US 20060057726A1

(19) **United States**

(12) **Patent Application Publication**

Sharpe

(10) **Pub. No.: US 2006/0057726 A1**

(43) **Pub. Date: Mar. 16, 2006**

(54) **METHOD FOR THE POSITIVE SELECTION OF CHROMOSOMAL MUTATIONS IN C1 METABOLIZING BACTERIA VIA HOMOLOGOUS RECOMBINATION**

(76) Inventor: **Pamela L. Sharpe**, Newark, DE (US)

Correspondence Address:
E I DU PONT DE NEMOURS AND COMPANY
LEGAL PATENT RECORDS CENTER
BARLEY MILL PLAZA 25/1128
4417 LANCASTER PIKE
WILMINGTON, DE 19805 (US)

(21) Appl. No.: **10/997,308**

(22) Filed: **Nov. 24, 2004**

Related U.S. Application Data

(60) Provisional application No. 60/527,877, filed on Dec. 8, 2003. Provisional application No. 60/527,083, filed on Dec. 3, 2003.

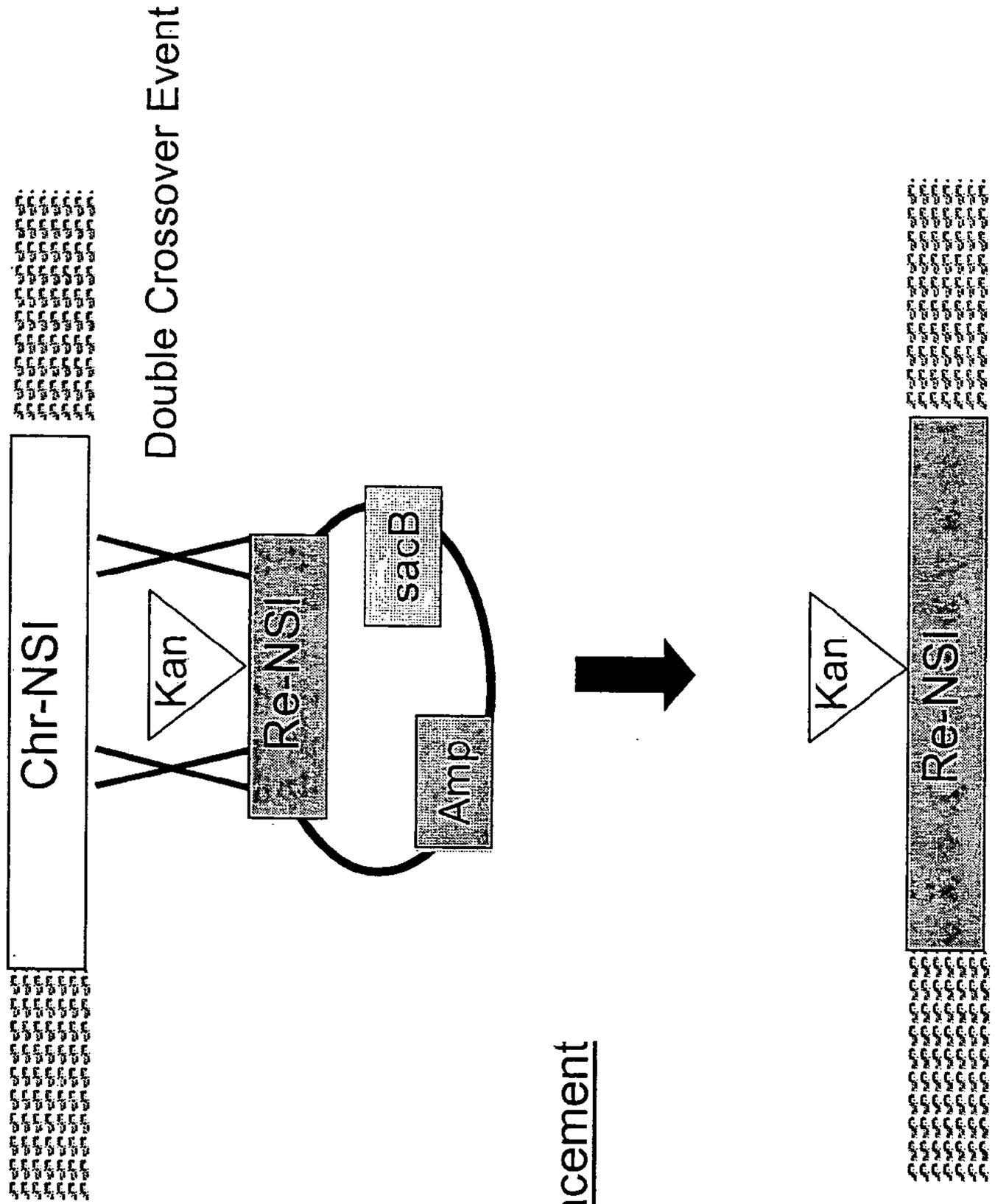
Publication Classification

(51) **Int. Cl.**
C12N 15/74 (2006.01)

(52) **U.S. Cl.** **435/471**

(57) **ABSTRACT**

A method for the positive selection of allelic exchange mutants is provided for C1 metabolizing bacteria. The chromosomal integration vectors, based on the pGP704 suicide vector, comprise at least one genetic selectable marker and the sacB gene, encoding levansucrase. A one- and two-step selection strategy is provided for the facile identification of double-crossover mutations in C1 metabolizing bacteria. This methodology enables production of "markerless" transformants, such that multiple rounds of mutation can be performed. Optimized conditions for conjugal transfer, homologous recombination, transformant purification, and screening are also presented.



Gene Replacement

Figure 1

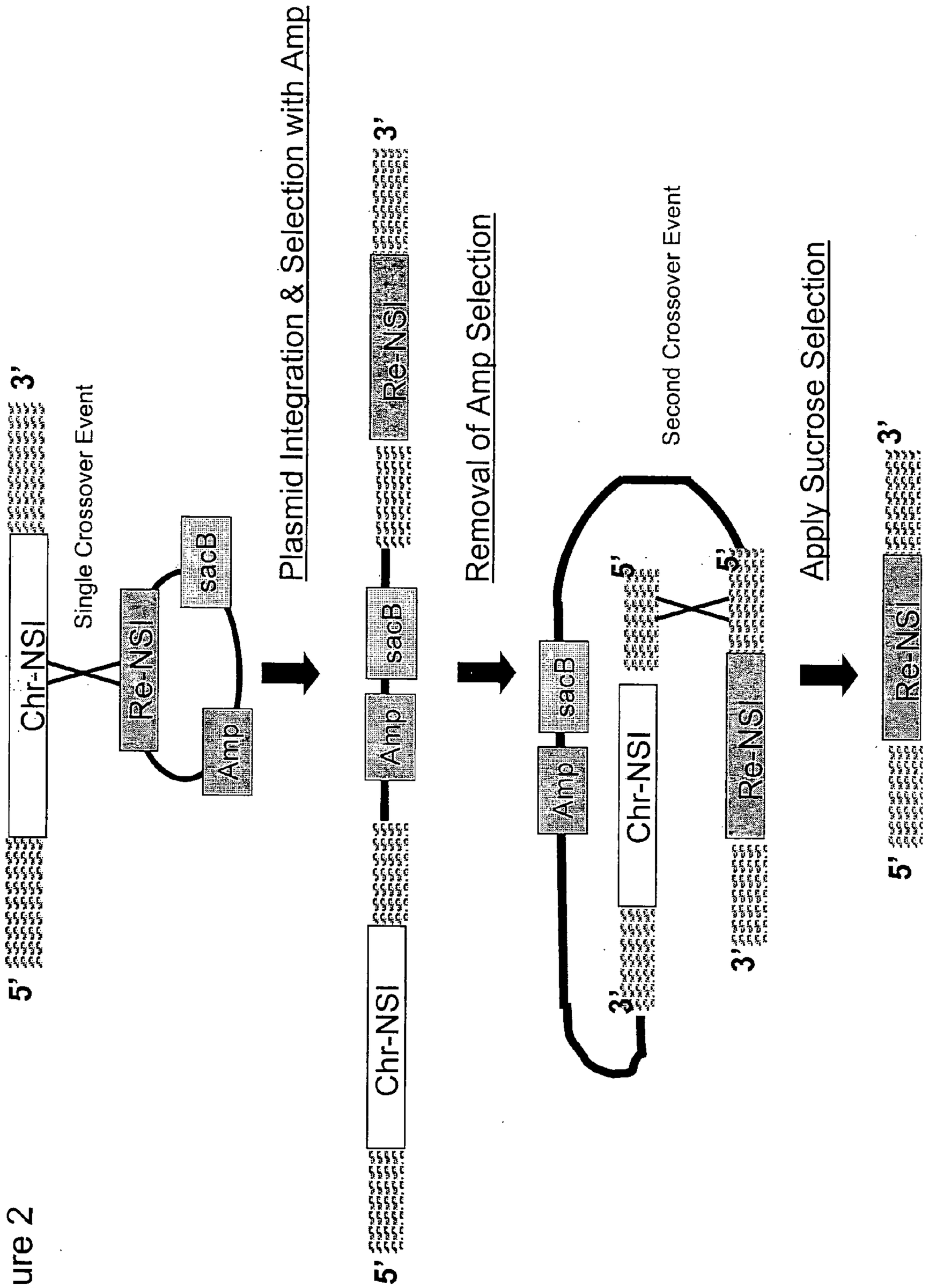

5'
Chr-NSI
3'
Single Crossover Event
Re-NSI
sacB
Amp
Plasmid Integration & Selection with Amp
Chr-NSI
Amp
sacB
Re-NSI
Removal of Amp Selection
Second Crossover Event
Apply Sucrose Selection

Figure 2

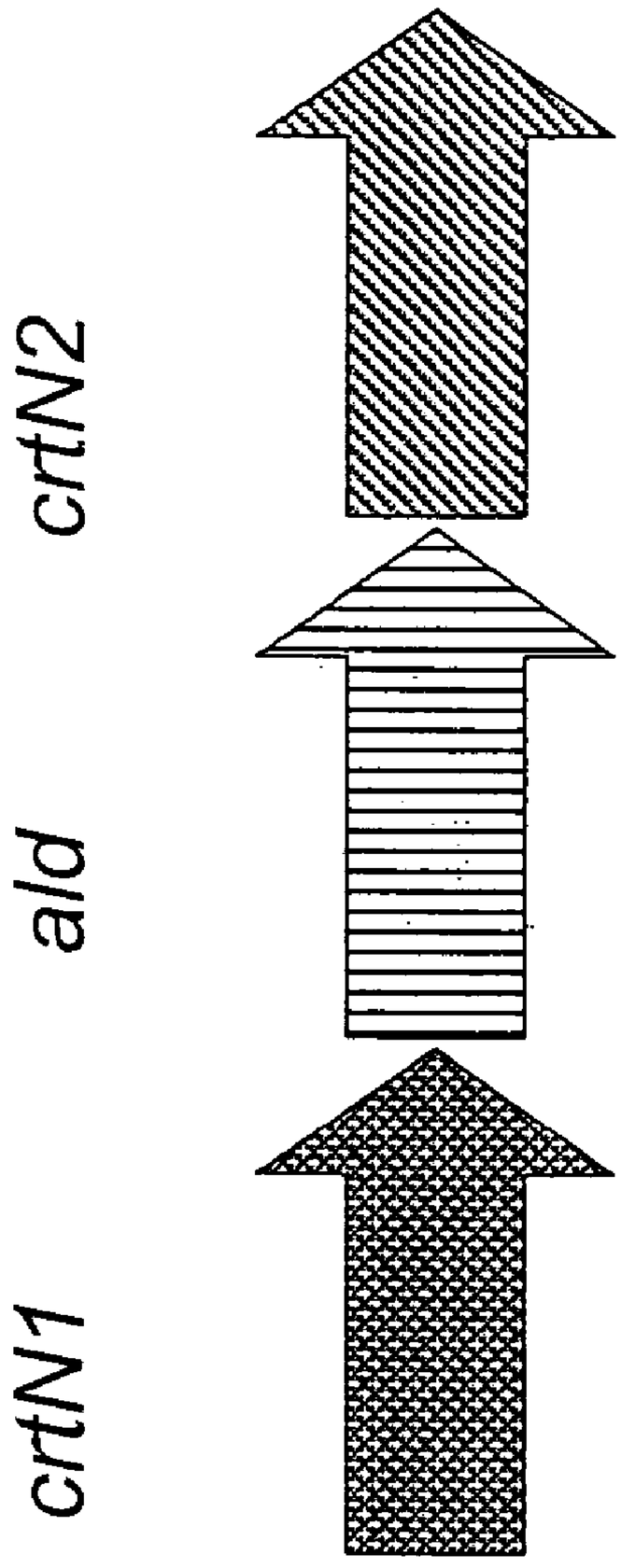


Figure 3

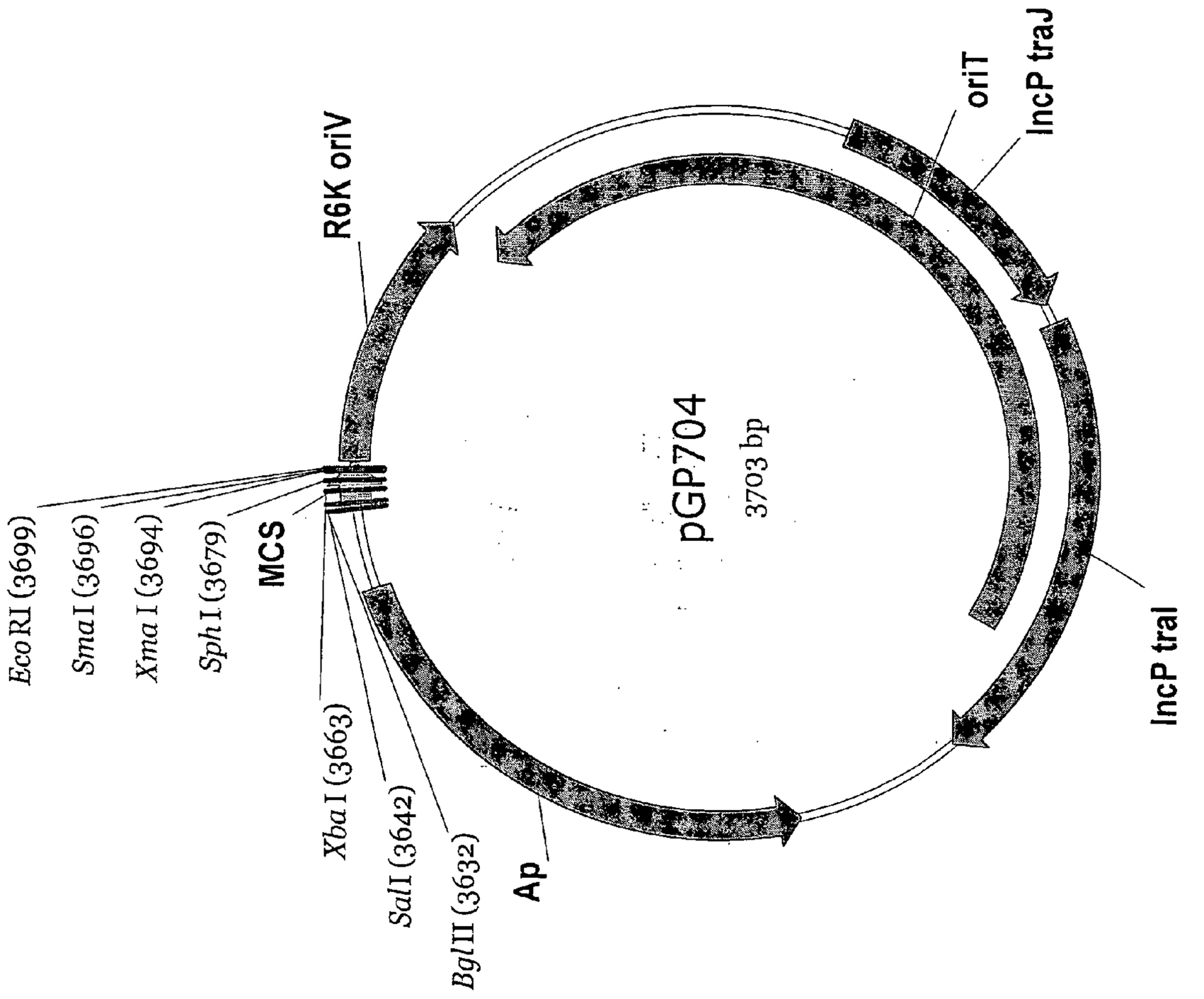

R6K oriV
oriT
IncP traJ
pGP704
3703 bp
IncP traI
Eco RI (3699)
Sma I (3696)
Xma I (3694)
Sph I (3679)
MCS
Xba I (3663)
Sal I (3642)
Bgl II (3632)
Ap

Figure 4

METHOD FOR THE POSITIVE SELECTION OF CHROMOSOMAL MUTATIONS IN C1 METABOLIZING BACTERIA VIA HOMOLOGOUS RECOMBINATION

[0001] This application claims the benefit of U.S. Provisional Application No. 60/527,877 filed Dec. 8, 2003 and U.S. Provisional Application No. 60/527,083 filed Dec. 3, 2003.

FIELD OF THE INVENTION

[0002] The present invention relates to the field of molecular biology and the use of selection vectors for the identification of recombinant bacteria containing heterologous DNA. More specifically, vectors have been constructed for the positive selection of double-crossover mutants in C1 metabolizing bacteria.

BACKGROUND INFORMATION

[0003] There are a number of microorganisms that utilize single carbon substrates as their sole energy source. Such microorganisms are referred to herein as "C1 metabolizers". These organisms are characterized by the ability to use carbon substrates lacking carbon to carbon bonds as a sole source of energy and biomass. All C1 metabolizing microorganisms are generally classified as methylotrophs. Methylotrophs may be defined as any organism capable of oxidizing organic compounds that do not contain carbon-carbon bonds. However, facultative methylotrophs, obligate methylotrophs, and obligate methanotrophs are all various subsets of methylotrophs. Specifically:

[0004] Facultative methylotrophs have the ability to oxidize organic compounds which do not contain carbon-carbon bonds, but may also use other carbon substrates such as sugars and complex carbohydrates for energy and biomass;

[0005] Obligate methylotrophs are those organisms which are limited to the use of organic compounds that do not contain carbon-carbon bonds for the generation of energy; and

[0006] Obligate methanotrophs are those obligate methylotrophs that have the distinct ability to oxidize methane.

