Recombinational cloning is provided by the use of nucleic acids, vectors and methods, in vitro and in vivo, for moving or exchanging segments of DNA molecules using engineered recombination sites and recombination proteins to provide chimeric DNA molecules that have the desired characteristic(s) and/or DNA segment(s).
Figure 1

A (insert) → B (Cloning vector) → C (repression cassette) → D (Subcloning vector) → Recombinase

A

Insert donor

B

Vector donor

C

Subclones

Product

Byproduct

Recombinase

Cointegrate

Recombinase
Figure 2B

pEZC705
3007 bps

amp

ori

attB

loxP

2309, FspI

2203, BglII

2567, ScaI

2756, ApaI

2891, SspI

BanII, 158

SspI, 439

BglII, 469

FspI, 480

BanII, 653

SstI, 653

NotI, 692

XbaI, 700

BamHI, 712

AvaI, 718

SmaI, 718

PstI, 724

EcoRI, 730

EcoRV, 736

HindIII, 742

ClaI, 748

HindII, 767

SalI, 757

KpnI, 798

SapI, 1073

1510, ApaI
Figure 2C

`pEZC726`

4570 bps
Figure 2F

Intbypro
4761 bps

attL
amp
tetR
loxP

F1

4737, SstI
4737, BanII
4564, FspI
4553, BglII
4523, SspI

4242, BanII
3968, SspI
3833, ApaLI
3644, Scal
3386, FspI
3280, BglII

2587, ApaLI

NcoI, 144
NsiI, 158
NcoI, 921
SspI, 930
BamHI, 1067
AvaI, 1073
SmaI, 1073
PstI, 1079
MluI, 1183
XbaI, 1192

EcoRI, 620
NcoI, 1767
EcoRV, 1813
HindIII, 1819
ClaI, 1825
HindII, 1834
Sall, 1834
KpnI, 1875

SspI, 2150
Figure 3C

EZC6prod
2628 bps

HindIII,109
Bgll,376
Ssp,2542
Smal,2491
Aval,2491

[kan]
tetOP

loxP

lox511

BanII,560
Sstl,560
NotI,600
XbaI,608
BamHI,620
Aval,626
Sml,626
PstI,632
EcoRI,638
EcoRV,644
HindIII,650
ClaI,656
HindII,665
SalI,665
KpnI,706

SapI,981

1418,ApaLI

2309,ClaI
2270,BanII
2217,XhoI
2217,Aval
2182,BglII
2111,BglII
Figure 3D
Figure 4F
Figure 5B

pEZC1305
4608 bps
Figure 5E

![Diagram of pEZX1317](image-url)
Figure 5G

pEZC1405
4698 bps
Figure 6B

pEZC1706
4039 bps
Figure 7A

pEZC2901

4019 bps
Figure 7B

pEZC2913 4695 bps

[kan]
tetOP
T2
T1
tetR
ori
attR112
T1
T2
cam

attR1
Figure 8C

pEZC3601
2737 bps
Figure 8E

pEZC3617
2739 bps

attL3
CAT rf3
ori
attL1
KmR
Figure 8F

pEZC3606
3418 bps
Figure 8G

pEZC3613
3419 bps

attL1
ori
KmR
phoA rf2
attL3
Figure 8H

pEZC3621
3420 bps
Figure 8I

GST-CAT
5770 bps
Figure 8J

GST-phoA

6451 bps
Figure 9A

5.2kbPCRprod (5327 bps)
Figure 9C

5.2kb clone
8278 bps

7303, Aval
7577, Aval
8277, Aval
Aval, 236
BamH1, 377
BamH1, 1835
BamH1, 2598
BamH1, 2770
Aval, 992
Aval, 2932
Aval, 1707
attR0
[kan]
tetOP
ori
loxP
5446, Aval

8000
Figure 10B

pEYZC6701

3827 bps
Figure 10D

attR Product

3794 bps
Figure 11B

Diagram of pEZC7501:
- Ori
- 5231, SalI
- CMV P
- attB1
- GFP
- XbaI, 874
- attB3
- lox P
- f1
- Amp
- incA
- 2689, SalI

5256 bps
Figure 12A

attB1 AMPattB3 (987 bps)
Figure 12B

attB1TETattB3 (1442 bps)
amp7102
2931 bps
RECOMBINATIONAL CLONING USING NUCLEIC ACIDS HAVING RECOMBINATION SITES

CROSS-REFERENCE TO RELATED APPLICATIONS


BACKGROUND OF THE INVENTION

[0002] 1. Field of the Invention

[0003] The present invention relates to recombinant DNA technology. DNA and vectors having engineered recombination sites are provided for use in a recombinational cloning method that enables efficient and specific recombination of DNA segments using recombination proteins. The DNAs, vectors and methods are useful for a variety of DNA exchanges, such as subcloning of DNA, in vitro or in vivo.

[0004] 2. Related Art

[0005] Site-specific recombinases. Site-specific recombinases are proteins that are present in many organisms (e.g., viruses and bacteria) and have been characterized to have both endonuclease and ligation properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in DNA and exchange the DNA segments flanking those segments. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, e.g., Landy, A., *Current Opinion in Biotechnology* 3:699-707 (1993)).


[0007] Many of these belong to the integrase family of recombinases (Argos et al. *EMBO J.* 5:433-440 (1986)). Perhaps the best studied of these are the Integrate/att system from bacteriophage λ (Landy, A. *Current Opinions in Genetics and Devel.* 3:699-707 (1993)), the Cre/loxP system from bacteriophage P1 (Hoess and Abermski 1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-100), and the FLP/FRT system from the Saccharomyces cerevisiae 2μ circle plasmid (Broach et al. *Cell* 29:227-234 (1982)).

[0008] Backman (U.S. Pat. No. 4,673,640) discloses the in vivo use of λ recombinase to recombine a protein producing DNA segment by enzymatic site-specific recombination using wild-type recombination sites attB and attP.

[0009] Hasan and Szybalski (Gene 56:145-151 (1987)) discloses the use of λ Int recombinase in vivo for intramolecular recombination between wild type attP and attB sites which flank a promoter. Because the orientations of these sites are inverted relative to each other, this causes an irreversible flipping of the promoter region relative to the gene of interest.

[0010] Palazzolo et al. *Gene* 88:25-36 (1990), discloses phage lambda vectors having bacteriophage λ arms that contain restriction sites positioned outside a cloned DNA sequence and between wild-type loxP sites. Infection of *E. coli* cells that express the Cre recombinase with these phage vectors results in recombination between the loxP sites and the in vivo excision of the plasmid replicon, including the cloned cDNA.

[0011] Pösflai et al. (*Nucl. Acids Res.* 22:2392-2398 (1994)) discloses a method for inserting into genomic DNA partial expression vectors having a selectable marker, flanked by two wild-type FRT recognition sequences. FLP site-specific recombinase as present in the cells is used to integrate the vectors into the genome at predetermined sites. Under conditions where the replicon is functional, this cloned genomic DNA can be amplified.

[0012] Rebece et al. (U.S. Pat. No. 5,434,066) discloses the use of site-specific recombinases such as Cre for DNA containing two loxP sites is used for in vivo recombinant DNA exchanges between the sites.

[0013] Boyd (Nucl. Acids Res. 21:817-821 (1993)) discloses a method to facilitate the cloning of blunt-ended DNA using conditions that encourage intramolecular ligation to a dephosphorylated vector that contains a wild-type loxP site acted upon by a Cre site-specific recombinase present in *E. coli* host cells.

[0014] Waterhouse et al. (PCT No. 93/19172 and Nucl. Acids Res. 21 (9):2265 (1993)) disclose an in vivo method where light and heavy chains of a particular antibody were cloned in different phage vectors between loxP and loxP 511 sites and used to transform new *E. coli* cells. Cre, acting in the host cells on the two parental molecules (one plasmid, one phage), produced four products in equilibrium: two different cointegrates (produced by recombination at either loxP or loxP 511 sites), and two daughter molecules, one of which was the desired product.

[0015] In contrast to the other related art, Schlake & Bode (*Biochemistry* 33:12746-12751 (1994)) discloses an in vivo method to exchange expression cassettes at defined chromosomal locations, each flanked by a wild type and a spacer-mutated FRT recombination site. A double-reciprocal crossover was mediated in cultured mammalian cells by using this FLP/FRT system for site-specific recombination.

[0016] Transposases. The family of enzymes, the transposases, has also been used to transfer genetic information between replicons. Transposons are structurally variable, being described as simple or compound, but typically encode the recombinase gene flanked by DNA sequences organized in inverted orientations. Integration of trans-
posons can be random or highly specific. Representatives such as Tn7, which are highly site-specific, have been applied to the in vivo movement of DNA segments between replicons (Luckow et al., *J. Virol.* 67:4566-4579 (1993)).

[0017] Devine and Boeke *Nucleic Acids Res.* 22:3765-3772 (1994), discloses the construction of artificial transposons for the insertion of DNA segments, in vitro, into recipient DNA molecules. The system makes use of the integrate of yeast TY1 virus-like particles. The DNA segment of interest is cloned, using standard methods, between the ends of the transposon-like element TY1. In the presence of the TY1 integrase, the resulting element integrates randomly into a second target DNA molecule.

[0018] DNA cloning. The cloning of DNA segments currently occurs as a daily routine in many research labs and as a prerequisite step in many genetic analyses. The purpose of these clonings is various, however, two general purposes can be considered: (1) the initial cloning of DNA from large DNA or RNA segments (chromosomes, YACs, PCR fragments, mRNA, etc.), done in a relative handful of known vectors such as pUC, pGem, pBlueScript, and (2) the subcloning of these DNA segments into specialized vectors for functional analysis. A great deal of time and effort is expended both in the transfer of DNA segments from the initial cloning vectors to the more specialized vectors. This transfer is called subcloning.

[0019] The basic methods for cloning have been known for many years and have changed little during that time. A typical cloning protocol is as follows:

- [0020] (1) digest the DNA of interest with one or two restriction enzymes;
- [0021] (2) gel purify the DNA segment of interest when known;
- [0022] (3) prepare the vector by cutting with appropriate restriction enzymes, treating with alkaline phosphohase, gel purify etc., as appropriate;
- [0023] (4) ligate the DNA segment to the vector, with appropriate controls to eliminate background of uncut and self-ligated vector;
- [0024] (5) introduce the resulting vector into an *E. coli* host cell;
- [0025] (6) pick selected colonies and grow small cultures overnight;
- [0026] (7) make DNA minipreps; and
- [0027] (8) analyze the isolated plasmid on agarose gels (often after diagnostic restriction enzyme digestions) or by PCR.

[0028] The specialized vectors used for subcloning DNA segments are functionally diverse. These include but are not limited to: vectors for expressing genes in various organisms; for regulating gene expression; for providing tags to aid in protein purification or to allow tracking of proteins in cells; for modifying the cloned DNA segment (e.g., generating deletions); for the synthesis of probes (e.g., riboprobes); for the preparation of templates for DNA sequencing; for the identification of protein coding regions; for the fusion of various protein-coding regions; to provide large amounts of the DNA of interest, etc. It is common that a particular investigation will involve subcloning the DNA segment of interest into several different specialized vectors.

[0029] As known in the art, simple subclonings can be done in one day (e.g., the DNA segment is not large and the restriction sites are compatible with those of the subcloning vector). However, many other subclonings can take several weeks, especially those involving unknown sequences, long fragments, toxic genes, unsuitable placement of restriction sites, high backgrounds, impure enzymes, etc. Subcloning DNA fragments is thus often viewed as a chore to be done as few times as possible. Several methods for facilitating the cloning of DNA segments have been described, e.g., as in the following references.

[0030] Ferguson, J., et al. *Gene* 16:191 (1981), discloses a family of vectors for subcloning fragments of yeast DNA. The vectors encode kanamycin resistance. Clones of longer yeast DNA segments can be partially digested and ligated into the subcloning vectors. If the original cloning vector conveys resistance to ampicillin, no purification is necessary prior to transformation, since the selection will be for kanamycin.


[0032] Accordingly, traditional subcloning methods, using restriction enzymes and ligase, are time consuming and relatively unreliable. Considerable labor is expended, and if two or more days later the desired subclone can not be found among the candidate plasmids, the entire process must then be repeated with alternative conditions attempted. Although site specific recombinases have been used to recombine DNA in vivo, the successful use of such enzymes in vitro was expected to suffer from several problems. For example, the site specificities and efficiencies were expected to differ in vitro; topologically-linked products were expected; and the topology of the DNA substrates and recombination proteins was expected to differ significantly in vitro (see, e.g., Adams et al., *J. Mol. Biol.* 226:661-73 (1992)). Reactions that could go on for many hours in vitro were expected to occur in significantly less time in vitro before the enzymes became inactive. Multiple DNA recombination products were expected in the biological host used, resulting in unsatisfactory reliability, specificity or efficiency of subcloning. Thus, in vitro recombination reactions were not expected to be sufficiently efficient to yield the desired levels of product.

[0033] Accordingly, there is a long felt need to provide an alternative subcloning system that provides advantages over the known use of restriction enzymes and ligases.

**SUMMARY OF THE INVENTION**

[0034] The present invention provides nucleic acids, vectors and methods for obtaining amplified, chimeric or recombinant nucleic acid molecules using recombination proteins and at least one recombination site, in vitro or in vivo. These methods are highly specific, rapid, and less labor intensive than standard cloning or subcloning techniques. The improved specificity, speed and yields of the present
invention facilitates DNA or RNA cloning or subcloning, regulation or exchange useful for any related purpose.

[0035] The present invention relates to nucleic acids, vectors and methods for moving or exchanging nucleic acid segments (preferably DNA segments or fragments) using at least one recombination site and at least one recombination protein to provide chimeric DNA molecules which have the desired characteristic(s) and/or DNA segment(s). Use of the invention thus allows for cloning or subcloning such nucleic acid molecules into a variety of vectors. Generally, one or more parent nucleic acid molecules (preferably DNA molecules) are recombined to give one or more daughter molecules, at least one of which is the desired Product molecule, which is preferably a vector comprising the desired nucleic acid segment. The invention thus relates to nucleic acid molecules, vectors and methods to effect the exchange and/or to select for one or more desired products.

[0036] One embodiment of the present invention relates to a method of making chimeric molecule, which comprises

[0037] (a) combining in vitro or in vivo

[0038] (i) one or more Insert Donor molecules comprising a desired nucleic acid segment flanked by a first recombination site and a second recombination site, wherein the first and second recombination sites do not substantially recombine with each other;

[0039] (ii) one or more Vector Donor molecules comprising a third recombination site and a fourth recombination site, wherein the third and fourth recombination sites do not substantially recombine with each other; and

[0040] (iii) one or more site specific recombination proteins capable of recombining the first and third recombinational sites and/or the second and fourth recombinational sites;

[0041] thereby allowing recombination to occur, so as to produce at least one chimeric nucleic acid molecule, at least one desired Product nucleic acid molecule which comprises said desired segment, and optionally a Byproduct nucleic acid molecule; and then, optionally,

[0042] (b) selecting for the Product or Byproduct DNA molecule.

[0043] In another embodiment, the present invention relates to a method of making chimeric molecule, which comprises

[0044] (a) combining in vitro or in vivo

[0045] (i) one or more Insert Donor molecules comprising a desired nucleic acid segment flanked by two or more recombination sites wherein said recombination sites do not substantially recombine with each other;

[0046] (ii) one or more Vector Donor molecules comprising two or more recombination sites, wherein said recombination sites do not substantially recombine with each other; and

[0047] (iii) one or more site specific recombination proteins;

[0048] (b) incubating said combination under conditions sufficient to transfer one or more said desired segments into one or more of said Vector Donor molecules, thereby producing one or more Product molecules. The resulting Product molecules may optionally be selected or isolated away from other molecules such as cointegrate molecules, Byproduct molecules, and unreacted Vector Donor molecules or Insert Donor molecules. In a preferred aspect of the invention, the Insert Donor molecules are combined with one or more different Vector Donor molecules, thereby allowing for the production of different Product molecules in which the nucleic acid of interest is transferred into any number of different vectors in the single step.

[0049] In accordance with the invention, the above methods may be reversed to provide the original Insert Donor molecules which may then be used in combination with one or more different Vector Donor molecules to produce new Product or Byproduct molecules. Alternatively, the Product molecules produced by the method of the invention may serve as the Insert Donor molecules which may be used directly in combination with one or more different Vector Donor molecules, thereby producing new Product or Byproduct molecules. Thus, nucleic acid molecules of interest may be transferred or moved to any number of desired vectors, thereby providing an efficient means for subcloning molecules of interest.

[0050] Thus, the invention relates to combining a Product molecule with a second Vector Donor molecules to produce a second Product molecule. The second Product DNA molecule may then be utilized in combination with a third Vector Donor molecule to produce a third Product molecule. This process of the invention may be repeated any number of times to transfer or move the insert of interest into any number of different vectors. In this aspect of the invention, a combination of two or more different Vector Donor molecules may be combined with the Product molecule to produce in a single step different Product molecules in which the desired nucleic acid segment (derived from the Product DNA molecule) is transferred into any number of different vectors.

[0051] In particular, the present invention relates to a method for cloning or subcloning one or more desired nucleic acid molecules comprising

[0052] (a) combining in vitro or in vivo

[0053] (i) one or more Insert Donor molecules comprising one or more desired nucleic acid segments flanked by at least two recombination sites, wherein said recombination sites do not substantially recombine with each other;

[0054] (ii) one or more Vector Donor molecules comprising at least two recombination sites, wherein said recombination sites do not substantially recombine with each other; and

[0055] (iii) one or more site specific recombination proteins;


[0056] (b) incubating said combination under conditions sufficient to allow one or more of said desired segments to be transferred into one or more of said Vector Donor molecules, thereby producing one or more Product molecules;

[0057] (c) optionally selecting for or isolating said Product molecule;

[0058] (d) combining in vitro or in vivo

[0059] (i) one or more of said Product molecules comprising said desired segments flanked by two or more recombination sites, wherein said recombination sites do not substantially recombine with each other;

[0060] (ii) one or more different Vector Donor molecules comprising two or more recombination sites, wherein said recombination sites do not substantially recombine with each other; and

[0061] (iii) one or more site specific recombination protein; and

[0062] (e) incubating said combination under conditions sufficient to transfer one or more of said desired segments into one or more of said different Vector Donor molecules, thereby producing one or more different Product molecules.

