic exchange to delete a significant portion of the *sigK* ORF as confirmed by PCR (Fig. 5C) and Sanger sequencing. Similarly, of the 10 colonies screened for the *ca_p0167* mutant, eight
- 20 contained the desired allelic exchange event to delete 285 bp of the *ca_p0167* ORF as confirmed by PCR (Fig. 5D), as well as sequencing.
- Isolation of *sigF* and *ca_p0157* disruption mutants with small (<300 bp), internal regions of homology.** To demonstrate the strength of this method, gene disruption mutants of the sporulation specific sigma factor F (σ^F) and the predicted megaplasmid sigma factor CA_P0157 were generated by utilizing the Th^R marker flanked
- 25 by two internal regions of homology of <300 bp (Fig. 6A). Previous work in attempting to isolate allelic exchange disruption mutants of the *sigF* gene yielded only mutants, which had undergone a Campbell-like single-crossover event that resulted in the entire plasmid backbone rolling into the targeted region of the chromosome. As described above, the replica plating protocol was employed to enrich for mutants that have undergone the
- 30 allelic exchange recombination with the target gene. All 5 putative mutants isolated contained the desired recombination event (Fig. 6B). Similarly, we aimed to isolate allelic exchange disruption mutant of the *ca_p0157* gene. The pKOCAP0157_mazF plasmid was methylated and electrotransformed into wild-type *C. acetobutylicum* and the replica plating protocol was carried out as above. All 5 colonies screened had undergone the
- 35 desired allelic exchange event (Fig. 6B). It was previously reported that the rate of recombination significantly decreases when the regions of homology are less than 400 bp in various unicellular organisms as well as in mammalian cells. Notwithstanding, the

~ 22 ~

selection protocol described herein is capable of identifying rare allelic exchange mutants generated by small regions of homology (<300 bp) for the target gene(s).

A generic illustration of the design of the KO plasmid containing the *mazF* under the control of the lactose inducible promoter is shown in Figure 7. The regions of

5 homology can be designed for each individual gene KO target.

Unmarked genomic integration of the 2.1 kb formate dehydrogenase gene from *C. ljungdahlii*.

To further demonstrate the utility of our method, integration of the 2.1 kb formate dehydrogenase (*fdh*) gene (CLJU_c08930) from *C. ljungdahlii* was undertaken.

The expected sizes of the wild-type (WT) and integrant loci are shown in (Fig. 8A). The

10 *fdh* KI plasmid was methylated and electrotransformed into the asporogenous M5 *C. acetobutylicum* strain which has lost the pSOL1 megaplasmid. Due to the relatively large size of the pKISIGK::FDH_mazF plasmid (~12.3 kb) and in order to increase the transformation efficiency, the plasmid was incubated with competent cells on ice for 20 min before electrotransformation. Following 4 hrs of outgrowth in liquid 2X YTG, cells

15 were plated onto solid media containing 5 µg/mL Th and 40 mM of β-lactose aiming to isolate putative integrants immediately following plasmid transformation. Of the estimated 5 X 10³ transformants, we were immediately able to isolate three (i.e., 0.06% of the total estimated 5 X 10³ transformants) that had undergone the desired allelic exchange event, namely to integrate the ~3.6 kb "cargo" DNA (2.1 kb *fdh* and a 1.5 kb

20 Th^R-FRT marker) as confirmed by PCR (Fig. 8B), and sequencing. To test active expression of the newly integrated *fdh* gene, sq-RT-PCR was carried out as described in the materials and methods section using primers that amplified ~ 700 bp of the reverse transcribed *fdh* cDNA. Two biological replicates from each strain (M5 *sigK::fdh* and M5 control strain) were used for RNA isolation and cDNA generation. Both M5 *sigK::fdh*

25 biological replicates showed the expected size band while none of the control strains did (Fig. 8C), thus, confirming the expression of the integrated *fdh* gene.

Easy marker removal: expression of the FLP recombinase from a plasmid with inducible-segregational instability.

Due to the limited number antibiotic resistance markers that are suitable for use in *C. acetobutylicum*, we set out to "flip-out" the Th^R

30 marker from the M5 *sigK::fdh* integration strain and the 824 Δ*sigK* strain via the expression of FLP recombinase. Anecdotal evidence suggested once excision of the FRT-flanked marker is confirmed, stimulating plasmid loss is laborious and time consuming due to the relatively good stability of *repL* Gram⁺ replicon. There are no known temperature sensitive replicons in *Clostridium*. There are few established replication origins suitable for

35 use on plasmids to be propagated in *Clostridium*. However, these origins are *segregationally* very stable with an estimated retention frequency of >90% after 60 generations without any selective pressure. Hence, a means was developed to easily and readily cure the FLP recombinase expression plasmid once the excision of the antibiotic

~ 23 ~

marker was confirmed. Once again, the *C. perfringens* lactose inducible promoter was utilized and its divergent regulator (*bgaR*-P_{bgaL}) and fused it immediately upstream of a 100 bp antisense-RNA (asRNA) that targets the first 100 bp of the *repL* transcript coding for the replication protein (RepL) creating plasmid p94FLP_asRepL. This plasmid was

5 transformed into the M5 *sigK::fdh* and the 824 Δ *sigK* strains, and the following procedure was used to stimulate the FLP activity. First, a single colony was grown overnight in 10 mL of CGM supplemented with 100 μ g/mL of Em. One mL of this culture was used to inoculate a new 10 mL tube of CGM which was incubated in an anaerobic chamber for 6 hrs at 30 °C to ensure optimal FLP activity (7). Serial dilutions were plated onto solid 2X

10 YTG media containing 40 μ g/ml Em and incubated at 37 °C until colonies were observed. Subsequently, a replica of the plate was made onto 2X YTG plates supplemented with 5 μ g/ml Th. Colonies that did not exhibit growth on the Th supplemented plate were screened to confirm the excision of the Th^R gene by PCR confirmation. For the *fdh* integration mutant, 3 of the 14 colonies screened had the Th^R excised while all 14 colonies

15 screened of the *sigK* deletion mutant had lost the Th^R gene. The resulting unmarked *sigK* deletion mutant was named 824 Δ *sigK*_um (Fig. 5B) and the unmarked integration mutant was renamed M5 *sigK::fdh*_um (Fig. 8B). To stimulate the loss of the p94FLP_asRepL plasmid, a single colony was grown overnight in liquid CGM. A 5% inoculum from the overnight culture was used to inoculate two fresh 10 mL CGM tubes

20 simultaneously (with and without lactose) and allowed to grow for 8-12 hours. Lactose was added to induce the expression of the asRNA fragment to stimulate plasmid loss. This process was repeated a total of six times. One hundred microliters was then plated onto solid 2X YTG plates with Em (encoded on the plasmid backbone) and allowed to grow for 48 hrs to confirm loss of the plasmid. Following incubation, the plates were visually

25 inspected for growth. The culture that was not induced with lactose showed growth (indicating the presence of the plasmid) while that was induced with lactose did not (indicating plasmid loss) (Fig 7D). Thus, the strategy of inducing segregational instability by inducing an asRNA against the *repL* origin of replication is sound and effective.

Isolation of *sigF* and *sigK* deletion mutants with a KO plasmid lacking the

30 **resolvase (*recU*) gene.** Previous work in our group aimed to enhance the homologous recombination machinery by expressing the *B. subtilis* resolvase gene (*recU*). Resolvases are a class of proteins responsible for Holliday-junction resolution during homologous recombination. However, these attempts yielded mutants that were only single-crossover integrants similar to what others have reported without the expression of *recU*. Thus far,

35 all our KO/KI plasmids were designed to express the *recU* gene. To test whether *recU* was dispensable in this counter-selection strategy, *sigF* and *sigK* deletion plasmid were designed without including the *recU* gene. The *sigF* deletion mutant was designed to remove 521 bp of the 759 bp *sigF* ORF (Fig. 9A). Upon transforming *C. acetobutylicum*

~ 24 ~

cells with the pKOSIGF2_mazF plasmid as described above, 14 colonies were immediately isolated after transformation and outgrowth from the Th+lactose plates. PCR confirmation revealed that 8 of those colonies had undergone the desired gene replacement via allelic exchange to generate an in-frame deletion of the *sigF* gene (Fig. 9B). No colonies were observed on the Th + lactose plates upon transforming plasmid pKOSIGK2_mazF that is designed to generate a *sigK* in-frame deletion mutant as described above, albeit without the expression of *recU*. Thus, the *sigK* deletion mutant was generated by utilizing a liquid transfer protocol described in the material and methods section (Fig. 10). In conclusion, the *sigF* mutant was immediately isolated without using the extended replica plating protocol described above in the absence of *recU*. Likewise, the *sigK* deletion mutant was isolated without the expression of *recU* on the vector backbone.

DISCUSSION

The lack of a generally applicable counter-selection marker for use in *Clostridium*, without the need to rely on auxotrophic strains, has been described as a major limitation for efficiently and readily isolating allelic exchange mutants. Unlike yeast or other naturally competent prokaryotes like *B. subtilis*, which can be transformed with linear DNA or suicide plasmids, respectively, isolating rare homologous recombination mutants is a major obstacle due to difficulty in identifying desirable gene replacement mutants via allelic exchange, low transformation efficiency in clostridia and the poor recombination systems of most *Clostridium* organisms. However, these deficiencies can be overcome if a suitable counter-selection marker is placed on the vector backbone to select against cells bearing the plasmid, enabling the isolation of gene replacement mutants. Nonetheless, identifying suitable counter-selection markers and optimizing the conditions for their use is a challenging and laborious task.

Here, the development of an easy-to-implement counter-selection method based on an inducible *mazF* gene for isolating rare allelic exchange chromosomal insertions and deletions in *C. acetobutylicum* is reported. The work flow and the estimated effort involved in designing and isolating a desired mutant is shown in Figure 11. This method does not resort to generating auxotrophic strains nor does it limit the isolation of integrants in certain genomic loci. Furthermore, the method developed here can be utilized to study the conditions that can potentially enhance native homologous recombination in the host organism (e.g., culture age, temperature, etc.). It is anticipated that this new method will prove invaluable for easily and efficiently creating mutant strains for studying gene function or for integrating DNA sequences that confer desirable traits in *Clostridium* spp.

Unlike the group II intron system, where target sequences smaller than 400 bp may not contain optimal group II intron target sites, it is shown here that specific DNA

~ 25 ~

sequences that are smaller than 300 bp can be targeted and that homologous recombination mutants can be readily isolated.

Due to the good stability of the *repL* origin of replication and in order to swiftly cure the cells of the FLP-recombinase bearing plasmid, an inducible segregational
 5 instability element was utilized. Antisense RNA (asRNA) against target transcript(s) has been successfully utilized to downregulate the expression of a number of genes in *Clostridium*. The asRNA is a single-stranded RNA molecule that is complementary to the target gene's synthesized mRNA molecule. Therefore, upon transcription of the asRNA, it binds to the target mRNA molecule and inhibits its translation. To exploit this feature for
 10 curing the plasmid, a lactose inducible asRNA fragment against the *repL* transcript was designed and implemented to promote inducible segregational instability of the replicon. Consequently, this system enables a fast turnaround time if the desired mutant is required for further genetic manipulations.