Additionally, the ability to utilize single carbon substrates is not limited to bacteria but extends also to yeasts and fungi. A number of yeast genera are able to use single carbon substrates as energy sources in addition to more complex materials (i.e., the methylotrophic yeasts).

[0007] The ability of obligate methanotrophic bacteria to use methane as their sole source of carbon and energy under ambient conditions, in conjunction with the abundance of methane, makes the biotransformation of methane a potentially unique and valuable process. As such, several have attempted to harness the unique natural abilities of these organisms for commercial applications.

[0008] Historically, the commercial applications for the biotransformation of methane have fallen broadly into three categories: 1.) production of single cell protein; 2.) epoxidation of alkenes for production of chemicals; and 3.) biodegradation of chlorinated pollutants. Only epoxidation of alkenes has experienced slight commercial success due to low product yields, toxicity of products, and the large amount of cell mass required to generate products. Large-scale protein production from methane, termed single cell protein (or "SCP"), has been technically feasible and commercialized at large scale (Villadsen, J., Recent Trends Chem. React. Eng., [Proc. Int. Chem. React. Eng. Conf.], 2nd ed.; Kulkarni, B. D., Mashelkar, R. A., and Sharma, M. M., Eds.; Wiley East: New Delhi, India (1987); Vol 2, pp 320-33). However, SCP has not been economically successful thus far due to the relatively high cost of producing microbial protein, as compared to agriculturally derived protein (i.e., soy protein). This makes SCP a relatively low value product whose economic production cannot tolerate heavy bioprocessing costs. Thus, the yield of the methanotrophic strain in the process may be critical to the overall economic viability of the process. Microbial biomass produced by methanotrophic bacteria is typically very high in protein content (~70-80% by weight), which can restrict the direct use of this protein to certain types of animal feed.

[0009] In addition to the synthesis of SCP, methanotrophic cells can further build the oxidation products of methane (i.e., methanol and formaldehyde) into more complex molecules such as carbohydrate, lipids, and other proteins. For example, under certain conditions methanotrophs are known to produce exopolysaccharides (U.S. Pat. No. 6,537,786; U.S. Pat. No. 6,689,601; Ivanova et al., *Mikrobiologiya,* 57(4):600-5 (1988); Kilbane, John J., II, Gas, *Oil, Coal, Environ. Biotechnol.* 3, [Pap. IGT's Int. Symp.], 3rd ed., Meeting Date 1990; Akin, C. and J. Smith, Eds; IGT: Chicago, Ill. (1991); pp 207-26). Similarly, methanotrophs are known to accumulate both isoprenoid compounds and carotenoid pigments of various carbon lengths (U.S. Pat. No. 6,660,507; U.S. Pat. No. 6,689,601; Urakami et al., *J. Gen. Appl. Microbiol.,* 32(4):317-41 (1986)).

[0010] Most recently, the natural abilities of methanotrophic organisms have been stretched by the advances of genetic engineering. Odom et al. has investigated *Methylomonas* sp. 16a as a microbial platform of choice for production of a variety of materials beyond single cell protein, including carbohydrates, pigments, terpenoid compounds and aromatic compounds (U.S. Pat. No. 6,689,601 and U.S. Pat. No. 09/941947, herein incorporated entirely by reference). This particular methanotrophic bacterial strain is capable of efficiently using either methanol or methane as a carbon substrate, is metabolically versatile in that it contains multiple pathways for the incorporation of carbon from formaldehyde into 3-carbon units, and is capable of genetic exchange with donor species such as *Escherichia coli* via bacterial conjugation. Thus, *Methylomonas* sp. 16a can be engineered to produce new classes of products other than those naturally produced from methane and/or methanol. Further advancement in the metabolic engineering of C1 metabolizing microorganisms such as *Methylomonas* sp. 16a, however, is currently limited by the lack of efficient tools for generating defined mutants by homologous recombination.

[0011] The ability to produce specific defined mutations in a microorganism frequently relies on exploitation of the native homologous recombination properties of the cell to replace a nucleotide sequence of interest with a modified copy. Most frequently, the nucleotide sequence of interest is a particular functional gene of interest, which is then inactivated by the insertion of an antibiotic-resistance marker. In

theory, this type of recombination event is easily detected on a selective medium; however, performing allelic exchange in C1 metabolizing microorganisms has been relatively cumbersome due to the organisms' slow growth rates and the rarity of double-crossover events (which require extensive screening to isolate an allelic-exchange mutant). Despite these difficulties, there are examples of successful identification of allelic exchange in slow-growing C1 metabolizing microorganisms using suicide vectors (Naumov, G. N. et al., *Mol. Gen Mikrobiol. Virusol.,* 11:44-48 (1986) (abstract); Toyama and Lidstrom, *Microbiol.,* 144:183-191 (1998); Stolyar, S. et al., *Microbiology,* 145(5):1235-1244 (1999)). None of these previous studies, however, have utilized positive selection as a method for the rapid identification of allelic exchange mutants. Clearly, the identification of allelic exchange mutants obtained by targeted homologous recombination would greatly benefit from a system allowing for the positive selection of mutants.

[**0012**] The positive selection (or direct genetic selection) of mutant bacteria is possible whenever survival of the recombinant bacteria depends upon the presence or absence of a particular function encoded by the DNA that is introduced into the organism. The advantage of a selection method over a screening method is that growth of bacteria with the specific desired mutation is greatly favored over bacteria lacking that specific mutation, thus facilitating the identification of the preferred mutants.

[**0013**] Direct or positive selection vectors containing genes that convey lethality to the host are well known. For example, expression of the *Bacillus subtilis* or the *B. amyloliquefaciens* sacB gene in the presence of sucrose is lethal to *E. coli* and a variety of other Gram-negative and Gram-positive bacteria. The sacB gene encodes levansucrase, which catalyzes both the hydrolysis of sucrose and the polymerization of sucrose to form the lethal product levan. Although the basis for the lethality of levansucrase in the presence of sucrose is not fully understood, the inability of *E. coli* and many other gram negative bacteria to grow when sacB is expressed can be exploited to directly select for cells that have lost the sacB gene via homologous recombination. Numerous methods have been developed for the selection of various bacterial mutants, based on sacB. See for example: U.S. Pat. No. 6,048,694 (Bramucci et al.) concerning *Bacillus;* U.S. Pat. No. 5,843,664 (Pelicic et al.) concerning mycobacterium; U.S. Pat. No. 5,380,657 (Schaefer et al.) concerning *Coryneform bacteria;* Hoang et al. (*Gene,* 212(1):77-86 (1998)) concerning *Pseudomonas aeruginosa;* Copass et al. (*Infection and Immun.,* 65(5):1949-1952 (1997)) concerning *Helicobacter pylon;* and Kamoun et al. (*Mol. Microbiol.,* 6(6):809-816 (1992)) concerning *Xanthomonas.* However, no one has previously attempted to use sacB in a positive selection method for C1 metabolizing bacteria.

[**0014**] The problem to be solved, therefore, is to develop an efficient means of generating defined mutants by homologous recombination in C1 metabolizing bacteria. This method must permit transformants that have undergone allelic exchange to be selected for growth.

[**0015**] The present problem has been solved through the development of positive selection vectors for C1 metabolizing bacteria. These vectors, comprise at least one genetic selectable marker and the sacB gene, encoding levansucrase. Thus, the selectable marker can be used for the primary selection of transformants, while the sacB gene has a conditionally dominant lethal effect, to enable counter-selection against those clones which have chromosomally integrated the vector DNA. This eliminates the need for extensive screening. A one- and two-step selection strategy is provided for the facile identification of allelic exchange mutations in C1 metabolizing bacteria. Further, this methodology enables production of "markerless" transformants, such that multiple rounds of mutation can be performed.

SUMMARY OF THE INVENTION

[**0016**] The present invention provides methods for the integration of nucleotide sequences into the chromosome of a C1 metabolizing host microbe. The method employs a positive selection vector comprising the necessary genetic elements for replication of the plasmid and transfer of the plasmid into the host cell. Additionally, the vector comprises a sacB coding region and an variety or genetic selectable markers. The vector may be used in a number of conformations providing for targeted gene replacement as well as transformants free of any vector or marker sequences.

[**0017**] Accordingly the invention provides a method for the positive selection of double-crossover events in a C1 metabolizing bacterial host cell comprising:

[**0018**] a) providing a C1 metabolizing bacteria selected from the group consisting of methanotrophs and methylotrophs;

[**0019**] b) providing a positive selection vector comprising:

[**0020**] (i) a first genetic selectable marker;

[**0021**] (ii) an origin of transfer for a Cl metabolizing bacteria;

[**0022**] (iii) plasmid mobilization genes;

[**0023**] (iv) a sacB coding region encoding a levansucrase enzyme, under the control of a suitable promoter;

[**0024**] (v) a replacement nucleotide sequence of interest having at least two regions homology to a chromosomal nucleotide sequence of interest in the C1 metabolizing bacteria, said replacement nucleotide sequence of interest being disrupted by a second genetic selectable marker;

[**0025**] wherein the first and second genetic selectable markers are different and the vector is unable to replicate in the Cl metabolizing bacteria;

[**0026**] c) transforming the C1 metabolizing bacteria of (a) with the vector of (b)

[**0027**] d) selecting the transformants of (c) on the basis of the genetic selectable markers and sacB gene expression wherein, those transformants which are positive for the second selectable marker and grow in the presence of sucrose, but are negative for the first selectable marker have undergone double-crossover events; and

[**0028**] e) recovering the selected transformants of (d) which have undergone double-crossover events.

[0029] In similar fashion, the invention provides a method for the positive selection of double-crossover events in C1 metabolizing bacterial host cells comprising:

[0030] a) providing a C1 metabolizing bacteria selected from the group consisting of methanotrophs and methylotrophs;

[0031] b) providing a positive selection vector comprising:

[0032] (i) a genetic selectable marker;

[0033] (ii) an origin of transfer for a C1 metabolizing bacteria;

[0034] (iii) plasmid mobilization genes;

[0035] (iv) a sacB coding region encoding a levansucrase enzyme under the control of a suitable promoter;

[0036] (v) a mutant replacement nucleotide sequence of interest having at least two regions of homology to a chromosomal nucleotide sequence of interest in the C1 metabolizing bacteria;

[0037] the vector being unable to replicate in the C1 metabolizing bacteria chromosome;

[0038] c) transforming the C1 metabolizing bacteria of (a) with the vector of (b);

[0039] d) isolating the transformants of (c) on the basis of the genetic selectable marker; and

[0040] e) isolating the transformants of (d) which grow in the presence of sucrose, wherein said transformants contain double-crossover events.

[0041] C1 metabolizing bacteria transformants selected by the methods of the invention are also provided.

[0042] Additionally, the invention provides a positive selection vector for the positive selection of double-crossover events in a C1 metabolizing bacterial host cell comprising:

[0043] (i) at least one genetic selectable marker;

[0044] (ii) an origin of transfer for a C1 metabolizing bacteria;

[0045] (iii) plasmid mobilization genes;

[0046] (iv) a sacB coding region encoding a levansucrase enzyme under the control of a suitable promoter;

[0047] (v) a replacement nucleotide sequence of interest having at least two regions of homology to a chromosomal nucleotide sequence of interest in the C1 metabolizing bacteria; the vector being unable to replicate in the C1 metabolizing bacteria chromosome, wherein the vector of the invention may optionally comprise an additional genetic selectable marker disrupting the replacement nucleotide sequence of interest.

BRIEF DESCRIPTION OF THE FIGURES. SEQUENCE DESCRIPTIONS, AND BIOLOGICAL DEPOSITS

[0048] **FIG. 1** is a schematic diagram illustrating the single-step selection protocol, useful for identifying double-crossover mutants containing a marker in C1 metabolizing bacteria.

[0049] **FIG. 2** is a schematic diagram illustrating the two-step selection protocol, useful for identifying "marker-less" double-crossover mutants in C1 metabolizing bacteria.

[0050] **FIG. 3** illustrates the organization of the crtN1 gene cluster in *Methylomonas* sp. 16a.

[0051] **FIG. 4** is a plasmid map of pGP704.

[0052] The invention can be more fully understood from the following detailed description and the accompanying sequence descriptions, which form a part of this application.

[0053] The following sequences conform with 37 C.F.R. 1.821-1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures—the Sequence Rules") and are consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (1998) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a-bis), and Section 208 and Annex C of the Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

[0054] SEQ ID NOs:1-3 are the full-length *Methylomonas* sp. 16a genes known as crtN1, ald, and crtN2, respectively.

[0055] SEQ ID NOs:4 and 5 encode primers Drdl/npr-sacB and Tthlll/npr-sacB, respectively, used for amplification of the npr-sacB cassette from plasmid pBE83.

[0056] SEQ ID NOs:6 and 7 encode primers crtN#2-ctg288+1 kB/Bg/ll and crtN#2-ctg 288+1kB/Xbal, respectively, used for amplification of the crtN2 gene from *Methylomonas* sp. 16a.

[0057] SEQ ID NOs:8 and 9, respectively, encode the transposon-specific primers KAN-2FP-1 and KAN-2RP-1.

[0058] SEQ ID NOs:10 and 11, respectively, encode primers OP#1crtN#2 and OP#2crtN#2, used for confirming the double-crossover crtN2 mutant.

[0059] SEQ ID NOs:12 and 13 encode primers Bg/ll/ald (deletion)#1 and Sphl-Notl/ald (deletion)#1, respectively, used for the amplification of aldehyde deletion fragment #1 when constructing a deletion of the ald gene in *Methylomonas* sp. 16a chromosomal DNA.

[0060] SEQ ID NOs:14 and 15 encode primers Notl/ald (deletion)#2 and Sphl/ald (deletion)#2, respectively, used for the amplification of aldehyde deletion fragment #2 when constructing a deletion of the ald gene in *Methylomonas* sp. 16a chromosomal DNA.

[0061] Applicants made the following biological deposit under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for the Purposes of Patent Procedure:

Depositor Identification Reference	International Depository Designation	Date of Deposit
Methylomonas 16a	ATCC PTA 2402	Aug. 22, 2000

[0062] As used herein, "ATCC" refers to the American Type Culture Collection International Depository Authority located at ATCC, 10801 University Blvd., Manassas, Va. 20110-2209, U.S.A. The "International Depository Designation" is the accession number to the culture on deposit with ATCC.

[0063] The listed deposit will be maintained in the indicated international depository for at least thirty (30) years and will be made available to the public upon the grant of a patent disclosing it. The availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by government action.

DETAILED DESCRIPTION OF THE INVENTION

[0064] The present invention describes a process for the positive selection of allelic exchange mutants in a C1 metabolizing bacteria. Thus, using the native homologous recombination machinery of the C1 metabolizing host cell it is possible to replace a specific chromosomal nucleotide sequence of interest in a host organism's genome with a modified sequence (replacement nucleotide sequence of interest) via a double-crossover event.

[0065] In a single-step protocol, a suicide vector containing at least two genetic selectable markers (one marker positioned within the replacement nucleotide sequence of interest), a replacement nucleotide sequence of interest having homology to the chromosomal nucleotide sequence of interest, and the sacB gene are used to produce allelic exchange mutants that can be directly selected on sucrose and the selectable marker located within the disrupted nucleotide sequence.

[0066] In a two-step selection procedure, the suicide vector comprises at least one genetic selectable marker, a replacement nucleotide sequence of interest having homology to the chromosomal nucleotide sequence of interest, and a sacB coding sequence under the control of a suitable promoter. Following transformation of the host bacteria, transformants that underwent a single-crossover event are first selected based on the genetic selectable marker. Then, by plating on sucrose, the mutants that have lost the sacB gene during a second crossover event can be positively selected.

[0067] Thus, depending on the specific vector construct used and the choice of selection strategy, it is possible to produce both allelic exchange transformants containing a marker at the site of DNA replacement and markerless transformants.

DEFINITIONS

[0068] In this disclosure, a number of terms and abbreviations are used. The following definitions are provided.

[0069] As used herein, the term "comprising" means the presence of the stated features, integers, steps, or components as referred to in the claims, but that it does not preclude the presence or addition of one or more other features, integers, steps, components or groups thereof.

[0070] "Polymerase chain reaction" is abbreviated PCR.

[0071] "Kanamycin" is abbreviated Kan.

[0072] "Ampicillin" is abbreviated Amp.

[0073] The term "positive selection" means a selection method that enables only those cells that carry a DNA insert integrated at a specific chromosomal location to grow under particular conditions. In contrast, negative selection is based on selection methods whereby only those individuals that do not possess a certain character (e.g., cells that do not carry a DNA insert integrated at a specific chromosomal location) are selected.

[0074] The term "homologous recombination" refers to the exchange of DNA fragments between two DNA molecules (during cross over). The fragments that are exchanged are flanked by sites of identical nucleotide sequences between the two DNA molecules (i.e., "regions of homology"). Homologous recombination is the most common means for generated genetic diversity in microbes.

[0075] The term "regions of homology" refer to stretches of nucleotide sequence on nucleic acid fragments that participate in homologous recombination that have homology to each other. Effective homologous recombination will take place where these regions of homology are at least about 10 bp in length where at least about 50 bp in length is preferred. Typically, fragments that are intended for recombination contain at least two regions of homology where targeted gene disruption or replacement is desired.

[0076] The term "chromosomal integration" means that a chromosomal integration vector becomes congruent with the chromosome of a microorganism through recombination between homologous DNA regions on the chromosomal integration vector and within the chromosome.

[0077] The term "chromosomal integration vector" means an extra-chromosomal vector that is capable of integrating into the host's genome through homologous recombination.

[0078] The term "suicide vector" or "positive selection vector" refers to a type of chromosomal integration vector that is capable of replicating in one host but not in another. Thus, the vector is conditional for its replication.