[0063] In accordance with the invention, Vector Donor molecules may comprise vectors which may function in a variety of systems or host cells. Preferred vectors for use in the invention include plasmid vectors, eukaryotic vectors or vectors which may shuttle between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred plasmid vectors for use in the invention include vectors which propagate and/or replicate in gram negative and/or gram positive bacteria, including bacteria of the genus Escherichia, Salmonella, Proteus, Clostridium, Klebsiella, Bacillus, Streptomyces, and Pseudomonas and preferably in the species E. coli. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate and yeast cells, plant cells, mammalian cells, (particularly human), fungal cells, insect cells, fish cells and the like. In particular, vectors of interest include but are not limited to cloning vectors, sequencing vectors, expression vectors, fusion vectors, two-hybrid vectors, gene therapy vectors, and reverse two-hybrid vectors. Such vectors may be used in plasmidic and/or eukaryotic systems depending on the particular vector.

[0064] The Insert Donor molecules used in accordance with the invention preferably comprise two or more recombination sites which allow the insert (e.g. the nucleic acid segment of interest) of the Donor molecules to be transferred or moved into one or more Vector Donor molecules in accordance with the invention. The Insert Donor molecules of the invention may be prepared by any number of techniques by which two or more recombination sites are added to the molecule of interest. Such means for including recombination sites to prepare the Insert Donor molecules of the invention includes mutation of a nucleic acid molecule (e.g. random or site specific mutagenesis), recombinant techniques (e.g. ligation of adapters or nucleic acid molecules comprising recombination sites to linear molecules), amplification (e.g. using primers which comprise recombination sites or portions thereof) transposition (e.g. using transposons which comprise recombination sites), recombination (e.g. using one or more homologous sequences comprising recombination sites), nucleic acid synthesis (e.g. chemical synthesis of molecules comprising recombination sites or enzymatic synthesis using various polymerases or reverse transcriptases) and the like. In accordance with the invention, nucleic acid molecules to which one or more recombination sites are added may be any nucleic acid molecule derived from any source and may include non naturally occurring nucleic acids (e.g. RNA's; see U.S. Pat. Nos. 5,539,082 and 5,482,836). Particularly preferred nucleic acid molecules are DNA molecules (single stranded or double stranded). Additionally, the nucleic acid molecules of interest for producing Insert Donor molecules may be linear or circular and further may comprise a particular sequence of interest (e.g. a gene) or may be a population of molecules (e.g. molecules generated from a genomic or cDNA libraries).

[0065] Thus, the invention relates to a number of methods for preparing Insert Donor molecules and the Insert Donor molecules produced by such methods. In one aspect of the invention, primers comprising one or more recombination sites or portions thereof are used in the nucleic acid synthesis or nucleic acid amplification to prepare the Insert Donor molecules of the invention. Thus, the invention relates to a method of synthesizing a nucleic acid molecule comprising:

[0066] (a) mixing one or more nucleic acid templates with a polypeptide having polymerase activity and one or more primers comprising one or more recombination sites or portions thereof; and

[0067] (b) incubating said mixture under conditions sufficient to synthesize one or more nucleic acid molecules which are complementary to all or a portion of said templates and which comprises one or more recombination sites. In accordance with the invention, the synthesized nucleic acid molecule comprising one or more recombination sites may be used as templates under appropriate conditions to synthesize nucleic acid molecules complementary to all or a portion of the recombination site containing templates, thereby forming double stranded molecules comprising one or more recombination sites. Preferably, such second synthesis step is performed in the presence of one or more primers comprising one or more recombination sites. In yet another aspect, the synthesized double stranded molecules may be amplified using primers which may comprise one or more recombination sites.

[0068] In another aspect of the invention, one or more recombination sites may be added to nucleic acid molecules by any of a number of nucleic acid amplification techniques. In particular, such method comprises:

[0069] (a) contacting a first nucleic acid molecule with a first primer molecule which is complementary to a portion of said first nucleic acid molecule and a second nucleic acid molecule with a second primer molecule which is complementary to a portion of said second nucleic acid molecule in the presence of one or more polypeptides having polymerase activity;

[0070] (b) incubating said molecules under conditions sufficient to form a third nucleic acid molecule
complementary to all or a portion of said first nucleic acid molecule and the fourth nucleic acid molecule complementary to all or a portion of said second nucleic acid molecule;

(c) denaturing said first and third and said second and fourth nucleic acid molecules; and

(d) repeating steps (a) through (c) one or more times,

wherein said first and/or said second primer molecules comprise one or more recombination sites or portions thereof.

In yet another aspect of the invention, a method for adding one or more recombination sites to nucleic acid molecules may comprise:

(a) contacting one or more nucleic acid molecules with one or more adapters or nucleic acid molecules which comprise one or more recombination sites or portions thereof; and

(b) incubating said mixture under conditions sufficient to add one or more recombination sites to said nucleic acid molecules. Preferably, linear molecules are used for adding such adapters or molecules in accordance with the invention and such adapters or molecules are preferably added to one or more termini of such linear molecules. The linear molecules may be prepared by any technique including mechanical (e.g., sonication or shearing) or enzymatic (e.g., nuclease such as restriction endonucleases). Thus, the method of the invention may further comprise digesting the nucleic acid molecule with one or more nuclease (preferably any restriction endonucleases) and ligating one or more of the recombination site containing adapters or molecules to the molecule of interest. Ligation may be accomplished using blunt ended or sticky ended molecules. Alternatively, topoisomerases may be used to introduce recombination sites in accordance with the invention. Topoisomerases cleave and rejoin nucleic acid molecules and therefore may be used in place of nuclease and ligases.

In another aspect, one or more recombination sites may be added to nucleic acid molecules by de novo synthesis. Thus, the invention relates to such a method which comprises chemically synthesizing one or more nucleic acid molecules in which recombination sites are added by adding the appropriate sequence of nucleotides during the synthesis process.

In another embodiment of the invention, one or more recombination sites may be added to nucleic acid molecules of interest by a method which comprises:

(a) contacting one or more nucleic acid molecules with one or more integration sequences which comprise one or more recombination sites or portions thereof; and

(b) incubation of said mixture under conditions sufficient to incorporate said recombination site containing integration sequences into said nucleic acid molecules. In accordance with this aspect of the invention, integration sequences may comprise any nucleic acid molecules which through recombination or by integration become a part of the nucleic acid molecule of interest. Integration sequences may be introduced in accordance with this aspect of the invention by in vivo or in vitro recombination (homologous recombination or illegitimate recombination) or by in vivo or in vitro installation by using transposons, insertion sequences, integrating viruses, homing introns, or other integrating elements.

In another aspect, the invention relates to kits for carrying out the methods of the invention and more specifically relates to cloning or subcloning kits and kits for making Insert Donor molecules of the invention. Such kits may comprise a carrier or receptacle being compartmentalized to receive and hold therein any number of containers. Such containers may contain any number of components for carrying out the methods of the invention or combinations of such components. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins or recombinases, one or more Vector Donor molecules, one or more Insert Donor molecules and one or more host cells (e.g., competent cells).

Kits for making the Insert Donor molecules of the invention may comprise any or a number of components and the composition of such kits may vary depending on the specific method involved. Kits for synthesizing Insert Donor molecules by amplification may comprise one or more components (or combinations thereof) selected from the group consisting of one or more polypeptides having polymerase activity (preferably DNA polymerases and most preferably thermostable DNA polymerases), one or more nucleotides, and one or more primers comprising one or more recombination sites. Kits for inserting or adding recombination sites to nucleic acid molecules of interest may comprise one or more nuclease (preferably restriction endonuclease), one or more ligases, one or more topoisomerases one or more polymerases, and one or more nucleic acid molecules or adapters comprising one or more recombination sites. Kits for integrating recombination sites into one or more nucleic acid molecules of interest may comprise one or more components (or combinations thereof) selected from the group consisting of one or more integration sequences comprising one or more recombination sites. Such integration sequences may comprise one or more transposons, integrating viruses, homologous recombination sequences, one or more host cells and the like.

The invention also relates to compositions for carrying out the methods of the invention or compositions which are produced from carrying out the methods of the invention. In particular, such compositions may comprise one or more Insert Donor molecules, one or more Vector Donor molecules and one or more recombination proteins (or combinations thereof). In a further aspect, the compositions of the invention may comprise one or more co-integrate molecules, one or more Product molecules and one or more Byproduct molecule (or combinations thereof).

Compositions related to preparing Insert Donor molecules may vary depending on the particular method utilized in preparing the desired Insert Donor molecules. Compositions for preparing such molecules by amplification
may comprise one or more polypeptides having polymerase activity, one or more primers comprising one or more recombination sites, one or more nucleotides and one or more nucleic acid molecule to be amplified (or combinations thereof). Compositions related to inserting or adding recombination sites in a desired nucleic acid molecule may comprise one or more nucleic acid molecules or adapters comprising one or more recombination sites, one or more ligases, one or more restriction endonucleases, one or more topoisomerases, and one or more nucleic acid molecules desired to contain such recombination sites (or combinations thereof). Compositions related to integration of recombination sites in a desired nucleic acid molecule may comprise one or more integration sequences comprising one or more recombination sites and one or more nucleic acid molecules desired to contain the recombination sites.

[0085] In a particularly preferred aspect of the invention, libraries (e.g., populations of genomic DNA or cDNA, or populations of nucleic acid molecules, produced by de novo synthesis such as random sequences or degenerate oligonucleotides) are utilized in accordance with the present invention. By inserting or adding recombination sites to such populations of nucleic acid molecules, a population of Insert Donor molecules are produced. By the recombinant methods of the invention, the library may be easily moved into different vectors (or combinations of vectors) and thus into different host systems (prokaryotic and eukaryotic) to evaluate and analyze the library or a particular sequences or clones derived from the library. Alternatively, the vectors containing the desired molecule may be used in vitro systems such as in vitro expression systems for production of RNA and/or protein. In a particularly preferred aspect, one or more recombination sites are added to nucleic acid molecules of the library by method comprising:

[0086] (a) mixing a population of linear nucleic acid molecules with one or more adapters comprising one or more recombination sites; and

[0087] (b) incubating said mixture under conditions sufficient to add one or more of said adapters to one or more termini of said linear molecules. In a preferred aspect, the population of nucleic acid molecules are double stranded DNA molecules (preferably genomic DNA or cDNA). A population of linear fragments for use in the invention may be prepared by cleaving (by mechanical or enzymatic means) the genomic or cDNA. In a preferred aspect, the adapters are added to one or more termini of the linear molecules.

[0088] In another particularly preferred aspect of the invention, cDNA libraries are used to prepare a population of Insert Donor DNA molecules of the invention. In particular, this aspect of the invention relates to a method which comprises:

[0089] (a) contacting a population of RNA, mRNA or polyA+ RNA templates with one or more polypeptides having reverse transcriptase activity and one or more primers which comprises one or more recombination sites;

[0090] (b) incubating said mixture under conditions sufficient to synthesize a first population of DNA molecules complementary to said templates, wherein said DNA molecules comprise one or more recombination sites. This aspect of the invention may further comprise incubating said synthesized DNA under conditions sufficient to make a second population of DNA molecules complementary to all or a portion of said first population of DNA molecules, thereby forming a population of double stranded DNA molecules comprising one or more recombination sites.

[0091] In a particularly preferred aspect, the Insert Donor molecules of the invention comprise at least two recombination sites and where the Insert Donor molecules are linear, such two or more recombination sites are preferably located at or near both termini of the molecules. In accordance with the invention, the use of additional recombination sites (i.e., more than two) may be used to facilitate subcloning of different inserts within the Insert Donor molecule, depending on the type and placement of such recombination sites.

Other embodiments include DNA and vectors useful in the methods of the present invention. In particular, Vector Donor molecules are provided in one embodiment, wherein DNA segments within the Vector Donor are separated either by, (i) in a circular Vector Donor, at least two recombination sites, or (ii) in a linear Vector Donor, at least one recombination site, where the recombination sites are preferably engineered to enhance specificity or efficiency of recombination. One Vector Donor embodiment comprises a first DNA segment and a second DNA segment, the first or second segment comprising a selectable marker. A second Vector Donor embodiment comprises a first DNA segment and a second DNA segment, the first or second DNA segment comprising a toxic gene. A third Vector Donor embodiment comprises a first DNA segment and a second DNA segment comprising an inactive fragment of at least one selectable marker, wherein the inactive fragment of the Selectable marker is capable of reconstituting a functional Selectable marker when recombined across the first or second recombination site with another inactive fragment of at least one Selectable marker.

[0093] Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of what is known in the art, in light of the following drawings and description of the invention, and in light of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0094] FIG. 1 depicts one general method of the present invention, wherein the starting (parent) DNA molecules can be circular or linear. The goal is to exchange the original cloning vector B. It is desirable in one embodiment to select for AD and against all the other molecules, including the Cointegrate. The square and circle are sites of recombination: e.g., loxP sites, att sites, etc. For example, segment D can contain expression signals, new drug markers, new origins of replication, or specialized functions for mapping or sequencing DNA.

[0095] FIG. 2A depicts an in vitro method of recombining an Insert Donor plasmid (here, pEZC705) with a Vector Donor plasmid (here, pEZC726), and obtaining Product DNA and Byproduct daughter molecules. The two recombination sites are attP and loxP on the Vector Donor. On one segment defined by these sites is a kanamycin resistance
gene whose promoter has been replaced by the tetOP operator/promoter from transposon Tn10. See, e.g., Sizemore et al., *Nucl. Acids Res.* 18(10):2875 (1990). In the absence of tet repressor protein, *E. coli* RNA polymerase transcribes the kanamycin resistance gene from the tetOP. If tet repressor is present, it binds to tetOP and blocks transcription of the kanamycin resistance gene. The other segment of pEZC726 has the tet repressor gene expressed by a constitutive promoter. Thus cells transformed by pEZC726 are resistant to chloramphenicol, because of the chloramphenicol acetyltransferase gene on the same segment as tetR, but are sensitive to kanamycin. The recombinant-mediated reactions result in separation of the tetR gene from the regulated kanamycin resistance gene. This separation results in kanamycin resistance only in cells receiving the desired recombination product. The first recombination reaction is driven by the addition of the recombinase called Integrate. The second recombination reaction is driven by adding the recombinase Cre to the Cointegrate (here, pEZC7 Cointegr).

**FIG. 2B** depicts a restriction map of pEZC705.
**FIG. 2C** depicts a restriction map of pEZC726.
**FIG. 2D** depicts a restriction map of pEZC7 Coint.
**FIG. 2E** depicts a restriction map of Intprod.
**FIG. 2F** depicts a restriction map of Intbypro.
**FIG. 3A** depicts an in vitro method of recombining an Insert Donor plasmid (here, pEZC602) with a Vector Donor plasmid (here, pEZC629), and obtaining Product (here, EZC6prod) and Byproduct (here, EZC6Bypr) daughter molecules. The two recombination sites are loxP and loxP 511. One segment of pEZC629 defined by these sites is a kanamycin resistance gene whose promoter has been replaced by the tetOP operator/promoter from transposon Tn10. In the absence of tet repressor protein, *E. coli* RNA polymerase transcribes the kanamycin resistance gene from the tetOP. If tet repressor is present, it binds to tetOP and blocks transcription of the kanamycin resistance gene. The other segment of pEZC629 has the tet repressor gene expressed by a constitutive promoter. Thus cells transformed by pEZC629 are resistant to chloramphenicol, because of the chloramphenicol acetyltransferase gene on the same segment as tetR, but are sensitive to kanamycin. The reactions result in separation of the tetR gene from the regulated kanamycin resistance gene. This separation results in kanamycin resistance only in cells receiving the desired recombination product. The first and the second recombination events are driven by the addition of the same recombinase, Cre.

**[0108]** **FIG. 4B** depicts a restriction map of pEZC843.
**[0109]** **FIG. 4C** depicts a restriction map of pEZC1003.
**[0110]** **FIG. 4D** depicts a restriction map of CMVBypro.
**[0111]** **FIG. 4E** depicts a restriction map of CMVprod.
**[0112]** **FIG. 4F** depicts a restriction map of CMVcoint.
**[0113]** **FIG. 5A** depicts a vector diagram of pEZC1301.
**[0114]** **FIG. 5B** depicts a vector diagram of pEZC1305.
**[0115]** **FIG. 5C** depicts a vector diagram of pEZC1309.
**[0116]** **FIG. 5D** depicts a vector diagram of pEZC1313.
**[0117]** **FIG. 5E** depicts a vector diagram of pEZC1317.
**[0118]** **FIG. 5F** depicts a vector diagram of pEZC1321.
**[0119]** **FIG. 5G** depicts a vector diagram of pEZC1405.
**[0120]** **FIG. 5H** depicts a vector diagram of pEZC1502.
**[0121]** **FIG. 6A** depicts a vector diagram of pEZC1603.
**[0122]** **FIG. 6B** depicts a vector diagram of pEZC1706.
**[0123]** **FIG. 7A** depicts a vector diagram of pEZC2901.
**[0124]** **FIG. 7B** depicts a vector diagram of pEZC2913.
**[0125]** **FIG. 7C** depicts a vector diagram of pEZC3101.
**[0126]** **FIG. 7D** depicts a vector diagram of pEZC1802.
**[0127]** **FIG. 8A** depicts a vector diagram of pGEX-2TK.
**[0128]** **FIG. 8B** depicts a vector diagram of pEZC3501.
**[0129]** **FIG. 8C** depicts a vector diagram of pEZC3601.
**[0130]** **FIG. 8D** depicts a vector diagram of pEZC3609.
**[0131]** **FIG. 8E** depicts a vector diagram of pEZC3617.
**[0132]** **FIG. 8F** depicts a vector diagram of pEZC3606.
**[0133]** **FIG. 8G** depicts a vector diagram of pEZC3613.
**[0134]** **FIG. 8H** depicts a vector diagram of pEZC3621.
**[0135]** **FIG. 8I** depicts a vector diagram of GST-CAT.
**[0136]** **FIG. 8J** depicts a vector diagram of GST-phoA.
**[0137]** **FIG. 8K** depicts a vector diagram of pEZC3201.
**[0138]** **FIG. 9A** depicts a diagram of 5.2 kb PCR prod.
**[0139]** **FIG. 9B** depicts a vector diagram of pEZC1202.
**[0140]** **FIG. 9C** depicts a vector diagram of 5.2 kb clone.
**[0141]** **FIG. 10A** depicts a vector diagram of pEZC5601.
**[0142]** **FIG. 10B** depicts a vector diagram of pEZC6701.
**[0143]** **FIG. 10C** depicts a vector diagram of attR product.
**[0144]** **FIG. 10D** depicts attR product.
FIG. 11A depicts a vector diagram of pEZC7102.