Although this example focused on *C. acetobutylicum*, this method could be adapted
 15 to most, if not all, *Clostridium* organisms and other organisms with underdeveloped genetic tools. Potentially, the method described herein may prove invaluable for Clostridia geneticists aiming to generate chromosomal fusion proteins or for creating whole genome KO libraries in *Clostridium* spp. enabling a functional genomics approach in studying these important and poorly understood organisms.

20 **Example 2. Isolation of in-frame markerless deletions**

The newly developed method can be used to KO genes in a markerless fashion (i.e. no antibiotic marker flanking the two regions of homology). The KO plasmid for the gene of interest (GOI) is transformed into host cells and transformants are initially selected for on Th plates. Once the plasmid is confirmed, the cells can be vegetatively transferred as
 25 described above. Mutants with the desired in-frame KO are screened for on lactose plates without any antibiotic. Colonies that are resistant to lactose are selected and confirmed. A schematic for this technique is shown in Figure 12.

Example 3. Gene knock-in (KI)

A gene or DNA of interest (*gene A* or DNA A), flanked by two regions of homology,
 30 can be KI in the chromosome of a host cell. Essentially, the antibiotic resistance gene is replaced by a gene (e.g., *gene A*) or any DNA (e.g., DNA A) to be delivered into the chromosome of the host cell (Fig. 13).

Example 4. Multiple sequential KO (or KIs) in either a marked or markerless fashion

35 Multiple sequential gene deletions or gene insertions can be accomplished using the methods described above. For example, a first gene may be knocked out as described in examples 1-3. Once the first KO mutant is confirmed (e.g., for gene A), another target (gene B) may be targeted for deletion as described above. This way, the presently

~ 26 ~

disclosed screening method utilizing *mazF* may be utilized for multiple sequential KOs. The genes can be KO either in a marked fashion or in a markerless fashion (Fig. 12). Alternatively, sequential gene KI may be accomplished in the same fashion (Fig. 13). Similarly, the present method may be adopted in sequential gene deletion and gene
 5 insertion.

Example 5. Multiple antibiotic selection markers, flanked by the target gene's regions of homology, can be employed

Multiple antibiotic markers can be employed. Erythromycin (Em) and tetracycline (Tc) can be used instead of thiamphenicol. The target regions of homology can then be
 10 designed to flank any of the aforementioned antibiotic resistance genes. The location of each marker can be swapped in the plasmid backbone depending on the user's desire and application. For example, Em can be flanked by the regions of homology and Th can be placed on the plasmid backbone. After a successful KO, Em will be integrated into the chromosome of the target cell while Th will be lost along with the plasmid backbone.
 15 Alternatively, Tc can be used instead of Em and so forth.

All documents, books, manuals, papers, patents, published patent applications, guides, abstracts, and/or other references cited herein are incorporated by reference in their entirety. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein.

20 It is intended that the specification and examples be considered as exemplary only, with the true scope and spirit of the invention being indicated by the following claims.

References

1. Awad, M. M., A. E. Bryant, D. L. Stevens, and J. I. Rood. 1995. Virulence Studies on Chromosomal Alpha-Toxin and Theta-Toxin Mutants Constructed by Allelic
 25 Exchange Provide Genetic-Evidence for the Essential Role of Alpha-Toxin in *Clostridium Perfringens*-Mediated Gas-Gangrene. *Molecular microbiology* 15:191-202.
2. Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2.
- 30 3. Bi, C. H., S. W. Jones, D. R. Hess, B. P. Tracy, and E. T. Papoutsakis. 2011. SpoIIE Is Necessary for Asymmetric Division, Sporulation, and Expression of sigma(F), sigma(E), and sigma(G) but Does Not Control Solvent Production in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* 193:5130-5137.
- 35 4. Boeke, J. D., J. Trueheart, G. Natsoulis, and G. R. Fink. 1987. 5-Fluoroorotic Acid as a Selective Agent in Yeast Molecular-Genetics. *Methods in enzymology* 154:164-175.
5. Brüggemann, H., S. Baumer, W. F. Fricke, A. Wiezer, H. Liesegang, I. Decker, C. Herzberg, R. Martinez-Arias, R. Merkl, A. Henne, and G. Gottschalk. 2003. The
 40 genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proceedings of the National Academy of Sciences of the United States of America* 100:1316-1321.
6. Brüggemann, H., and G. Gottschalk. 2009. *Clostridia: Molecular Biology in the Post-Genomic Era*. Caister Academic Press.

~ 27 ~

7. Buchholz, F., L. Ringrose, P. O. Angrand, F. Rossi, and A. F. Stewart. 1996. Different thermostabilities of FLP and Cre recombinases: Implications for applied site-specific recombination. *Nucleic acids research* 24:4256-4262.
8. Cartman, S. T., M. L. Kelly, D. Heeg, J. T. Heap, and N. P. Minton. 2012. Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association between the *tcdC* Genotype and Toxin Production. *Appl Environ Microb* 78:4683-4690.
9. Cornillot, E., R. V. Nair, E. T. Papoutsakis, and P. Soucaille. 1997. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *J Bacteriol* 179:5442-5447.
10. Desai, R. P., and E. T. Papoutsakis. 1999. Antisense RNA strategies for metabolic engineering of *Clostridium acetobutylicum*. *Appl Environ Microb* 65:936-945.
11. Gietz, D., A. St Jean, R. A. Woods, and R. H. Schiestl. 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic acids research* 20:1425.
12. Green, E. M., and G. N. Bennett. 1996. Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *Appl Biochem Biotech* 57-8:213-221.
13. Green, E. M., Z. L. Boynton, L. M. Harris, F. B. Rudolph, E. T. Papoutsakis, and G. N. Bennett. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiol-Uk* 142:2079-2086.
14. Harris, L. M., N. E. Welker, and E. T. Papoutsakis. 2002. Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *Journal of Bacteriology* 184:3586-3597.
15. Hartman, A. H., H. Liu, and S. B. Melville. 2011. Construction and characterization of a lactose-inducible promoter system for controlled gene expression in *Clostridium perfringens*. *Appl Environ Microb* 77:471-478.
16. Heap, J. T., M. Ehsaan, C. M. Cooksley, Y. K. Ng, S. T. Cartman, K. Winzer, and N. P. Minton. 2012. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic acids research* 40:e59.
17. Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Carter, and N. P. Minton. 2007. The ClosTron: A universal gene knock-out system for the genus *Clostridium*. *J Microbiol Meth* 70:452-464.
18. Heap, J. T., O. J. Pennington, S. T. Cartman, and N. P. Minton. 2009. A modular system for *Clostridium* shuttle plasmids. *J Microbiol Meth* 78:79-85.
19. Higuchi, R., B. Krummel, and R. K. Saiki. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic acids research* 16:7351-7367.
20. Itaya, M., and T. Tanaka. 1990. Gene-directed mutagenesis on the chromosome of *Bacillus subtilis* 168. *Molecular & general genetics* : MGG 223:268-272.
21. Jones, S. W., B. P. Tracy, S. M. Galda, and E. T. Papoutsakis. 2011. Inactivation of *sigma(F)* in *Clostridium acetobutylicum* ATCC 824 Blocks Sporulation Prior to Asymmetric Division and Abolishes *sigma(E)* and *sigma(G)* Protein Expression but Does Not Block Solvent Formation. *J Bacteriol* 193:2429-2440.
22. Kennedy, C. L., E. O. Krejany, L. F. Young, J. R. O'Connor, M. M. Awad, R. L. Boyd, J. J. Emmins, D. Lyras, and J. I. Rood. 2005. The alpha-toxin of *Clostridium septicum* is essential for virulence. *Molecular microbiology* 57:1357-1366.
23. Khan, S. A. 1997. Rolling-circle replication of bacterial plasmids. *Microbiol Mol Biol R* 61:442-8.
24. Klein, J. R., Y. Q. Chen, D. A. Manias, J. Zhuo, L. Zhou, C. L. Peebles, and G. M. Dunny. 2004. A conjugation-based system for genetic analysis of group II intron splicing in *Lactococcus lactis*. *J Bacteriol* 186:1991-1998.
25. Klein, J. R., and G. M. Dunny. 2002. Bacterial group II introns and their association with mobile genetic elements. *Front Biosci* 7:d1843-1856.

~ 28 ~

26. Lee, S. Y., L. D. Mermelstein, G. N. Bennett, and E. T. Papoutsakis. 1992. Vector construction, transformation, and gene amplification in *Clostridium acetobutylicum* ATCC 824. *Ann N Y Acad Sci* 665:39-51.
- 5 27. Lee, S. Y., L. D. Mermelstein, and E. T. Papoutsakis. 1993. Determination of plasmid copy number and stability in *Clostridium acetobutylicum* ATCC 824. *FEMS Microbiol Lett* 108:319-323.
28. Mermelstein, L. D., and E. T. Papoutsakis. 1993. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol* 59:1077-1081.
- 10 29. Mermelstein, L. D., N. E. Welker, G. N. Bennett, and E. T. Papoutsakis. 1992. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Biotechnology (N Y)* 10:190-195.
30. Minton, N. P., M. L. Mauchline, M. J. Lemmon, J. K. Brehm, M. Fox, N. P. Michael, A. Giaccia, and J. M. Brown. 1995. Chemotherapeutic Tumor Targeting Using Clostridial Spores. *Fems Microbiology Reviews* 17:357-364.
- 15 31. Nair, R. V., and E. T. Papoutsakis. 1994. Expression of Plasmid-Encoded Aad in *Clostridium-Acetobutylicum* M5 Restores Vigorous Butanol Production. *J Bacteriol* 176:5843-5846.
- 20 32. Nariya, H., S. Miyata, M. Suzuki, E. Tamai, and A. Okabe. 2011. Development and Application of a Method for Counterselectable In-Frame Deletion in *Clostridium perfringens*. *Appl Environ Microb* 77:1375-1382.
33. Nelson, R. T., B. A. Pryor, and J. K. Lodge. 2003. Sequence length required for homologous recombination in *Cryptococcus neoformans*. *Fungal Genet Biol* 38:1-9.
- 25 34. Nickoloff, J. A. 1995. Electroporation protocols for microorganisms. Humana Press, Totowa, N.J.
- 35 35. O'Connor, J. R., D. Lyras, K. A. Farrow, V. Adams, D. R. Powell, J. Hinds, J. K. Cheung, and J. I. Rood. 2006. Construction and analysis of chromosomal *Clostridium difficile* mutants. *Molecular microbiology* 61:1335-1351.
- 30 36. Pan, G., K. Luetke, and P. D. Sadowski. 1993. Mechanism of cleavage and ligation by FLP recombinase: classification of mutations in FLP protein by in vitro complementation analysis. *Mol Cell Biol* 13:3167-3175.
37. Papadopoulou, B., and C. Dumas. 1997. Parameters controlling the rate of gene targeting frequency in the protozoan parasite *Leishmania*. *Nucleic acids research* 25:4278-4286.
- 35 38. Papoutsakis, E. T. 2008. Engineering solventogenic clostridia. *Curr Opin Biotechnol* 19:420-429.
39. Plante, I., and B. Cousineau. 2006. Restriction for gene insertion within the *Lactococcus lactis* LI.LtrB group II intron. *RNA* 12:1980-1992.
- 40 40. Raju, D., P. Setlow, and M. R. Sarker. 2007. Antisense-RNA-mediated decreased synthesis of small, acid-soluble spore proteins leads to decreased resistance of *Clostridium perfringens* spores to moist heat and UV radiation. *Appl Environ Microb* 73:2048-2053.
41. Reytrat, J. M., V. Pelicic, B. Gicquel, and R. Rappuoli. 1998. Counterselectable markers: Untapped tools for bacterial genetics and pathogenesis. *Infect Immun* 66:4011-4017.
- 45 42. Roberts, A. P., C. Hennequin, M. Elmore, A. Collignon, T. Karjalainen, N. Minton, and P. Mullany. 2003. Development of an integrative vector for the expression of antisense RNA in *Clostridium difficile*. *J Microbiol Meth* 55:617-624.
- 50 43. Rubnitz, J., and S. Subramani. 1984. The Minimum Amount of Homology Required for Homologous Recombination in Mammalian-Cells. *Mol Cell Biol* 4:2253-2258.
44. Schlake, T., and J. Bode. 1994. Use of Mutated FLP Recognition Target (Frt) Sites for the Exchange of Expression Cassettes at Defined Chromosomal Loci. *Biochemistry-Us* 33:12746-12751.