[0079] The terms "single-crossover event" and "plasmid integration" are used interchangeably and mean the incorporation of a chromosomal integration vector into the genome of a host via homologous recombination between regions of homology between DNA present within the chromosomal integration vector and the host's chromosomal DNA. A "5single-crossover mutant" refers to a cell that has undergone a single-crossover event.

[0080] The terms "double-crossover event", "allelic exchange" and "gene replacement" are used interchangeably and mean the homologous recombination between a DNA region within the chromosomal integration vector and a region within the chromosome that results in the replacement of the functional chromosomal nucleotide sequence of interest (i.e., chr-NSI) with a homologous plasmid region (i.e., the replacement nucleotide sequence of interest, or re-NSI). A "double-crossover mutant" or "allelic exchange mutant" is the result of a double-crossover event. This mutant can be generated by two simultaneous reciprocal breakage and reunion events between the same two DNA fragments; alternatively, a double-crossover mutant can be the result of two single-crossovers that occur non-simultaneously.

[0081] The term "chromosomal nucleotide sequence of interest" is abbreviated "chr-NSI" and refers to a specific chromosomal sequence that is targeted for homologous recombination. In a preferred embodiment, the chr-NSI encodes a native gene of the C1 metabolizing bacteria.

[0082] The term "replacement nucleotide sequence of interest" is abbreviated "re-NSI" and refers to a nucleotide sequence of interest that is cloned into a chromosomal integration vector for the purpose of inducing homologous recombination with a chromosomal sequence. The re-NSI is modified with respect to chr-NSI by the addition, deletion, or substitution of at least one nucleotide. Sufficient homology must exist, however, between the two nucleotide sequences of interest to enable homologous recombination to occur. Similarly the term "mutant replacement nucleotide sequence of interest" refers to a nucleic acid sequence having homology to a sequence in a microbial genome, that has nucleotide additions, deletions or substitutions in the sequence that differ from the chromosomal sequence.

[0083] The terms "genetic selectable marker" or "selectable marker" will be used interchangeably and means a phenotypic trait, encoded by a genetic element that can be visualized under special conditions. For example, an antibiotic resistance marker serves as a useful selectable marker, since it enables detection of cells which are resistant to the antibiotic, when cells are grown on media containing that particular antibiotic.

[0084] The term "markerless mutants" or "markerless transformants" refers to an allelic exchange mutant, wherein the mutant allele does not carry a genetic marker.

[0085] The term "SacB" means a *Bacillus* encoded protein that catalyzes the conversion of sucrose into levan, a product that is toxic to most Gram negative microorganisms. The term "sacB" means a gene that encodes the "SacB" protein.

[0086] The terms "transposon" and "transposable element" are used interchangeably and refer to a relatively small DNA transposable or movable genetic element that has the ability to move from one chromosomal position to another within a DNA molecule, using a process known as transposition. Transposition in bacteria does not require extensive DNA homology between the transposon and the target DNA. A well-known example of a bacterial transposon is Tn5, which carries the genes for resistance to the antibiotics neomycin and kanamycin and a transposase. The term "transposase" means a protein that catalyses the chemical steps, i.e., breakage and joining, of a transposition reaction.

[0087] The term "in vitro transposition" means a biochemical reaction that is initiated outside the cell that catalyzes the movement of a transposable element from one site into a different site within a DNA molecule.

[0088] As used herein, the term "C_1 carbon substrate" refers to any carbon-containing molecule that lacks a carbon-carbon bond. Non-limiting examples are methane, methanol, formaldehyde, formic acid, formate, methylated amines (e.g., mono-, di-, and tri-methyl amine), methylated thiols, and carbon dioxide. In one embodiment, the C_1 carbon substrate is selected from the group consisting of methanol and/or methane.

[0089] The term "C_1 metabolizer" refers to a microorganism that has the ability to use a single carbon substrate as its sole source of energy and biomass. C_1 metabolizers will typically be methylotrophs and/or methanotrophs.

[0090] The term "C_1 metabolizing bacteria" refers to bacteria that have the ability to use a single carbon substrate as their sole source of energy and biomass. C_1 metabolizing bacteria, a subset of C_1 metabolizers, will typically be methylotrophs and/or methanotrophs.

[0091] The term "methylotroph" means an organism capable of oxidizing organic compounds that do not contain carbon-carbon bonds. Where the methylotroph is able to oxidize CH_4, the methylotroph is also a methanotroph. In one embodiment, the methylotroph organism is a methylotrophic bacteria capable of using methanol and/or methane as single carbon substrate.

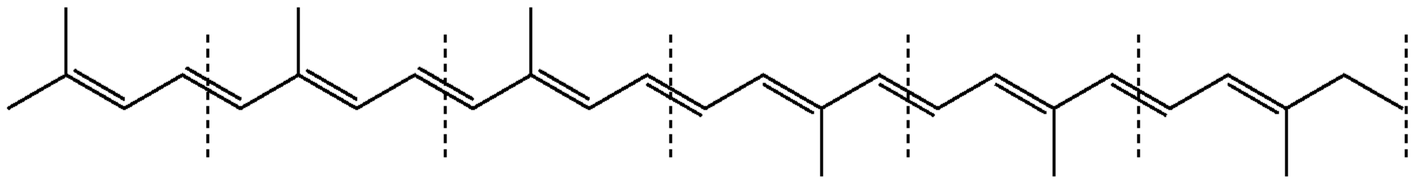
[0092] The term "methanotroph" or "methanotrophic bacteria" means a prokaryote capable of utilizing methane and/or methanol as a source of carbon and energy. Complete oxidation of methane to carbon dioxide occurs by aerobic degradation pathways. Typical examples of methanotrophs useful in the present invention include (but are not limited to) the genera *Methylomonas, Methylobacter, Methylococcus,* and *Methylosinus.*

[0093] The term "high growth methanotrophic bacterial strain" refers to a bacterium capable of growth with methane or methanol as the sole carbon and energy source and which possesses a functional Embden-Meyerof carbon flux pathway, resulting in a high rate of growth and yield of cell mass per gram of C_1 substrate metabolized (U.S. Pat. No. 6,689,601). The specific "high growth methanotrophic bacterial strain" described herein is referred to as "*Methylomonas* 16a", "16a" or "*Methylomonas* sp. 16a", which terms are used interchangeably and which refer to the *Methylomonas* strain used in the present invention.

[0094] The term "carotenoid" refers to a class of hydrocarbons having a conjugated polyene carbon skeleton formally derived from isoprene. This class of molecules is composed of C_{30} diapocarotenoids and C_{40} carotenoids and their oxygenated derivatives; and, these molecules typically have strong light absorbing properties.

[0095] "C_{30} diapocarotenoids" consist of six isoprenoid units joined in such a manner that the arrangement of isoprenoid units is reversed at the center of the molecule so that the two central methyl groups are in a 1,6-positional relationship and the remaining nonterminal methyl groups are in a 1,5-positional relationship. All C_{30} carotenoids may be formally derived from the acyclic $C_{30}H_{42}$ structure (Formula I below, hereinafter referred to as "diapophytoene"), having a long central chain of conjugated double bonds, by: (i) hydrogenation (ii) dehydrogenation, (iii) cyclization, (iv) oxidation, (v) esterification/glycosylation, or any combination of these processes.

Formula I



[0096] The term "CrtN1" refers to an enzyme encoded by the crtN1 gene, active in the native carotenoid biosynthetic pathway of *Methylomonas* sp. 16a. This gene is located within an operon comprising crtN2 and ald.

[0097] The term "CrtN2" refers to an enzyme encoded by the crtN2 gene, active in the native carotenoid biosynthetic pathway of *Methylomonas* sp. 16a. This gene is located within an operon comprising CrtN1 and ald.

[0098] The term "ALD" refers to an enzyme (an aldehyde dehydrogenase) encoded by the ald gene, active in the native carotenoid biosynthetic pathway of *Methylomonas* sp. 16a. This gene is located within an operon comprising CrtN1 and CrtN2.

[0099] A "nucleic acid" is a polymeric compound comprised of covalently linked subunits called nucleotides. Nucleic acids include polyribonucleic acid (RNA) and polydeoxyribonucleic acid (DNA), both of which may be single-stranded or double-stranded. DNA includes cDNA, genomic DNA, synthetic DNA, and semi-synthetic DNA.

[0100] As used herein, an "isolated nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

[0101] A nucleic acid fragment is "hybridizable" to another nucleic acid fragment, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid fragment can anneal to the other nucleic acid fragment under the appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual,* 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor (1989), particularly Chapter 1 1 and Table 11.1 therein (entirely incorporated herein by reference). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. Stringency conditions can be adjusted to screen for moderately similar fragments (such as homologous sequences from distantly related organisms) to highly similar fragments (such as genes that duplicate functional enzymes from closely related organisms). Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6×SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2×SSC, 0.5% SDS at 45° C. for 30 min, and then repeated twice with 0.2×SSC, 0.5% SDS at 50° C. for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2×SSC, 0.5% SDS was increased to 60° C. Another preferred set of highly stringent conditions uses two final washes in 0.1×SSC, 0.1% SDS at 65° C. An additional set of stringent conditions include hybridization at 0.1×SSC, 0.1% SDS, 65° C. and washed with 2×SSC, 0.1% SDS followed by 0.1×SSC, 0.1% SDS, for example.

[0102] Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of Tm for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher Tm) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating Tm have been derived (see Sambrook et al., supra, 9.50-9.51). For hybridizations with shorter nucleic acids, i.e., oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook et al., supra, 11.7-11.8). In one embodiment the length for a hybridizable nucleic acid is at least about 10 nucleotides. Preferable a minimum length for a hybridizable nucleic acid is at least about 15 nucleotides; more preferably at least about 20 nucleotides; and most preferably the length is at least 30 nucleotides. Furthermore, the skilled artisan will recognize that the temperature and wash solution salt concentration may be adjusted as necessary according to factors such as the length of the probe.

[0103] The term "oligonucleotide" refers to a nucleic acid, generally of at least about 18 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an mRNA molecule.

[0104] "Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A

"transgene" is a gene that has been introduced into the genome by a transformation procedure.

[0105] "Chemically synthesized", as related to a sequence of DNA, means that the component nucleotides were assembled in vitro. Manual chemical synthesis of DNA may be accomplished using well-established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines.

[0106] "Synthetic genes" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments that are then enzymatically assembled to construct the entire gene. Accordingly, the genes can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

[0107] As used herein, the term "homolog", as applied to a gene, means any gene derived from the same or a different microbe having the same function. A homologous gene may have significant sequence similarity.

[0108] "Coding sequence" or "coding region of interest" refers to a DNA sequence that codes for a specific amino acid sequence.

[0109] "Suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, RNA processing sites, effector binding sites and stem-loop structures.

[0110] "Transcriptional and translational control sequences" are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. "Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene at different stages of development, or in response to different environmental or physiological conditions. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

[0111] The "3' non-coding sequences" refer to DNA sequences located downstream of a coding sequence and include sequences encoding regulatory signals capable of affecting mRNA processing or gene expression.

[0112] The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

[0113] "Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. As used herein, the host cell's genome is comprised of chromosomal and extrachromosomal (e.g. plasmid) genes. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" organisms.

[0114] "Conjugation" refers to a particular type of transformation in which a unidirectional transfer of DNA (e.g., from a bacterial plasmid) occurs from one bacterium cell (i.e., the "donor") to another (i.e., the "recipient"). The process involves direct cell-to-cell contact.

[0115] The terms "plasmid", "vector" and "cassette" refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction. "Transformation cassette" refers to a specific vector containing a foreign gene and having elements (in addition to the foreign gene) that facilitate transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

[0116] The term "altered biological activity" or "altered activity" will refer to an activity, associated with a protein encoded by a nucleotide sequence which can be measured by an assay method, where that activity is either greater than or less than the activity associated with the native sequence. "Enhanced biological activity" refers to an altered activity that is greater than that associated with the native sequence. "Diminished biological activity" is an altered activity that is less than that associated with the native sequence.

[0117] The term "sequence analysis software" refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. "Sequence analysis software" may be commercially available or independently developed. Typical sequence analysis software will include, but is not limited to: the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, Wis.); BLASTP, BLASTN, BLASTX (Altschul et al., *J. Mol. Biol.* 215:403410 (1990)); DNASTAR (DNASTAR, Inc., Madison, Wis.); and the FASTA program incorporating the Smith-Waterman algorithm (W. R. Pearson, *Comput. Methods Genome Res.,* [Proc. Int. Symp.], Meeting Date 1992,111-20. Suhai, Sandor, Ed.; Plenum: New York, N.Y. (1994)). Within the context of this application it will be understood that where sequence analysis software is used for analysis, the results of

the analysis are based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters which originally load with the software when first initialized.

[0118] Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual,* 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y. (1989) (hereinafter "Maniatis"); and by Silhavy, T. J., Bennan, M. L. and Enquist, L. W., *Experiments with Gene Fusions,* Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y. (1984); and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology,* published by Greene Publishing Assoc. and Wiley-Interscience (1987).

Overview of the One-Step and Two-Step Selection Strategy

[0119] The principles of the one-step and two-step selection strategies of the present invention are illustrated in **FIGS. 1 and 2**, respectively. Each of these methods relies on the application of a positive selection vector, in a preferred embodiment derived from the suicide delivery vector pGP704, which is able to integrate into the chromosome of C1 metabolizing bacteria to produce mutations that are the result of both single- or double-crossover events. The positive selection vector must contain an origin of transfer suitable for a C1 metabolizing bacteria, as well as plasmid mobilization genes.

[0120] The one-step selection strategy for the positive selection of the double-crossover events so produced in C1 metabolizing bacterial host cells is illustrated in **FIG. 1**. Specifically, the positive selection vector comprises:

[0121] (i) at least one gene encoding resistance to a first genetic selectable marker (e.g., Amp);

[0122] (ii) a sacB coding region encoding a levansucrase enzyme under the control of a suitable promoter; and

[0123] (iii) a replacement nucleotide sequence of interest (i.e., re-NSI) that is disrupted by a gene encoding resistance to a second genetic selectable marker, wherein the second genetic selectable marker is different than the first genetic selectable marker (e.g., Kan). The re-NSI in the positive selection vector must have at least two regions homology to a nucleotide sequence of interest in the chromosome of the C1 metabolizing bacteria (i.e., chr-NSI), in order to undergo allelic exchange.

[0124] Upon transformation of C1 metabolizing bacteria with the positive selection vector described above, a double-crossover event by homologous recombination may occur between chr-NSI and re-NSI, such that the chr-NSI is replaced with re-NSI from the vector. Direct selection of these allelic exchange transformants is possible by growing the transformants in the presence of the genetic selectable markers and sucrose. Specifically, Table 1 illustrates the specific type of transformants that would be identified using various combinations of selection, according to the illustration shown in **FIG. 1** (although the selection markers illustrated therein are by no means to be construed as a limitation to the present invention).

TABLE 1

Identification of Single and Double Crossover Mutants using Various Strategies of Selection

Selection	Bacteria Identified by Selection
sucrose only	wildtype and double-crossover mutants
Kan only	single- and double-crossover mutants
Amp only	single-crossover mutants
Amp + Kan	single-crossover mutants
Kan + sucrose	double-crossover mutants

[0125] Clearly, facile identification of transformants which have undergone double-crossover events is possible by growing transformants in the presence of the second genetic selectable marker (i.e., that disrupted re-NSI) and sucrose.

[0126] In **FIG. 2**, the two-step selection strategy for the positive selection of the double-crossover events produced in C1 metabolizing bacterial host cells is illustrated. For this strategy, the positive selection vector comprises:

[0127] (i) at least one gene encoding resistance to a first genetic selectable marker (e.g., Amp);

[0128] (ii) a sacB coding region encoding a levansucrase enzyme under the control of a suitable promoter; and

[0129] (iii) a replacement nucleotide sequence of interest (i.e., re-NSI), which one desires to insert into the chromosome of the C1 metabolizing bacteria as a replacement to an existing nucleotide sequence of interest in the bacterial chromosome (i.e., chr-NSI). Thus, re-NSI is a mutant re-NSI modified with respect to chr-NSI by the addition, substitution, or deletion of at least one nucleotide.

[0130] Upon transformation of C1 metabolizing bacteria with the positive selection vector described above, a single-crossover event by homologous recombination occurs between chr-NSI and re-NSI, such that the entire positive selection vector is integrated into the bacterial chromosome at the site of crossover. These events can be selected by growth on the first genetic selectable marker (e.g., Amp), whereby a complete copy of chr-NSI and a complete copy of re-NSI are present in the chromosome. Upon removal of selection by the first genetic selectable marker, a second crossover event may occur, resulting in the "looping out" of the positive selection vector, to yield transformants containing either the chr-NSI or the re-NSI in the chromosome. Direct selection of these allelic exchange transformants is possible by growing the transformants in the presence of sucrose, since single-crossover mutants will be killed under these conditions.

[0131] Although both the one-step versus two-step selection strategy will result in the facile identification allelic exchange mutants, use of the one-step selection strategy will produce allelic exchange transformants containing a marker at the site of DNA replacement while use of the two-step selection strategy will produce allelic exchange transformants that are markerless. Each methodology will therefore have advantages that must be considered with respect to subsequent applications.

Chromosomal Integration Vectors for Generating and Isolating Single- and Double-Crossover Mutants

[0132] Suicide plasmids can be used to create defined modifications within a targeted chromosome of a host organism's genome. Frequently, these defined modifications result in disruptions, insertions, and replacements upstream, downstream, and within native genes via homologous recombination. A single-or double-crossover event can occur via targeted homologous recombination using the following methodology. First, a specific nucleotide sequence of interest (i.e., re-NSI) is cloned into a chromosomal integration vector and the vector is subsequently introduced into the specific host organism (the "recipient"). The transformed host is then placed under conditions where the plasmid cannot replicate.