FIG. 11B depicts a vector diagram of pEZC7501.

FIG. 11C depicts the attL product.

FIG. 12A depicts an amp PCR product with terminal attB sites.

FIG. 12B depicts a tet PCR product with terminal attB sites.

FIG. 12C depicts a restriction map of amp7102.

FIG. 12D depicts a restriction map of tet 7102.

DETAILED DESCRIPTION OF THE INVENTION

It is unexpectedly discovered in the present invention that reversible and/or repeatable cloning and subcloning reactions can be used to manipulate nucleic acids to form chimeric nucleic acids using recombination proteins and recombination sites. Recombinational cloning according to the present invention thus uses recombination proteins with recombinant nucleic acid molecules having at least one selected recombination site for moving or exchanging segments of nucleic acid molecules, in vitro and in vivo.

These methods use recombination reactions to generate chimeric DNA or RNA molecules that have the desired characteristic(s) and/or nucleic acid segment(s). The methods of the invention provide a means in which nucleic acid molecules of interest may be moved or transferred into any number of vector systems. In accordance with the invention, such transfer to various vector systems may be accomplished separately, sequentially or in mass (e.g. into any number of different vectors in one step). The improved specificity, speed and/or yields of the present invention facilitates DNA or RNA cloning, subcloning, regulation or exchange useful for any related purpose. Such purposes include in vitro recombination of DNA or RNA segments and in vitro or in vivo insertion or modification of transcribed, replicated, isolated or genomic DNA or RNA.

Definitions

In the description that follows, a number of terms used in recombinant DNA technology are utilized extensively. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

Byproduct: a daughter molecule (a new clone produced after the second recombination event during the recombinational cloning process) lacking the segment which is desired to be cloned or subcloned.

Cointegrate: at least one recombination intermediate nucleic acid molecule of the present invention that contains both parental (starting) molecules. It will usually be circular. In some embodiments it can be linear.

Host: any prokaryotic or eukaryotic organism that can be a recipient of the recombinational cloning Product. A "host," as the term is used herein, includes prokaryotic or eukaryotic organisms that can be genetically engineered. For examples of such hosts, see Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1982).

Insert or inserts: include the desired nucleic acid segment or a population of nucleic acid segments (segment A of FIG. 1) which may be manipulated by the methods of the present invention. Thus, the terms Insert(s) are meant to include a particular nucleic acid (preferably DNA) segment or a population of segments. Such Insert(s) can comprise one or more genes.

Insert Donor: is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the Insert. The Insert Donor molecule comprises the Insert flanked on both sides with recombination sites. The Insert Donor can be linear or circular. In one embodiment of the invention, the Insert Donor is a circular DNA molecule and further comprises a cloning vector sequence outside of the recombination signals (see FIG. 1). When a population of Inserts or population of nucleic acid segments are used to make the Insert Donor, a population of Insert Donors result and may be used in accordance with the invention.

Product: is one the desired daughter molecules comprising the A and D sequences which is produced after the second recombination event during the recombinational cloning process (see FIG. 1). The Product contains the nucleic acid which was to be cloned or subcloned. In accordance with the invention, when a population of Insert Donors are used, the resulting population of Product molecules will contain all or a portion of the population of Inserts of the Insert Donors and preferentially will contain a representative population of the original molecules of the Insert Donors.

Promoter: is a DNA sequence generally described as the S-region of a gene, located proximal to the start codon. The transcription of an adjacent DNA segment is initiated at the promoter region. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions.

Recognition sequence: Recognition sequences are particular sequences which a protein, chemical compound, DNA, or RNA molecule (e.g., restriction endonuclease, a modification methylase, or a recombinase) recognizes and binds. In the present invention, a recognition sequence will usually refer to a recombination site. For example, the recognition sequence for Cre recombinase is loxP which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See FIG. 1 of Sauer, B., Current Opinion in Biotechnology 5:521-527 (1994). Other examples of recognition sequences are the attB, attP, attL, and attR sequences which are recognized by the recombinase enzyme λ Integrase. attB is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. attP is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IIH), FIS and excisionase (Xis). See Landy, Current Opinion in Biotechnology 3:699-707 (1993). Such sites may also be engineered according to the present invention to enhance pro-
duction of products in the methods of the invention. When such engineered sites lack the PI or HI domains to make the recombination reactions irreversible (e.g., attR or attP), such sites may be designated attR' or attP' to show that the domains of these sites have been modified in some way.

[0164] Recombinease: is an enzyme which catalyzes the exchange of DNA segments at specific recombination sites.

[0165] Recombinational Cloning: is a method described herein, whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, in vitro or in vivo.

[0166] Recombination proteins: include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites. See, Landy (1994), infra.

[0167] Repression cassette: is a nucleic acid segment that contains a repressor of a Selectable marker present in the subcloning vector.

[0168] Selectable marker: is a DNA segment that allows one to select for or against a molecule or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to, production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like. Examples of Selectable markers include but are not limited to: (1) DNA segments that encode products which provide resistance against otherwise toxic compounds (e.g., antibiotics); (2) DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); (3) DNA segments that encode products which suppress the activity of a gene product; (4) DNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β-galactosidase, green fluorescent protein (GFP), and cell surface proteins); (5) DNA segments that bind products which are otherwise detrimental to cell survival and/or function; (6) DNA segments that otherwise inhibit the activity of any of the DNA segments described in Nos. 1-5 above (e.g., antisense oligonucleotides); (7) DNA segments that bind products that modify a substrate (e.g., restriction endonucleases); (8) DNA segments that can be used to isolate or identify a desired molecule (e.g., specific protein binding sites); (9) DNA segments that encode a specific nucleotide sequence which can be otherwise non-functional (e.g., for PCR amplification of subpopulations of molecules); (10) DNA segments, which when absent, directly or indirectly confer resistance or sensitivity to particular compounds; and/or (11) DNA segments that encode products which are toxic in recipient cells.

[0169] Selection scheme: is any method which allows selection, enrichment, or identification of a desired Product or Product(s) from a mixture containing the Insert Donor, Vector Donor, any intermediates (e.g., a CoinTEGRATE), and/or Byproducts. The selection schemes of one preferred embodiment have at least two components that are either linked or unlinked during recombiningation cloning. One component is a Selectable marker. The other component controls the expression in vitro or in vivo of the Selectable marker, or survival of the cell harboring the plasmid carrying the Selectable marker. Generally, this controlling element will be a repressor or inducer of the Selectable marker, but other means for controlling expression of the Selectable marker can be used. Whether a repressor or activator is used will depend on whether the marker is for a positive or negative selection, and the exact arrangement of the various DNA segments, as will be readily apparent to those skilled in the art. A preferred requirement is that the selection scheme results in selection of or enrichment for only one or more desired Products. As defined herein, selecting for a DNA molecule includes (a) selecting or enriching for the presence of the desired DNA molecule, and (b) selecting or enriching against the presence of DNA molecules that are not the desired DNA molecule.

[0170] In one embodiment, the selection schemes (which can be carried out in reverse) will take one of three forms, which will be discussed in terms of FIG. 1. The first, exemplified herein with a Selectable marker and a repressor therefore, selects for molecules having segment D and lacking segment C. The second selects against molecules having segment C and for molecules having segment D. Possible embodiments of the second form would have a DNA segment carrying a gene toxic to cells into which the in vitro reaction products are to be introduced. A toxic gene can be a DNA that is expressed as a toxic gene product (a toxic protein or RNA), or can be toxic in and of itself. (In the latter case, the toxic gene is understood to carry its classical definition of “heritable trait”.)

[0171] Examples of such toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (e.g., DpnII), apoptosis-related genes (e.g., ASK1 or members of the bcl-2/ced-9 family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic DNA sequences, bacteriophage lytic genes such as those from φX174 or bacteriophage T4; antibiotic sensitivity genes such as rpsL, antimicrobial sensitivity genes such as pheS, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GAlA-1, and genes that kill hosts in the absence of a suppressing function, e.g., kicB or ccdB. A toxic gene can alternatively be selectable in vitro, e.g., a restriction site.

[0172] Many genes coding for restriction endonucleases operably linked to inducible promoters are known, and may be used in the present invention.

[0173] See, e.g., U.S. Pat. Nos. 4,960,707 (DpnI and DpnII); 5,000,333, 5,082,784 and 5,192,675 (KpnI); 5,147,800 (NgoAlIII and NgoAI); 5,179,015 (FspI and HaellIII); 5,200,333 (Haell and TaqI); 5,248,605 (HpaII); 5,312,746 (Clai); 5,231,021 and 5,304,480 (Xhol and Xholl); 5,334,526 (AluI); 5,470,740 (NsiI); 5,534,428 (SstI/Sacl); 5,502,248 (NcoI); 5,139,942 (Ndel); and 5,098,839 (PacI). See also Wilson, G. G., Nucl. Acids Res. 19:2539-2566 (1991); and Lunnan, K. D., et al., Gene 74:25-32 (1988).

[0174] In the second form, segment D carries a Selectable marker. The toxic gene would eliminate transformants harboring the Vector Donor, CoinTEGRATE, and Byproduct molecules, while the Selectable marker can be used to select for cells containing the Product and against cells harboring only the Insert Donor.

[0175] The third form selects for cells that have both segments A and D in cis on the same molecule, but not for
cells that have both segments in trans on different molecules. This could be embodied by a Selectable marker that is split into two inactive fragments, each one on segments A and D.

[0176] The fragments are so arranged relative to the recombination sites that when the segments are brought together by the recombination event, they reconstitute a functional Selectable marker. For example, the recombination event can link a promoter with a structural gene, can link two fragments of a structural gene, or can link genes that encode a heterodimeric gene product needed for survival, or can link portions of a replicon.

[0177] Site-specific recombinase: is a type of recombinase which typically has at least the following four activities (or combinations thereof): (1) recognition of one or two specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase activity to resell the cleaved strands of nucleic acid. See Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994). Conservative site-specific recombinase is distinguished from homologous recombinase and transposition by a high degree of specificity for both partners. The strand exchange mechanism involves the cleavage and rejoicing of specific DNA sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

[0178] Subcloning vector: is a cloning vector comprising a circular or linear nucleic acid molecule which includes preferably an appropriate replicon. In the present invention, the subcloning vector (segment D in FIG. 1) can also contain functional and/or regulatory elements that are desired to be incorporated into the final product to act upon or with the cloned DNA insert (segment A in FIG. 1). The subcloning vector can also contain a Selectable marker (preferably DNA).

[0179] Vector: is a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an Insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated in vitro or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A Vector can have one or more restriction endonuclease recognition sites at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites, e.g., for PCR, transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, Selectable markers, etc. Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of homologous recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Pat. No. 5,334,375, entirely incorporated herein by reference), T:A cloning, and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers suitable for use in the identification of cells transformed with the cloning vector.

[0180] Vector Donor: is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the DNA segments comprising the DNA vector which is to become part of the desired Product. The Vector Donor comprises a subcloning vector D (or it can be called the cloning vector if the Insert Donor does not already contain a cloning vector) and a segment C flanked by recombination sites (see FIG. 1). Segments C and/or D can contain elements that contribute to selection for the desired Product daughter molecule, as described above for selection schemes. The recombinase signals can be the same or different, and can be acted upon by the same or different recombinases. In addition, the Vector Donor can be linear or circular.

[0181] Primer: refers to a single stranded or double stranded oligonucleotide that is extended by covalent bonding of nucleotide monomers during amplification or polymerization of a nucleic acid molecule (e.g. a DNA molecule). In a preferred aspect, the primer comprises one or more recombination sites or portions of such recombination sites. Portions of recombination sites comprise at least 2 bases, at least 5 bases, at least 10 bases or at least 20 bases of the recombination sites of interest. When using portions of recombination sites, the missing portion of the recombination site may be provided by the newly synthesized nucleic acid molecule. Such recombination sites may be located within and/or at one or both termini of the primer. Preferably, additional sequences are added to the primer adjacent to the recombination site(s) to enhance or improve recombination and/or to stabilize the recombination site during recombination. Such stabilization sequences may be any sequences (preferably G/C rich sequences) of any length. Preferably, such sequences range in size from 1 to about 100 bases, 1 to about 500 bases, and 1 to about 1000 bases, 1 to about 60 bases, 1 to about 25, 1 to about 10, 2 to about 10 and preferably about 4 bases. Preferably, such sequences are greater than 1 base in length and preferably greater than 2 bases in length.

[0182] Template: refers to double stranded or single stranded nucleic acid molecules which are to be amplified, synthesized or sequenced. In the case of double stranded molecules, denaturation of its strands to form a first and a second strand is preferably performed before these molecules will be amplified, synthesized or sequenced, or the double stranded molecule may be used directly as a template. For single stranded templates, a primer complementary to a portion of the template is hybridized under appropriate conditions and one or more polypeptides having polymerase activity (e.g. DNA polymerases and/or reverse transcriptases) may then synthesize a nucleic acid molecule complementary to all or a portion of said template. Alternatively, for double stranded templates, one or more promoters may be used in combination with one or more polypeptides to make nucleic acid molecules complementary to all or a portion of the template. The newly synthesized molecules, according to the invention, may be equal or shorter in length than the original template. Additionally, a population of nucleic acid templates may be used during synthesis or amplification to produce a population of nucleic acid molecules typically representative of the original template population.

[0183] Adapter: is an oligonucleotide or nucleic acid fragment or segment (preferably DNA) which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added
to a circular or linear Insert Donor linear molecule, although the adapters are preferably added at or near one or both termini of a linear molecule. Preferably, adapters are positioned to be located on both sides (flanking) a particularly nucleic acid molecule of interest. In accordance with the invention, adapters may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g., restriction digest and ligation). For example, adapters may be added to a circular molecule by first digesting the molecule with an appropriate restriction enzyme, adding the adapter at the cleavage site and reforming the circular molecule which contains the adapter(s) at the site of cleavage. Alternatively, adapters may be ligated directly to one or more and preferably both termini of a linear molecule thereby resulting in linear molecule(s) having adapters at one or both termini. In one aspect of the invention, adapters may be added to a population of linear molecules, (e.g. a cDNA library or genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adapters at one and preferably both termini of all or substantial portion of said population.

[0184] Library: refers to a collection of nucleic acid molecules (circular or linear). In one preferred embodiment, a library is representative of all or a significant portion of the DNA content of an organism (a "genomic" library), or a set of nucleic acid molecules representative of all or a significant portion of the expressed genes (a cDNA library) in a cell, tissue, organ or organism. A library may also comprise random sequences made by de novo synthesis, mutagenesis of one or more sequences and the like. Such libraries may or may not be contained in one or more vectors.

[0185] Amplification: refers to any in vitro method for increasing a number of copies of a nucleotide sequence with the use of a polymerase. Nucleic acid amplification results in the incorporation of nucleotides into a DNA and/or RNA molecule or primer thereby forming a new molecule complementary to a template. The formed nucleic acid molecule and its template can be used as templates to synthesize additional nucleic acid molecules. As used herein, one amplification reaction may consist of many rounds of replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 5-100 "cycles" of denaturation and synthesis of a DNA molecule.

[0186] Oligonucleotide: refers to a synthetic or natural molecule comprising a covalently linked sequence of nucleotides which are joined by a phosphodiester bond between the 3’ position of the deoxyribose or ribose of one nucleotide and the 5’ position of the deoxyribose or ribose of the adjacent nucleotide.

[0187] Nucleotide: refers to a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid sequence (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTP, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dGTP, dUTP, and dTTP. Such derivatives include, for example, [α-32P]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to deoxyribonucleoside triphosphates (dNTPs) and their derivatives. Illustrated examples of deoxyribonucleoside triphosphates include, but are not limited to, dATP, dCTP, dGTP, dUTP, and dTTP. According to the present invention, a “nucleotide” may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.

[0188] Hybridization: The terms “hybridization” and “hybridizing” refers to base pairing of two complementary single-stranded nucleic acid molecules (RNA and/or DNA) to give a double stranded molecule. As used herein, two nucleic acid molecules may be hybridized, although the base pairing is not completely complementary. Accordingly, mismatched bases do not prevent hybridization of two nucleic acid molecules provided that appropriate conditions, well known in the art, are used.

[0189] Other terms used in the fields of recombinant DNA technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable art.