~ 29 ~

45. Shao, L., S. Hu, Y. Yang, Y. Gu, J. Chen, W. Jiang, and S. Yang. 2007. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. *Cell Research* 17:963-965.
- 5 46. Sillers, R., A. Chow, B. Tracy, and E. T. Papoutsakis. 2008. Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance. *Metab Eng* 10:321-332.
47. Steiner, E., A. E. Dago, D. I. Young, J. T. Heap, N. P. Minton, J. A. Hoch, and M. Young. 2011. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Molecular microbiology* 80:641-654.
- 10 48. Tracy, B. P., S. W. Jones, A. G. Fast, D. C. Indurthi, and E. T. Papoutsakis. 2012. Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. *Curr Opin Biotechnol* 23:364-381.
- 15 49. Tracy, B. P., S. W. Jones, and E. T. Papoutsakis. 2011. Inactivation of sigma(E) and sigma(G) in *Clostridium acetobutylicum* Illuminates Their Roles in Clostridial-Cell-Form Biogenesis, Granulose Synthesis, Solventogenesis, and Spore Morphogenesis. *J Bacteriol* 193:1414-1426.
50. Tripathi, S. A., D. G. Olson, D. A. Argyros, B. B. Miller, T. F. Barrett, D. M. Murphy, J. D. McCool, A. K. Warner, V. B. Rajgarhia, L. R. Lynd, D. A. Hogsett, and N. C. Caiazza. 2010. Development of pyrF-Based Genetic System for Targeted Gene Deletion in *Clostridium thermocellum* and Creation of a pta Mutant. *Appl Environ Microb* 76:6591-6599.
- 20 51. Tummala, S. B., S. G. Junne, and E. T. Papoutsakis. 2003. Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic *Clostridium acetobutylicum* fermentations. *J Bacteriol* 185:3644-3653.
- 25 52. Tummala, S. B., N. E. Welker, and E. T. Papoutsakis. 2003. Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. *J Bacteriol* 185:1923-1934.
- 30 53. Tummala, S. B., N. E. Welker, and E. T. Papoutsakis. 1999. Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol* 65:3793-3799.
54. Webb, C. D., P. L. Graumann, J. A. Kahana, A. A. Teleman, P. A. Silver, and R. Losick. 1998. Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28:883-892.
- 35 55. Wilkinson, S. R., and M. Young. 1994. Targeted Integration of Genes into the *Clostridium-Acetobutylicum* Chromosome. *Microbiol-Uk* 140:89-95.
- 40 56. Yamaguchi, Y., and M. Inouye. 2009. mRNA interferases, sequence-specific endoribonucleases from the toxin-antitoxin systems. *Progress in molecular biology and translational science* 85:467-500.
57. Yao, J., J. Zhong, Y. Fang, E. Geisinger, R. P. Novick, and A. M. Lambowitz. 2006. Use of targetrons to disrupt essential and nonessential genes in *Staphylococcus aureus* reveals temperature sensitivity of Ll.LtrB group II intron splicing. *Rna-a Publication of the Rna Society* 12:1271-1281.
- 45 58. Zhang, X. Z., X. Yan, Z. L. Cui, Q. Hong, and S. P. Li. 2006. mazF, a novel counter-selectable marker for unmarked chromosomal manipulation in *Bacillus subtilis*. *Nucleic acids research* 34:e71.
- 50 59. Zhang, Y. L., M. R. Yu, and S. T. Yang. 2012. Effects of ptb knockout on butyric acid fermentation by *Clostridium tyrobutyricum*. *Biotechnol Progr* 28:52-59.
60. Zhong, J., M. Karberg, and A. M. Lambowitz. 2003. Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic acids research* 31:1656-1664.
- 55

~ 30 ~

61. Zhu, X. D., and P. D. Sadowski. 1995. Cleavage-Dependent Ligation by the Flp Recombinase - Characterization of a Mutant Flp Protein with an Alteration in a Catalytic Amino-Acid. *J Biol Chem* 270:23044-23054.
- 5 62. Melville, S. B., R. Labbe, and A. L. Sonenshein. 1994. Expression from the *Clostridium perfringens* cpe promoter in *C. perfringens* and *Bacillus subtilis*. *Infect Immun* 62:5550-5558.
63. Mitchell, W. J., K. A. Albasheri, and M. Yazdanian. 1995. Factors Affecting Utilization of Carbohydrates by Clostridia. *Fems Microbiol Rev* 17:317-329.
- 10 64. Papoutsakis, E. T., S. W. Jones, B. P. Tracy, and S. M. Gaida. 2011. Inactivation of sigma(F) in *Clostridium acetobutylicum* ATCC 824 Blocks Sporulation Prior to Asymmetric Division and Abolishes sigma(E) and sigma(G) Protein Expression but Does Not Block Solvent Formation. *Journal of Bacteriology* 193:2429-2440.
- 15 65. Tracy, B. P., and E. T. Papoutsakis. 2010. Methods and compositions for genetically engineering clostridia species. U.S. Patent Application Publication No. 20100075424.

~ 31 ~

What

1. A *Clostridium* organism comprising a recombinant DNA, wherein the recombinant DNA comprises
 - (a) a first nucleic acid sequence encoding a first counter-selection marker,
 - (b) a first promoter operably linked to the first nucleic acid sequence, and
 - (c) a second nucleic acid sequence flanked by a first region at least 90% identical to an upstream region of a first target genomic locus in the recombinant *Clostridium* organism and a second region at least 90% identical to a downstream region of the first target genomic locus.
2. The recombinant *Clostridium* organism of claim 1, wherein the first counter-selection marker is a heterologous toxin.
3. The recombinant *Clostridium* organism of claim 1, wherein the first promoter is a heterologous inducible promoter.
4. The recombinant *Clostridium* organism of claim 1, wherein the *Clostridium* organism is selected from the group consisting of *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetanii*, *C. sporogenes*, *C. novyi*, *C. acetobutylicum*, *C. pasteurianum*, *C. beijerinckii*, *C. ljungdahlii*, *C. carboxidivorans*, *C. thermocellum*, *C. cellulolyticum*, *C. phytofermentans*, *C. tyrobutyricum*, *C. butyricum*, *C. kluyveri*, *C. sticklandii*, *Moorella thermoacetica*, *C. autoethanogenum*, *C. ragsdalei*, *Moorella thermoautotrophica*, *C. propionicum*, *C. thermobutyricum*, *C. thermoaceticum*, *C. formicoaceticum*, and *Butyrivacterium methylotrophicum*.
5. The recombinant *Clostridium* organism of claim 1, wherein the *Clostridium* organism is *C. acetobutylicum* ATCC 824.
6. The recombinant *Clostridium* organism of claim 1, wherein the second nucleic acid sequence is a heterologous nucleic acid sequence.
7. The recombinant *Clostridium* organism of claim 1, wherein the second nucleic acid sequence encodes a positive selection marker.
8. The recombinant *Clostridium* organism of claim 1, wherein the second nucleic acid sequence comprises a gene, a series of genes, a gene transcription promoter, an antisense RNA, or a combination thereof.
9. A method for generating a mutant of the recombinant *Clostridium* organism of claim 1, comprising growing the recombinant *Clostridium* organism, whereby a mutant having double-crossover allelic exchange of the first target genomic locus with the second nucleic acid sequence is generated.
10. The method of claim 9, further comprising isolating the mutant.
11. The method of claim 9, wherein the first target genomic locus comprises a first target gene, whereby the first target gene is disrupted in the mutant.

~ 32 ~

12. claim 9, wherein the first target genomic locus comprises a first target gene, wherein the first target gene is deleted in the mutant.

13. claim 9, wherein the second nucleic acid sequence is a heterologous nucleic acid sequence, whereby the heterologous nucleic acid sequence is integrated into the first target genomic locus in the mutant.

14. The method of claim 9, wherein the second nucleic acid sequence encodes a positive selection marker, whereby the positive selection marker is integrated into the mutant.

15. The method of claim 14, further comprising removing the positive selection marker from the mutant.

16. The method of claim 9, wherein the second nucleic acid sequence comprises a gene, a series of genes, a gene transcription promoter, an antisense RNA, or a combination thereof.

17. The method of claim 9, wherein the first counter-selection marker is a heterologous toxin.

18. The method of claim 9, wherein the first promoter is a heterologous inducible promoter.

19. The method of claim 9, wherein the *Clostridium* organism is selected from the group consisting of *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetanii*, *C. sporogenes*, *C. novyi*, *C. acetobutylicum*, *C. pasteurianum*, *C. beijerinckii*, *C. ljungdahlii*, *C. carboxidivorans*, *C. thermocellum*, *C. cellulolyticum*, *C. phytofermentans*, *C. tyrobutyricum*, *C. butyricum*, *C. kluyveri*, *C. sticklandii*, *Moorella thermoacetica*, *C. autoethanogenum*, *C. ragsdalei*, *Moorella thermoautotrophica*, *C. propionicum*, *C. thermobutylicum*, *C. thermoaceticum*, *C. formicoaceticum*, and *Butyribacterium methylotrophicum*.