[0133] Single-crossover mutants can be identified by selection for a selectable marker on the plasmid. These transformants have integrated the plasmid into the host chromosome via homology between re-NSI and chr-NSI. In one embodiment, it is possible to produce a single-crossover mutant whereby two non-functional copies of the nucleotide sequence of interest are present in the recipient chromosome. This requires one to clone an internal portion of the chr-NSI (i.e., wherein the 5' and 3' portion of chr-NSI are not present) into the chromosomal integration vector as re-NSI. Upon homologous recombination, the functional 5' portion of chr-NSI is joined with the non-functional 3' portion of re-NSI; in like manner, the non-functional 5' portion of re-NSI is joined with the functional 3' portion of chr-NSI. Thus, both chromosomal copies of the nucleotide sequence of interest are non-functional.

[0134] In contrast to single-crossover mutants, double-crossover mutants occur with less frequency and are typically much more difficult to isolate. Identification of double-crossover events, whereby two crossovers between re-NSI and chr-NSI must occur, requires either the one-step or two-step selection procedure of the present invention. In one embodiment, re-NSI must possess an easily selectable phenotype (e.g., when disrupted with a Kan^R insertion), thus enabling a one-step procedure for selection (**FIG. 1**). Alternatively, when direct selection of the mutation is not possible (e.g., when re-NSI is marked with only a point mutation) a two-step procedure is used instead (**FIG. 2**). First, the entire plasmid is integrated into the chromosome by a single-crossover between the homologous nucleotide sequences of interest, producing a chromosomal duplication. Wild-type cells are selected against while those transformants possessing a single-crossover are selected for according to the genetic selectable marker present on the suicide vector. Then, the chromosomal duplication must undergo a second homologous recombination event between the two copies of the nucleotide sequences of interest. Depending on the specific site of this second recombination event, a single copy of the nucleotide sequence of interest remains on the chromosome—either the wild-type copy (i.e., chr-NSI) or the mutant replacement copy (i.e., re-NSI). In all cases where a second recombination event occurs, however, the integrated plasmid is excised from the recient's chromosome. Based on the rare occurrence of this second homologous recombination event, a second selection strategy must exist to differentiate between single- and double-crossover mutants. One way of selecting against the integrated plasmid is to use a plasmid that also carries the sacB gene. This enables a direct selection for loss of the plasmid.

[0135] Based on the description above for the generation and selection of allelic exchange mutants, chromosomal integration vectors utilized for the generation of double-crossover mutations in plasmids will generally possess the following characteristics: 1.) The plasmid will be conditional for replication to permit integration into the chromosome. This can be achieved by using a plasmid that is able to replicate autonomously only in permissive hosts or by using conditional replicons (e.g., a plasmid that is temperature-sensitive for replication; 2.) the plasmid will typically carry at least one genetic selectable marker; 3.) it is desirable if the plasmid is transferable to a variety of other bacteria; and 4.) preferably, the plasmid will possess a variety of unique cloning sites.

[0136] In addition to the above characteristics of a chromosomal integration vector, the vector in the present invention must be suitable for integration into C1 metabolizing bacterial host cells. Thus, the chromosomal integration vector suitable for positive selection in the present invention will generally possess the following characteristics:

[0137] 1. The plasmid will typically be conditional for replication to permit integration into the chromosome;

[0138] 2. The plasmid will generally possess at least one genetic selectable marker;

[0139] 3. The plasmid will have an origin of transfer for a C1 metabolizing bacteria;

[0140] 4. The plasmid will generally possess mobilization genes, such that it is transferable to other bacteria (e.g., *Escherchia coli*);

[0141] 5. The plasmid will possess a sacB coding region (encoding levansucrase) under the control of a suitable promoter; and,

[0142] 6. The plasmid will typically contain a nucleotide sequence of interest having homology to a nucleotide sequence of interest in the chromosome of the C1 metabolizing bacteria.

[0143] 7. Preferably, the plasmid will contain a variety of unique cloning sites.

Each of these components of the positive selection vector will be discussed in detail below.

[0144] 1. Conditional Replication and Chromosomal Integration

[0145] In general, the act of plasmid replication requires two components: a DNA segment on the vector DNA known as an origin of replication (i.e., the site on the plasmid where DNA replication initiates) and vector-encoded proteins that control replication through interactions with the DNA sequences. Conditional replication occurs when a plasmid replicates only under a specific set of conditions. There are several types of conditional replication systems known in the art. For example, some vectors have the replication origin of R6K, which requires the trans-acting factor (Π protein) encoded by the pir gene to be present for replication to occur. By selecting the particular host background in which the vector is located (i.e., by selecting hosts that do or do not possess the Π protein), it is therefore possible to

control replication of the vector. In other conditional replication vectors, replication is controlled by temperature. Specifically, under one set of growth conditions, the temperature-sensitive vector is capable of replication; however, under a second set of growth conditions, replication is impaired.

[0146] The conditional replication of some vectors has been exploited to allow for the targeted alteration of the host's genome. When conditions prevent vector replication, vector integration is stimulated. Thus, if DNA sequences are present on the vector (i.e., re-NSI sequences) that share sequence homology with a region within the chromosome (i.e., chr-NSI sequences), the vector sequences can be integrated into the host genome via homologous recombination.

[0147] 2. A Selectable Or Screenable Genetic Marker

[0148] The chromosomal integration vector will contain at least one genetic selectable or screenable genetic marker that is suitable for use in the C1 metabolizing bacterial host cell. Genetic selectable markers of the invention are those that are encoded by a genetic element and result in a phenotypic change in the transformant such that a genetic event may be identified and tracked. Genetic selectable markers may encoded polypeptides that serve as markers in and of them selves (i.e. lac), or may encode polypeptides that result in a specific a phenotype under specific conditions (antibiotic resistance). A person of skill in the art will be able to quickly determine suitable markers. In a preferred embodiment, antibiotic markers are used, such as those markers for kanamycin, tetracycline, ampicillin, chloramphenicol, and streptinomycin resistance. Additionally, metal resistance markers, substrate-utilization markers, genes encoding fluorescent and bioluminescent proteins, lacZ, gfp, cat, galK, inaZ, luc, luxAB, bgaB, nptll, phoA, uidA and xylE are also considered suitable selectable or screenable genetic marker(s), although this list is not intended to be limiting.

[0149] 3. An Origin Of Transfer ("oriT") For A C1 Metabolizing Bacteria

[0150] The mobilization of vector DNA into a C1 metabolizing bacteria typically occurs via conjugation. Conjugation is a process by which DNA is transferred from a donor cell to a recipient cell, mediated by physical contact between the cells. Once the cells are brought into close contact, the donor vector is nicked at a specific site within the origin of transfer (oriT) by a plasmid-encoded protein (i.e., a DNA relaxase). The vector DNA is passed from the donor cell into the recipient cell as a single strand via a process known as rolling circle replication. Once the single strand of vector has been transferred into the recipient cell, replication of the single strand occurs to re-generate a double-stranded molecule.

[0151] The oriT itself is not influenced by the host background; thus, any oriT that is recognized by the DNA relaxase encoded by the chromosomal integration vector of the present invention should be functional in C1 metabolizing bacteria.

[0152] 4. Plasmid Mobilization Genes

[0153] Most plasmids are classified on the basis of their intrinsic properties, according to their "incompatibility groups" or "Inc" groups. This classification reflects the similarities in sequence, function, and the nature of the replicon (as replicons of the same type are unable to co-exist in a cell, while replicons from different incompatibility groups may exist simultaneously in a single cell). Natural plasmid isolates of gram-negative bacteria that belong to incompatibility groups C, N, P, Q and W display replication and maintenance proficiency in a diversity of bacterial species.

[0154] The chromosomal integration vector must encode mobilization proteins, known as "Mob" proteins such that it is transferable from donor to recipient bacteria cells via conjugation; however, the specific Inc group of the vector affects the mob genes that are required. A single gene or multiple genes may encode the nickase and helicase functions imparted by the Mob proteins.

[0155] 5. A sacB Coding Region (Encoding Levansucrase) Under The Control Of A Suitable Promoter

[0156] It has been demonstrated that the *Bacillus subtilis* and *B. amyloliquefaciens* sacB gene (GenBank® Accession No.s X02730 and X52988, respectively), encoding the secreted enzyme levansucrase (EC 2.4.1.10), confers sucrose sensitivity upon various Gram-negative bacteria. Other likely *Bacillus* species which may be used as a source of the sacB gene include, for example: *B. pumilis, B. brevis, B. licheniformis, B. thurgiensis, B. sphearicus,* and *B. stearothenmophilus.* The *B. subtilis* and *B. amyloliquefaciens* sacB nucleic acid fragments may be used to isolate genes encoding homologous proteins from the same or other microbial species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to: 1.) methods of nucleic acid hybridization; 2.) methods of DNA and RNA amplification, as exemplified by various uses of nucleic acid amplification technologies [e.g., polymerase chain reaction (PCR), Mullis et al., U.S. Pat. No. 4,683,202; ligase chain reaction (LCR), Tabor, S. et al., *Proc. Acad. Sci. USA,* 82:1074 (1985); or strand displacement amplification (SDA), Walker, et al., *Proc. Natl. Acad. Sci. U.S.A.,* 89:392 (1992)]; and 3.) methods of library construction and screening by complementation.

[0157] For example, sacB genes encoding similar proteins or polypeptides to *B. subtilis* and *B. amyloliquefaciens* could be isolated directly by using all or a portion of the instant sacB nucleic acid fragments as DNA hybridization probes to screen libraries from any desired bacteria using methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis, supra). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primers DNA labeling, nick translation, end-labeling techniques, or RNA probes using available in vitro transcription systems. In addition, specific primers can be designed and used to amplify a part of or full-length of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length DNA fragments under conditions of appropriate stringency.

[0158] Alternatively the instant *B. subtilis* and *B. amyloliquefaciens* sacB sequences may be employed as hybridization reagents for the identification of homologs. The basic components of a nucleic acid hybridization test include a

probe, a sample suspected of containing the gene or gene fragment of interest, and a specific hybridization method. Hybridization methods are well defined; see, for example, Maniatis (supra).

[**0159**] The expression of the sacB gene requires a suitable promoter. Initiation control regions or promoters, which are useful to drive expression of the sacB gene in the desired host cell are numerous and familiar to those skilled in the art. Virtually any promoter capable of driving these genes is suitable for the present invention. Preferred promoters for the present invention are, but not limited to: CYC1, HIS3, GAL1, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, and TPI (useful for expression in Saccharomyces); AOX1 (useful for expression in Pichia); and lac, ara, tet, trp, IP_L, IP_R, T7, tac, and trc (useful for expression in *Escherichia coli*), as well as the amy, apr, npr promoters and various phage promoters (useful for expression in *Bacillus*). Additionally, the deoxyxylulose phosphate synthase or methanol dehydrogenase operon promoter (Springer et al., *FEMS Microbiol Lett,* 160:119-124 (1998)), the promoter for polyhydroxyalkanoic acid synthesis (Foellner et al., *Appl. Microbiol. Biotechnol.,* 40:284-291 (1993)), promoters identified from native plasmids in methylotrophs (EP 296484), Plac (Toyama et al., *Microbiology,* 143:595-602 (1997); EP 62971), Ptrc (Brosius et al., *Gene,* 27:161-172 (1984)), and promoters associated with antibiotic resistance [e.g., kanamycin (Springer et al., *FEMS Microbiol Lett,* 160:119-124 (1998); Ueda et al., *Appl. Environ. Microbiol.,* 57:924-926 (1991)) or tetracycline (U.S. Pat. No. 4,824,786)] are suitable for expression in C1 metabolizers. Most preferred in the present invention is the *Bacillus* npr promoter, whose native function in *Bacillus* is to drive expression of the neutral protease (npr) gene.

[**0160**] Termination control regions may also be derived from various genes native to the preferred hosts. Optionally, a termination site may be unnecessary, however, it is most preferred if included.

[**0161**] It would be readily apparent to one skilled in the art that a variety of alternative markers, other than sacB, could be utilized in the construction chromosomal integration vectors having positive selection attributes, without significantly altering the spirit or scope of the present positive selection method for C1 metabolizing bacteria. For a review of other genes and the construction of positive selection vectors generally see: Maniatis, supra; Alexeyev, Mikhail F., *BioTechniques,* 26(5):824-828 (1999); Matin et al., *Anal. Biochem.,* 278(1):46-51 (2000); and Bramucci et al., WO 97/16558.

[**0162**] 6. A Replacement Nucleotide Sequence Of Interest, Having Homology to a Chromosomal Nucleotide Sequence of Interest

[**0163**] The particular strategy to be used for selection of double-crossover mutants (i.e., the one-step or two-step selection process) affects the construction of the re-NSI that is cloned into the chromosomal integration vector. As discussed previously, the one-step selection process requires the presence of a genetic selectable marker within re-NSI. In contrast, the two-step selection process is limited only such that re-NSI is modified with respect to chr-NSI by the addition, substitution, or deletion of at least one nucleotide.

[**0164**] One factor to consider regardless of the specific type of re-NSI generated is the overall homology between the re-NSI and the chr-NSI. Typically the re-NSI and chr-NSI will have at least two regions of homology facilitating homologous recombination. These regions will be at least about 10 bp in length where at least about 50 bp in length is preferred and wherein lengths of up to about 2000 bp are suitable.

[**0165**] When preparing a mutant re-NSI for use in the selection processes described herein, it is preferable to have regions homologous to the chr-NSI flanking (both 5' and 3') the site of mutation. More preferably, a 1 kB region of homology is preferred on both sides of the addition, substitution, or deletion. In contrast, re-NSI is not expected to be limited in length, beyond the limitations inherent to homologous recombination. Thus, it is contemplated that the present invention is suitable for recombination of re-NSI fragments that range in size from about 0.10-20 kB; more preferred from about 0.5-10 kB; and most preferably from about 1-5 kB.

[**0166**] Re-NSI may, for example, represent a portion of a gene, the complete coding sequence of a gene, a series of genes or an operon, a promoter, a non-coding sequence of DNA, or any combination thereof. The nucleic acids may be derived from any source including microbial sources (e.g., bacterial, yeast, fungi or algae), as well as from plant, animal or human sources and may have medical, veterinary, environmental, nutritional or industrial significance. Additionally, re-NSI's or mutant re-NSI's can be produced synthetically.

[**0167**] Replacement Nucleotide Sequences Of Interest for One-Step Selection

[**0168**] Insertion of a genetic selectable marker within the re-NSI for use in the one-step selection process can be readily generated using methods well-known to one of skill in the art (e.g., ligation of selectable markers into specific restriction sites, use of overlap PCR methods to insert selectable markers into the re-NSI, or in vitro transposition to randomly insert into the re-NSI).

[**0169**] In a preferred embodiment, the present method employs a method of in vitro transposition to disrupt the re-NSI (this approach is referred as "targeted insertion"). In vitro transposition involves the use of a transposable element in combination with a transposase enzyme. For a discussion of transposase use and function see Haren et al. (*Annu. Rev. Microbiol.,* 53:245-281 (1999)). When the transposable element or transposon is contacted with a nucleic acid fragment in the presence of the transposase, the transposable element will randomly insert into the nucleic acid fragment. The technique is useful for rapid disruption of a gene, when the specific site of disruption within the gene is not relevant; nonetheless, it is possible to quickly determine the site of disruption on the basis of the sequence of the transposable element.

[**0170**] Kits for in vitro transposition are commercially available. For example, see: 1.) The Primer Island Transposition Kit (available from Perkin Elmer Applied Biosystems, Branchburg, N.J.), based upon the yeast Ty1 element (including the AT2 transposon); 2.) The Genome Priming System (available from New England Biolabs, Beverly, Mass.), based upon the bacterial transposon Tn7; and 3.) the EZ::TN Transposon Insertion Systems (available from Epicentre Technologies, Madison, Wis.), based upon the Tn5 bacterial transposable element).

[0171] Although not available commercially, several other in vitro transposition reactions have been reported in the literature. A DNA sequencing tool was developed based on the bacteriophage Mu in vitro transposition reaction (Haapa et. al., *Genome Research,* 9:308-315 (1999)). Similarly, a tool for DNA sequencing and mutagenesis was developed based on the bacterial transposon Tn552 in vitro transposition system (Griffin et. al., *Nucleic Acid Research,* 27(19):3859-3865 (1999)). The mariner transposable element Himar1 also has an in vitro transposition reaction used for mutagenesis and DNA sequencing (Lampe et. al., *Proc. Natl. Acad. Sci.,* 96:11428-11433 (1999)).

[0172] Thus, transposons suitable in the present invention include, but are not limited to those based upon: the yeast Ty1 element, the bacterial transposon Tn7, EZ::TN, bacteriophage Mu, bacterial transposon Tn552, and the mariner transposable element Himar1. A number of transposons and methods of identifying and isolating transposons are reviewed by Dyson, Paul J. (*Methods Microbiol.* 29 (Genetic Methods for Diverse Prokaryotes):133-167 (1999)). Although these specific transposon systems have been developed for use in in vitro systems, it is contemplated that many of the transposon systems, currently only available for in vivo transposition, may be modified and developed for in vitro work. With appropriate development and characterization, these in vivo transposon systems will also be suitable within the context of the present invention.

[0173] Although any commercially available in vitro transposition system can be used as a mutagenizing tool, the Tn7-based in vitro transposition system (New England Biolabs, Beverly, Mass.) is preferred. In addition to being able to customize the Tn7-based element with an appropriate selectable or screenable genetic marker for the microorganism under study, the insertion of the Tn7-based transposon into any DNA target molecule renders that molecule "immune" to further insertions by the Tn7-based element (Stellwagen and Craig, *The EMBO Journal,* 16(22):6823-6834 (1997)). The genetic marker is used to identify transformants having the transposon disrupted re-NSI. The genetic marker may be associated with the transposon in any fashion; however, it is particularly suitable if the genetic marker is located between the ends of the transposon. Most common are antibiotic resistance markers encoding, e.g., ampicillin-resistance, kanamycin-resistance, tetracycline-resistance, etc. Also suitable are genetic markers encoding metal resistance, substrate-utilization, and genes encoding fluorescent and bioluminescent proteins (e.g., green fluorescent proteins, Lux genes), as well as lacZ, gfp, cat, galK, inaZ, luc, luxAB, bgaB, nptll, phoA, uidA, and xylE.