[0190] Recombination Schemes

[0191] One general scheme for an in vitro or in vivo method of the invention is shown in FIG. 1, where the Insert Donor and the Vector Donor can be either circular or linear DNA, but is shown as circular. Vector D is exchanged for the original cloning vector B. The Insert Donor need not comprise a vector. The method of the invention allows the Inserts A to be transferred into any number of vectors. According to the invention, the Inserts may be transferred to a particular Vector or may be transferred to a number of vectors in one step. Additionally, the Inserts may be transferred to any number of vectors sequentially, for example, by using the Product DNA molecule as the Insert Donor in combination with a different Vector Donor. The nucleic acid molecule of interest may be transferred into a new vector thereby producing a new Product DNA molecule. The new Product DNA molecule may then be used as starting material to transfer the nucleic acid molecule of interest into a new vector. Such sequential transfers can be performed a number of times in any number of different vectors. Thus the invention allows for cloning or subcloning nucleic acid molecules and because of the ease and simplicity, these methods are particularly suited for high throughput applications. In accordance with the invention, it is desirable to select for the daughter molecule containing elements A and D and against other molecules, including one or more Cointegrate(s). The square and circle are different sets of recombination sites (e.g., lox sites or att sites). Segment A or D can contain at least one Selection Marker, expression signals, origins of replication, or specialized functions for detecting, selecting, expressing, mapping or sequencing DNA, where D is used in this example. This scheme can also be reversed according to the present invention, as described herein. The resulting product of the reverse reaction (e.g. the Insert Donor) may then be used in combination with one or a number of vectors to produce new product molecules in which the Inserts are contained by any number of vectors.

[0192] Examples of desired DNA segments that can be part of Element A or D include, but are not limited to, PCR products, large DNA segments, genomic clones or fragments, cDNA clones or fragments, functional elements, etc., and genes or partial genes, which encode useful nucleic acids or proteins. Moreover, the recombination cloning of the present invention can be used to make ex vivo and in
vivo gene transfer vehicles for protein expression (native or fusion proteins) and/or gene therapy.

[0193] In FIG. 1, the scheme provides the desired Product as containing A and Vector D, as follows. The Insert Donor (containing A and B) is first recombined at the square recombination sites by recombination proteins, with the Vector Donor (containing C and D), to form a Co-integrate having each of A-D-C-B. Next, recombination occurs at the circle recombination sites to form Product DNA (A and D) and Byproduct DNA (C and B). However, if desired, two or more different Co-integrates can be formed to generate two or more Products.

[0194] In one embodiment of the present in vitro or in vivo recombinational cloning method, a method for selecting at least one desired Product DNA is provided. This can be understood by consideration of the map of plasmid pEZC726 depicted in FIG. 2. The two exemplary recombination sites are attP and loxP. On one segment defined by these sites is a kanamycin resistance gene whose promoter has been replaced by the tetOP operator/promoter from transposon Tn10. In the absence of tet repressor protein, E. coli RNA polymerase transcribes the kanamycin resistance gene from the tetOP. If tet repressor is present, it binds to tetOP and blocks transcription of the kanamycin resistance gene. The other segment of pEZC726 has the tet repressor gene expressed by a constitutive promoter. Thus cells transformed by pEZC726 are resistant to chloramphenicol, because of the chloramphenicol acetyl transferase gene on the same segment as tetR, but are sensitive to kanamycin. The recombination reactions result in separation of the tetR gene from the regulated kanamycin resistance gene. This separation results in kanamycin resistance in cells receiving the desired recombination Product.

[0195] Two different sets of plasmids were constructed to demonstrate the in vitro method. One set, for use with Cre recombinase only (cloning vector 602 and subcloning vector 629 (FIG. 3)) contained loxP and loxP 511 sites. A second set, for use with Cre and integrate (cloning vector 705 and subcloning vector 726 (FIG. 2)) contained loxP and att sites. The efficiency of production of the desired daughter plasmid was about 60 fold higher using both enzymes than using Cre alone. Nineteen of twenty four colonies from the Cre-only reaction contained the desired product, while thirty eight of thirty eight colonies from the integrate plus Cre reaction contained the desired product plasmid.

[0196] A variety of other selection schemes can be used that are known in the art as they can suit a particular purpose for which the recombinational cloning is carried out. Depending upon individual preferences and needs, a number of different types of selection schemes can be used in the recombinational cloning or subcloning methods of the present invention. The skilled artisan can take advantage of the availability of the many DNA segments or methods for making them and the different methods of selection that are routinely used in the art. Such DNA segments include but are not limited to those which encode an activity such as, but not limited to, production of RNA, peptide, or protein, or providing a binding site for such RNA, peptide, or protein. Examples of DNA molecules used in devising a selection scheme are given above, under the definition of “selection scheme”

[0197] Additional examples include but are not limited to:

[0198] (i) Generation of new primer sites for PCR (e.g., juxtaposition of two DNA sequences that were not previously juxtaposed);

[0199] (ii) Inclusion of a DNA sequence acted upon by a restriction endonuclease or other DNA modifying enzyme, chemical, ribozyme, etc.;

[0200] (iii) Inclusion of a DNA sequence recognized by a DNA binding protein, RNA, DNA, chemical, etc.) (e.g., for use as an affinity tag for selecting for or excluding from a population (Davis, Nucl. Acids Res. 24:702-706 (1996); J. Virol. 69: 8027-8034 (1995)) or for juxtaposing a promoter for in vitro transcription;

[0201] (iv) In vitro selection of RNA ligands for the ribosomal L22 protein associated with Epstein-Barr virus-expressed RNA by using randomized and cDNA-derived RNA libraries;

[0202] (vii) The positioning of functional elements whose activity requires a specific orientation or juxtaposition (e.g., (a) a recombination site which reacts poorly in trans, but when placed in cis, in the presence of the appropriate proteins, results in recombination that destroys certain populations of molecules; (e.g., reconstitution of a promoter sequence that allows in vitro RNA synthesis). The RNA can be used directly, or can be reverse transcribed to obtain the desired DNA construct;

[0203] (vii) Selection of the desired product by size (e.g., fractionation) or other physical property of the molecule(s); and

[0204] (viii) Inclusion of a DNA sequence required for a specific modification (e.g., methylation) that allows its identification.

[0205] After formation of the Product and Byproduct in the method of the present invention, the selection step can be carried out in vivo or in vitro or in vivo depending upon the particular selection scheme which has been optionally devised in the particular recombinational cloning procedure.

[0206] For example, an in vitro method of selection can be devised for the Insert Donor and Vector Donor DNA molecules. Such scheme can involve engineering a rare restriction site in the starting circular vectors in such a way that after the recombination events the rare cutting sites end up in the Byproduct. Hence, when the restriction enzyme which binds and cuts at the rare restriction site is added to the reaction mixture in vitro, all of the DNA molecules carrying the rare cutting site, i.e., the starting DNA molecules, the Cointegrate, and the Byproduct, will be cut and rendered nonreplicable in the intended host cell. For example, cutting sites in segments B and C (see FIG. 1) can be used to select against all molecules except the Product. Alternatively, only a cutting site in C is needed if one is able to select for segment D, e.g., by a drug resistance gene not found on B.

[0207] Similarly, an in vitro selection method can be devised when dealing with linear DNA molecules. DNA sequences complementary to a PCR primer sequence can be so engineered that they are transferred, through the recombinational cloning method, only to the Product molecule. After the reactions are completed, the appropriate primers
are added to the reaction solution and the sample is subjected to PCR. Hence, all or part of the Product molecule is amplified.

[0208] Other in vivo selection schemes can be used with a variety of host cells, particularly E. coli lines. One is to put a repressor gene on one segment of the subcloning plasmid, and a drug marker controlled by that repressor on the other segment of the same plasmid. Another is to put a killer gene on segment C of the subcloning plasmid (FIG. 1). Of course a way must exist for growing such a plasmid, i.e., there must exist circumstances under which the killer gene will not kill. There are a number of these genes known which require particular strains of E. coli. One such scheme is to use the restriction enzyme DpnII, which will not cleave unless its recognition sequence GAIC is methylated. Many popular common E. coli strains methylate GAIC sequences, but there are mutants in which cloned DpnII can be expressed without harm. Other restriction enzyme genes may also be used as a toxic gene for selection. In such cases, a host containing a gem encoding the corresponding methylase gene provides protected host for use in the invention. Similarly, the ccdB protein is a potent poison of DNA gyrase, efficiently trapping gyrase molecules in a covalent complex, resulting in DNA strand breakage and cell death. Mutations in the gyrA subunit of DNA gyrase, specifically the gyrA462 mutation, confers resistance to ccdB (Bernard and Couturier, J. Mol. Bio. 226 (1992) 735-745). An E. coli strain, DB2, has been constructed that contains the gyrA462 mutation. DB2 cells containing plasmids that express the ccdB gene are not killed by ccd B. This strain is available from Life Technologies and has been deposited on Oct. 14, 1997 with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, IL 61604 as deposit number NRRL B-21852.

[0209] Of course analogous selection schemes can be devised for other host organisms. For example, the tet repressor/operator of Tn10 has been adapted to control gene expression in eukaryotes (Gossen, M., and Bujard, H., Proc. Natl. Acad. Sci. USA 89:5547-5551 (1992)). Thus the same control of drug resistance by the tet repressor exemplified herein or other selection schemes described herein can be applied to select for Product in eukaryotic cells.

[0210] Recombination Proteins

[0211] In the present invention, the exchange of DNA segments is achieved by the use of recombinase proteins, including recombinases and associated co-factors and proteins. Various recombinase proteins are described in the art. Examples of such recombinases include:

[0212] Cre: A protein from bacteriophage P1 (Abreuksi and Hoess, J. Biol. Chem. 259(3):1509-1514 (1984)) catalyzes the exchange (i.e., causes recombination) between 34 bp DNA sequences called loxP (locus of crossover) sites (See Hoess et al., Nucl. Acids Res. 14(5):2287 (1986)). Cre is available commercially (Novagen, Catalog No. 69247-1). Recombination mediated by Cre is freely reversible. From thermodynamic considerations it is not surprising that Cre-mediated integration (recombination between two molecules to form one molecule) is much less efficient than Cre-mediated excision (recombination between two loxP sites in the same molecule to form two daughter molecules). Cre works in simple buffers with either magnesium or spermidine as a co-factor, as well known in the art. The DNA substrates can be either linear or supercoiled. A number of mutant loxP sites have been described (Hoess et al., supra). One of these, loxP 511, recombines with another loxP 511 site, but will not recombine with a loxP site.

[0213] Integrate: A protein from bacteriophage lambda that mediates the integration of the lambda genome into the E. coli chromosome. The bacteriophage lambda Int recombinational proteins promote recombination between its substrate att sites as part of the formation or induction of a lysogenic state. Reversibility of the recombination reactions results from two independent pathways for integrative and excisive recombination. Each pathway uses a unique, but overlapping, set of the 15 protein binding sites that comprise att site DNAs. Cooperative and competitive interactions involving four proteins (Int, Xis, IIF and FIS) determine the direction of recombination.

[0214] Integrative recombination involves the Int and IIF proteins and sites attP (240 bp) and attB (25 bp). Recombination results in the formation of two new sites: attL and attR. Excisive recombination requires Int, IIF, and Xis, and sites attL and attR to generate attP and attB. Under certain conditions, FIS stimulates excisive recombination. In addition to these normal reactions, it should be appreciated that attP and attB, when placed on the same molecule, can promote excisive recombination to generate two excision products, one with attL and one with attR. Similarly, intramolecular recombination between molecules containing attL and attR, in the presence of Int, IIF and Xis, can result in integrative recombination and the generation of attP and attB. Hence, by flanking DNA segments with appropriate combinations of engineered att sites, in the presence of the appropriate recombinase proteins, one can direct excisive or integrative recombination, as reverse reactions of each other.

[0215] Each of the att sites contains a 15 bp core sequence; individual sequence elements of functional significance lie within, outside, and across the boundaries of this common core (Landy, A., Annu. Rev. Biochem. 58:913 (1989)). Efficient recombination between the various att sites requires that the sequence of the central common region be identical between the recombining partners, however, the exact sequence is now found to be modifiable. Consequently, derivatives of the att site with changes within the core are now discovered to recombine as least as efficiently as the native core sequences.

[0216] Integrate acts to recombine the attP site on bacteriophage lambda (about 240 bp) with the attB site on the E. coli genome (about 25 bp) (Weisberg, R. A. and Landy, A. in Lambda II, p. 211 (1983), Cold Spring Harbor Laboratory), to produce the integrated lambda genome flanked by attL (about 100 bp) and attR (about 160 bp) sites. In the absence of Xis (see below), this reaction is essentially irreversible. The integration reaction mediated by integrate and IIF works in vitro, with simple buffer containing spermidine. Integrate can be obtained as described by Nash, H. A., Methods of Enzymology 100:210-216 (1983). IIF can be obtained as described by Filutowicz, M., et al., Gene 147:149-150 (1994).

[0217] Numerous recombinase systems from various organisms can also be used, based on the teaching and guidance provided herein. See, e.g., Hoess et al., Nucl. Acids Research 14(6):2287 (1986); Abreuksi et al., J. Biol.
Members of a second family of site-specific recombinases, the resolvase family (e.g., β, Tn3 resolvase, Hin, Gin, and Cin) are also known. Members of this highly related family of recombinases are typically constrained to intramolecular reactions (e.g., inversions and excisions) and can require host-encoded factors. Mutants have been isolated that relieve some of the requirements for host factors (Maeser and Kahmann (1991) Mol. Gen. Genet. 230:170-176), as well as some of the constraints of intramolecular recombination.

Other site-specific recombinases similar to λ Int and similar to P1 Cre can be substituted for Int and Cre. Such recombinases are known. In many cases the purification of such other recombinases has been described in the art. In cases when they are not known, cell extracts can be used or the enzymes can be partially purified using procedures described for Cre and Int.

While Cre and Int are described in detail for reasons of example, many related recombinase systems exist and their application to the desired invention is also provided according to the present invention. The integrase family of site-specific recombinases can be used to provide alternative recombination proteins and recombination sites for the present invention, as site-specific recombination proteins encoded by, for example bacteriophage λ, phi 80, P22, P2, 186, P4 and P1. This group of proteins exhibits an unexpectedly large diversity of sequences. Despite this diversity, all of the recombinases can be aligned in their C-terminal halves. A 40-residue region near the Cr terminus is particularly well conserved in all the proteins and is homologous to a region near the C terminus of the yeast 2 μ plasmid Flp protein. Three positions are perfectly conserved within this family: histidine, arginine and tyrosine are found at respective alignment positions 396, 399 and 433 within the well-conserved C-terminal region. These residues contribute to the active site of this family of recombinases, and suggest that tyrosine-433 forms a transient covalent linkage to DNA during strand cleavage and rejoining. See, e.g., Argos, P. et al., EMBO J. 5:433-40 (1986).


Alternatively, IS231 and other Bacillus thuringiensis transposable elements could be used as recombination proteins and recombination sites. Bacillus thuringiensis is an entomopathogenic bacterium whose toxicity is due to the presence in the sporangia of delta-endotoxin crystals active against agricultural pests and vectors of human and animal diseases. Most of the genes coding for these toxin proteins are plasmid-borne and are generally structurally associated with insertion sequences (IS231, IS232, IS240, ISBT1 and ISBT2) and transposons (Tn4430 and Tn5401). Several of these mobile elements have been shown to be active and participate in the crystal gene mobility, thereby contributing to the variation of bacterial toxicity.

[0223] Structural analysis of the iso-IS231 elements indicates that they are related to IS1151 from Clostridium perfringens and distantly related to IS4 and IS186 from Escherichia coli. Like the other IS4 family members, they contain a conserved transposase-integrase motif found in other IS families and retroviruses. Moreover, functional data gathered from IS231A in Escherichia coli indicate a non-replicative mode of transposition, with a preference for specific targets. Similar results were also obtained in Bacillus subtilis and B. thuringiensis. See, e.g., Mahillon, J. et al., Genetica 93:13-26 (1994); Campbell, J Bacteriol. 7495-7499 (1992).

[0224] An unrelated family of recombinases, the transposases, have also been used to transfer genetic information between replicons. Transposons are structurally variable, being described as simple or compound, but typically encode the recombinase gene flanked by DNA sequences organized in inverted orientations. Integration of transposons can be random or highly specific. Representatives such as Tn7, which are highly site-specific, have been applied to the efficient movement of DNA segments between replicons (Lucklow et al. 1993. J. Virol 67:4566-4579).

[0225] A related element, the integron, are also translocable-promoting movement of drug resistance cassettes from one replicon to another. Often these elements are defective transposon derivatives. Transposon Tn21 contains a class I integron called In2. The integrase (Int1) from Tn2 is common to all integrons in this class and mediates recombination between two 59-bp elements or between a 59-bp element and an att site that can lead to insertion into a recipient integron. The integrase also catalyzes excisional recombination. (Hall, 1997. Ciba Found Symp 207:192; Francia et al., 1997. J Bacteriol 179:4419).

[0226] Group II introns are mobile genetic elements encoding a catalytic RNA and protein. The protein component possesses reverse transcriptase, maturase and an endonuclease activity, while the RNA possesses endonuclease activity and determines the sequence of the target site into which the intron integrates. By modifying portions of the RNA sequence, the integration sites into which the element integrates can be defined. Foreign DNA sequences can be incorporated between the ends of the intron, allowing targeting to specific sites. This process, termed retrohoming, occurs via a DNA:RNA intermediate, which is copied into cDNA and ultimately into double stranded DNA (Matsura et al., Genes and Dev 1997; Guo et al, EMBO J, 1997). Numerous intron-encoded homing endonucleases have been identified (Belfort and Roberts, 1997. NAR 25:3379). Such systems can be easily adopted for application to the described subcloning methods.

[0227] The amount of recombinase which is added to drive the recombination reaction can be determined by using
known assays. Specifically, titration assay is used to determine the appropriate amount of a purified recombinase enzyme, or the appropriate amount of an extract.

[0228] Engineered Recombination Sites

[0229] The above recombinases and corresponding recombinase sites are suitable for use in recombination cloning according to the present invention. However, wildtype recombination sites may contain sequences that reduce the efficiency or specificity of recombination reactions or the function of the Product molecules as applied in methods of the present invention. For example, multiple stop codons in attB, attR, attP, attL and 

[0230] Accordingly, the present invention also provides engineered recombination sites that overcome these problems. For example, att sites can be engineered to have one or multiple mutations to enhance specificity or efficiency of the recombination reaction and the properties of Product DNAs (e.g., attP, attR, attL, and att3 sites); to decrease reverse reaction (e.g., removing P1 and H1 from attR). The testing of these mutants determines which mutants yield sufficient recombinational activity to be suitable for recombination cloning according to the present invention.