20. The method of claim 9, wherein the *Clostridium* organism is *C. acetobutylicum* ATCC 824.

1/15

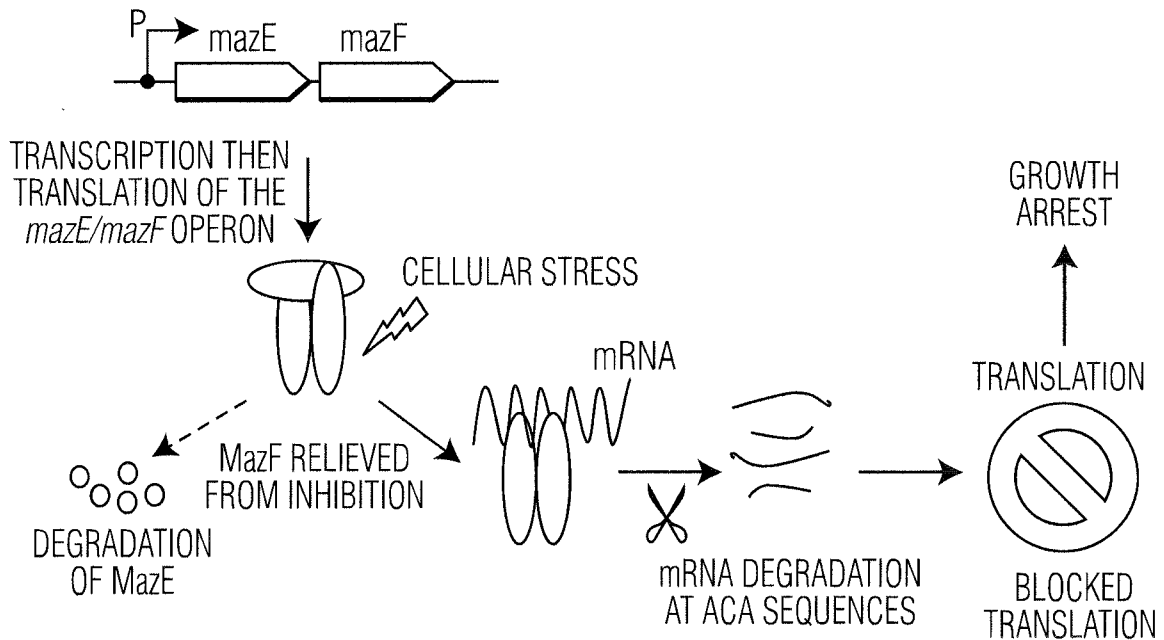


FIG. 1

OPTIMIZED *mazF* DNA SEQUENCE (SEQ ID NO: 1):

ATG GTA AGC AGG TAC GTA CCT GAT ATG GGA GAT TTA ATA TGG GTT GAT TTT GAT CCT
 ACA AAA GGA AGT GAG CAA GCT GGA CAT AGA CCA GCT GTT GTT TTA AGT CCT TTT ATG
 TAT AAT AAT AAA ACA GGA ATG TGT TTA TGT GTT CCT TGT ACA ACA CAA TCA AAA GGA
 TAT CCT TTT GAA GTT GTT TTA TCA GGA CAA GAA AGA GAT GGA GTA GCA TTA GCT GAT
 CAA GTA AAA AGT ATA GCA TGG AGA GCA AGA GGA GCA ACA AAA AAA GGA ACA GTT GCA
 CCA GAG GAA TTA CAA TTA ATA AAA GCA AAA ATA AAT GTA TTA ATA GGA TAG