[0174] Markers may be used singly or in association with other agents to identify allelic exchange. For example, tetracycline can also be used in combination with detergents for the detection of cells that have lost the non-homologous region of the chromosomal integration vector (Maloy and Nunn, *J. Bact.,* 145(2):1110-1112 (1981)).

[0175] Targeted insertion may be performed directly upon an isolated re-NSI. The library of DNA fragments containing insertions may then be cloned into the chromosomal insertion vector. Or, the re-NSI may be first cloned into the chromosomal integration vector and subjected to the in vitro transposition reaction. In either case, the chromosomal integration vector is then used in homologous recombination to remove the chr-NSI from the bacterial host and insert the re-NSI.

[0176] Replacement Nucleotide Sequences Of Interest for Two-Step Selection

[0177] Generation of a mutant re-NSI (wherein the mutation does not constitute a genetic selectable marker, as discussed above for the one-step selection method) containing a mutation (e.g. addition, substitution, or deletion of at least one nucleotide with respect to the chr-NSI) can be accomplished using numerous techniques known to a skilled artisan in the field of molecular biology. In all cases, the advantage of the two-step selection methodology as compared to that for the one-step selection method is that the double-crossover mutant thus generated is "markerless". This enables subsequent mutations to be created using the same technique (i.e., since there is no need for a different selectable marker corresponding to each mutation created).

[0178] Although not intended to be limiting, deletions and additions may be generated by the use of restriction endonucleases, in vitro transposition reactions, or PCR methodologies, as is well known to one of skill in the art.

[0179] Substitutions may be generated by mutagenesis of the re-NSI. Two suitable approaches include error-prone PCR (Leung et al., *Techniques,* 1:11-15 (1989); Zhou et al., *Nucleic Acids Res.,* 19:6052-6052 (1991); and Spee et al., *Nucleic Acids Res.,* 21:777-778 (1993)) and in vivo mutagenesis. The principal advantage of error-prone PCR is that all mutations introduced by this method will be within the re-NSI, and any change may be easily controlled by changing the PCR conditions. Alternatively, in vivo mutagenesis may be employed using commercially available materials such as *E. coli* XL1-Red strain, and the *Epicurian coli* XL1-Red mutator strain from Stratagene (La Jolla, Calif.; Greener and Callahan, *Strategies,* 7:32-34 (1994)). This strain is deficient in three of the primary DNA repair pathways (mutS, mutD, and mutT), resulting in a mutation rate 5000-fold higher than that of wild-type. In vivo mutagenesis does not depend on ligation efficiency (as with error-prone PCR); however, a mutation may occur at any region of the vector and the mutation rates are generally much lower.

[0180] It is also contemplated that it may be desirable to replace a wild type gene of interest (i.e., chr-NSI) in the C1 metabolizing bacteria with a mutant enzyme (i.e., re-NSI) that has been constructed using the method of "gene shuffling" (U.S. Pat. No. 5,605,793; U.S. Pat. No. 5,811,238; U.S. Pat. No. 5,830,721; and U.S. Pat. No. 5,837,458). The method of gene shuffling is particularly attractive due to its facile implementation, and high rate of mutagenesis. The process of gene shuffling involves the restriction of a gene of interest into fragments of specific size in the presence of additional populations of DNA regions of both similarity to, or difference to, the gene of interest. This pool of fragments is then denatured and reannealed to create a mutated gene. The mutated gene is subsequently screened for altered activity.

[0181] Any wild type C1 genomic sequence may be mutated and screened for altered or enhanced activity by this method. The sequences should be double stranded and can be of various lengths ranging from about 50 bp to about 10

kB. The sequences may be randomly digested into fragments ranging from about 10 bp to about 1000 bp, using restriction endonucleases well known in the art (Maniatis, supra). In addition to the full-length sequences, populations of fragments that are hybridizable to all or portions of the sequence may be added. Similarly, a population of fragments that are not hybridizable to the wild type sequence may also be added. Typically these additional fragment populations are added in about a 10- to 20-fold excess by weight as compared to the total nucleic acid. Generally this process will allow generation of about 100 to 1000 different specific nucleic acid fragments in the mixture. The mixed population of random nucleic acid fragments are denatured to form single-stranded nucleic acid fragments and then reannealed. Only those single-stranded nucleic acid fragments having regions of homology with other single-stranded nucleic acid fragments will reanneal. The random nucleic acid fragments may be denatured by heating. One skilled in the art could determine the conditions necessary to completely denature the double-stranded nucleic acid. Preferably, the temperature is from about 80° C. to about 100° C. The nucleic acid fragments may be reannealed by cooling. Preferably the temperature is from about 20° C. to about 75° C. Renaturation can be accelerated by the addition of polyethylene glycol ("PEG") or salt. The salt concentration is preferably from about 0 mM to about 200 mM. The annealed nucleic acid fragments are next incubated in the presence of a nucleic acid polymerase and dNTP's (i.e., dATP, dCTP, dGTP and dTTP). The nucleic acid polymerase may be the Klenow fragment, the Taq polymerase or any other DNA polymerase known in the art. The polymerase may be added to the random nucleic acid fragments prior to annealing, simultaneously with annealing or after annealing. The cycle of denaturation, renaturation and incubation in the presence of polymerase is repeated for a desired number of times. Preferably the cycle is repeated from about 2 to about 50 times, more preferably the sequence is repeated from about 10 to about 40 times. The resulting nucleic acid is a larger double-stranded polynucleotide of from about 50 bp to about 100 kB and may be screened for expression and altered activity by standard cloning and expression protocols (Maniatis, supra).

[**0182**] One factor to consider during the preparation of a re-NSI for use in the two-step selection strategy concerns the placement of the mutation within the re-NSI. Specifically, the re-NSI is first inserted into the chromosome by integration of the chromosomal integration vector (a single-crossover event). The second crossover event that occurs can result in either a mutant or wildtype sequence in the chromosome, since the single-crossover contains two copies of the nucleotide sequence of interest. In order to increase the percentage of segregants that retain the re-NSI, as opposed to reverting to the wildtype encoded by the chr-NSI, it is desirable to "center" the mutation with respect to the flanking DNA that has homology to the chr-NSI. For example, if a point mutation was perfectly centered within a re-NSI, about 50% of the segregants would be expected to retain the mutation in the chromosome (thus producing a 1:1 ratio of double-crossover mutants to wild-type cells). However, it is contemplated that the present invention is suitable for homologous recombination wherein the mutation (i.e., insertion, deletion, or substitution) is positioned within about 5 to about 95% of the total length of re-NSI; more preferred when positioned within about 25 to about 75% of the total length of re-NSI; and most preferred when positioned within about 45 to about 55% of the total length of re-NSI.

Transformation of C1 Metabolizing Bacteria

[**0183**] Techniques for the transformation of C1 metabolizing bacteria are not well developed, although general methodology mimics that utilized for other bacteria and well known to those of skill in the art.

[**0184**] Electroporation has been used successfully for the transformation of: *Methylobacterium extorquens* AM1 (Toyama, H., et al., *FEMS Microbiol Lett.* 166:1-7 (1998)), *Methylophilus methylotrophus* AS1 (Kim, C. S., and T. K. Wood. *Appl Microbiol Biotechnol* 48: 105-108 (1997)), and *Methylobacillus* sp. strain 12S (Yoshida, T., et al., *Biotechnol. Lett.,* 23: 787-791 (2001)). Extrapolation of specific electroporation parameters from one specific C1 metabolizing utilizing organism to another may be difficult, however, as is well to known to one of skill in the art.

[**0185**] Bacterial conjugation, relying on the direct contact of donor and recipient cells, is frequently more readily amenable for the transfer of genes into C1 metabolizing bacteria. Simplistically, this bacterial conjugation process involves mixing together "donor" and "recipient" cells in close contact with one another. Conjugation occurs by formation of cytoplasmic connections between donor and recipient bacteria, with direct transfer of newly synthesized donor DNA into the recipient cells. As is well known in the art, the recipient in a conjugation is defined as any cell that can accept DNA through horizontal transfer from a donor bacterium; the donor in conjugative transfer is a bacterium that contains a conjugative plasmid, conjugative transposon, or mobilizable plasmid. The physical transfer of the donor plasmid can occur in one of two fashions, as described below:

[**0186**] 1. In some cases, only a donor and recipient are required for conjugation. This occurs when the plasmid to be transferred is a self-transmissible plasmid that is both conjugative and mobilizable (i.e., carrying both tra genes and genes encoding the Mob proteins). In general, the process involves the following steps: 1.) Double-strand plasmid DNA is nicked at a specific site in onT; 2.) A single-strand DNA is released to the recipient through a pore or pilus structure; 3.) A DNA relaxase enzyme cleaves the double-strand DNA at onT and binds to a release 5' end (forming a relaxosome as the intermediate structure); and 4.) Subsequently, a complex of auxiliary proteins assemble at onT to facilitate the process of DNA transfer.

[**0187**] 2. Alternatively, a "triparental" conjugation is required for transfer of the donor plasmid to the recipient. In this type of conjugation, donor cells, recipient cells, and a "helper" plasmid participate. The donor cells carry a mobilizable plasmid or conjugative transposon. Mobilizable vectors contain an onT, a gene encoding a nickase, and have genes encoding the Mob proteins; however, the Mob proteins alone are not sufficient to achieve the transfer of the genome. Thus, mobilizable plasmids are not able to promote their own transfer unless an appropriate conjugation system is provided by a helper plasmid (located within the donor or within a "helper" cell). The conjugative plasmid is needed for the formation of the mating pair and DNA

transfer, since the plasmid encodes proteins for transfer (Tra) that are involved in the formation of the pore or pilus.

[**0188**] Examples of successful conjugations involving C1 metabolizing bacteria include the work of:

[**0189**] Stolyar et al. (*Mikrobiologiya,* 64(5): 686-691 (1995)), wherein Inc P-1 incompatibility group plasmids were transferred by conjugation from *Eschenchia coli* and *Pseudomonas putida* into *Methylomonas rubra* 15sh and *Methylococcus thermophilus* 111p;

[**0190**] Motoyama, H. et al. (*Appl. Micro. Biotech.,* 42(1): 67-72 (1994)), wherein genes coding for the L-threonine biosynthetic enzymes were transferred by conjugation into the Gram-negative obligate methylotroph, *Methylobacillus glycogenes;*

[**0191**] Lloyd, J. S. et al. (*Archives of Microbiology,* 171(6): 364-370 (1999)), wherein broad-host range plasmids containing heterologous soluble methane monooxygenase (sMMO) genes were transferred into *Methylocystis parvus* OBBP and *Methylomicrobium album* BG8; and

[**0192**] Odom, J. M. et al. (WO 02/18617), wherein the broad-host range plasmid pBHR1 containing carotenoid genes of *Pantoea stewartii* were transferred into *Methylomonas* sp. 16a.

Screening Methods-SacB

[**0193**] Methods of screening in microbiology are discussed at length in Brock, supra. In one preferred embodiment, a one-step selection process permits the identification of double-crossover mutations in C1 metabolizing bacterial cells by applying positive selection pressure. In this embodiment, the positive selection vector comprises a first genetic selectable marker, a sacB marker, and a second genetic selectable marker that disrupts the sequence of the re-NSI. For selection, cells that have undergone a double-crossover event are selected by plating onto medium that permits selection according to the sacB marker on the chromosomal integration vector and the second genetic selectable marker disrupting the re-NSI. This selection strategy efficiently identifies only double-crossover events, as it relies on two selectable and screenable markers. First, the sacB marker is able to select against all single-crossover mutants, since SacB expression (which occurs when the entire chromosomal integration vector, comprising the vector plus the re-NSI, merely integrates into the host's genome) enables conversion of sucrose into levan, which is toxic to most Gram-negative bacteria. Those bacterial cells possessing a double-crossover, in contrast, do not possess the sacB gene within their chromosome; the lethal marker is present on the non-replicating chromosomal plasmid integration vector. Secondly, the second genetic selectable marker disrupting the re-NSI ensures allelic exchange has occurred between the chr-NSI and the re-NSI introduced into the cell on the chromosomal integration vector. To confirm the double-crossover event, the cells may be transferred to media containing the positive selection vector's first genetic selectable marker (and optionally containing the second genetic selectable marker); those cells unable to grow have undergone allelic exchange.

[**0194**] In an alternative preferred embodiment, a two-step selection process permits the identification of double-crossover mutations in C1 metabolizing bacterial cells by applying positive selection pressure. Using this strategy, the positive selection vector should comprise a first genetic selectable marker and a sacB marker. Selection involves first growing the transformants on media containing the first genetic selectable marker, to identify those cells that have undergone a single-crossover (i.e., wherein the entire chromosomal integration vector has integrated into the host cell's genome). Then, the selection pressure is removed and a second crossover event may occur. Selection for allelic exchange mutants requires growth of the cells on sucrose, since SacB expression will be lethal to all single-crossover mutants. Differentiation between allelic exchange mutants containing the wildtype and mutant allele is then possible using standard molecular techniques (e.g., PCR), well known to one of skill in the art.

C1 Metabolizing Bacterial Hosts

[**0195**] In one embodiment, heterologous host cells for use with the present positive selection methodologies are C1 metabolizing bacteria that are able to use single carbon substrates (i.e., those substrates lacking carbon to carbon bonds) as a sole energy source. Many C1 metabolizing microorganisms are known in the art which are able to use a variety of single carbon substrates. Single carbon substrates useful in the present invention include, but are not limited to: methane, methanol, formate, formaldehyde, formic acid, methylated amines (e.g., mono-, di- and tri-methyl amine), methylated thiols, carbon dioxide, and various other reduced carbon compounds which lack any carbon-carbon bonds. In another embodiment, the single carbon substrates are methane and/or methanol.

[**0196**] All C1 metabolizing microorganisms are generally classified as methylotrophs. Methylotrophs may be defined as any organisms capable of oxidizing organic compounds that do not contain carbon-carbon bonds. However, facultative methylotrophs, obligate methylotrophs, and obligate methanotrophs are all various subsets of methylotrophs. Specifically:

[**0197**] Facultative methylotrophs have the ability to oxidize organic compounds which do not contain carbon-carbon bonds, but may also use other carbon substrates such as sugars and complex carbohydrates for energy and biomass. Facultative methylotrophic bacteria are found in many environments, but are isolated most commonly from soil, landfill and waste treatment sites. Many facultative methylotrophs are members of the β and γ subgroups of the Proteobacteria (Hanson et al., *Microb. Growth C*1 *Compounds.,* [*Int. Symp.*], 7th (1993), pp 285-302. Murrell, J. Collin and Don P. Kelly, Eds. Intercept: Andover, UK; Madigan et al., *Brock Biology of Microorganisms,* 8th ed., Prentice Hall: UpperSaddle River, N.J. (1997)).

[**0198**] Obligate methylotrophs are those organisms that are limited to the use of organic compounds that do not contain carbon-carbon bonds for the generation of energy.

[**0199**] Obligate methanotrophs are those obligate methylotrophs that have the distinct ability to oxidize methane.

[**0200**] For the purposes of the present invention herein, suitable facultative methylotrophic bacteria include, but are

not limited to: *Methylophilus, Methylobacillus, Methylobacterium, Hyphomicrobium, Xanthobacter, Bacillus, Paracoccus, Nocardia, Arthrobacter, Rhodopseudomonas,* and *Pseudomonas*. Exemplary methanotrophs are included in, but not limited to, the genera *Methylomonas, Methylobacter, Methylococcus, Methylosinus, Methylocyctis, Methylomicrobium,* and *Methanomonas,* although this is not intended to be limiting.

[0201] Of particular interest in the present invention are high growth obligate methanotrophs having an energetically favorable carbon flux pathway. For example, Applicants have discovered a specific strain of methanotroph having several pathway features that make it particularly useful for carbon flux manipulation (U.S. Pat. No. 6,689,601; hereby incorporated by reference). This particular strain possessing an energetically favorable carbon flux pathway was used as a "model" organism in the present application and is known as *Methylomonas* 16a (ATCC PTA 2402).

Applications for Positive Selection Methods

[0202] Knowledge of the sequence of the present chromosomal integration vector and methodologies for one-step and two-step selection of double-crossover C1 metabolizing mutant bacteria as disclosed herein will be useful in manipulating various genetic pathways in any C1 metabolizing organism, particularly in methanotrophs. As is well known to those of skill in the art, efforts to genetically engineer a microorganism for high-level production of a specific product frequently require substantial manipulation to the native host machinery and biosynthetic pathways.

[0203] In the past, metabolic engineering of methanotrophs has remained relatively undeveloped due to the general lack of efficient genetic engineering tools, as compared to other industrial bacteria such as *E. coli* and the yeasts. Most methanotrophic molecular biology has focused on the engineering of the methane monooxygenase, such that it is directed toward more useful co-metabolic products (and wherein methane is not directly incorporated into the product molecule). In contrast, pathway engineering for net synthesis of carbon-containing compounds from methane requires both: 1.) the ability to inactivate genes already present in the methanotroph; and 2.) the ability to introduce foreign genes into the organism, to create new metabolic capabilities. In this way, new metabolic networks can be constructed to produce products that could not be made by the genetic engineering of methane monooxygenase alone.

[0204] The present invention provides a method to make mutations in specific genes and/or competing pathways when the nucleotide sequence of the gene to be altered is known. This permits reduction or elimination of the expression of certain genes in the target pathway or in competing pathways that may serve as competing sinks for energy or carbon. The gene alterations can be "tagged" with a genetic selectable marker; or alternatively, the mutations can be markerless, thus allowing subsequent mutations to be introduced into the chromosome using the same methodology and genetic markers for selection.

[0205] In an alternative embodiment (or in addition to the selective reduction or elimination of specific genes), it may be desirable to introduce multiple copies of a gene into a specific site of the C1 metabolizing host genome, to promote that gene's enhanced biological activity. In this instance, the gene could be foreign or endogenous to the cell. This methodology would be especially useful for industrial processes, where it is frequently undesirable to require constant selection pressure for maintenance of a specific expression plasmid. Additionally, the ability to target the insertion event to specific genes within the host chromosome is advantageous since the controlled insertion event will not unknowingly disrupt other essential genes and pathways of the host organism.