[0231] Mutations can therefore be introduced into recombination sites for enhancing site specific recombination. Such mutations include, but are not limited to: recombination sites without translation stop codons that allow fusion proteins to be encoded; recombination sites recognized by the same proteins but differing in base sequence such that they react largely or exclusively with their homologous partners allowing multiple reactions to be contemplated; and mutations that prevent hairpin formation of recombination sites. Which particular reactions take place can be specified by which particular partners are present in the reaction mixture. For example, a tripartite protein fusion could be accomplished with parental plasmids containing recombination sites attR1 and attL, and attB3; attR1; attP3 and 10xP; and/or attR3 and 10xP; and/or attR3 and attL.

[0232] There are well known procedures for introducing specific mutations into nucleic acid sequences. A number of these are described in Ausubel, F. M. et al., Current Protocols in Molecular Biology, Wiley Interscience, New York (1989-1996). Mutations can be designed into oligonucleotides, which can be used to modify existing cloned sequences, or in amplification reactions. Random mutagenesis can also be employed if appropriate selection methods are available to isolate the desired mutant DNA or RNA. The presence of the desired mutations can be confirmed by sequencing the nucleic acid by well known methods.

[0233] The following non-limiting methods can be used to modify or mutate a core region of a given recombination site to provide mutated sites that can be used in the present invention:

[0234] 1. By recombination of two parental DNA sequences by site-specific (e.g. attL and attR to give attB) or other (e.g. homologous) recombination mechanisms where the parental DNA segments contain one or more base alterations resulting in the final mutated core sequence;

[0235] 2. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) directly of the desired core sequence;

[0236] 3. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) of parental DNA sequences, which are recombined to generate a desired core sequence;

[0237] 4. By reverse transcription of an RNA encoding the desired core sequence; and

[0238] 5. By de novo synthesis (chemical synthesis) of a sequence having the desired base changes.

[0239] The functionality of the mutant recombination sites can be demonstrated in ways that depend on the particular characteristic that is desired. For example, the lack of translation stop codons in a recombination site can be demonstrated by expressing the appropriate fusion proteins. Specificity of recombination between homologous partners can be demonstrated by introducing the appropriate molecules into in vitro reactions, and assaying for recombination products as described herein or known in the art. Other desired mutations in recombination sites might include the presence or absence of restriction sites, translation or transcription start signals, protein binding sites, and other known functionalities of nucleic acid base sequences. Genetic selection schemes for particular functional attributes in the recombination sites can be used according to known method steps. For example, the modification of sites to provide (from a pair of sites that do not interact) partners that do interact could be achieved by requiring deletion, via recombination between the sites, of a DNA sequence encoding a toxic substance. Similarly, selection for sites that remove translation stop sequences, the presence or absence of protein binding sites, etc., can be easily devised by those skilled in the art.

[0240] Accordingly, the present invention provides a nucleic acid molecule, comprising at least one DNA segment having at least two engineered recombination sites flanking a Selectable marker and/or a desired DNA segment, wherein at least one of said recombination sites comprises a core region having at least one engineered mutation that enhances recombination in vitro in the formation of a Cointegrate DNA or a Product DNA.

[0241] While in the preferred embodiment the recombination sites differ in sequence and do not interact with each other, it is recognized that sites comprising the same sequence can be manipulated to inhibit recombination with each other. Such conceptions are considered and incorporated herein. For example, a protein binding site can be engineered adjacent to one of the sites. In the presence of the protein that recognizes said site, the recombinase fails to access the site and the other site is therefore used preferentially. In the cointegrate this site can no longer react since it has been changed e.g. from attB to attL. In resolution of the cointegrate, the protein can be inactivated (e.g. by antibody, heat or a change of buffer) and the second site can undergo recombination.

[0242] The nucleic acid molecule can have at least one mutation that confers at least one enhancement of said
recombination, said enhancement selected from the group consisting of substantially (i) favoring integration; (ii) favoring recombination; (iii) relieving the requirement for host factors; (iii) increasing the efficiency of said Cointegrate DNA or Product DNA formation; and (iv) increasing the specificity of said Cointegrate DNA or Product DNA formation.

[0243] The nucleic acid molecule preferably comprises at least one recombination site derived from attP, attB, attL, or attR, such as attR or attP. More preferably the att site is selected from attL, att2, or att3, as described herein.

[0244] In a preferred embodiment, the core region comprises a DNA sequence selected from the group consisting of:

(a) RYXCGCTTTTTATCACAAATSSGB; (SEQ ID NO:1) (n-att)
(b) AGCCGWCTTTTTATCACAAATSSGB; (SEQ ID NO:2) (m-attB)
(c) GTCGAGCTTCTCTTACAAAGTSSGB; (SEQ ID NO:3) (m-attR)
(d) AGCCGWCTTTTTATCACAAAGTSSGB; (SEQ ID NO:4) (m-attL1)
(e) GTCGAGCTTTCTTACAAAGTSSGB; (SEQ ID NO:5) (m-attP1)
(f) RYXCGCTTTTTATTACAAATSSGD; (SEQ ID NO:39) (n-att)
(g) AGCCGWCTTTTTATTACAAATSSGD; (SEQ ID NO:40) (n-attB)
(h) AGCCGWCTTTTTATTACAAATSSGD; (SEQ ID NO:41) (n-attL2)
(i) GTCGAGCTTTCTTACAAAGTSSGD; (SEQ ID NO:42) (n-attL3)
(j) GTCGAGCTTTCTTACAAAGTSSGD; (SEQ ID NO:43) (n-attP)

[0245] or a corresponding or complementary DNA or RNA sequence, wherein R=A or G; K=G or T/U; Y=C or T/U; W=A or T/U; N=A or C or G or T/U; S=C or G; and B=C or G or T/U, as presented in 37 C.F.R. §1.822, which is entirely incorporated herein by reference, wherein the core region does not contain a stop codon in one or more reading frames.

[0246] The core region also preferably comprises a DNA sequence selected from the group consisting of:

(a) ACCTCGCTTTTTATCACAAACTTTG; (SEQ ID NO:6) (attB1)
(b) ACCTCGCTTTTTATCACAAACTTTG; (SEQ ID NO:7) (attB2)
(c) ACCTCGCTTTTTATCACAAACTTTG; (SEQ ID NO:8) (attB3)
(d) GTCGAGCTTTCTTACAAACTTTG; (SEQ ID NO:9) (attB1)

[0247] NO:16); or a corresponding or complementary DNA or RNA sequence.

[0248] The present invention thus also provides a method for making a nucleic acid molecule, comprising providing a nucleic acid molecule having at least one engineered recombination site comprising at least one DNA sequence having at least 80-99% homology (or any range or value therein) to at least one of the above sequences, or any suitable recombination site, or which hybridizes under stringent conditions thereto, as known in the art.

[0249] Clearly, there are various types and permutations of such well-known in vitro and in vivo selection methods, each of which are not described herein for the sake of brevity. However, such variations and permutations are contemplated and considered to be the different embodiments of the present invention.

[0250] It is important to note that as a result of the preferred embodiment being in vitro recombination reactions, non-biological molecules such as PCR products can be manipulated via the present recombinational recombination method. In one example, it is possible to clone linear molecules into circular vectors.

[0251] There are a number of applications for the present invention. These uses include, but are not limited to, changing vectors, appending promoters with genes, constructing genes for fusion proteins, changing copy number, changing replicons, cloning into phages, and cloning, c.g., PCR products (with an attB site at one end and a loXP site at the other end), genomic DNAs, and cDNAs.

[0252] Vector Donors

[0253] In accordance with the invention, any vector may be used to construct the Vector Donors of the invention. In particular, vectors known in the art and those commercially available (and variants or derivatives thereof) may in accordance with the invention be engineered to include one or more recombination sites for use in the methods of the invention. Such vectors may be obtained from, for example, Vector Laboratories Inc., Invitrogen, Promega, Novagen, NIB, Clontech, Boehringer Mannheim, Pharmacia, Epicenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc., and Research Genetics. Such vectors may then for example be used for cloning or subcloning nucleic acid molecules of interest. General
classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, expression vectors, fusion vectors, two-hybrid or reverse two-hybrid vectors, shuttle vectors for use in different hosts, mutagenesis vectors, transcription vectors, vectors for receiving large inserts and the like.

[0254] Other vectors of interest include viral origin vectors (M13 vectors, bacterial phase λ vectors, adenovirus vectors, and retrovirus vectors), high, low and adjustable copy number vectors, vectors which have compatible replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCMDS).


[0259] Polymerases

[0260] Polypeptides having reverse transcriptase activity (i.e., those polypeptides able to catalyze the synthesis of a DNA molecule from an RNA template) include, but are not limited to Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase, Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, Myeloblastosis Associated Virus (MAV) reverse transcriptase, Human Immunodeficiency Virus (HIV) reverse transcriptase, retroviral reverse transcriptase, retrotcoponse reverse transcriptase, hepatitis B reverse transcriptase, cauliﬂower mosaic virus reverse transcriptase and bacterial reverse transcriptase. Particularly preferred are those polypeptides having reverse transcriptase activity that are also substantially reduced in RNase H activity (i.e., “RNase H’” polypeptides). By a polypeptide that is “substantially reduced in RNase H activity” is meant that the polypeptide has less than about 20%, more preferably less than about 15%, 10% or 5%, and most preferably less than about 2%, of the RNase H II activity of a wildtype or RNase H’ enzyme such as wildtype M-MLV reverse transcriptase.
The RNase H activity may be determined by a variety of assays, such as those described, for example, in U.S. Pat. No. 5,244,797, in Kotewicz, M. L. et al., *Nucl. Acids Res.* 16:265 (1988) and in Gerard, G. F., et al., *FOCUS* 14(S):91 (1992), the disclosures of all of which are fully incorporated herein by reference. Suitable RNase H polypeptides for use in the present invention include, but are not limited to, M-MLV H' reverse transcriptase, RSV H' reverse transcriptase, AMV H' reverse transcriptase, RAV H' reverse transcriptase, MAV H' reverse transcriptase, HIV H' reverse transcriptase, and SUPERSCRIPT™ II reverse transcriptase and SUPERSCRIPT™ II reverse transcriptase which are available commercially, for example from Life Technologies, Inc. (Rockville, Md.).


[0262] It will be understood by one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are readily apparent and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

**EXAMPLES**

[0263] The present recombinational cloning method accomplishes the exchange of nucleic acid segments to render something useful to the user, such as a change of cloning vectors. These segments must be flanked on both sides by recombination signals that are in the proper orientation with respect to one another. In the examples below the two parental nucleic acid molecules (e.g., plasmids) are called the Insert Donor and the Vector Donor. The Insert Donor contains a segment that will become joined to a new vector contributed by the Vector Donor. The recombination intermediate(s) that contain(s) both starting molecules is called the Cointegrate(s). The second recombination event produces two daughter molecules, called the Product (the desired new clone) and the Byproduct.

**Buffers**

[0264] Various known buffers can be used in the reactions of the present invention. For restriction enzymes, it is advisable to use the buffers recommended by the manufacturer. Alternative buffers can be readily found in the literature or can be devised by those of ordinary skill in the art.

[0265] Examples 1-3. One exemplary buffer for lambda integrase is comprised of 50 mM Tris-HCl, at pH 7.5-7.8, 70 mM KCl, 5 mM spermidine, 0.5 mM EDTA, and 0.25 mg/ml bovine serum albumin, and optionally, 10% glycerol.

[0266] One preferred buffer for PI Cre recombinase is comprised of 50 mM Tris-HCl at pH 7.5, 33 mM NaCl, 5 mM spermidine, and 0.5 mg/ml bovine serum albumin.

[0267] The buffer for other site-specific recombinases which are similar to lambda Int and PI Cre are either known in the art or can be determined empirically by the skilled artisans, particularly in light of the above-described buffers.

**Example 1**

Recombinalional Cloning Using Cre and Cre & Int

[0268] Two pairs of plasmids were constructed to do the in vitro recombinational cloning method in two different ways. One pair, pE7C075 and pE7C726 (FIG. 2A), was constructed with loxP and att sites, to be used with Cre and λ integrase. The other pair, pE7C602 and pE7C629 (FIG. 3A), contained the loxP (wild type) site for Cre, and a second mutant lox site, loxP 511, which differs from loxP in one base (out of 34 total). The minimum requirement for recombinational cloning of the present invention is two recombinational sites in each plasmid, in general X and Y, and X' and Y'. Recombinational cloning takes place if either or both types of site can recombine to form a Cointegrate (e.g. X and X'), if both can recombine to excise the Product and Byproduct plasmids from the Cointegrate (e.g. Y and Y'). It is important that the recombinational sites on the same plasmid do not recombine. It was found that the present recombinational cloning could be done with Cre alone.

[0269] Cre-Only

[0270] Two plasmids were constructed to demonstrate this conception (see FIG. 3A). pE7C629 was the Vector Donor plasmid. It contained a constitutive drug marker (chloramphenicol resistance), an origin of replication, loxP and loxP 511 sites, a conditional drug marker (kanamycin resistance whose expression is controlled by the operator/promoter of the tetracycline resistance operon of transposon Tn10), and a constitutively expressed gene for the tet repressor protein, tetR. E. coli cells containing pE7C629 were resistant to chloramphenicol at 30 μg/ml, but sensitive to kanamycin at 100 μg/ml. pE7C602 was the Insert Donor plasmid which contained a different drug marker (ampicillin resistance), an origin, and loxP and loxP 511 sites flanking a multiple cloning site.

[0271] This experiment was comprised of two parts as follows:

[0272] Part I: About 75 ng each of pE7C602 and pE7C629 were mixed in a total volume of 30 μl of Cre buffer (50 mM Tris-HCl pH 7.5, 33 mM NaCl, 5 mM spermidine-HCl, 500 μg/ml bovine serum albumin). Two 10 μl aliquots were transferred to new tubes. One tube received 0.5 μl of Cre protein (approx. 4 units per μl; partially purified according to Albremski and Hoess, *J. Biol. Chem.* 259:1509 (1984)). Both tubes were incubated at 37°C for 30 minutes,
then 70°C for 10 minutes. Aliquots of each reaction were diluted and transformed into DH5α. Following expression, aliquots were plated on 30 μg/ml chloramphenicol; 100 μg/ml ampicillin plus 200 μg/ml methicillin; or 100 μg/ml kanamycin. Results: See Table 1. The reaction without Cre gave 1.1x10⁸ ampicillin resistant colonies (from the Insert Donor plasmid pEZC602); 7.8x10⁷ chloramphenicol resistant colonies (from the Vector Donor plasmid pEZC620); and 140 kanamycin resistant colonies (background). The reaction with added Cre gave 7.5x10⁸ ampicillin resistant colonies (from the Insert Donor plasmid pEZC602); 6.1x10⁷ chloramphenicol resistant colonies (from the Vector Donor plasmid pEZC620); and 760 kanamycin resistant colonies (mixture of background colonies and colonies from the recombinational cloning Plasmid plasmid). Analysis: Because the number of colonies on the kanamycin plates was much higher in the presence of Cre, many or most of them were predicted to contain the desired Product plasmid.

### TABLE 1

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Ampicillin</th>
<th>Chloramphenicol</th>
<th>Kanamycin</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1.1 x 10⁸</td>
<td>7.8 x 10⁷</td>
<td>140</td>
<td>140/7.8 x 10⁷ = 0.02%</td>
</tr>
<tr>
<td>Cre</td>
<td>7.5 x 10⁸</td>
<td>6.1 x 10⁷</td>
<td>760</td>
<td>760/6.1 x 10⁷ = 0.12%</td>
</tr>
</tbody>
</table>

[0273] Part II: Twenty four colonies from the “+Cre” kanamycin plates were picked and inoculated into medium containing 100 μg/ml kanamycin. Minipreps were done, and the miniprep DNAs, uncut or cut with Smal or HindIII, were electrophoresed. Results: 19 of the 24 minipreps showed supercoiled plasmid of the size predicted for the Product plasmid. All 19 showed the predicted Smal and HindIII restriction fragments. Analysis: The Cre only scheme was demonstrated. Specifically, it was determined that the Cre plus scheme has yielded about 70% (19 of 24) Product clones. The efficiency was about 0.1% (760 kanamycin resistant colonies resulted from 6.1x10⁷ chloramphenicol resistant colonies).

[0274] Cre Plus Integrase

[0275] The plasmids used to demonstrate this method are exactly analogous to those used above, except that pEZC726, the Vector Donor plasmid, contained an attP site in place of loxP 511, and pEZC705, the Insert Donor plasmid, contained an attB site in place of loxP 511 (FIG. 2A).

[0276] This experiment was comprised of three parts as follows:

[0277] Part I: About 500 ng of pEZC705 (the Insert Donor plasmid) was cut with ScaI, which linearized the plasmid within the ampicillin resistance gene. (This was done because the λ integrase reaction has been historically done with the attB plasmid in a linear state (H. Nash, personal communication). However, it was found later that the integrase reaction proceeds well with both plasmids supercoiled.) Then, the linear plasmid was ethanol precipitated and dissolved in 20 μl of λ integrase buffer (50 mM Tris-HCl, about pH 7.8, 70 mM KCl, 5 mM spermidine-HCl, 0.5 mM EDTA, 250 μg/ml bovine serum albumin). Also, about 500 ng of the Vector Donor plasmid pEZC726 was ethanol precipitated and dissolved in 20 μl λ integrase buffer. Just before use, λ integrase (2 μl, 393 μg/ml) was thawed and diluted by adding 18 μl cold λ integrase buffer. One μl λ integrase buffer (λ integrase, 2.4 mg/ml, an accessory protein) was diluted into 150 μl cold λ integrase buffer. Aliquots (2 μl) of each DNA were mixed with λ integrase buffer, with or without 1 μl each λ integrase and IHF, in a total of 10 μl. The mixture was incubated at 25°C for 45 minutes, then at 70°C for 10 minutes. Half of each reaction was applied to an agarose gel. Results: In the presence of integrase and IHF, about 5% of the total DNA was converted to a linear Complementation form. Analysis: Activity of integrase and IHF was confirmed.