FIG. 2

2/15

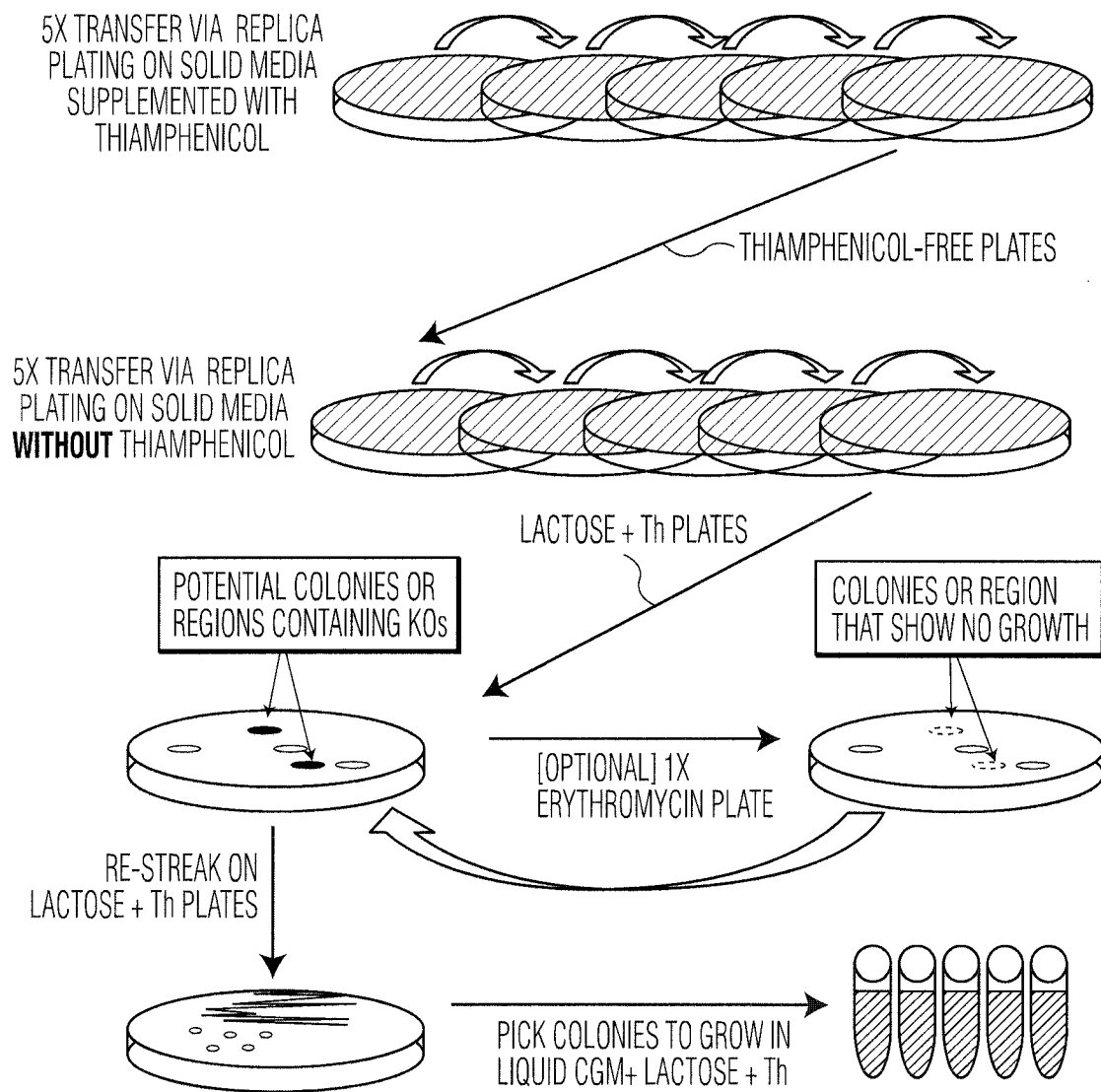


FIG. 3

3/15

bgaR FUSED UPSTREAM OF REPPORTER GENE ON PLASMID pAH2

FIG. 4A

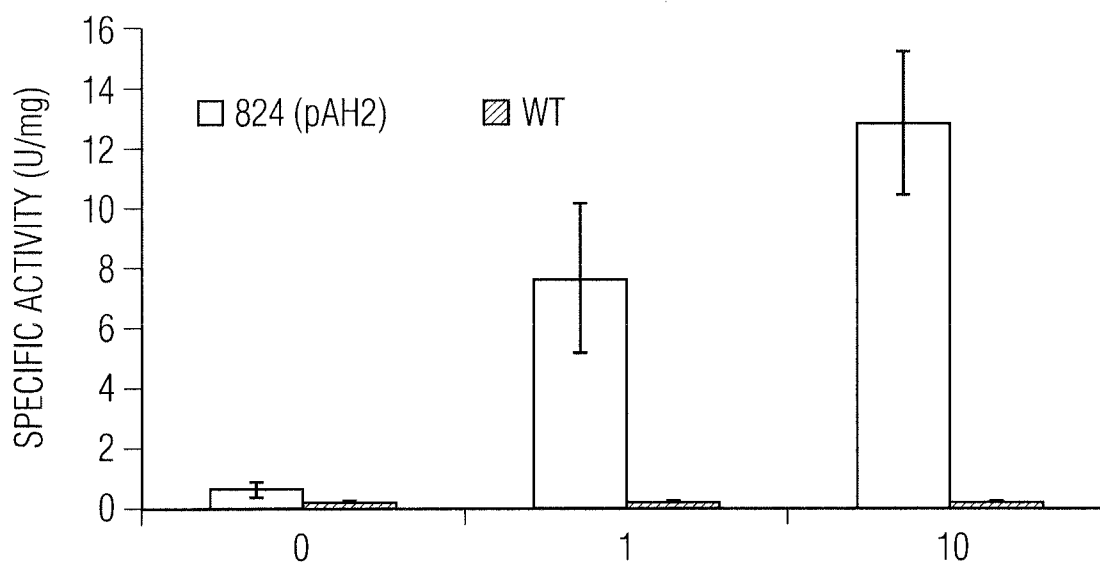


FIG. 4B

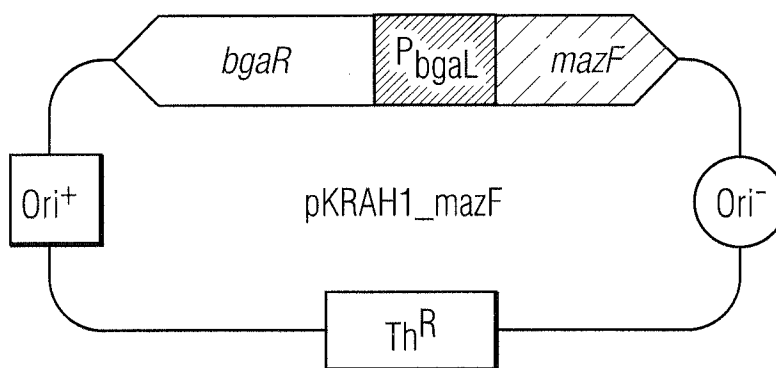


FIG. 4C

4/15

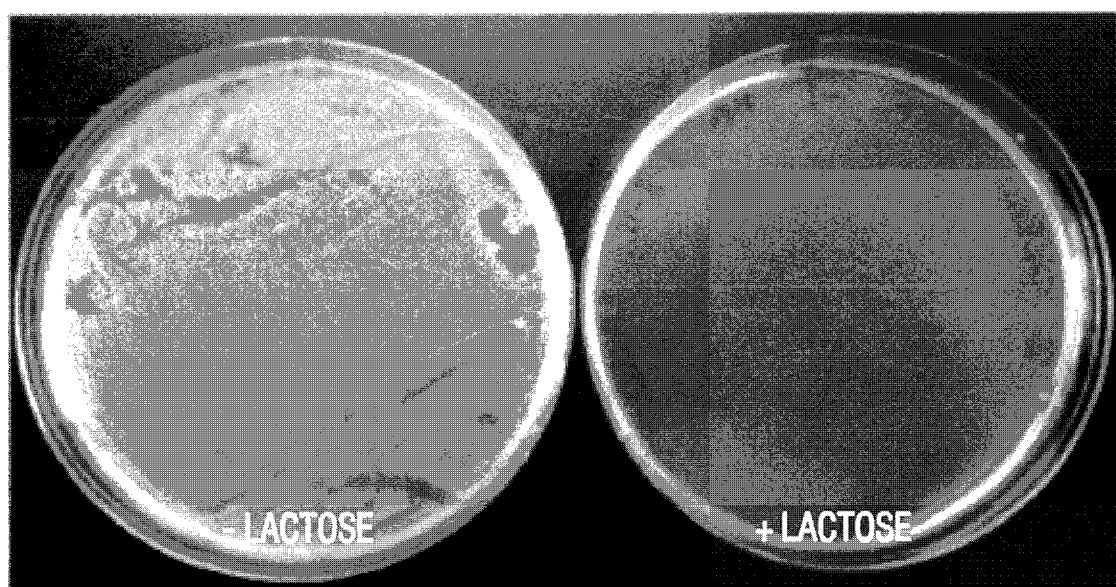


FIG. 4D

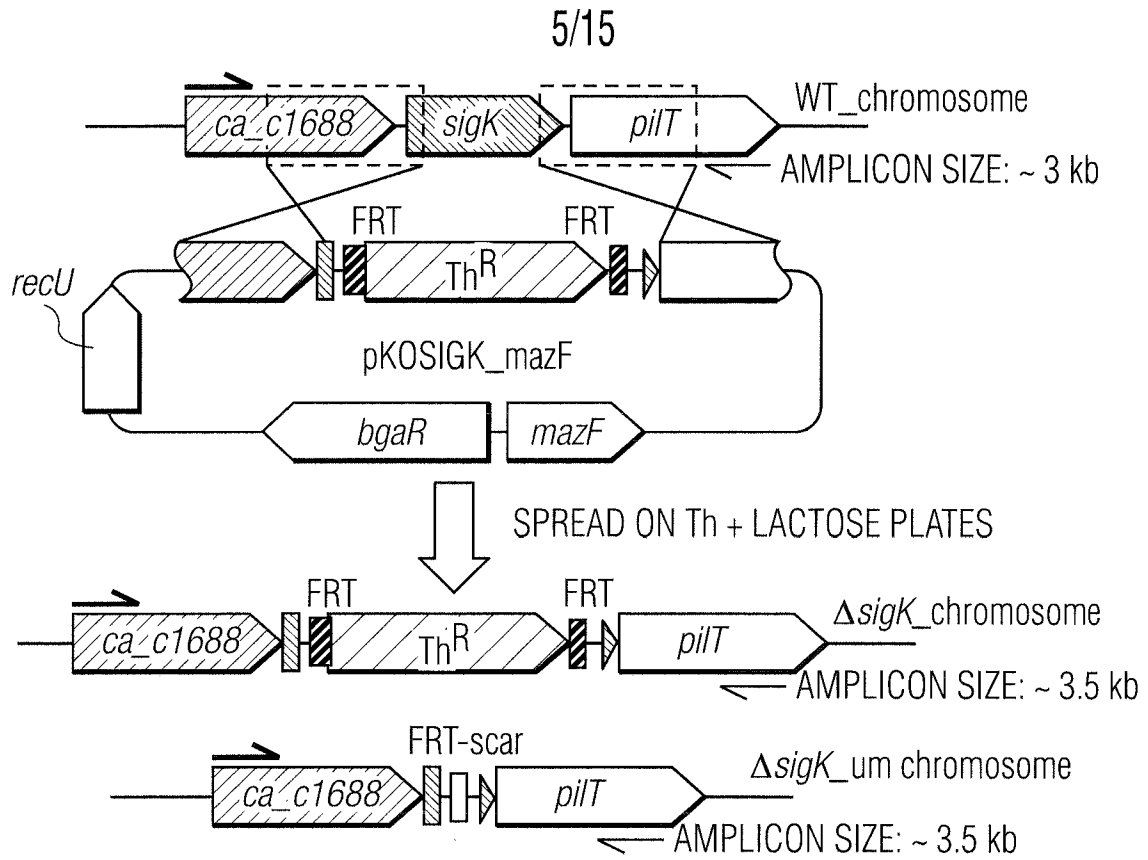