[0206] Once a key genetic pathway has been identified and sequenced, it may also be necessary to optimize specific genes to increase the output of the pathway. For example, a gene encoding a rate-limiting enzyme within the host organism may be subjected to a "gene shuffling" reaction or error-prone PCR reaction. Upon identification of a mutant gene having enhanced biological activity for the rate-limiting reaction, it may be desirable to insert this mutated gene containing point mutations into the chromosome of the C1 metabolizing host in place of the wildtype gene.

[0207] As such, the positive selection methods described herein are expected to be useful to create allelic exchange mutations in a variety of metabolic pathways within methanotrophic hosts. Useful metabolic pathways include, but are not limited to: glycolytic pathways for the production of C_3 and C_2 molecules, pentose phosphate pathways for the production of C_3, C_4, C_5, C_6 and C_7 sugars, aromatic amino acid biosynthesis pathways for the production of a variety of aromatics, pathways leading to the intra- and extra-cellular production of polysacharrides, and isoprenoid biosynthesis pathways.

[0208] Finally, one skilled in the art will recognize that the use of a two-step selection strategy allows the construction of defined unmarked mutants with changes that are usually undetectable, such as point mutations or in-frame deletions. These "markerless" mutants may offer significant advantages during industrial development of a C1 metabolizing host bacterium, in terms of required regulatory approvals. For example, isoprenoids are actively made by many methanotrophs and metabolic engineering of this pathway offers the ability to produce a wide variety of carotenoid compounds. Using the positive selection methods described herein, it is possible to create allelic exchange mutations in genetically engineered hosts expressing carotenoid compounds that will be useful in various food and feed applications, in which the presence of antibiotic resistance genes would be undesirable.

DESCRIPTION OF PREFERRED EMBODIMENTS

[0209] *Methylomonas* sp. 16a is a C1 metabolizing bacteria that is capable of efficiently using either methanol and/or methane as a carbon substrate. *Methylomonas* is also metabolically versatile in that it contains multiple pathways for the incorporation of carbon from formaldehyde into 3-carbon units, and is capable of genetic exchange with donor species such as *Escherichia coli* via bacterial conjugation. Despite these advantageous qualities, the utility of this organism as a microbial platform of choice for the production of a variety of materials was hampered prior to the development of a method for the generation and identification of defined mutations within the organism by homologous recombination. Previous attempts to create

mutations via a double-crossover event were unsuccessful, largely due to the organism's slow growth rate, difficulties associated with identification of a suitable chromosomal integration vector, and the rarity of double-crossover events. Furthermore, no prior studies had confirmed the ability of *Methylomonas* sp. 16a to undergo homologous recombination.

[**0210**] The present study was initiated to determine if the sacB gene could be employed as a positive selectable marker in C1 metabolizing bacterium, such as *Methylomonas*. Thus, the effect of sucrose on the growth of *Methylomonas* sp. 16a was first examined by growing the wildtype bacteria on media containing 1%, 2.5%, 5%, and 10% sucrose. Since sucrose did not appear to interfere with normal growth, work aimed at the development of a positive selection vector based on sacB was pursued.

[**0211**] Although a variety of metabolic pathways within *Methylomonas* sp. 16a could have been chosen as a model pathway to investigate the utility of the positive selection methods described herein, it was desirable to focus work on the native C_{30} carotenoid biosynthetic pathway of *Methylomonas* sp. 16a since it was anticipated that creation of particular allelic exchange mutations within this pathway would result in a non-pigmented bacterial strain. The transformation from wildtype (pink-pigmented) host to a transformed non-pigmented host would provide visual confirmation of allelic exchange.

[**0212**] A carotenoid biosynthetic operon has been identified within the genomic sequence of *Methylomonas* sp. 16a, which encodes three genes (**FIG. 3**):

[**0213**] The first gene (designated crtN1; SEQ ID NO:1) encodes a putative diapophytoene dehydrogenase with the highest BLAST hit to a diapophytoene dehydrogenase from *Heliobacillus mobilis* (34% identity and 58% similarity).

[**0214**] The middle gene (designated ald; SEQ ID NO:2) encodes a putative aldehyde dehydrogenase with the highest BLAST hit to a betaine aldehyde dehydrogenase from *Arabidopsis thaliana* (33% identity and 50% similarity).

[**0215**] The third gene (designated crtN2; SEQ ID NO:3) also encodes a putative diapophytoene dehydrogenase with the highest BLAST hit to a hypothetical protein of phytoene dehydrogenase family from *Staphylococcus aureus* (51 % identity and 67% similarity).

The interaction between the proteins encoded by these three genes in *Methylomonas* sp. 16a is not well understood at this time.

[**0216**] The ability of the one-step selection process to identify allelic exchange mutants was first investigated herein, by creating a re-NSI wherein the native crtN2 gene of *Methylomonas* was disrupted via in vitro transposition using the EZ::TN™ <Kan-2> transposon Insertion Kit (Epicentre Technologies; Madison, Wis.). Following triparental conjugation to mobilize the chromosomal integration vector pGP704::sacB::crtN2::EZ::TN™ <Kan-2> into Methylomonas, it was found that crtN2 mutants resulting from allelic-exchange events were readily identified by the one-step selection process.

[**0217**] Subsequently, allelic-exchange mutants wherein the native *Methylomonas* ald(crtN1) gene was deleted from the genome were also identified using the two-step selection process. The allelic-exchange mutants generated contained a major deletion within the ald gene (all but ~80 nucleotides were removed) and a small deletion within the carboxyl terminus of the crtN1 gene (~30 nucleotides remained). Specifically, two DNA fragments flanking ald(crtN1) were prepared by PCR and ligated together within the chromosomal integration vector pGP704::sacB to create a markerless re-NSI. Upon mobilization of the pGP704::sacB::Δa/d(crtN1) construct into Methylomonas, application of the two-step selection strategy permitted isolation of transformant clones lacking the ald gene and lacking any residual selective markers.

[**0218**] The results described herein and in the following Examples clearly illustrate the utility of sacB, as a means to facilitate experiments designed to identify allelic-exchange events in C1 metabolizing bacteria using positive selection.

EXAMPLES

[**0219**] The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

General Methods

[**0220**] Standard recombinant DNA and molecular cloning techniques used in the Examples are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual;* Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y. (1989) (Maniatis); by T. J. Silhavy, M. L. Bennan, and L. W. Enquist, *Experiments with Gene Fusions,* Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y. (1984); and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology,* published by Greene Publishing Assoc. and Wiley-Interscience (1987).

[**0221**] The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "hr" means hour(s), "d" means day(s), "μL" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "μM" means micromolar, "mM" means millimolar, "M" means molar, "mmol" means millimole(s), "μmole" mean micromole(s), "g" means gram(s), "μg" means microgram(s), "ng" means nanogram(s), "U" means unit(s), "bp" means base pair(s), and "kB" means kilobase(s).

Molecular Biology Techniques:

[**0222**] Methods for agarose gel electrophoresis were performed as described in Maniatis (supra). Polymerase Chain Reactions (PCR) techniques were found in White, B., *PCR Protocols: Current Methods and Applications,* Humana: Totowa, N.J. (1993), Vol. 15.

Media and Culture Conditions:

[**0223**] General materials and methods suitable for the maintenance and growth of bacterial cultures are found in:

Experiments in Molecular Genetics (Jeffrey H. Miller), Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y. (1972); *Manual of Methods for General Bacteriology* (Phillip Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology: Washington, D.C., pp 210-213; or, Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology,* 2nd ed. Sinauer Associates: Sunderland, Mass. (1989). All reagents and materials used for the growth and maintenance of bacterial cells were obtained from Aldrich Chemicals (Milwaukee, Wis.), DIFCO Laboratories (Detroit, Mich.), Gibco/BRL Life Technologies (Gaithersburg, Md.), or Sigma Chemical Company (St. Louis, Mo.), unless otherwise specified.

Example 1

Growth of Methylomonas 16a

[0224] Example 1 summarizes the standard conditions used for growth of *Methylomonas* sp. 16a (ATCC# PTA-2402), as described in U.S. Pat. No. 6,689,601.

Methylomonas Strain and Culture Media

[0225] The growth conditions described below were used throughout the experimental Examples for treatment of *Methylomonas* 16a, unless conditions were specifically described otherwise.

[0226] *Methylomonas* 16a is typically grown in serum stoppered Wheaton bottles (Wheaton Scientific, Wheaton, Ill.) using a gas/liquid ratio of at least 8:1 (i.e., 20 mL of Nitrate liquid "BTZ" media in 160 mL total volume). The standard gas phase for cultivation contained 25% methane in air, although methane concentrations can vary ranging from about 5-50% by volume of the culture headspace. These conditions comprise growth conditions and the cells are referred to as growing cells. In all cases, the cultures were grown at 30° C. with constant shaking in a Lab-Line rotary shaker unless otherwise specified.

[0227] Ammonium and Nitrate Medium for Methylomonas 16a

[0228] *Methylomonas* 16a typically grows in a defined medium comprised of only minimal salts; no organic additions such as yeast extract or vitamins are required to achieve grow. This defined medium known as BTZ medium (also referred to herein as "ammonium liquid medium") was comprised of various salts mixed with Solution 1 as indicated below (Tables 2 and 3). Alternatively, the ammonium was replaced with 10 mM sodium nitrate, where specified. Solution 1 provides the composition for a 100 fold concentrated stock solution of trace minerals.

TABLE 2

Solution 1*	MW	Conc. (mM)	g per L
Nitriloacetic acid	191.10	66.90	12.80
$CuCl_2 \times 2H_2O$	170.48	0.15	0.0254
$FeCl_2 \times 4H_2O$	198.81	1.50	0.30
$MnCl_2 \times 4H_2O$	197.91	0.50	0.10
$CoCl_2 \times 6H_2O$	237.90	1.31	0.312
$ZnCl_2$	136.29	0.73	0.10
H_3BO_3	61.83	0.16	0.01
$Na_2MoO_4 \times 2H_2O$	241.95	0.04	0.01
$NiCl_2 \times 6H_2O$	237.70	0.77	0.184

*Mix the gram amounts designated above in 900 mL of H_2O, adjust to pH=7, and add H_2O to an end volume of 1 L. Keep refrigerated.

[0229]

TABLE 3

Ammonium liquid medium (BTZ)**	MW	Conc. (mM)	g per L
NH_4Cl	84.99	10	0.85
KH_2PO_4	136.09	3.67	0.5
Na_2SO_4	142.04	3.52	0.5
$MgCl_2 \times 6H_2O$	203.3	0.98	0.2
$CaCl_2 \times 2H_2O$	147.02	0.68	0.1
1 M HEPES (pH 7)	238.3		50 mL
Solution 1			10 mL

**Dissolve in 900 mL H_2O. Adjust to pH=7, and add H_2O to give 1 L. For agar plates: Add 15 g of agarose in 1 L of medium, autoclave, cool liquid solution to 50° C., mix, and pour plates.

Example 2

Construction of a Positive-Selective Suicide Vector for *Methylomonas* sp. 16a

[0230] The construction of chromosomal mutations within the *Methylomonas* genome required the use of suicide vectors. Thus, a modified version of the conditional replication vector pGP704 was created, comprising a npr-sacB cassette.

pGP704 as a Vector Backbone for the C1 Chromosomal Integration Vector

[0231] The plasmid pGP704 (Miller and Mekalanos, *J. Bacteriol.,* (170): 2575-2583 (1988); **FIG. 4**) was chosen as a suitable vector backbone for the C1 chromosomal integration vector, since it could be used as a vehicle to transfer replacement nucleotide sequences of interest into *Methylomonas* sp. 16a via conjugation. Plasmid pGP704 is a derivative of pBR322 that is Amp^R but has a deletion of the pBR322 origin of replication (oriE1). Instead, the plasmid contains a cloned fragment containing the origin of replication of plasmid R6K. The R6K origin of replication (oriR6K) requires the Π protein, encoded by the pir gene. In *E. coli*, the Π protein can be supplied in trans by a prophage (γ pir) that carries a cloned copy of the pir gene. The pGP704 plasmid also contains a 1.9 kB BamHl fragment encoding the mob region of RP4. Thus, pGP704 can be mobilized into recipient strains by transfer functions provided by a derivative of RP4 integrated in the chromosome of *E. coli* strain SM10 or SY327. Once the plasmid is transferred, however, it is unable to replicate in recipients that lack the Π protein (e.g., recipients such as *Methylomonas* and other C1

metabolizing bacteria). This inability permits homologous recombination to occur between nucleotide sequences of interest on pGP704 and the intact chromosomal nucleotide sequences of interest.

[0232] Thus, on the basis of the above characteristics, the pGP704 vector backbone met the following conditions for a chromosomal integration vector suitable C1 metabolizing bacteria: 1.) it was conditional for replication, thus allowing selection for integration into the chromosome; 2.) it possessed at least one selectable marker; 3.) it had an origin of transfer that was expected to be suitable for C1 metabolizing bacteria; 4.) it possessed mobilization genes; and 5.) it contained a variety of unique cloning sites. Other alternative chromosomal integration vectors having the characteristics above are expected to be suitable for use in the present invention, as described herein.

[0233] Plasmid pGP704 did not, however, permit easy detection and identification of clones that had undergone allelic exchange. Thus, pGP704 was modified to permit the positive selection of double-crossover events within *Methylomonas* and other C1 metabolizing bacteria.

Cloning of the npr-sacB Cassette

[0234] Plasmid pBE83 contained a *Bacillus amyloliquifaciens* sacB gene under the control of the neutral protease (npr) promoter (gift from V. Nagarajan, E.I. du Pont de Nemours and Co., Inc.). The npr-sacB cassette was PCR amplified from pBE83 using DNA primers DrdI/npr-sacB and TthIII/npr-sacB. The DNA primers were constructed to include unique restriction sites at each terminus of the PCR product, to facilitate subsequent cloning (as indicated by the underlined sequences below):

```
DrdI/npr-sacB:
5'-GACATCGATGTCGAATTCGAGCTCGGTACCGATC-3'    (SEQ ID NO:4)

TthIII/npr-sacB:
5'-GACCTCGTCGCTGTTATTAGTTGACTGTCAGC-3'      (SEQ ID NO:5)
```

[0235] The PCR reaction mixture was composed of the following: 10 μL of 10× PCR buffer; 16 μL (4 μL each) of dNTPs (320 mM stock); 1 μL *Methylomonas* chromosomal DNA (~500 ng/μL); 8 μl $MgCl_2$ (25 mM); 0.5 μL Taq polymerase (5 U/μL); 1 μL DrdI/npr-sacB primer (~36 nmoles); 1 μL TthIII/npr-sacB primer (~35 nmoles); and 71 μL sterile nanopure water. The PCR protocol was the performed on a 9600 GeneAmp PCR System (Perkin Elmer), according to the thermocycling parameters below:

[0236] 1 cycle: 94° C. (5 min);

[0237] 1 cycle: 94° C. (5 min), 60° C. (2 min), 72° C. (3 min);

[0238] 35 cycles: 94° C. (1 min), 60° C. (2 min), 72° C. (3 min);

[0239] 1 cycle: 94° C. (1 min), 60° C. (2 min), 72° C. (10 min); and

[0240] Hold −4° C. (∞).

Afterward, the PCR product was ligated into the pCR2.1TOPO vector per the manufacturer's instructions (Invitrogen, Carlsbad, Calif.). The ligation mixture was transformed into TOP10 One Shot™ calcium chloride competent cells and transformants were screened as recommended by Invitrogen.

[0241] Plasmid DNA was isolated from positive clones (white colonies in a blue/white screen) using the QIAprep® Spin Mini-prep Kit (Qiagen, Valencia, Calif.) and the DNA was digested according to the manufacturer's instructions with restriction endonucleases DrdI and TthIII (New England Biolabs; Boston, Mass.). Initially, this PCR product was to be inserted into pGP704 digested with DrdI and TthIII; however, there were difficulties in cloning the DrdI/TthIII PCR product.

[0242] A modified cloning strategy was adopted, such that the PCR reaction of above was "repeated" using the PFU polymerase (Stratagene; LaJolla, Calif.). Specifically, the PCR reaction and protocol were performed exactly as described above, with the exception that PFU polymerase and buffers from Stratagene were used. A PCR product having flush or blunt ends was produced. This PCR product was ligated directly into the XcaI site of pGP704. The ligation mixture was transformed into calcium chloride competent *E. coli* SY327 cells (Miller, V., L., and Mekalanos, J., *J., Proc. Natl. Acad. Sci.,* 81(11):3471-3475 (1984)).

[0243] The transformants were screened using the DrdI/npr-sacB and TthIII/npr-sacB PCR primers (SEQ ID NOs: 4 and 5, respectively) to identify vectors containing the npr-sacB insert. The PCR products were analyzed on a 0.8% agarose gel. Plasmid DNA was isolated from cells containing the pGP704::sacB vector.

Example 3

Construction of pGP704::sacB::crtN2::EZ::TN™ <Kan-2>

[0244] The native *Methylomonas* gene encoding crtN2 was disrupted by EZ::TN™ <Kan-2> (Epicenter Technologies, Madison, Wis.) to generate an appropriate replacement nucleotide sequence of interest (i.e., re-NSI) in the chromosomal integration vector pGP704::sacB. This construct, pGP704::sacB::crtN2::EZ::TN™ <Kan-2> was then utilized in Examples 4 and 5 to test the homologous recombination abilities of *Methylomonas* sp. 16a, using the one-step selection strategy to identify allelic exchange mutants.