[0278] Part II: Three microfilters of each reaction (i.e., with or without integrase and IHF) were diluted into 27 μl of Cre buffer (above), then each reaction was split into two 10 μl aliquots (four altogether). To two of these reactions, 0.5 μl of Cre protein (above) were added, and all reactions were incubated at 37°C for 30 minutes, then at 70°C for 10 minutes. TE buffer (90 μl; TE: 10 mM Tris-HCl, pH 7.5, 1 mM EDTA) was added to each reaction, and 1 μl each was transformed into E. coli DH5α. The transformation mixtures were plated on 100 μg/ml ampicillin plus 200 μg/ml methicillin; 30 μg/ml chloramphenicol; or 100 μg/ml kanamycin. Results: See Table 2.

### TABLE 2

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Ampicillin</th>
<th>Chloramphenicol</th>
<th>Kanamycin</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>990</td>
<td>20000</td>
<td>4</td>
<td>4/2 x 10⁷ = 0.02%</td>
</tr>
<tr>
<td>Cre</td>
<td>290</td>
<td>3640</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Integrase*</td>
<td>1040</td>
<td>27000</td>
<td>9</td>
<td>9/2.7 x 10⁷ = 0.03%</td>
</tr>
<tr>
<td>Integrase* + Cre</td>
<td>110</td>
<td>1110</td>
<td>76</td>
<td>76/1.1 x 10⁷ = 6.9%</td>
</tr>
</tbody>
</table>

*Integrase reactions also contained IHF.

[0279] Analysis: The Cre protein impaired transformation. When adjusted for this effect, the number of kanamycin resistant colonies, compared to the control reactions, increased more than 100 fold when both Cre and Integrase were used. This suggests a specificity of greater than 99%.

[0280] Part III: 38 colonies were picked from the Integrase plus Cre plates, miniprep DNAs were made and cut with HindIII to give diagnostic mapping information. Result: All 38 had precisely the expected fragment sizes. Analysis: The Cre plus λ integrase method was observed to have much higher specificity than Cre-alone. Conclusion: The Cre plus λ integrase method was demonstrated. Efficiency and specificity were much higher than for Cre only.

**Example 2**

Using in vitro Recombinational Cloning to Subclone the Chloramphenicol Acetyl Transferase Gene into a Vector for Expression in Eukaryotic Cells (FIG. 4A)

[0281] An Insert Donor plasmid, pEZC843, was constructed, comprising the chloramphenicol acetyl transferase gene of E. coli, cloned between loxP and attB sites such that the loxP site was positioned at the 5'-end of the gene (FIG. 4B). A Vector Donor plasmid, pEZC1003, was constructed, which contained the cytomegalovirus eukaryotic promoter apposed to a loxP site (FIG. 4C). One microfilter aliquots of
each supercoiled plasmid (about 50 ng crude miniprep DNA) were combined in a ten microliter reaction containing equal parts of lambda integrase buffer (50 mM Tris-HCl, pH 7.5, 70 mM KCl, 5 mM spermidine, 0.5 mM EDTA, 0.25 mg/ml bovine serum albumin) and Cre recombinase buffer (50 mM Tris-HCl, pH 7.5, 33 mM NaCl, 5 mM spermidine, 0.5 mg/ml bovine serum albumin), two units of Cre recombinase, 16 ng integration host factor, and 32 ng lambda integrase. After incubation at 30°C for 30 minutes and 75°C for 10 minutes, one microliter was transformed into competent E. coli strain DH5α (Life Technologies, Inc.). Aliquots of transformations were spread on agar plates containing 200 μg/ml kanamycin and incubated at 37°C overnight. An otherwise identical control reaction contained the Vector Donor plasmid only. The plate receiving 10% of the control reaction transformation gave one colony; the plate receiving 10% of the recombinational cloning reaction gave 144 colonies. These numbers suggested that greater than 99% of the recombinational cloning colonies contained the desired product plasmid. Miniprep DNA made from six recombinational cloning colonies gave the predicted size plasmid (5026 base pairs), CMVPro. Restriction digestion with NcoI gave the fragments predicted for the chloramphenicol acetyl transferase cloned downstream of the CMV promoter for all six plasmids.

**Example 3**

Subcloned DNA Segments Flanked by attB Sites Without Stop Codons

[0282] Part I: Background

[0283] The above examples are suitable for transcriptional fusions, in which transcription crosses recombination sites. However, both attR and loxP sites contain multiple stop codons on both strands, so translational fusions can be difficult, where the coding sequence must cross the recombination sites, (only one reading frame is available on each strand of loxP sites) or impossible (in attR or attL).

[0284] A principal reason for subcloning is to fuse protein domains. For example, fusion of the glutathione S-transferase (GST) domain to a protein of interest allows the fusion protein to be purified by affinity chromatography on glutathione agarose (Pharmacia, Inc., 1995 catalog). If the protein of interest is fused to runs of consecutive histidines (for example Hs6), the fusion protein can be purified by affinity chromatography on chelating resins containing metal ions (Qiagen, Inc.). It is often desirable to compare amino terminal and carboxy terminal fusions for activity, solubility, stability, and the like.

[0285] The attB sites of the bacteriophage λ integration system were examined as an alternative to loxP sites, because they are small (25 bp) and have some sequence flexibility (Nashi, H. A. et al., *Proc. Natl. Acad. Sci. USA* 84:4049-4053 (1987)). It was not previously suggested that multiple mutations to remove all stop codons would result in useful recombination sites for recombinational subcloning.

[0286] Using standard nomenclature for site specific recombination in lambda bacteriophage (Weisberg, in *Lambda III*, Hendrix, et al., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989)), the nucleotide regions that participate in the recombination reaction in an E. coli host cell are represented as follows:  

\[
\text{attB} \rightarrow B \rightarrow O \rightarrow B' \\
\text{attP} \rightarrow P_1 \rightarrow H_1 \rightarrow P_2 \rightarrow X \rightarrow H_2 \rightarrow C \rightarrow O \rightarrow B' \\
\text{attR} \rightarrow P_1 \rightarrow H_1 \rightarrow P_2 \rightarrow X \rightarrow H_2 \rightarrow C \rightarrow O \rightarrow B' \\
\text{attL} \rightarrow B \rightarrow O \rightarrow B' \rightarrow P_1 \rightarrow P_2 \rightarrow P_3 \\
\]

[0287] where: O represents the 15 bp core DNA sequence found in both the phage and E. coli genomes; B and B' represent approximately 5 bases adjacent to the core in the E. coli genome; and P1, H1, P2, X, H2, C, C', H, P2, and P3 represent known DNA sequences encoding protein binding domains in the bacteriophage λ genome.

[0288] The reaction is reversible in the presence of the protein Xis (excisionase); recombination between attL and attR precisely excise the λ genome from its integrated state, regenerating the circular λ genome containing attP and the linear E. coli genome containing attB.

[0289] Part II: Construction and Testing of Plasmids Containing Mutant att Sites

[0290] Mutant attL and attR sites were constructed. Importantly, Landy et al. (*Annu. Rev. Biochem.* 58:913 (1989)) observed that deletion of the P1 and H1 domains of attP facilitated the excision reaction and eliminated the integration reaction, thereby making the excision reaction irreversible. Therefore, as mutations were introduced in attR, the P1 and H1 domains were also deleted. attR' sites in the present example lack the P1 and H1 regions and have the NdeI site removed (base 27630 changed from C to G), and contain sequences corresponding to bacteriophage λ coordinates 27619-27738 (GenBank release 92.0, bg:LMCG, “Complete Sequence of Bacteriophage Lambda”).

[0291] The sequence of attB produced by recombination of wild type attL and attR sites is:

\[
\text{attB} \rightarrow B \rightarrow O \rightarrow B' \\
\text{attP} \rightarrow P_1 \rightarrow H_1 \rightarrow P_2 \rightarrow X \rightarrow H_2 \rightarrow C \rightarrow O \rightarrow B' \\
\text{attR} \rightarrow P_1 \rightarrow H_1 \rightarrow P_2 \rightarrow X \rightarrow H_2 \rightarrow C \rightarrow O \rightarrow B' \\
\text{attL} \rightarrow B \rightarrow O \rightarrow B' \rightarrow P_1 \rightarrow P_2 \rightarrow P_3 \\
\]

[0292] The stop codons are italicized and underlined. Note that sequences of attL, attR, and attP can be derived from the attB sequence and the boundaries of bacteriophage λ contained within attL and attR (coordinates 27619 to 27818).

[0293] When mutant attR1 (attR') and attL1 sites were recombined the sequence attB1 was produced (mutations in bold, large font):
B O B'  
attB1: 5' AGCCTGCTTTTTTGTACAAA CTTGT 3' (SEQ. ID NO: 6)  
3' TCGGA CGAAAAAACATGTTT GAACA 5' (SEQ. ID NO: 45)  

[0294] Note that the four stop codons are gone.

[0295] When an additional mutation was introduced in the attR1 (attR') and attL1 sequences (bold), attR2 (attR) and attL2 sites resulted. Recombination of attR2 and attL2 produced the attB2 site:

B O B'  
attB2: 5' AGCCTGCTTTTTTGTACAAA CTTGT 3' (SEQ. ID NO: 7)  
3' TCGGA CGAAAAAACATGTTT GAACA 5' (SEQ. ID NO: 46)  

[0296] The recombination activities of the above attL and attR' sites were assayed as follows. The attB site of plasmid pEZC705 (FIG. 2B) was replaced with attB1, attL1, or attL2. The attP site of plasmid pEZC726 (FIG. 2C) was replaced with attR1 (attR', lacking regions P1 and H1) or attR2 (attR', lacking regions P1 and H1). Thus, the resulting plasmids could recombine via their loxP sites, mediated by Cre, and via their attR' and attL sites, mediated by Int, Xis, and IHF. Pairs of plasmids were mixed and reacted with Cre, Int, Xis, and IHF, transformed into E. coli competent cells, and plated on agar containing kanamycin. The results are presented in Table 3:

<table>
<thead>
<tr>
<th>Vector donor site</th>
<th>Insert donor site</th>
<th># of kanamycin resistant colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>attR' wt (pEZC1301)</td>
<td>None</td>
<td>1 (background)</td>
</tr>
<tr>
<td>*</td>
<td>attLwt (pEZC1313)</td>
<td>147</td>
</tr>
<tr>
<td>*</td>
<td>attL1 (pEZC1317)</td>
<td>47</td>
</tr>
<tr>
<td>*</td>
<td>attL2 (pEZC1321)</td>
<td>0</td>
</tr>
<tr>
<td>attR' 1 (pEZC1305)</td>
<td>None</td>
<td>1 (background)</td>
</tr>
<tr>
<td>*</td>
<td>attLwt (pEZC1313)</td>
<td>4</td>
</tr>
<tr>
<td>*</td>
<td>attL1 (pEZC1317)</td>
<td>128</td>
</tr>
<tr>
<td>*</td>
<td>attL2 (pEZC1321)</td>
<td>0</td>
</tr>
<tr>
<td>attR' 2 (pEZC1309)</td>
<td>None</td>
<td>0 (background)</td>
</tr>
<tr>
<td>*</td>
<td>attLwt (pEZC1313)</td>
<td>0</td>
</tr>
<tr>
<td>*</td>
<td>attL1 (pEZC1317)</td>
<td>0</td>
</tr>
<tr>
<td>*</td>
<td>attL2 (pEZC1321)</td>
<td>209</td>
</tr>
</tbody>
</table>

*1% of each transformation was spread on a kanamycin plate.

[0297] The above data show that whereas the wild type att and attL sites recombine to a small extent, the attL and attR sites do not recombine detectably with each other.

[0298] Part III. Recombination was demonstrated when the core region of both attB sites flanking the DNA segment of interest did not contain stop codons. The physical state of the participating plasmids was discovered to influence recombination efficiency.

[0299] The appropriate att sites were moved into pEZC705 and pEZC726 to make the plasmids pEZC1405 (FIG. 5G) (attR'1 and attR'2) and pEZC1502 (FIG. 5H) (attL1 and attL2). The desired DNA segment in this experiment was a copy of the chloramphenicol resistance gene cloned between the two attL sites of pEZC1502. Pairs of plasmids were recombined in vitro using Int, Xis, and IHF (no Cre because no loxP sites were present). 100 ng of each plasmid was incubated in 10 µl reactions of 50 mM Tris HCl pH about 7.8, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 0.25 mM EDTA, 0.375 mg/ml BSA, 3% glycerol that contained 8.1 ng HGF, 43 ng Lint, 4.3 ng Xis, and 2 units Cre. Reactions were incubated at 25°C for 45 min., 65°C for 10 min, and 1 µl aliquots were transformed into DH5α cells, and spread on kanamycin plates. The yield of desired kanamycin resistant colonies determined whether both parental plasmids were circular, or when one plasmid was circular and the other linear as presented in Table 4:

### TABLE 4

<table>
<thead>
<tr>
<th>Vector donor site</th>
<th>Insert donor site</th>
<th>Kanamycin resistant colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circular pEZC1405</td>
<td>None</td>
<td>30</td>
</tr>
<tr>
<td>Circular pEZC1405</td>
<td>Circular pEZC1502</td>
<td>2680</td>
</tr>
<tr>
<td>Linear pEZC1405</td>
<td>None</td>
<td>90</td>
</tr>
<tr>
<td>Linear pEZC1405</td>
<td>Circular pEZC1502</td>
<td>17200</td>
</tr>
<tr>
<td>Circular pEZC1505 Linear pEZC1502</td>
<td>73000</td>
<td></td>
</tr>
</tbody>
</table>

*DNA was purified with QiaGen columns, concentrations determined by A260, and linearized with XbaI (pEZC1405) or AlwNI (pEZC1502). Each reaction contained 100 ng of the indicated DNA. All reactions (10 µl total) contained 3 µl of enzyme mix (Xis, Int, and IHF). After incubation (45 minutes at 25°C, 10 minutes at 65°C), one µl was used to transform E. coli DH5α cells.

*1Number of colonies expected if the entire transformation reaction (1 ml) had been plated. Either 100 µl or 1 µl of the transformations were actually plated.

[0300] Analysis: Recombinational cloning using mutant attR and attL sites was confirmed. The desired DNA segment was subcloned between attB sites that do not contain any stop codons in either strand. The enhanced yield of Product DNA (when one parent was linear) was unexpected because of earlier observations that the excision reaction was more efficient when both participating molecules were supercoiled and proteins were limiting (Nunes-Duby et al., Cell 50:779-788 (1987)).

### Example 4

Demonstration of Recombinational Cloning Without Inverted Repeats

[0301] Part I: Rationale

[0302] The above Example 3 showed that plasmids containing inverted repeats of the appropriate recombination sites (for example, attL1 and attL2 in plasmid pEZC1502) (FIG. 5H) could recombine to give the desired DNA segment flanked by attB sites without stop codons, also in inverted orientation. A concern was the in vivo and in vitro influence of the inverted repeats. For example, transcription of a desired DNA segment flanked by attB sites in inverted orientation could yield a single stranded RNA molecule that might form a hairpin structure, thereby inhibiting translation.

[0303] Inverted orientation of similar recombination sites can be avoided by placing the sites in direct repeat arrangement at sites. If parental plasmids each have a wild type attL and wild type attR site, in direct repeat the Int, Xis, and IHF proteins will simply remove the DNA segment flanked by those sites in an intramolecular reaction. However, the
mutant sites described in the above Example 3 suggested that it might be possible to inhibit the intramolecular reaction while allowing the intermolecular recombination to proceed as desired.

[0304] Part II: Structure of Plasmids Without Inverted Repeats for Recombinational Cloning

[0305] The attR2 sequence in plasmid pEZC1405 (FIG. 5G) was replaced with attL2, in the opposite orientation, to make pEZC1603 (FIG. 6A). The attL2 sequence of pEZC1502 (FIG. 5H) was replaced with attR2, in the opposite orientation, to make pEZC1706 (FIG. 6B). Each of these plasmids contained mutations in the core region that make intramolecular reactions between att1 and att2 cores very inefficient (see Example 3, above).

[0306] Plasmids pEZC1405, pEZC1502, pEZC1603 and pEZC1706 were purified on Qiagen columns (Qiagen, Inc.). Aliquots of plasmids pEZC1405 and pEZC1603 were linearized with XbaI. Aliquots of plasmids pEZC1502 and pEZC1706 were linearized with AlwNI. One hundred ng of plasmids were mixed in buffer (50 mM Tris HCl pH about 7.5, 16.5 mM NaCl, 55 mM KCl, 5 mM spermidine, 0.25 mM EDTA, 0.375 mg/ml BSA, 3% glycerol) containing Int (43.5 ng), Xis (4.3 ng) and IFF (6.1 ng) in a final volume of 10 µl. Reactions were incubated for 45 minutes at 25°C, 10 minutes at 65°C, and 1 µl was transformed into E. coli DH5α. After expression, aliquots were spread on agar plates containing 200 µg/ml kanamycin and incubated at 37°C.

[0307] Results, expressed as the number of colonies per 1 µl of recombination reaction are presented in Table 5:

<table>
<thead>
<tr>
<th>Vector Donor</th>
<th>Gene Donor</th>
<th>Colonies</th>
<th>Predicted % product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circular 1405</td>
<td>—</td>
<td>100</td>
<td>—</td>
</tr>
<tr>
<td>Circular 1405</td>
<td>Circular 1502</td>
<td>3740</td>
<td>3740/3740 = 100%</td>
</tr>
<tr>
<td>Linear 1405</td>
<td>—</td>
<td>90</td>
<td>—</td>
</tr>
<tr>
<td>Linear 1405</td>
<td>Circular 1502</td>
<td>172,000</td>
<td>171,910/172,000 = 99.9%</td>
</tr>
<tr>
<td>Linear 1502</td>
<td>—</td>
<td>73,000</td>
<td>73,000/73,000 = 100%</td>
</tr>
<tr>
<td>Circular 1603</td>
<td>—</td>
<td>80</td>
<td>—</td>
</tr>
<tr>
<td>Linear 1603</td>
<td>Circular 1706</td>
<td>410</td>
<td>330/410 = 80%</td>
</tr>
<tr>
<td>Circular 1603</td>
<td>Circular 1706</td>
<td>7000</td>
<td>6730/7000 = 96%</td>
</tr>
<tr>
<td>Linear 1603</td>
<td>Linear 1706</td>
<td>10,800</td>
<td>10,530/10,800 = 97%</td>
</tr>
</tbody>
</table>

[0308] Analysis. In all configurations, i.e., circular or linear, the pEZC1405×pEZC1502 pair (with att sites in inverted repeat configuration) was more efficient than pEZC1603×pEZC1706 pair (with att sites mutated to avoid hairpin formation). The pEZC1603×pEZC1706 pair gave higher backgrounds and lower efficiencies than the pEZC1405×pEZC1502 pair. While less efficient, 80% or more of the colonies from the pEZC1603×pEZC1706 reactions were expected to contain the desired plasmid product. Making one partner linear stimulated the reactions in all cases.