FIG. 5A

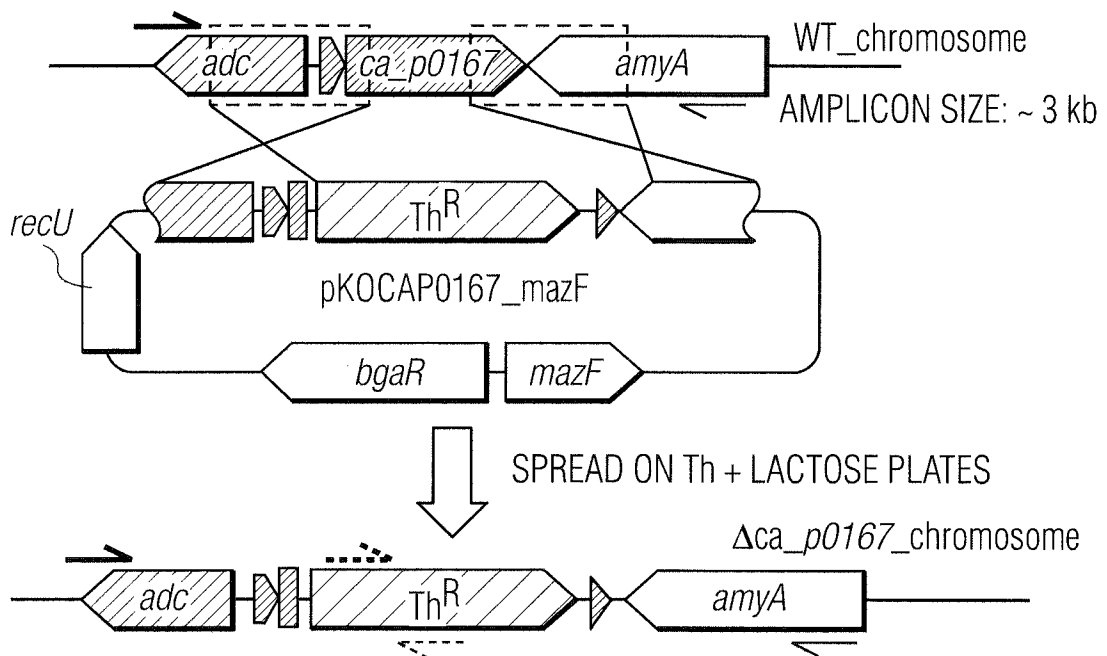


FIG. 5B

6/15

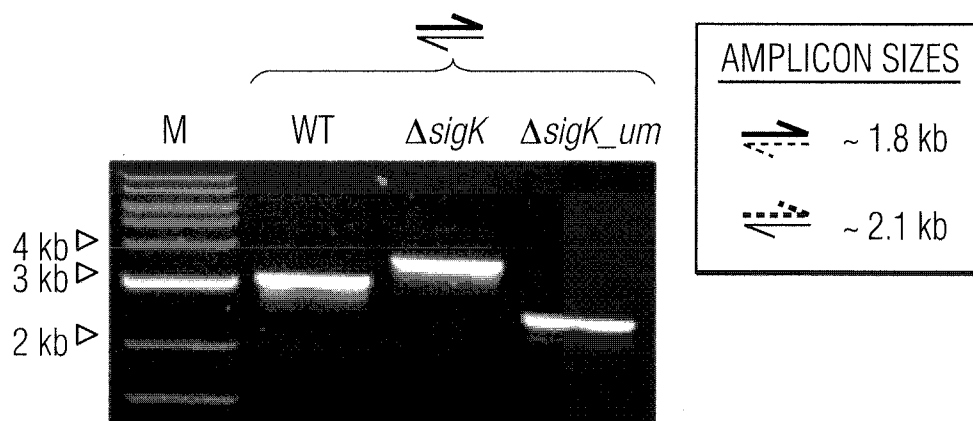


FIG. 5C

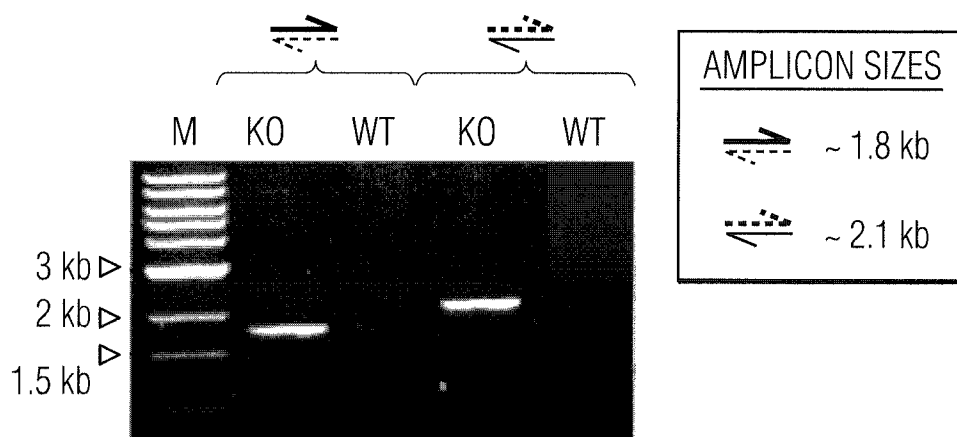


FIG. 5D

7/15

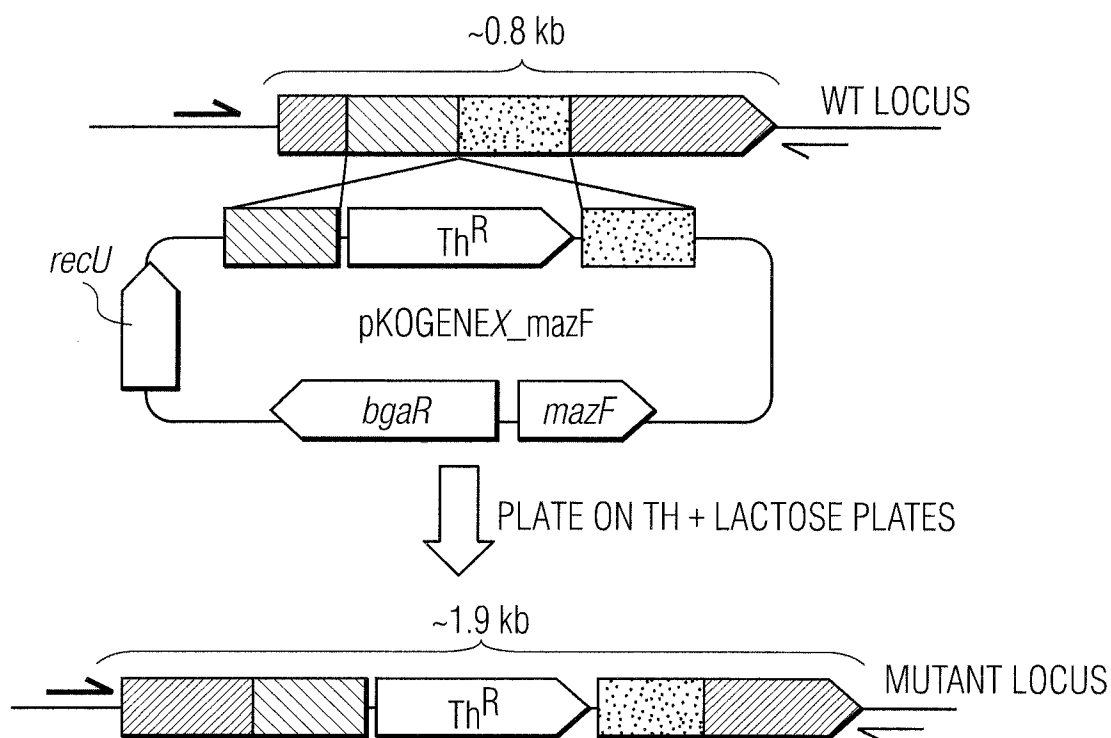


FIG. 6A

CONFIRMATION OF PUTATIVE MUTANTS WITH TARGET SPECIFIC PRIMERS