PCR Amplification and Cloning of the crtN2 DNA Fragment into pGP704::sacB

[0245] DNA primers crtN#2-ctg 288+1 kB/Bg/ll and crtN#2-ctg 288+1kB/XbaI were used to amplify a ~3.1 kB DNA fragment from *Methylomonas* sp. 16a chromosomal DNA, comprising SEQ ID NO:3 flanked by approximately 1 Kb of DNA.

crtN#2-ctg 288 + 1kB/Bg/II:
5'-AGATCTTTCCGGTCATGCTGGAGTTGGG-3' (SEQ ID NO:6)

crtN#2-ctg 288 + 1kB/XbaI:
5'-TCTAGATTGCAGCTCAAGCGATTCGG-3' (SEQ ID NO:7)

[0246] As described in Example 2, the underlined sequences represent restriction endonuclease recognition sites; additionally, the PCR reactions and cloning protocol were the same as described in Example 2. The PCR fragment generated above was cloned in the pCR2.1 TOPO vector and then transformed into calcium chloride competent *E. coli* TOP10 One Shot™ cells (Invitrogen, Carlsbad, Calif.). Following screening for positive clones, plasmid DNA was isolated from several positive clones using the QIAprep® Spin Mini-prep Kit and the DNA was digested with restriction endonucleases Bglll and Xbal according to manufacturer's instructions. The restriction digestion products were separated on a 0.8% agarose gel. The crtN2+1kB DNA fragment (~3.1 kB) was excised from the agarose gel and the DNA was extracted from the agarose using the QiaQuick® Gel Extraction Kit (Qiagen).

[0247] The purified crtN2+1 kB DNA fragment was ligated into pGP704-sacB (from Example 2) digested with Bg/ll and Xbal. After an overnight incubation at room temperature, the ligation mixture was transformed into calcium chloride competent *E. coli* SY327 (γ pir+) cells (Miller, V., L., and Mekalanos, J., *J., Proc. Natl. Acad. Sci.,* 81(11):3471-3475 (1984)). The transformation mixture was plated onto BTZ medium containing 25 μg/mL Amp. *E. coli* transformants containing the appropriate insert DNA fragment were detected using PCR amplification with DNA primers crtN#2-ctg 288+1 kB/Bg/ll (SEQ ID NO:6) and crtN#2-ctg 288+1 kB/Xbal (SEQ ID NO:7). Plasmid DNA was purified from positive clones using the QIAprep® Spin Mini-prep Kit (Qiagen) and was used as the "target DNA" in an in vitro transposition reaction.

The in vitro Transposition Reaction

[0248] The crtN2 gene was disrupted via in vitro transposition using the EZ::TN™ <Kan-2> transposon Insertion Kit (Epicentre Technologies; Madison, Wis.). The transposon donor DNA used in the transposition reactions was the EZ::TN™ Transposon (a linear DNA fragment), which carries the gene that confers resistance to kanamycin between the ends of the transposable element. As per the manufacturer's instructions, 0.2 μg of the target DNA was incubated with molar equivalents of the EZ::TN™ <Kan-2> transposon and 1 U of the EZ::TN™ <Kan-2> transposase. The reaction mixture was incubated for 2 hrs at 37° C. The transposition reaction was stopped by adding 1 μL of the 1× Stop Solution and incubating the mixture for 10 min at 70° C.

[0249] PCR amplification using gene-specific primers crtN#2-ctg 288+1 kB/Bg/ll (SEQ ID NO:6) or crtN#2-ctg 281+1 kB/Xbal (SEQ ID NO:7) and transposon-specific primers KAN-2FP-1 (SEQ ID NO:8) and KAN-2RP-1 (SEQ ID NO:9) were used to determine the location of the EZ::TN™ <Kan-2> transposon within the crtN2 gene. One clone was identified, in which the EZ::TN™ <Kan-2> transposon was located near the middle of the crtN2 gene. This crtN2:: EZ::TN™ <Kan-2> construct in pGP704:sacB served as the donor plasmid in a tri-parental mating reaction (see Example 4).

Example 4

Triparental Conjugation of pGP704:sacB:crtN2::EZ::TN™ <Kan-2> into *Methylomonas* sp. 16a

[0250] The pGP704::sacB::crtN2::EZ::TN™ <Kan-2> construct from Example 3 was transferred into *Methylomonas* sp. 16a, to test the homologous recombination abilities of the organism. A triparental conjugation was performed, using the following: *Methylomonas* sp. 16a (recipient), *E. coli* SY327 pGP704:sacB:crtN2::EZ::TN™ <Kan-2> (donor; from Example 3), and *E. coli* pRK2013 (helper; ATCC No. 37159).

Theory of the Conjugation

[0251] The mobilization of vector DNA into *Methylomonas* occurs through conjugation (tri-parental mating). The pGP704::sacB vector used to make chromosomal mutations in *Methylomonas* has a R6K origin of replication, which requires the FI protein. This vector can replicate in *E. coli* strain SY327, which expresses the Π protein. However, this protein is not present in the *Methylomonas* genome. Therefore, once the vector DNA has entered into *Methylomonas,* it is unable to duplicate itself. If the vector also contains a DNA segment that shares homology to a region of the *Methylomonas* genome, the vector can be integrated into the host's genome through homologous recombination. The homologous recombination system of *Methylomonas* appears to be similar to that of other Gram-negative organisms.

[0252] In the case of *Methylomonas,* the mobilizable plasmid (pGP704::sacB) was used to transfer re-NSI into this bacterium. The conjugative plasmid (pRK2013), which resided in a strain of *E. coli,* facilitated the DNA transfer.

Growth of *Methylomonas* sp. 16a

[0253] The growth of *Methylomonas* sp. 16a for tri-parental mating initiated with the inoculation of an –80° C. frozen stock culture into 20 mL of BTZ medium containing 25% methane. The culture was grown at 30° C. with aeration until the density of the culture was saturated. This saturated culture was in turn used to inoculate 100 mL of fresh BTZ medium containing 25% methane. The 100 mL culture was grown at 30° C. with aeration until the culture reached an OD_{600} between 0.7 to 0.8. To prepare the cells for the tri-parental mating, the *Methylomonas* sp. 16a cells were washed twice in an equal volume of BTZ medium. The *Methylomonas* cell pellets were re-suspended in the minimal volume needed (approximately 200 to 250 μL). Approximately 40 μL of the re-suspended *Methylomonas* cells were used in each tri-parental mating experiment.

Growth of the *Escherichia coli* Donor and Helper Cells

[0254] Isolated colonies of the *E. coli* donor (containing pGP704::sacB::crtN2::EZ::TN™ <Kan-2>) and helper (containing conjugative plasmid pRK2013) cells were used to inoculate 5 mL of LB broth containing 25 μg/μL Kan; these cultures were grown overnight at 30° C. with aeration. The following day, the *E. coli* donor and helper cells were

mixed together and incubated at 30° C. for ~2 hours. Subsequently, the cells were washed twice in equal volumes of fresh LB broth to remove the antibiotics.

Tri-Parental Mating: Mobilization of the Donor Plasmid into *Methylomonas* Strain 16a

[**0255**] Approximately 40 μL of the re-suspended *Methylomonas* cells were used to re-suspend the combined *E. coli* donor and helper cell pellets. After thoroughly mixing the cells, the cell suspension was spotted onto BTZ agar plates containing 0.05% yeast extract. The plates were incubated at 30° C. for 3 days in a jar containing 25% methane.

[**0256**] Following the third day of incubation, the cells were scraped from the plate and re-suspended in BTZ broth. The entire cell suspension was plated onto several BTZ agar plates containing 25 μg/μL Kan. The plates were incubated at 30° C. in a jar containing 25% methane until colonies were visible (~4-7 days).

[**0257**] Individual colonies were streaked onto fresh BTZ+ 25 μg/μL Kan agar plates and incubated at 30° C. in the presence of 25% methane. Cells growing on the BTZ+Kan agar plates were used to inoculate BTZ+Kan medium in bottles containing 25% methane. The bottles were incubated overnight at 30° C. with aeration. Then, to rid the cultures of *E. coli* cells that were introduced during the tri-parental mating, the cultures were inoculated into BTZ medium which contained nitrate (15 mM) as the nitrogen source and methanol (200 mM) as the carbon source, and incubated overnight at 30° C. with aeration. The cultures were monitored for *E. coli* killing by plating onto LB agar plates.

Example 5

Evaluation of *Methylomonas* Transconjugants Containing pGP704:sacB:crtN2::EZ::TN™ <Kan-2>

[**0258**] Following the mobilization of the pGP704::sacB::crtN2::EZ::TN™ <Kan-2> construct into *Methylomonas* sp. 16a, the one-step selection strategy to applied as described below to identify allelic exchange mutants (**FIG. 1**).

Preliminary Screening for Allelic Exchange Mutants

[**0259**] Cultures free of *E. coli* cells (from Example 4) were transferred back into BTZ+Kan medium and grown overnight at 30° C. with aeration. Cells were plated onto BTZ+Kan, BTZ+Amp (35 μg/μL), and BTZ+Kan+Sucrose (5%) agar plates. The ability of cells to grow on the BTZ +Kan +Sucrose agar plates suggested that the cells had undergone a double-crossover event, wherein the transposon-disrupted crtN2 gene (crtN2::EZ::TN™ <Kan-2>) had replaced the intact crtN2 gene in the *Methylomonas* chromosome.

[**0260**] To confirm the allelic exchange, individual colonies were streaked onto BTZ+Kan, BTZ+Amp, and BTZ+ Kan+Sucrose agar plates. Those clones identified as double-crossover mutants (i.e., those where crtN2 appeared to be disrupted with the EZ::TN™ <Kan-2> transposon) were able to grow on the BTZ+Kan and BTZ+Kan+Sucrose agar plates, but not on the BTZ+Amp agar plates.

Confirmation of Allelic Exchange Mutants

[**0261**] PCR methodology was used to confirm the construction of the Methylomonas crtN2 mutant using transposon-specific and outside primers (OP). OP primers were homologous to regions outside of the gene-specific primers. Thus, PCR reactions were performed using the following combination of primers:

```
Primer Set #1:   OP#1crtN#2;     (SEQ ID NO:10)
                 KAN-2-FP-1;     (SEQ ID NO:8)
and
                 KAN-2-RP-1.     (SEQ ID NO:9)

Primer Set #2:   OP#2crtN#2;     (SEQ ID NO:11)
                 KAN-2-FP-1;     (SEQ ID NO:8)
and
                 KAN-2-RP-1.     (SEQ ID NO:9)
```

[**0262**] A PCR fragment of ~2.6 kB was obtained when Primer Set #1 was used in the PCR reaction. In addition, a PCR fragment of ~2.8 kB was detected for the PCR reaction using Primer Set #2. Since neither OP#1crtN#2 (SEQ ID NO:10) nor OP#2crtN#2 (SEQ ID NO:11) primer sequences were present on the integration vector, PCR products should only be detected if re-NSI had integrated into the *Methylomonas* genome at the appropriate site. Thus, these PCR reactions confirmed the integration of the transposon-disrupted crtN2 gene into the genome of *Methylomonas* and the utility of the one-step selection method for identification of allelic exchange mutants.

Example 6

Construction of pGP704::sacB::Δald(crtN1)

[**0263**] The native *Methylomonas* ald gene (SEQ ID NO:2), normally located between the crtN1 (SEQ ID NO:1) and crtN2 (SEQ ID NO:3) genes within a single carotenoid gene cluster (**FIG. 3**), was deleted from the operon structure to generate an appropriate replacement nucleotide sequence of interest (i.e., re-NSI) in the chromosomal integration vector pGP704::sacB. Specifically, two DNA fragments flanking ald were prepared by PCR (i.e., crtN1/ald deletion fragment #1 and crtN1/ald deletion fragment #2) and ligated together within pGP704::sacB. The crtN1/ald deletion fragment #1 also caused the deletion of ~10 amino acids from the carboxyl terminus on the CrtN1 protein. The resulting construct, pGP704::sacB::Δaid(crtN1), was then utilized in Examples 7 and 8 to test the homologous recombination abilities of *Methylomonas* sp. 16a, using the two-step selection strategy to identify allelic exchange mutants in which ald(crtN1) had been deleted from the *Methylomonas* sp. 16a genome.

PCR Amplification and Cloning of the Δald(crtN1) DNA Fragment into pGP704::sacB

[**0264**] For amplification of crtN1/ald deletion fragment #1 (~1.1 kB in size) and crtN1/ald deletion fragment #2 (~1.1 kB in size), the following DNA primers were used, along with *Methylomonas* sp. 16a chromosomal DNA as template.

TABLE 4

Primers Utilized for Cloning of the Δald DNA fragment

Deletion Fragment	Forward Primer	Reverse Primer
crtN1/ald deletion fragment #1	Bg/II/ald (deletion) #1: 5'-AGATCTTTGCAACGGG TATTCGACGAAGG-3' (SEQ ID NO:12)	SphI-NotI/ ald (deletion) #1: 5'-CATGCGCGGCCGCCC CGATGTTTCTGGGAAATC AGC-3' (SEQ ID NO:13)
crtN1/ald deletion fragment #2	NotI/ald (deletion) #2: 5'-GCGGCCGCAATCAG CAAGCGCTGCAAGC-3' (SEQ ID NO:14)	SphI/ald (deletion) #2: 5'-GCATGCGTTGCGGA TACAGCCTGTCC-3' (SEQ ID NO:15)

**Underlined sequences represent restriction endonuclease recognition sites.

The methodology used for PCR reactions and cloning into *E. coli* TOP10 One Shot™ cells (Invitrogen) were the same as previously described in Example 3.

[**0265**] Several colonies from the transformation were screened for the proper insert DNA fragments using the QIAprep® Spin Mini-prep Kit for plasmid isolation. The crtN1/aid deletion fragment #1 was digested with Bglll and Sphl and the crtN1/ald deletion fragment #2 was digested with Notl and Sphl. The restriction digestion products were separated on an 0.8% agarose gel, subsequently excised from the agarose gel, and the DNA was extracted using the QiaQuick® Gel Extraction Kit (Qiagen).

[**0266**] Purified crtN1/ald deletion fragment #1 was ligated into pGP704::sacB digested with Bglll and Sphl. Following overnight room temperature incubation, the ligation mixture was used to transform calcium chloride competent *E. coli* SY327 cells (Miller, V., L., and Mekalanos, J., J., supra (1984)). The transformation mixture was plated onto LB+Amp (25 μg/mL) agar plates. PCR amplification using PCR primers Bglll/ald (deletion)#1 (SEQ ID NO:12) and Sphl-Notl/ald (deletion)#1 (SEQ ID NO:13) was used to detect *E. coli* transformants containing the correct DNA insert fragment. Plasmid DNA was purified from positive clones using the QIAprep® Spin Mini-prep Kit (Qiagen) and the resulting vector was digested with Notl and Sphl.

[**0267**] Purified crtN1/ald deletion fragment #2 was ligated into the linearized pGP704::sacB::crtN1/ald deletion fragment #1 vector described above. After overnight incubation at room temperature, the ligation mixture was transformed into *E. coli* SY327 cells and transformants were selected on LB+Amp (25 μg/mL) agar plates. PCR amplification using PCR primers Bg/ll/ald (deletion)#1 (SEQ ID NO:12) and Sphl/ald (deletion)#2 (SEQ ID NO:15) was used to detect *E. coli* transformants containing the appropriate DNA insert fragment. Subsequently, *E. coli* cells containing pGP704::sacB::Δaid were streaked onto LB+Amp (25 μg/mL) agar plates to obtain isolated colonies.

Example 7

Triparental Conjunction of pGP704::sacB::Δald(crtN1) into *Methylomonas* sp. 16a

[**0268**] The pGP704::sacB::Δald(crtN1) construct from Example 6 was transferred into *Methylomonas* sp. 16a via triparental conjugation. Specifically, the following were used as recipient, donor, and helper, respectively: *Methylomonas* sp. 16a, *E. coli* SY327 pGP704::sacB::Δald(crtN1) (from Example 6), and *E. coli* pRK2013. The mating was performed as described in Example 4; however, the cell suspension was plated onto BTZ agar plates containing Amp (35 μg/μL) instead of BTZ+Kan agar plates.

[**0269**] Isolated colonies were streaked onto BTZ+35 μg/μL Amp agar plates and grown 1-2 days at 30° C. in a jar containing 25% methane. These cells were used to inoculate bottles containing 20 mL of BTZ+25% methane. After overnight growth, 5 mL of the culture was concentrated by centrifugation using a tabletop centrifuge. The cells were inoculated into 20 mL of BTZ liquid medium containing nitrate (10 mM), methanol (200 mM), and 25% methane and grown overnight at 30° C. with aeration. Cells from the BTZ (NO_3) cultures were again inoculated into BTZ+25% methane and grown overnight growth at 30° C. with aeration. The cultures were monitored for *E. coli* killing by plating onto LB agar plates.

Example 8

Evaluation of *Methylomonas Transconjugants Containing pGP*704::sacB::Δald(crtN1)

[**0270**] Following the mobilization of the pGP704:sacB::Δald(crtN1) construct into *Methylomonas* sp. 16a, the two-step selection strategy was applied as described below to identify allelic exchange mutants (**FIG. 2**).

Preliminary Screening for Allelic Exchange Mutants

[**0271**] Cultures free of *E. coli* cells were passaged several times into fresh media (1 mL of culture into 20 mL of fresh BTZ medium), to increase the probability of occurrence of a second crossover event. Subsequently, cells were plated onto BTZ+Sucrose (5%) agar plates. Those cells growing on plates containing sucrose had lost the integration vector, which contained the sacB gene. However, the loss of the vector sequences could be due to the second crossover event occurring either on the same or opposite side of the ald-(crtN1) gene deletion that was present on the insert DNA. If the second crossover event had occurred on the same side of the ald(crtN1) gene deletion as the first crossover event, the wildtype ald gene would be regenerated. In contrast, if the second crossover event occurred on the opposite side of the

ald(crtN1) gene deletion as the first crossover event, the deletion of the ald(crtN1) gene would be established in the *Methylomonas* genome.