[0309] Part II: Confirmation of Product Plasmids' Structure

[0310] Six colonies each from the linear pEZC1405 (FIG. 5G)×circular pEZC1502 (FIG. 5H), circular pEZC1405× linear pEZC1502, linear pEZC1603 (FIG. 6A)×circular pEZC1706 (FIG. 6B), and circular pEZC1603×circular pEZC1706 reactions were picked into rich medium and miniprep DNAs were prepared. Diagnostic cuts with SspI gave the predicted restriction fragments for all 24 colonies.

[0311] Analysis. Recombination reactions between plasmids with mutant attL and attR sites on the same molecules gave the desired plasmid products with a high degree of specificity.

Example 5

Recombinational Cloning with a Toxic Gene

[0312] Part I: Background

[0313] Restriction enzyme DpnI recognizes the sequence GATC and cuts that sequence only if the A is methylated by the dam methylase. Most commonly used E. coli strains are dam+. Expression of DpnI in dam+ strains of E. coli is lethal because the chromosome of the cell is chopped into many pieces. However, in dam- cells expression of DpnI is innocuous because the chromosome is immune to DpnI cutting.

[0314] In the general recombinational cloning scheme, in which the vector donor contains two segments C and D separated by recombination sites, selection for the desired product depends upon selection for the presence of segment D, and the absence of segment C. In the original Example segment D contained a drug resistance gene (Km) that was negatively controlled by a repressor gene found on segment C. When C was present, cells containing D were not resistant to kanamycin because the resistance gene was turned off.

[0315] The DpnI gene is an example of a toxic gene that can replace the repressor gene of the above embodiment. If segment C expresses the DpnI gene product, transforming plasmid CD into a dam+ host kills the cell. If segment D is transferred to a new plasmid, for example by recombinational cloning, then selecting for the drug marker will be successful because the toxic gene is no longer present.

[0316] Part II: Construction of a Vector Donor Using DpnI as a Toxic Gene

[0317] The gene encoding DpnI endonuclease was amplified by PCR using primers 5'CCA CCA CCA ACG CTT CCA TGG AAT TAC ACT TTA ATT TAG3' (SEQ ID NO: 17) and 5'CCA CCA CAA GTC GAC GCA TGG CTC AAA TG T3' (SEQ ID NO: 18) and a plasmid containing the DpnI gene (derived from plasmids obtained from Sanford A. Lacks, Brookhaven National Laboratory, Upton, N.Y.; also available from American Type Culture Collection as ATCC 67494) as the template.

[0318] Additional mutations were introduced into the B and B' regions of attL and attR, respectively, by amplifying existing attL and attR domains with primers containing the desired base changes. Recombination of the mutant attL3 (made with oligo Xis15) and attR3 (attR, made with oligo Xis112) yielded attB3 with the following sequence (differences from attB1 in bold):

```
B'          B
ACCCA GCTTCTGCTCAAAC GTGGT (SEQ ID NO:18)
TGCGT CCAAGAACGAGTTT GGCAA (SEQ ID NO:47)
```

[0319] The attL3 sequence was cloned in place of attL2 of an existing Gene Donor plasmid to give the plasmid pEZC2901 (FIG. 7A). The attR3 sequence was cloned in place of attR2 in an existing Vector Donor plasmid to give plasmid pEZC2913 (FIG. 7B). The DpnI gene was cloned
into plasmid pEZC2913 to replace the tet repressor gene. The resulting Vector Donor plasmid was named pEZC3101 (FIG. 7C). When pEZC3101 was transformed into the dam- strain SCS110 (Stratagene), hundreds of colonies resulted. When the same plasmid was transformed into the dam+ strain DH5α, only one colony was produced, even though the DH5α cells were about 20 fold more competent than the SCS110 cells. When a related plasmid that did not contain the DpnI gene was transformed into the same two cell lines, 28 colonies were produced from the SCS110 cells, while 448 colonies resulted from the DH5α cells. This is evidence that the DpnI gene is being expressed on plasmid pEZC3101 (FIG. 7C), and that it is killing the dam- DH5α cells but not the dam+ SCS110 cells.

[0320] Part III: Demonstration of Recombinational Cloning Using DpnI Selection

[0321] A pair of plasmids was used to demonstrate recombinational cloning with selection for Product dependent upon the toxic gene DpnI. Plasmid pEZC3101 (FIG. 7C) was linearized with Msel and reacted with circular plasmid pEZC2901 (FIG. 7A). A second pair of plasmids using selection based on control of drug resistance by a repressor gene was used as a control: plasmid pEZC1802 (FIG. 7D) was linearized with XbaI and reacted with circular plasmid pEZC1502 (FIG. 5H). Eight microliter reactions containing buffer (50 mM Tris HCl pH about 7.8, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 0.375 mg/ml BSA, 0.25 mM EDTA, 2.5% glycerol) and proteins Xis (2.9 ng), Int (29 ng), and IHF (5.4 ng) were incubated for 45 minutes at 25° C, then 10 minutes at 75° C, and 1 μl aliquots were transformed into DH5α (i.e., dam+) competent cells, as presented in Table 6.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Vector donor</th>
<th>Basis of selection</th>
<th>Insert donor</th>
<th>Colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pEZC3101/Mse</td>
<td>DpnI I toxicity</td>
<td>Circular</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>pEZC3101/Mse</td>
<td>DpnI I toxicity</td>
<td>pEZC2901</td>
<td>4000</td>
</tr>
<tr>
<td>3</td>
<td>pEZC1802/Xba</td>
<td>Tet repressor</td>
<td>Circular</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>pEZC1802/Xba</td>
<td>Tet repressor</td>
<td>pEZC1502</td>
<td>12100</td>
</tr>
</tbody>
</table>

[0322] Miniprep DNAs were prepared from four colonies from reaction #2, and cut with restriction enzyme SspI. All gave the predicted fragments.

[0323] Analysis: Subcloning using selection with a toxic gene was demonstrated. Plasmids of the predicted structure were produced.

Example 6

Cloning of Genes with Uracil DNA Glycosylase and Subcloning of the Genes with Recombinational Cloning to Make Fusion Proteins

[0324] Part I: Converting an Existing Expression Vector to a Vector Donor for Recombinational Cloning

[0325] A cassette useful for converting existing vectors into functional Vector Donors was made as follows. Plasmid pEZC3101 (FIG. 7C) was digested with Apal and Kpnl, treated with T4 DNA polymerase and dNTPs to render the ends blunt, further digested with SmaI, Hpal, and AlwNI to render the undesirable DNA fragments small, and the 2.6 kb cassette containing the attR1-CmR-DpnI-attR2-3 domains was gel purified. The concentration of the purified cassette was estimated to be about 75 μg DNA/μl.

[0326] Plasmid pGEX-2TK (FIG. 8A) (Pharmacia) allows fusions between the protein glutathione S transferase and any second coding sequence that can be inserted in frame in its multiple cloning site. pGEX-2TK DNA was digested with SmaI and treated with alkaline phosphatase. About 75 μg of the above purified DNA cassette was ligated with about 100 ng of the pGEX-2TK vector for 2.5 hours in a 5 μl ligation, then 1 μl was transformed into competent E. coli BRL 3056 (a dam- derivative of DH10B; dam- strains commercially available include DMI from Life Technologies, Inc., and SCS 110 from Stratagene). Aliquots of the transformation mixture were plated on LB agar containing 100 μg/ml ampicillin (resistance gene present on pGEX-2TK) and 30 μg/ml chloramphenicol (resistance gene present on the DNA cassette). Colonies were picked and miniprep DNAs were made. The orientation of the cassette in pGEX-2TK was determined by diagnostic cuts with EcoRI. A plasmid with the desired orientation was named pEZC3501 (FIG. 8B).

[0327] Part II: Cloning Reporter Genes Into a Recombinational Cloning Gene Donor Plasmid in Three Reading Frames

[0328] Uracil DNA glycosylase (UDG) cloning is a method for cloning PCR amplification products into cloning vectors (U.S. Pat. No. 5,334,515, entirely incorporated herein by reference). Briefly, PCR amplification of the desired DNA segment is performed with primers that contain uracil bases in place of thymidine bases in their 5' ends. When such PCR products are incubated with the enzyme UDG, the uracil bases are specifically removed. The loss of these bases weakens base pairing in the ends of the PCR product DNA, and when incubated at a suitable temperature (e.g., 37° C), the ends of such products are largely single stranded. If such incubations are done in the presence of linear cloning vectors containing protruding 3' tails that are complementary to the 3' ends of the PCR products, base pairing efficiently anneals the PCR products to the cloning vector. When the annealed product is introduced into E. coli cells by transformation, in vivo processes efficiently convert it into a recombinant plasmid.

[0329] UDG cloning vectors that enable cloning of any PCR product in all three reading frames were prepared from pEZC3201 (FIG. 8K) as follows. Eight oligonucleotides were obtained from Life Technologies, Inc. (all written 5'→3': r1 forward (GCCG ATG TAC GAT ATC CCA ACG ACC GAA AAC CTC TAT TTT CAG GGT) (SEQ. ID NO:19), r1 forward (CAG GGT TTC TGG GTG AGT ATC GTA ATC) (SEQ. ID NO:20), r2 forward (GGCGA GAT TAC GAT ATC CCA ACC ACC GAA AAC CTC TAT TTT CAG GGT) (SEQ. ID NO:21), r2 forward (CAG GGT TTC TGG GTG AGT ATC GTA ATC) (SEQ. ID NO:20), r3 forward (GGCGA GAT TAC GAT ATC CCA ACC ACC GAA AAC CTC TAT TTT CAG GGT) (SEQ. ID NO:23), r3 forward (CAG GGT TTC TGG GTG AGT ATC GTA ATC) (SEQ. ID NO:20), carboxy top (ACC GGT TAC GTG GAC) (SEQ. ID NO:25) and carboxy bottom (TGG GTA CAC GAA AAC GGT TCC CAC TTA TTA) (SEQ. ID NO:27).
NO:26). The r1,2, and 3 top strands and the carboxy bottom strand were phosphorylated on their 5' ends with T4 poly
ucleotid kinase, and then the complementary strands of each pair were hybridized. Plasmid pEZC3201 (FIG. 8K)
was cut with Ncol and SalI, and aliquots of cut plasmid were mixed with the carboxy-oligo duplex (SalI end) and either
the r1, r2, or r3 duplexes (Ncol ends) (10 μg cut plasmid (about 5 pmol) mixed with 250 pmol carboxy oligo duplex,
split into three 20 μl volumes, added 5 μl (250 pmol) of r1, r2, or r3 duplex and 2 μl (2 units) T4 DNA ligase to each
reaction). After 90 minutes of ligation at room temperature, each reaction was applied to a preparative agarose gel
and the 2.1 kb vector bands were eluted and dissolved in 50 μl of TE.

[0330] Part III: PCR of CAT and phoA Genes

[0331] Primers were obtained from Life Technologies, Inc., to amplify the chloramphenicol acetyl transferase
(CAT) gene from plasmid pACYC184, and phoA, the alkaline phosphatase gene from E. coli. The primers had 12-base
5' extensions containing uracil bases, so that treatment of PCR products with uracil DNA glycosylase (UDG) would
weaken base pairing at each end of the DNAs and allow the 3' strands to anneal with the protruding 3' ends of the r1, 2,
and 3 vectors described above. The sequences of the primers (all written 5'→3') were: CAT left, UAU UU GAG GUG
ATG GAG AAA AAA AATC ACT GGA TAT ACC (SEQ. ID NO:27); CAT right, UCC CAC UUA UUA CGC CCC GCC
CTG CCA CTC ATC (SEQ. ID NO:28); phoA left, UAU UAU CAG GGU AGT CCT GTG CGG AAA AAC CGG
(SEQ. ID NO:29); and phoA right, UCC CAC UUA WUA TTT CAC CCC GAG GCC GGC TTC TCC (SEQ. ID NO:30).
The primers were then used for PCR reactions using known method steps (see, e.g., U.S. Pat. No. 5,334,515, entirely
incorporated herein by reference), and the polymerase chain reaction amplification products obtained with these primers
comprised the CAT or phoA genes with the initiating ATGs but without any transcriptional signals. In addition, the
uracil-containing sequences on the amino termini encoded the cleavage site for TEV protease (Life Technologies, Inc.),
and those on the carboxy terminal encoded consecutive TAA nonsense codons.

[0332] Unpurified PCR products (about 50 ng) were mixed with the gel purified, linear r1, r2, or r3 cloning vectors
(abount 50 ng) in a 10 μl reaction containing 1xREact 4 buffer (LTI) and 1 unit UDG (LTI). After 30 minutes at 37°C,
1 μl aliquots of each reaction were transformed into competent E. coli DH5α cells (LTI) and plated on agar containing
50 μg/ml kanamycin. Colonies were picked and analysis of miniprep DNA showed that the CAT gene had been
cloned in reading frame 1 (pEZC3601) (FIG. 8C), reading frame 2 (pEZC3609) (FIG. 8D) and reading frame
3 (pEZC3617) (FIG. 8E), and that the phoA gene had been cloned in reading frame 1 (pEZC606) (FIG. 8F), reading
frame 2 (pEZC3613) (FIG. 8G) and reading frame 3 (pEZC3621) (FIG. 8H).

[0333] Part IV: Subcloning of CAT or phoA from UDG Cloning Vectors into a GST Fusion Vector

[0334] Plasmids encoding fusions between GST and either CAT or phoA in all three reading frames were constructed
by recombinational cloning as follows. Miniprep DNA of GST vector donor pEZC3501 (FIG. 8B) (derived from Pharmacia
plasmid pGEX-2TK as described above) was linearized with Clal. About 5 μg of vector donor were mixed with about 10
ng each of the appropriate circular gene donor vectors containing CAT or phoA in 8 μl reactions containing buffer
and recombination proteins Int, Xis, and IHF (Example 5). After incubation, 1 μl of each reaction was transformed into
E. coli strain DH5α and plated on ampicillin, as presented in Table 7.

<table>
<thead>
<tr>
<th>DNA</th>
<th>Colonies (10% of each transformation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear vector donor (pEZC3501/Cla)</td>
<td>0</td>
</tr>
<tr>
<td>Vector donor + CAT r1</td>
<td>110</td>
</tr>
<tr>
<td>Vector donor + CAT r2</td>
<td>71</td>
</tr>
<tr>
<td>Vector donor + CAT r3</td>
<td>148</td>
</tr>
<tr>
<td>Vector donor + phoA r1</td>
<td>121</td>
</tr>
<tr>
<td>Vector donor + phoA r2</td>
<td>128</td>
</tr>
<tr>
<td>Vector donor + phoA r3</td>
<td>31</td>
</tr>
</tbody>
</table>

[0335] Part V: Expression of Fusion Proteins

[0336] Two colonies from each transformation were picked into 2 ml of rich medium (CircleGrow, Bio101 Inc.)
in 17×100 mm plastic tubes (Falcon 2059, Becton Dickinson)
containing 100 μg/ml ampicillin and shaken vigorously
for about 4 hours at 37°C, at which time the cultures were
visibly turbid. One ml of each culture was transferred to
a new tube containing 10 μl of 10% (w/v) IPTG to induce
expression of GST. After 2 hours additional incubation, all
cultures had about the same turbidity; the A600 of one
culture was 1.5. Cells from 0.35 ml each culture were
harvested and treated with sample buffer (containing SDS
and β-mercaptoethanol) and aliquots equivalent to about
0.15 A600 units of cells were applied to a Novex 4-20% gradient polyacrylamide gel. Following electrophoresis
the gel was stained with Coomassie blue.

[0337] Results: Enhanced expression of single protein bands was seen for all 12 cultures. The observed sizes of
these proteins correlated well with the sizes predicted for
GST being fused (through attB recombination sites without stop codons) to CAT (FIG. 8I) or phoA (FIG. 8J) in three
reading frames: CAT r1=269 amino acids; CAT r2=303
amino acids; CAT r3=478 amino acids; phoA r1=282
amino acids; phoA r2=280 amino acids; and phoA r3=705
amino acids.

[0338] Analysis: Both CAT and phoA genes were subcloned into a GST fusion vector in all three reading frames,
and expression of the six fusion proteins was demonstrated.

Example 7

Reverse Recombination and Subcloning by
Recombination

[0339] Two plasmids were constructed to demonstrate
reverse recombination according to the present invention.
The vector pEZC601 (FIG. 10A), containing attB
recombination sites and termed the attB parent plasmid
(this vector may correspond to the Product DNA), further
contained an ampicillin resistance gene, an origin of replication, an attB
site, a tetracycline resistance gene, and an attB site, as
described above. Plasmid pEZC6701 (FIG. 10B),
containing attP recombination sites and termed the attP
parent plasmid (this vector may correspond to the Byproduct DNA
or may correspond to a different Vector Donor DNA), also contained a kanamycin resistance gene, an origin of replication, an attP site, a gene encoding the toxic protein ccdB, and an attP0 site. Integrate buffer at 10x concentration comprised 0.25 M Tris HCl pH 7.5, 0.25 M Tris HCl pH 8.0, 0.7 M potassium chloride, 50 mM spermidine HCl, 5 mM EDTA, and 2.5 mg/ml BSA. Note that attP0 and attP2 contained the P1 and H1 domains. Integrate (1.5 µl of 435 ng/µl) and IIF (1.5 µl of 16 ng/µl in 1x Integrate buffer) were mixed with 13.5 µl of 1x Integrate buffer to make the recombinease mixture.