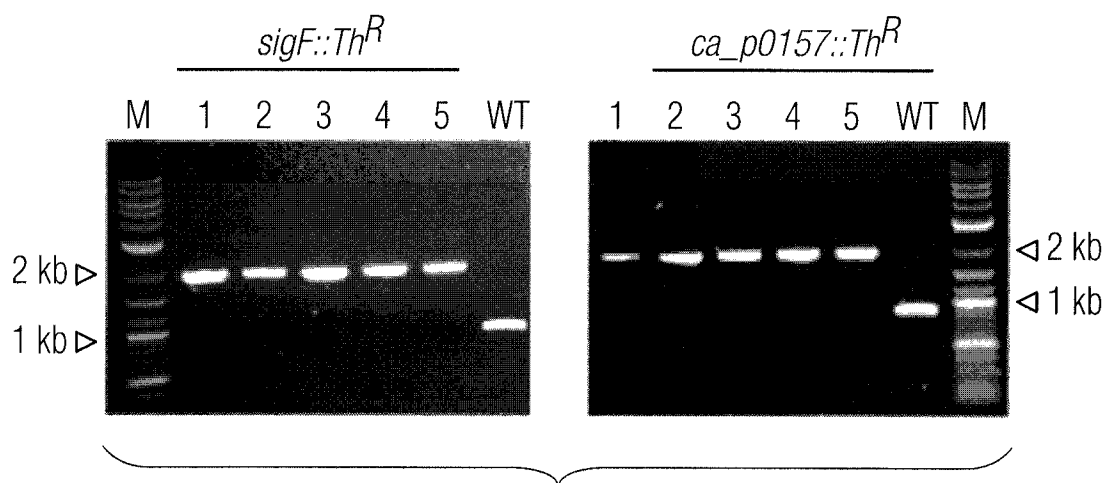


FIG. 6B

8/15

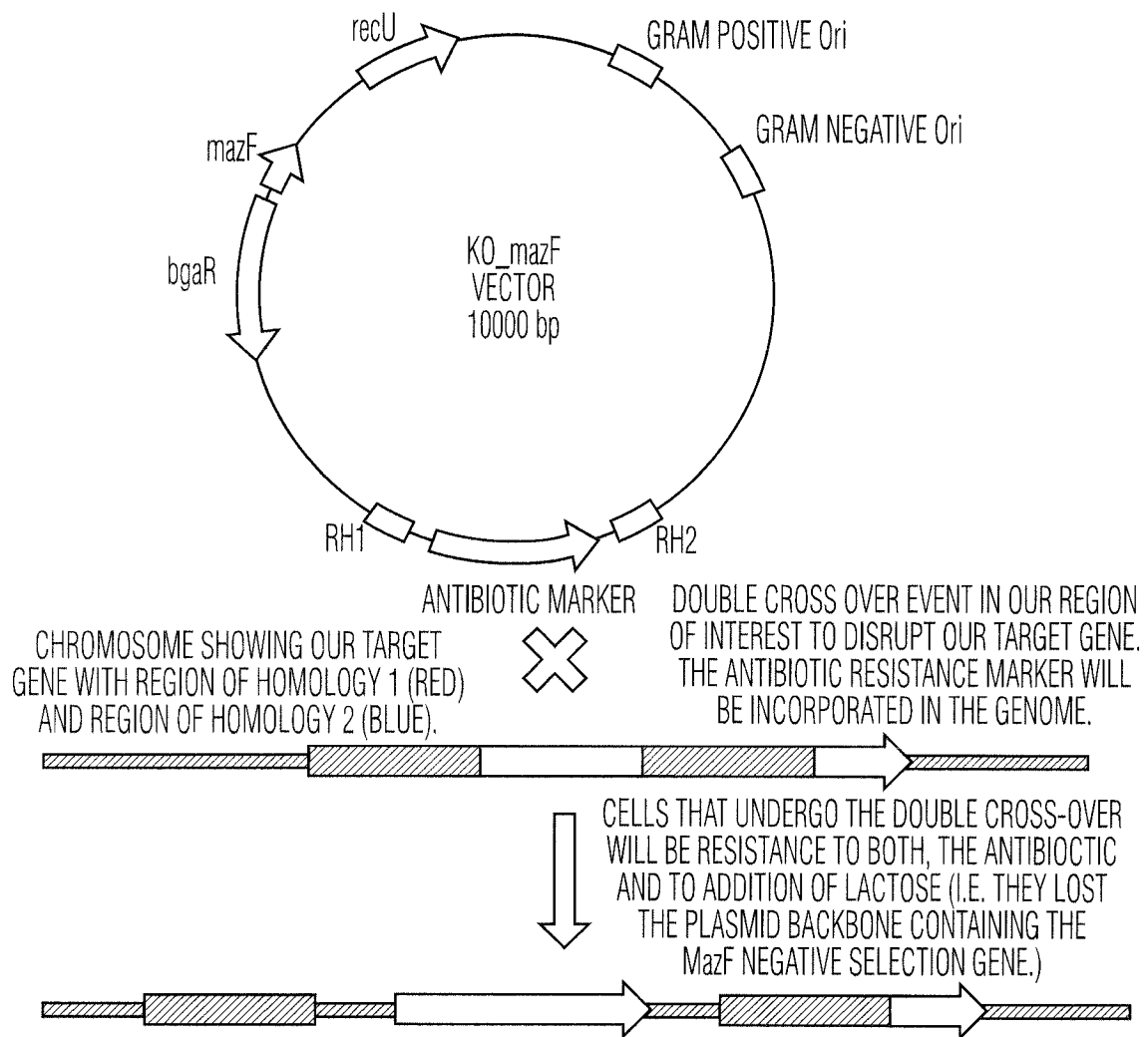


FIG. 7

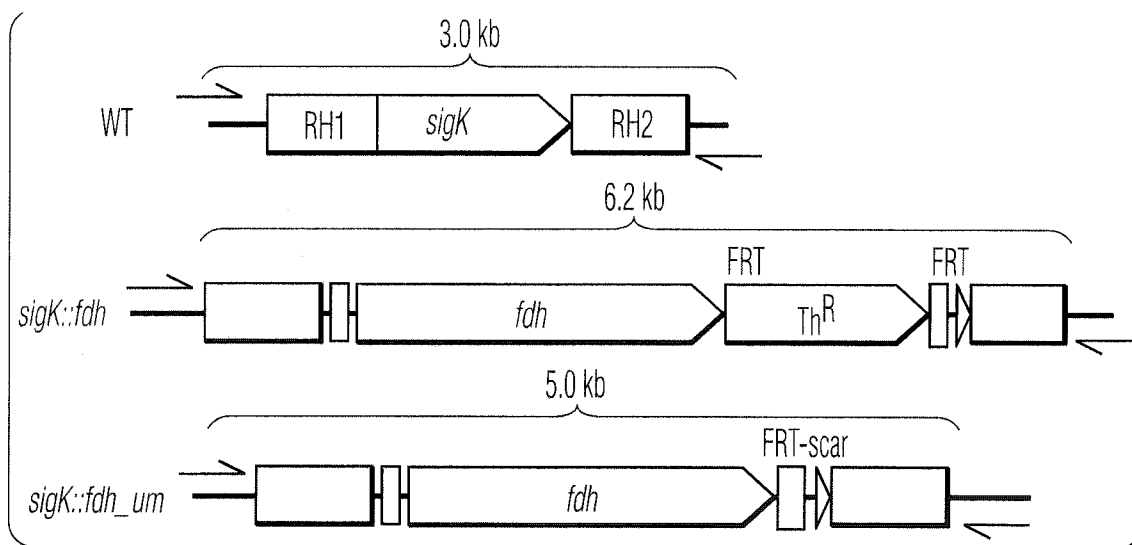


FIG. 8A

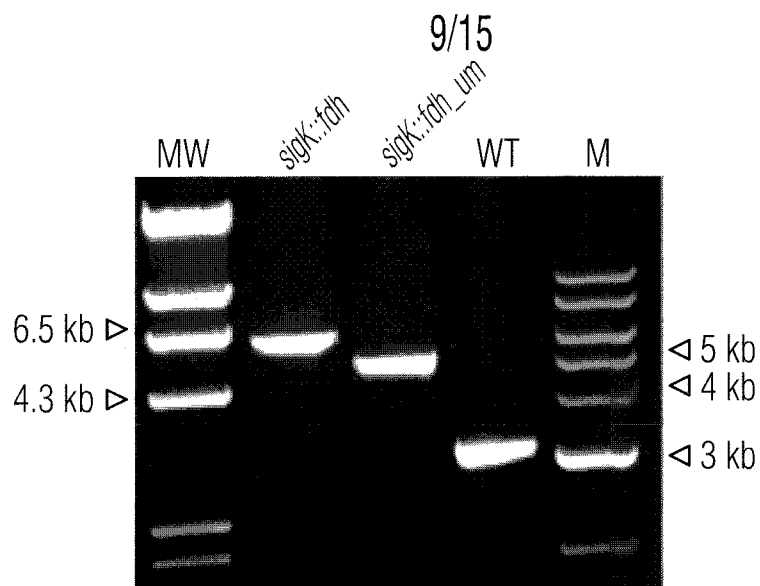


FIG. 8B

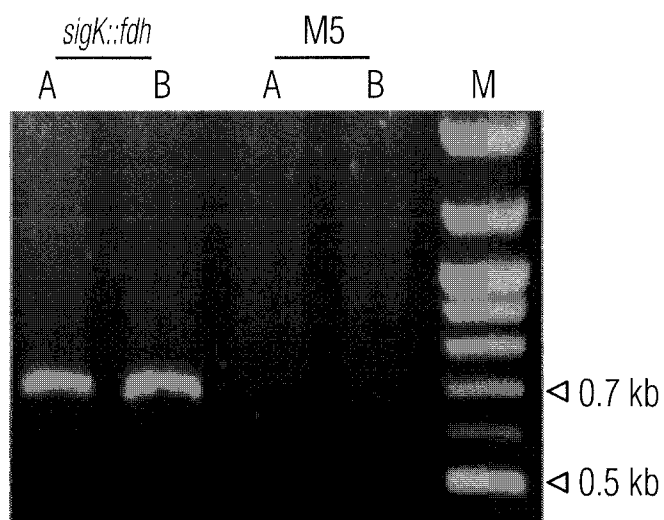


FIG. 8C

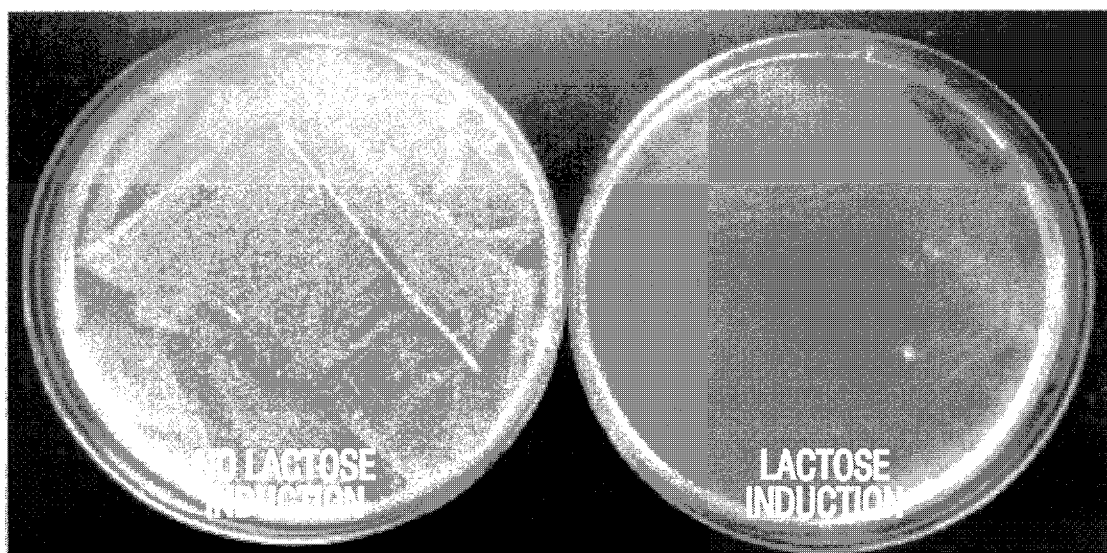


FIG. 8D

10/15