Confirmation of Allelic Exchange Mutants

[**0272**] Chromosomal DNA was purified from three cultures that had grown on the sucrose plates using the MasterPure DNA Purification Kit (Epicentre). PCR amplification using PCR primers Bg/ll/ald(deletion)#1 (SEQ ID NO:12) and Sphl/ald(deletion)#2 (SEQ ID NO:15) was used to distinguish between cells that contained the intact ald(crtN1) gene (wherein the expected size of the PCR product was ~3.7 kB) and those that contained the ald(crtN1) gene deletion (wherein the expected size of the PCR product was ~2.1 kB). Two of the three cultures produced a PCR product that was 2.1 kB, while the third culture appeared to have the wildtype gene present. Thus, these two PCR reactions confirmed the deletion of the ald(crtN1) gene in the genome of *Methylomonas*. Additionally, these results clearly demonstrated the utility of the two-step selection method for identification of allelic exchange mutants that are markerless, wherein the targeted chromosomal mutation does not require that an antibiotic resistance marker be left in the chromosome of the *Methylomonas* host.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 15

<210> SEQ ID NO 1
<211> LENGTH: 1536
<212> TYPE: DNA
<213> ORGANISM: Methylomonas sp.16a

<400> SEQUENCE: 1

atggccaaca ccaaacacat catcatcgtc ggcgcgggtc ccggcggact ttgcgccggc 60

atgttgctga gccagcgcgg cttcaaggta tcgattttcg acaaacatgc agaaatcggc 120

ggccgcaacc gcccgatcaa catgaacggc tttaccttcg ataccggtcc gacattcttg 180

ttgatgaaag gcgtgctgga cgaaatgttc gaactgtgcg agcgccgtag cgaggattat 240

ctggaattcc tgccgctaag cccgatgtac cgcctgctgt acgacgaccg cgacatcttc 300

gtctattccg accgcgagaa catgcgcgcc gaattgcaac gggtattcga cgaaggcacg 360

gacggctacg aacagttcat ggaacaggaa cgcaaacgct tcaacgcgct gtatccctgc 420

atcacccgcg attattccag cctgaaatcc tttttgtcgc tggacttgat caaggccctg 480

ccgtggctgg cttttccgaa aagcgtgttc aataatctcg gccagtattt caaccaggaa 540

aaaatgcgcc tggccttttg ctttcagtcc aagtatctgg gcatgtcgcc gtgggaatgc 600

ccggcactgt ttacgatgct gccctatctg gagcacgaat acggcattta tcacgtcaaa 660

ggcggcctga accgcatcgc ggcggcgatg gcgcaagtga tcgcggaaaa cggcggcgaa 720

attcacttga acagcgaaat cgagtcgctg atcatcgaaa acggcgctgc caagggcgtc 780

aaattacaac atggcgcgga gctgcgcggc gacgaagtca tcatcaacgc ggattttgcc 840

cacgcgatga cgcatctggt caaaccgggc gtcttgaaaa aatacacccc ggaaaacctg 900

aagcagcgcg agtattcctg ttcgaccttc atgctgtatc tgggtttgga caagatttac 960

gatctgccgc accataccat cgtgtttgcc aaggattaca ccaccaatat ccgcaacatt 1020

ttcgacaaca aaaccctgac ggacgatttt tcgttttacg tgcaaaacgc cagcgccagc 1080

gacgacagcc tagcgccagc cggcaaatcg gcgctgtacg tgctggtgcc gatgcccaac 1140

aacgacagcg gcctggactg gcaggcgcat tgccaaaacg tgcgcgaaca ggtgttggac 1200

acgctgggcg cgcgactggg attgagcgac atcagagccc atatcgaatg cgaaaaaatc 1260

atcacgccgc aaacctggga aacggacgaa cacgtttaca agggcgccac tttcagtttg 1320

tcgcacaagt tcagccaaat gctgtactgg cggccgcaca accgtttcga ggaactggcc 1380

aattgctatc tggtcggcgg cggcacgcat cccggtagcg gtttgccgac catctacgaa 1440

-continued

tcggcgcgga tttcggccaa gctgatttcc cagaaacatc gggtgaggtt caaggacata 1500

gcacacagcg cctggctgaa aaaagccaaa gcctga 1536

<210> SEQ ID NO 2
<211> LENGTH: 1593
<212> TYPE: DNA
<213> ORGANISM: Methylomonas sp.16a

<400> SEQUENCE: 2

atgacgacga tagcagccgt ctccccactg gatggccgct tgctgggaca ttttccagtc 60

agcaagccgg cgctcattca gcaacagctg acaaaatccc gccgcgccgc cctgctttgg 120

cgcgagctgc cggtcacgga acgggtcaaa cgcctgtcgc ccttgaaaaa acagctgctg 180

gataacctgg acagactctg cgaaaccatc cgcctcagca ccggcaaggt tcgcaccgag 240

gccttgctgg gggaaattta tccggtgctg gatttactgg cgtattacca aaagcgggcg 300

ccgcggattc tacgcacgcg cgccgtgtcc acctcgccgt tcgcgtttcc ggccgccacc 360

gcccgcatcg aacgccgccc ttacggcgtg gtcgcggtga tctcgccatg gaattacccg 420

tttcacctga gcgtcgcccc gctgctgacc gctttgctgg ccggcaatgc ggtaatcctg 480

aaaccctccg aactctgctt gccggtcggt cagttgatcg tcgatttgtt cgccacgctg 540

gatttgccgg acgggttggt gcaatgggtc atcggcgacg gccaaaccgg cgcggaactg 600

atagacgccc gccccgatct ggtgtttttc accggcggcc tgcagaccgg tcgggcggtc 660

atgcaacgcg ccgcccggca tccgattccg gtcatgctgg agttgggcgg taaagacacc 720

atgctggtgc tggccgacgc cgacctcaag cgcgccagcg ctgccgcgct gtacggcgcg 780

ttttgcaata gcggccaagt ctgcgtctcg gtcgaacgtc tgtacgtgca acaagcctgt 840

tttgcggaat tcctggccat gctgctgaag ggcctgtcca agctcaaggt cggccatgac 900

ccgcacggcg atgtgggagt gatgacgtcc gcccggcaaa tcgacatcgt ccaggcccat 960

tacgaggacg ccatcgccca gggcgccaag gcctccggcc cgctgctgcg cgacggcaat 1020

gtcgtgcaac ccgtggtgct ttgggacgtg caccacggca tgaaggtcat gcgcgaggaa 1080

accttcggtc cgttgctgcc ggtcatgccg ttcagcgacg aagccgaggc catcaagctc 1140

gccaacgaca gcgatctggg tctaaacgcc agcatctgga gccaggatat aatcaaggcc 1200

gagcgccttg ctggacaact agatgtcggc aactgggcga tcaacgacgt attgaaaaac 1260

gtgggccatt ccggcctgcc cttcggcggc gtcaagcaaa gcgggtttgg ccgttatcac 1320

ggcgccgaag gcttgctgaa cttcagctac ccggtatcgg gcctgaccaa tcgcagccgc 1380

ttgcccaaag aacccaactg gttcccttac agcgcatcag gctatgaaaa tttcaagggt 1440

ttcctcgatt ttatctacgg cgaagactcg atgctgcagc gcggtcgccg caatcagcaa 1500

gcgctgcaag ccttccgcga gttttccatt ttcgattgga cacaacgctg gcaaaacctg 1560

aaactgctgt tttcttggac acgggatgac taa 1593

<210> SEQ ID NO 3
<211> LENGTH: 1494
<212> TYPE: DNA
<213> ORGANISM: Methylomonas sp.16a

<400> SEQUENCE: 3

atgaactcaa atgacaacca acgcgtgatc gtgatcggcg ccggcctcgg cggcctgtcc 60

-continued

gccgctattt cgctggccac ggccggcttt tccgtgcaac tcatcgaaaa aaacgacaag 120

gtcggcggca agctcaacat catgaccaaa gacggcttta ccttcgatct ggggccgtcc 180

attttgacga tgccgcacat ctttgaggcc ttgttcacag gggccggcaa aaacatggcc 240

gattacgtgc aaatccagaa agtcgaaccg cactggcgca atttcttcga ggacggtagc 300

gtgatcgact tgtgcgaaga cgccgaaacc cagcgccgcg agctggataa acttggcccc 360

ggcacttacg cgcaattcca gcgctttctg gactattcga aaaacctctg cacggaaacc 420

gaagccggtt acttcgccaa gggcctggac ggcttttggg atttactcaa gttttacggc 480

ccgctccgca gcctgctgag tttcgacgtc ttccgcagca tggaccaggg cgtgcgccgc 540

tttatttccg atcccaagtt ggtcgaaatc ctgaattact tcatcaaata cgtcggctcc 600

tcgccttacg atgcgcccgc cttgatgaac ctgctgcctt acattcaata tcattacggc 660

ctgtggtacg tgaaaggcgg catgtatggc atggcgcagg ccatggaaaa actggccgtg 720

gaattgggcg tcgagattcg tttagatgcc gaggtgtcgg aaatccaaaa acaggacggc 780

agagcctgcg ccgtaaagtt ggcgaacggc gacgtgctgc cggccgacat cgtggtgtcg 840

aacatggaag tgattccggc gatggaaaaa ctgctgcgca gcccggccag cgaactgaaa 900

aaaatgcagc gcttcgagcc tagctgttcc ggcctggtgc tgcacttggg cgtggacagg 960

ctgtatccgc aactggcgca ccacaatttc ttttattccg atcatccgcg cgaacatttc 1020

gatgcggtat tcaaaagcca tcgcctgtcg gacgatccga ccatttatct ggtcgcgccg 1080

tgcaagaccg accccgccca ggcgccggcc ggctgcgaga tcatcaaaat cctgccccat 1140

atcccgcacc tcgaccccga caaactgctg accgccgagg attattcagc cttgcgcgag 1200

cgggtgctgg tcaaactcga acgcatgggc ctgacggatt tacgccaaca catcgtgacc 1260

gaagaatact ggacgccgct ggatattcag gccaaatatt attcaaacca gggctcgatt 1320

tacggcgtgg tcgccgaccg cttcaaaaac ctgggtttca aggcacctca acgcagcagc 1380

gaattatcca atctgtattt cgtcggcggc agcgtcaatc ccggcggcgg catgccgatg 1440

gtgacgctgt ccgggcaatt ggtgagggac aagattgtgg cggatttgca ataa 1494

<210> SEQ ID NO 4
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer DrdI/npr-sacB

<400> SEQUENCE: 4

gacatcgatg tcgaattcga gctcggtacc gatc 34

<210> SEQ ID NO 5
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer TthIII/npr-sacB

<400> SEQUENCE: 5

gacctcgtcg ctgttattag ttgactgtca gc 32

<210> SEQ ID NO 6
<211> LENGTH: 28
<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer crtN#2-ctg288 +1kb/BglII

<400> SEQUENCE: 6

agatctttcc ggtcatgctg gagttggg 28

<210> SEQ ID NO 7
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer crtN#2-ctg 288 + 1kb/XbaI

<400> SEQUENCE: 7

tctagattgc agctcaagcg attcgg 26

<210> SEQ ID NO 8
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer KAN-2FP-1

<400> SEQUENCE: 8

acctacaaca aagctctcat caacc 25

<210> SEQ ID NO 9
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer KAN-2RP-1

<400> SEQUENCE: 9

gcaatgtaac atcagagatt ttgag 25

<210> SEQ ID NO 10
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OP#1crtN#2

<400> SEQUENCE: 10

tcgtcgattt gttcgccacg c 21

<210> SEQ ID NO 11
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OP#2crtN#2

<400> SEQUENCE: 11

gccctatctt tttttatatt gtgtt 25

<210> SEQ ID NO 12
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer BglII/ald(deletion)#1

-continued

<400> SEQUENCE: 12

agatctttgc aacgggtatt cgacgaagg 29

<210> SEQ ID NO 13
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer SphI-NotI/ald (deletion)#1

<400> SEQUENCE: 13

gcatgcgcgg ccgccccgat gtttctggga aatcagc 37

<210> SEQ ID NO 14
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer NotI/ald (deletion) #2

<400> SEQUENCE: 14

gcggccgcaa tcagcaagcg ctgcaagc 28

<210> SEQ ID NO 15
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer SphI/ald (deletion) #2

<400> SEQUENCE: 15

gcatgcgttg cggatacagc ctgtcc 26

What is claimed is:

1. A method for the positive selection of double-crossover events in a C1 metabolizing bacterial host cell comprising:

a) providing a C1 metabolizing bacteria selected from the group consisting of methanotrophs and methylotrophs;

b) providing a positive selection vector comprising:

(i) a first genetic selectable marker;

(ii) an origin of transfer for a C1 metabolizing bacteria;

(iii) plasmid mobilization genes;

(iv) a sacB coding region encoding a levansucrase enzyme, under the control of a suitable promoter;

(v) a replacement nucleotide sequence of interest having at least two regions homology to a chromosomal nucleotide sequence of interest in the C1 metabolizing bacteria, said replacement nucleotide sequence of interest being disrupted by a second genetic selectable marker;

wherein the first and second genetic selectable markers are different and the vector is unable to replicate in the C1 metabolizing bacteria;

c) transforming the C1 metabolizing bacteria of (a) with the vector of (b)

d) selecting the transformants of (c) on the basis of the genetic selectable markers and sacB gene expression wherein, the transformants which are positive for the second selectable marker and grow in the presence of sucrose, but are negative for the first selectable marker have undergone double-crossover events; and

e) recovering the selected transformants of (d) which have undergone double-crossover events.

2. A method for the positive selection of double-crossover events in C1 metabolizing bacterial host cells comprising:

a) providing a C1 metabolizing bacteria selected from the group consisting of methanotrophs and methylotrophs;

b) providing a positive selection vector comprising:

(i) a genetic selectable marker;

(ii) an origin of transfer for a C1 metabolizing bacteria;

(iii) plasmid mobilization genes;

(iv) a sacB coding region encoding a levansucrase enzyme under the control of a suitable promoter;

(v) a mutant replacement nucleotide sequence of interest having at least two regions of homology to a chromosomal nucleotide sequence of interest in the C1 metabolizing bacteria;

the vector being unable to replicate in the C1 metabolizing bacteria chromosome;

c) transforming the C1 metabolizing bacteria of (a) with the vector of (b);

d) isolating the transformants of (c) on the basis of the genetic selectable marker; and

e) isolating the transformants of (d) which grow in the presence of sucrose, wherein said transformants contain double-crossover events.

3. A method according to either of claims **1** or 2 wherein the regions of homology are at least about 50 bp in length.

4. A method according to claim 2 wherein the mutation in the mutant nucleotide of interest is at least about 50 bp downstream of the 5' end of the sequence.

5. A method according to either of claims **1** or 2, wherein the C1 metabolizing bacteria is selected from the group consisting of *Methylomonas, Methylobacter, Methylococcus, Methylosinus, Methylocyctis, Methylomicrobium, Methanomonas, Methylophilus, Methylobacillus, Methylobacterium, Hyphomicrobium, Xanthobacter, Bacillus, Paracoccus, Nocardia, Arthrobacter, Rhodopseudomonas,* and *Pseudomonas.*

6. A method according to either of claims **1** or **2**, wherein the suitable promoter is a npr promoter.

7. A method according to either of claims **1** or 2, wherein said sacB coding region is derived from the group consisting of *Bacillus subtilis, Bacillus pumilus, Bacillus loceniformis, Bacillus amyloliquifaciens, Bacillus thurgiensis, Bacillus stearothermophilus,* and *Bacillus sphaericus.*

8. A method according to either of claims **1** or **2**, wherein the positive selection vector is pGP704::sacB.

9. A method according to either of claims **1** or **2** wherein any of the genetic selectable markers are independently selected from the group consisting of antibiotic resistance markers, metal resistance markers, substrate-utilization markers, genes encoding fluorescent and bioluminescent proteins, lacZ, gfp, cat, galK, inaZ, luc, luxAB, bgaB, nptll, phoA, uidA, and xylE.

10. A C1 metabolizing bacteria transformant selected by the process of either claim 1 or claim 2 employing a positive selection vector.

11. A recombinant C1 metabolizing bacteria transformant of claim 10 wherein the genome of said transformant lacks any sequence of the positive selection vector, other than the replacement nucleotide sequence of interest.

12. A recombinant C1 metabolizing bacteria transformant of claim 10 wherein the genome of said transformant lacks any genetic selectable marker sequence.

13. A positive selection vector for the positive selection of double-crossover events in a C1 metabolizing bacterial host cell comprising:

(i) at least one genetic selectable marker;

(ii) an origin of transfer for a C1 metabolizing bacteria;

(iii) plasmid mobilization genes;

(iv) a sacB coding region encoding a levansucrase enzyme under the control of a suitable promoter;

(v) a replacement nucleotide sequence of interest having at least two regions of homology to a chromosomal nucleotide sequence of interest in the C1 metabolizing bacteria;

the vector being unable to replicate in the C1 metabolizing bacteria chromosome.

14. A vector according to claim 13 wherein said sacB coding region is derived from the group consisting of *Bacillus subtilis, Bacillus pumilus, Bacillus loceniformis, Bacillus amyloliquifaciens, Bacillus thurgiensis, Bacillus stearothermophilus,* and *Bacillus sphearicus.*

15. A vector according to claim 13 wherein the suitable promoter is the npr promoter.

16. A vector according to claim 13 wherein the replacement nucleotide of interest is disrupted by at least one second genetic selectable marker.

17. A vector according to claim 13 or 16 wherein the genetic selectable marker is selected independently from the group consisting of antibiotic resistance markers, metal resistance markers, substrate-utilization markers, genes encoding fluorescent and bioluminescent proteins, lacZ, gfp, cat, galK, inaZ, luc, luxAB, bgaB, nptll, phoA, uidA, and xylE.

18. A vector according to claim 13 wherein the regions of homology are at least about 50 bp in length.

19. A vector according to claim 13 wherein the replacement nucleotide sequence of interest is a mutant replacement nucleotide sequence of interest.

20. A vector according to claim 13 wherein the vector is pGP704::sacB.

21. A transformed C1 metabolizing bacteria comprising the vector of claim 13.

22. A transformed C1 metabolizing bacteria of claim 21 selected from the group consisting of methanotrophs and methylotrophs.

23. A methylotroph of claim 22 selected from the group consisting of *Methylomonas, Methylobacter, Methylococcus, Methylosinus, Methylocyctis, Methylomicrobium, Methanomonas, Methylophilus, Methylobacillus, Methylobacterium, Hyphomicrobium, Xanthobacter, Bacillus, Paracoccus, Nocardia, Arthrobacter, Rhodopseudomonas,* and *Pseudomonas.*

* * * * *