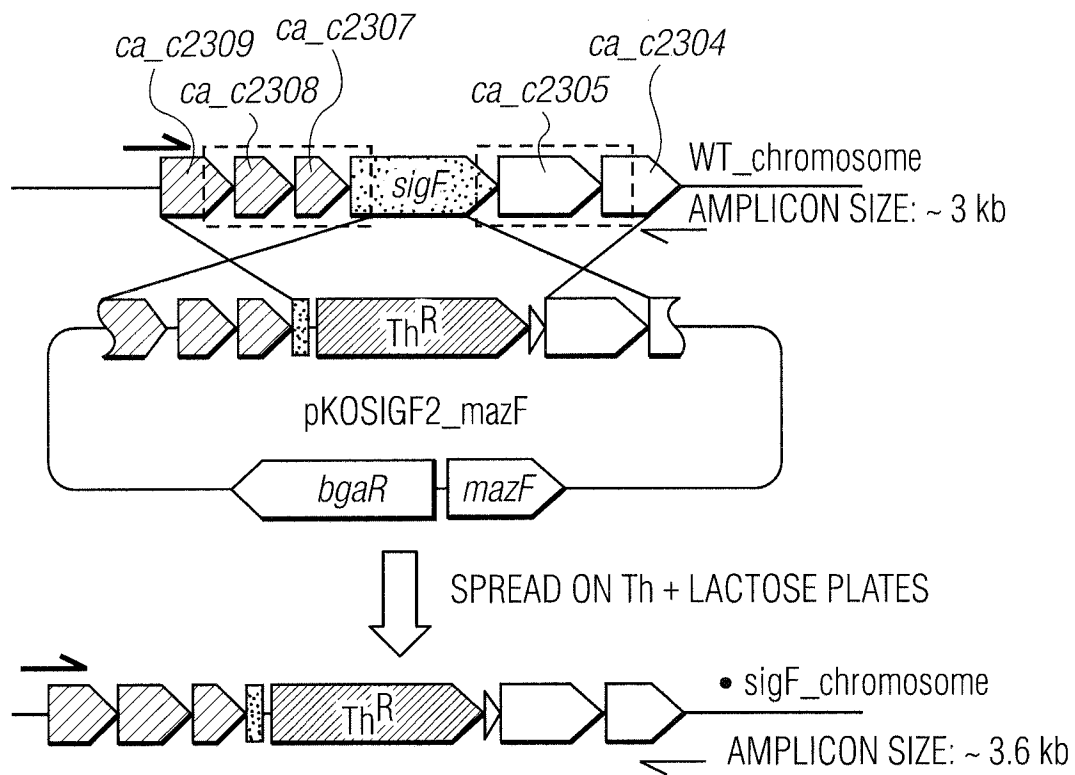


FIG. 9A

11/15

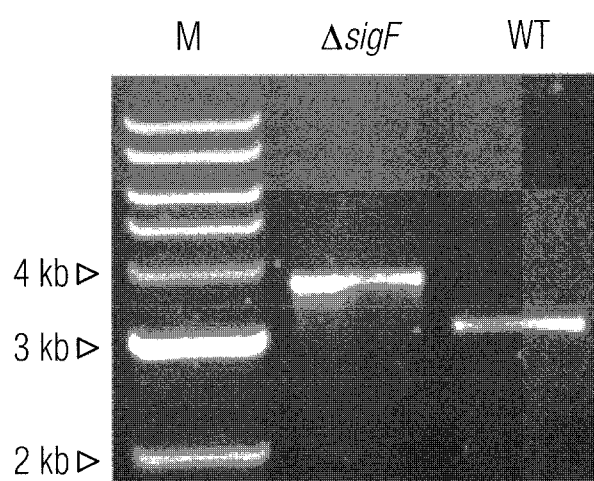


FIG. 9B

12/15

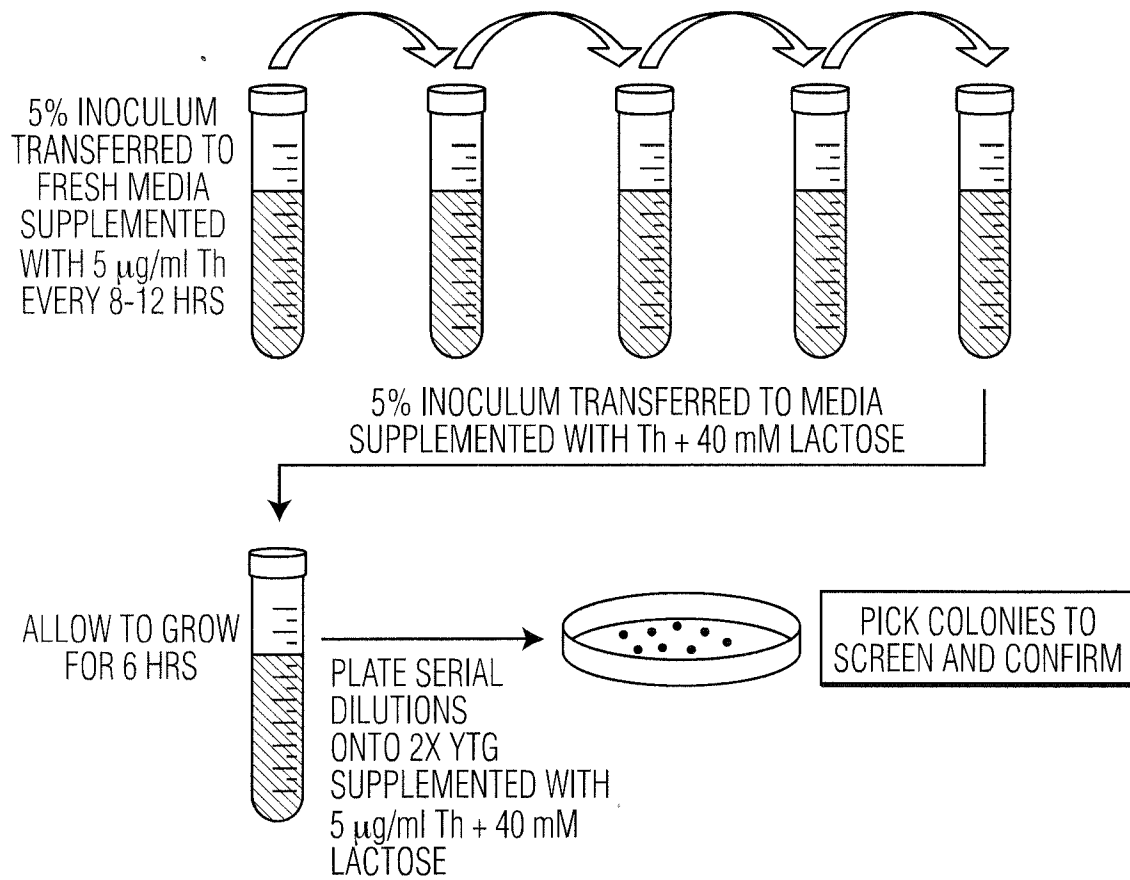


FIG. 10

13/15

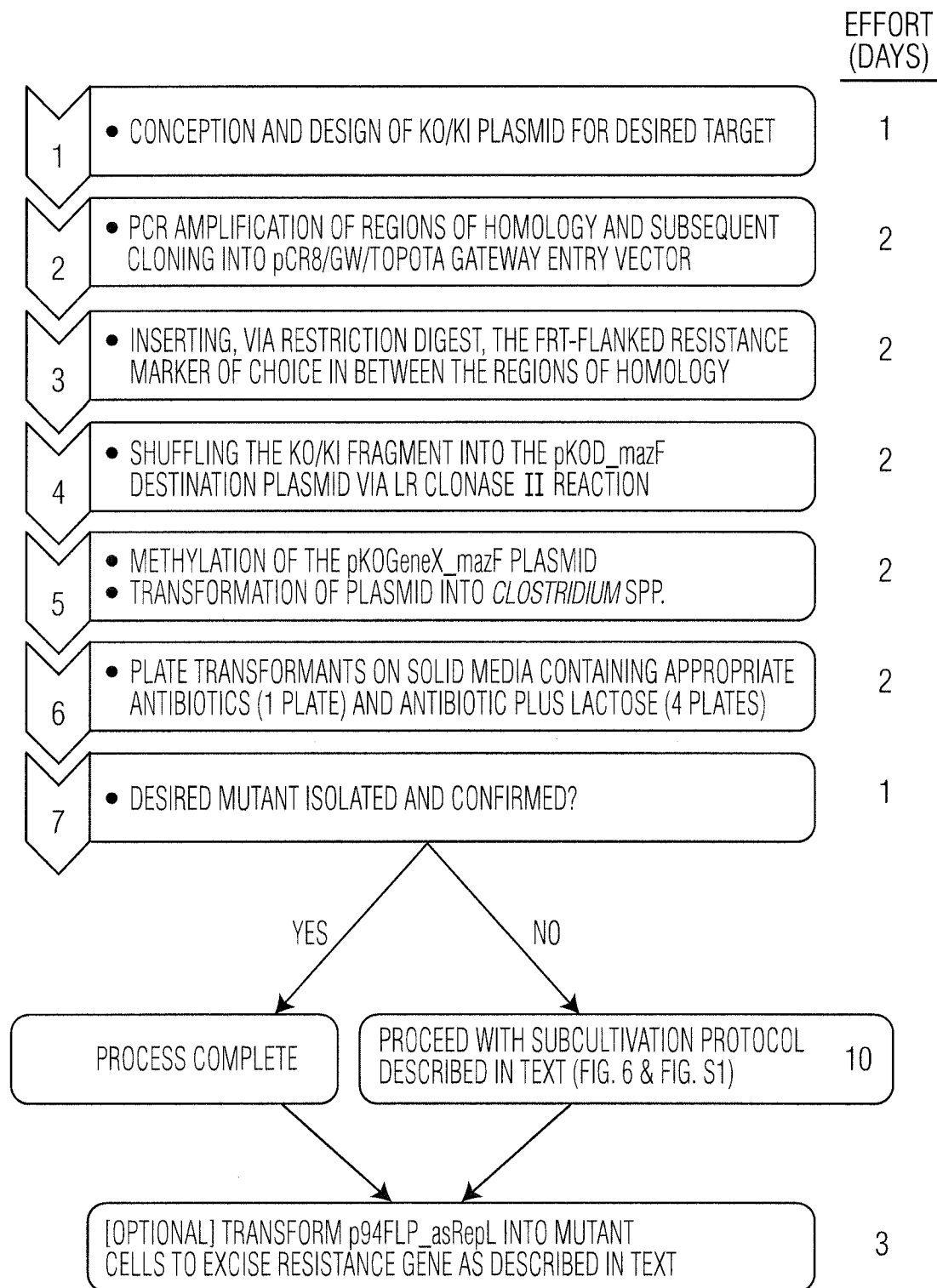


FIG. 11

14/15

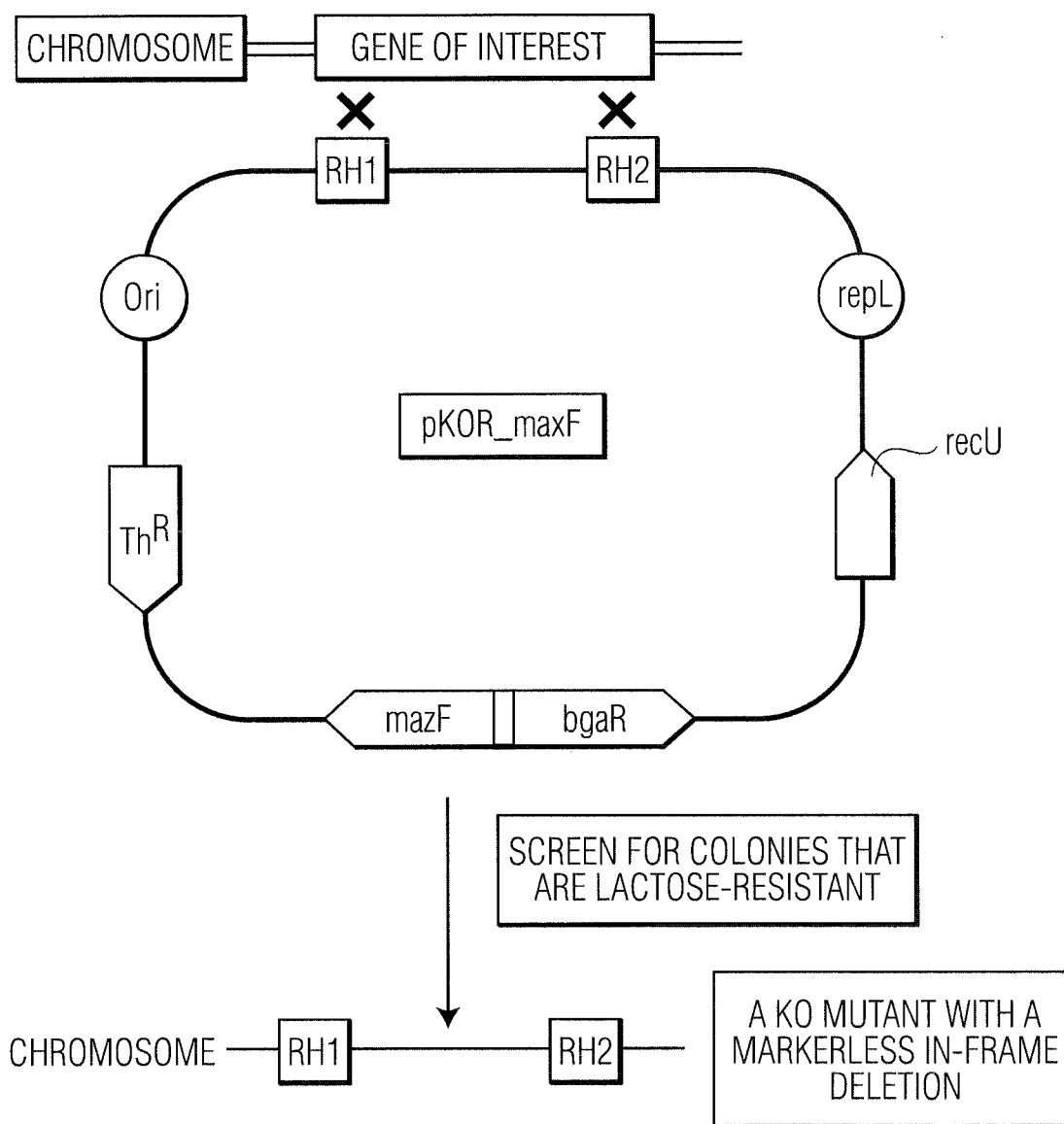


FIG. 12

15/15

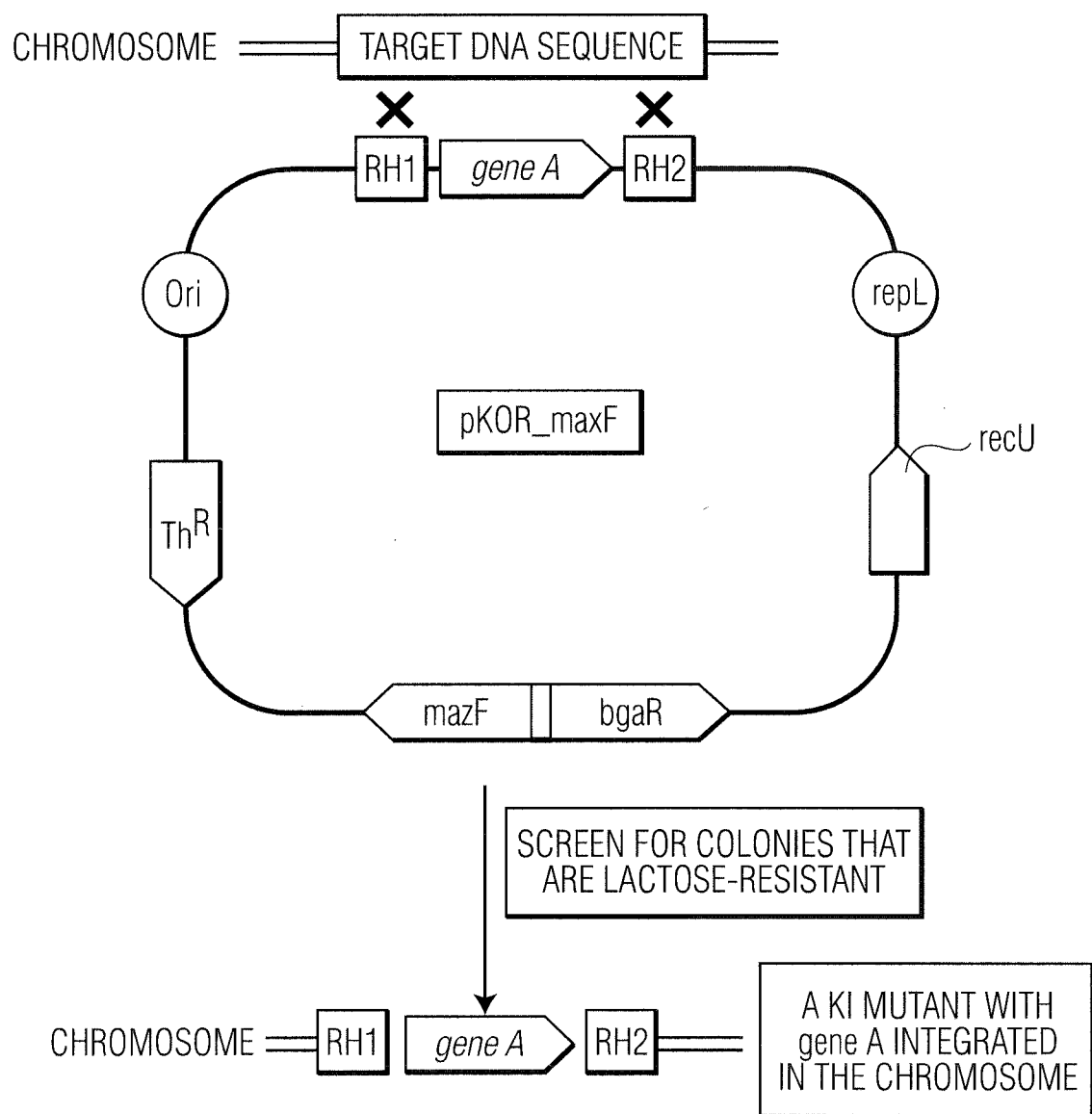


FIG. 13