[0340] Two 8 µl reactions were assembled. Reaction A contained 300 ng pEZC6701 plasmid and 2 µl of recombinease mixture in 1x Integrate buffer. Reaction B contained 300 ng pEZC6001, 300 ng pEZC6701, and 2 µl of recombinease mixture in 1x Integrate buffer. Both reactions were incubated at 25° C. for 45 minutes, then at 70° C. for 5 minutes, and then cooled. TE buffer (792 µl of 10 mM Tris HCl pH 7.5, 1 mM EDTA) was added to each reaction, and 1 µl of this diluted reaction was transformed into DH5α UltraMax competent E. coli cells (Life Technologies, Inc., Rockville, Md.). After 1 hour of expression in non-selective medium, one tenth (100 µl) of each transformation was spread onto agar plates containing 100 µg/ml kanamycin.

[0341] After overnight incubation at 37° C., the plate from reaction A contained 1 colony, while the plate from reaction B contained 392 colonies. Twelve colonies were picked from the reaction B plate into rich liquid medium and grown overnight. Miniprep DNAs prepared from these cultures were run uncut on an agarose gel and all 12 contained a plasmid of about 3.8 kb. Six of the miniprep DNAs were cut with restriction enzyme ClaI and run along with pEZC6701 (the kanamycin resistant parental plasmid) also cut with ClaI. Plasmid pEZC6701 was cut once with ClaI to give a fragment of about 3.8 kb. The six miniprep DNAs cut twice with ClaI to give fragments of about 900 base pairs and about 2900 base pairs.

[0342] Analysis: Recombination between the attP sites on pEZC6701 and the attB sites on pEZC6001 resulted in the production of two daughter plasmids, the attL product plasmid (FIG. 10C) (which may correspond to the Vector Donor DNA or a new Byproduct DNA) that contained the ampicillin resistance and ccdB genes, and the attR product plasmid (FIG. 10D) (which may also correspond to the Insert Donor DNA or a new Product DNA) that contained the kanamycin and tetracycline resistance genes. Competent E. coli cells that received the attL product plasmid, the attP parent plasmid pEZC6701, or recombination intermediates, were killed by the toxic ccdB gene product. Competent E. coli cells that received the attB parent plasmid pEZC6001 were killed by the kanamycin selection. Only competent E. coli cells that received the desired attR product plasmid, comprising the kanamycin and tetracycline resistance genes, survived to form colonies. The success of the selection strategy was indicated by the large number of colonies from the reaction that contained both parental plasmids, compared to the reaction that contained only one parental plasmid. The reaction mechanism predicted that the desired product plasmid would contain two ClaI restriction sites, one in the kanamycin resistance gene from the pEZC6701 attP parent plasmid and one in the tetracycline resistance gene from the pEZC6001 attB parent plasmid. The presence of the two sites and the sizes of the fragments resulting from the ClaI digestion confirmed the reaction mechanism.

[0343] Thus, the present invention relates to reversal of the recombination reaction shown in FIG. 1, in which the Product DNA and Byproduct DNA may be combined to produce the Insert Donor DNA and the Vector Donor DNA. Additionally, the invention provides for subcloning recombinations, in which a Product DNA (produced according to FIG. 1) may be combined with a new Vector Donor DNA to produce a new Product DNA (in a different Vector background) and a new Byproduct.

Example 8

Subcloning of Linearized Fragments

[0344] Plasmid pEZC7102 (FIG. 11A), the attP parent plasmid (which may correspond to the Vector Donor DNA), contained segments attPL, origin of replication, kanamycin resistance, attP3, chloramphenicol resistance, and the toxic gene ccdB, and in the experiment described here was supercoiled.

[0345] Plasmid pEZC7501 (FIG. 11B), the attB parent plasmid (which may correspond to the Insert Donor DNA or the Product DNA), contained the GFP gene cloned between attB1 and attB3 sites in a vector that comprised the functional domains of pCMVSPORT2.0 (Life Technologies, Inc.). The attP sites contained the P1 and H1 domains. Plasmid pEZC7501 was used uncut, or was linearized within the ampicillin resistance gene with Scal, or was cut with XbaI and SalI, to yield a fragment comprising the SalI end, 22 bp, the attB1 site, the GFP gene, the attB3 site, and 14 bp to the XbaI end:

SalI end -- 22 bp -- attB1 -- GFP -- attB3 -- 14 bp -- XbaI end

[0346] Reactions (8 µl final volume) contained about 40 ng of each DNA, 1x Integrate buffer (25 mM Tris HCl pH 7.5, 25 mM Tris HCl pH 8.0, 70 mM KCl, 5 mM spermidine HCl, 0.5 mM EDTA, and 0.25 mg/ml BSA), 12.5% glycerol, 8 ng IIF, and 43 ng lambda integrase. Reactions were incubated at 25° C. for 45 minutes, then at 70° C. for 5 minutes, and then cooled. Duplicate 1 µl aliquots of each reaction were transformed into DH5α UltraMax cells and plated in duplicate on kanamycin agar plates.

[0347] The reaction that contained only (supercoiled) pEZC7102 gave an average of 2 colonies (range 1 to 4). The reaction that contained both pEZC7102 and supercoiled pEZC7501 gave an average of 612 colonies (range 482-762). The reaction that contained pEZC7102 and linear (Scal-cut) pEZC7501 gave an average of 360 colonies (range 127-605). The reaction that contained pEZC7102 and the GFP gene on a fragment with attB sites and 22 bp and 14 bp beyond the attB sites (pEZC7501 cut with SalI and XbaI) gave an average of 274 colonies (range 243-308).
[0348] Miniprep DNAs were prepared from 4 colonies from the pEZC7102 supercoiled pEZC7510 reaction, and from 10 colonies from the pEZC7102+pEZC7510/Sall+XbaI reaction. All 14 DNAs were run uncut on a agarose gel, and the 10 DNAs from the pEZC7102+pEZC7510/Sall+XbaI reaction were cut with a mixture of NcoI and PstI and run on an agarose gel. All the uncut plasmids were about 2.8 kb in size. All ten plasmids cut with the mixture of NcoI and PstI gave fragments of about 700 and 2100 bp.

[0349] The results are presented in Table 8:

<table>
<thead>
<tr>
<th>attP Parent</th>
<th>attB Parent</th>
<th>Colonies of 4 plates</th>
<th>Miniprep done</th>
<th>Uncut product plasmid size</th>
<th>Fragment sizes, NcoI + PstI digest</th>
</tr>
</thead>
<tbody>
<tr>
<td>se 7102</td>
<td>se 7102</td>
<td>2</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>se 7102</td>
<td>7501</td>
<td>612</td>
<td>4</td>
<td>2.8 kb</td>
<td>—</td>
</tr>
<tr>
<td>se 7102</td>
<td>7501/Secl</td>
<td>360</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>se 7102</td>
<td>7501/Secl+</td>
<td>274</td>
<td>10</td>
<td>2.8 kb</td>
<td>ca. 2100 bp, 700 bp</td>
</tr>
</tbody>
</table>

Polymerase chain reactions were performed as follows. Reactions (50 microliters) contained 100 ng K562 human DNA (Life Technologies, Inc.), 0.2 μM of each primer, and 0.2 mM of each dNTP, in ELONGASE™ SuperMix (Life Technologies, Inc.). Reactions in thin wall tubes under mineral oil were denatured at 94°C for 1 minute, then cycled 35 times at 94°C for 30 seconds, 65°C for 30 seconds, and 68°C for 8 minutes 30 seconds. Following thermal cycling, reactions were maintained at 4°C. The 5.2 kb PCR product (FIG. 9A) was gel purified.

[0354] Plasmid pEZC1202 (FIG. 9B) contained a wild-type attP site, a chloramphenicol resistance gene, a gene encoding the tet repressor, a wild-type loxP site, an origin of replication, and a tet operator/promoter transcribing and controlling the transcription of a kanamycin resistance gene. This plasmid conferred chloramphenicol resistance but not kanamycin resistance, because the tet repressor made by one element of the plasmid kept the kanamycin resistance gene turned off. The pEZC1202 DNA used in this experiment was a miniprep whose concentration was estimated to be about 50 ng per microliter.

[0355] About 40 ng of the gel purified 5.2 kb PCR product were included in a 10 μl reaction that contained about 50 ng of supercoiled pEZC1202, 0.2 units of Cre recombinase, 3.6 ng IHF, and 11 ng of Int in 50 mM Tris-HCl pH about 7.8, 16 mM NaCl, 35 mM KCl, 0.25 mM EDTA, 0.375 mg/ml bovine serum albumin. A second reaction that did not contain the PCR product was included as a control. After incubating at 27°C for 45 min and then 70°C for 5 minutes, 1 μl aliquots were transformed into DH5α UltraMax competent E. coli cells (Life Technologies, Inc.). One fifth of each expression mix was plated on agar that contained 100 μg/ml kanamycin and the plates were incubated overnight at 37°C. The reaction that contained the PCR product gave 34 colonies, while the reaction that lacked the PCR product gave 31 colonies. After the plates sat at room temperature for four days, 26 additional small colonies were seen on the plate from the positive (+PCR product) reaction, while only one additional small colony was seen on the plate from the negative (no PCR product) reaction.

[0356] Twelve of the 26 small colonies were grown overnight in rich broth (CircleGrow) that contained 25 μg/ml kanamycin, and miniprep DNAs were prepared from these cultures, All twelve miniprep plasmids were about 8 kb in size, which corresponded to the size expected for replacement of the chloramphenicol resistance and tet repressor genes in pEZC1202 with the 5.2 kb PCR product. The predicted recombinant product is shown in FIG. 9C. Two of these plasmids were cut with Avai (8 sites predicted) and BamHI (4 sites predicted). All the predicted Avai fragments appeared to be present. One of the BamHI sites predicted in the PCR product (the one closest to the attB end) was absent from both minipreps, but the other BamHI fragments were consistent with the expected structure of the cloned 5.2 kb PCR product.

[0357] Analysis: The replacement of the chloramphenicol resistance and tet repressor genes in pEZC1202 with the 5.2 kb PCR product (part of the human myosin heavy chain) conferred a moderate resistance of the host E. coli cells to kanamycin, but this resistance was not sufficient to allow colonies to appear after overnight incubation. Thus, colonies containing the desired recombination product grew on kana-
mycin plates, but were not seen after overnight incubation, but only after an additional room temperature incubation. Of the 12 Avr1 and BamHI restriction sites predicted from the nucleotide sequence, 11 were confirmed experimentally. Thus the following three observations support the conclusion that the 5.2 kb PCR product was cloned by recombination: (a) small, slow growing colonies appeared only on the plate from the reaction that contained the PCR product; (b) the miniprep plasmids from these colonies were the expected size; and (c) diagnostic restriction cuts gave the expected fragments (with the one above noted exception).

Example 10

Cloning of PCR Fragments

Three sets of pairs of PCR primers (Table 9) were designed to amplify an 830 bp sequence within plasmid pEZC7501 (FIG. 11B) comprising: attB 1-GFP-attB3, with or without additional nucleotides at the outer ends of the 25 bp attB1 and attB3 recombination sites. (Here “outer” refers to the end of the attB sequence that is not adjacent to the GFP gene sequence.) Primer set A added 17 nucleotides upstream of attB1 and 15 nucleotides downstream of attB3; primer set B added 5 and 8 nucleotides to attB1 and attB3, respectively; and primer set C added no additional nucleotides to either attB recombination sequence.

The primer sequences are provided in Table 9:

<table>
<thead>
<tr>
<th>Primer Set</th>
<th>Sequences</th>
<th>(SEQ ID NO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>upper GFP A 5’-TCA CTA GTC GGC CCA CA</td>
<td>(SEQ ID NO:133)</td>
<td></td>
</tr>
<tr>
<td>lower GFP A 5’-GAG CGG CCC CGG ACC AC</td>
<td>(SEQ ID NO:134)</td>
<td></td>
</tr>
<tr>
<td>upper GFP B 5’-GCC CCA GAA GTT TOT ACA AAA</td>
<td>(SEQ ID NO:135)</td>
<td></td>
</tr>
<tr>
<td>lower GFP B 5’-CCC CGC GGA CCA CTT TOT AC</td>
<td>(SEQ ID NO:136)</td>
<td></td>
</tr>
<tr>
<td>upper GFP C 5’-ACA AGT TGT TAC AAA AAA GCA</td>
<td>(SEQ ID NO:137)</td>
<td></td>
</tr>
<tr>
<td>lower GFP C 5’-ACC ACT TGG TAC AAG AAA GCT</td>
<td>(SEQ ID NO:138)</td>
<td></td>
</tr>
</tbody>
</table>

PCR Reactions

Primer sets A and C were used first with the following PCR reactions, in 50 μl, in duplicate. Final concentrations were:

- 20 mM TrisHCl, pH 8.4
- 50 mM KCl
- 0.2 mM of all four deoxynucleotide triphosphates (dNTPs)
- 400 ng/ml pEZC7501 supercoiled DNA template
- 0.5 μM of each primer

Recombinant Taq DNA Polymerase (BRL-GIBCO) 100 U/ml

A duplicate set of the above reactions contained 1 M betaine.

The reactions were first heated for to 94° C. for 1', then cycled 25 times at 94° C. for 45', 55° C. for 30', and 72° C. for 1'.

The size of the PCR reaction products was analyzed on a 1% agarose gel in TAE buffer containing 0.5 μg/ml ethidium bromide. All reactions yielded products of the expected size, thus duplicate reactions were pooled. As the corresponding reactions with and without betaine were not significantly different, these also were pooled, giving a final pooled volume for reactions with primer sets A and C of 200 μl each.

Primer set B was then used with identical reactions to those above performed, except that the reaction volumes were increased to 100 μl. After duplicate reactions and reactions plus and minus betaine were pooled, the final volume of the reactions with primer set B was 400 μl.

The three pooled primer reaction products were stored at -20° C. for 4 weeks.

PCR Product Purification

Each of the three pooled PCR products was extracted once with an equal volume of a mixture of Tris-buffered phenol, isooamyl alcohol and chloroform. The aqueous supernatant was then extracted twice with an equal volume of isobutanol, and the aqueous layer ethanol precipitated with two volumes of ethanol, 0.1 M sodium acetate, pH 6.2. The ethanol precipitates were recovered by centrifugation at 13,000 rpm for 10' at room temperature, and the supernatant discarded. The dried pellets were dissolved in TE: 100 μl for reactions prepared with primer sets A and C; 200 μl for the reactions with primer set B.

To remove PCR primers and extraneous small PCR products, the PCR products were precipitated with polyethylene glycol (PEG) by adding ½ volume of a solution of 30% PEG 8000 (Sigma), 30 mM MgCl₂, mixing well, and centrifuging at 13,000 rpm for 10', all at room temperature. The supernatant was discarded, and the pellets were dissolved in their previous volume of TE buffer. 1 μl aliquots of each of the three PCR products were checked on a 1% agarose gel to quantitate the recovery, which was estimated to be over 90%. The concentration of each PCR product was adjusted using TE to 40 ng/μl.

Recombination Reaction with the PCR Products of Primer Sets A, B, and C

Five 8 μl reactions were assembled in 1XIntegrate buffer (25 mM Tris HCl pH 7.5, 25 mM Tris HCl pH 8.0, 80 mM KCl, 5 mM spermidine, 0.25 mM EDTA, 0.25 mg/ml BSA) containing: 40 ng of pEZC7102 DNA, 2 μl of recombinase mixture (8 ng/μl IHF, 22 ng/μl Int in 1XInt Buffer, 50% glycerol) the reactions differed by the addition of either the PCR product of primer set A (reaction A), primer set B (reaction B), or primer set C (reaction C), the addition of no PCR product (reaction D), or the addition of 40 ng of pEZC7501 SC (supercoiled) DNA(reaction E) as a positive control. All reactions were performed in duplicate.

The reactions were incubated for 45' at 25° C., for 10' at 70° C., then held at 0-5° C. 2 μl aliquots of each reaction were transformed into Max Efficiency DH5α in a
50 μl transformation reaction, and following expression in 50C medium, ½ (100 μl) and ½ (400 μl) of the reactions were plated on kanamycin-containing (50 μg/ml) LB culture plates. The results of the duplicate reactions are shown in Table 10.

### Table 10

<table>
<thead>
<tr>
<th>Transfection</th>
<th>No. Colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 100 μl</td>
<td>464, 668</td>
</tr>
<tr>
<td>A 400 μl</td>
<td>&gt;1000, &gt;1300</td>
</tr>
<tr>
<td>B 100 μl</td>
<td>980, 1229</td>
</tr>
<tr>
<td>B 400 μl</td>
<td>&gt;3000, &gt;3000</td>
</tr>
<tr>
<td>C 100 μl</td>
<td>2, 8</td>
</tr>
<tr>
<td>C 400 μl</td>
<td>13, 20</td>
</tr>
<tr>
<td>D 100 μl</td>
<td>0, 0</td>
</tr>
<tr>
<td>D 400 μl</td>
<td>0, 0</td>
</tr>
<tr>
<td>E 100 μl</td>
<td>56, 70</td>
</tr>
</tbody>
</table>

[0379] Analysis of the Colonies Obtained

[0380] Miniprep DNA was prepared from 8 colonies of each of the Recombination reactions with primer sets A, B or C. The supercoiled DNA was analyzed on a 1% agarose gel: all eight of colonies from the recombination products of primer sets A and B were of the predicted size (2791 bp) for correct recombination between the PCR products (about 824 bp) and the attB1 ori-kan’ attB3 sequence donated by pEZC7102 (1967 bp). Three of the eight reaction products of primer set C were of the predicted size; the other five all were slightly larger than 4 kb.

[0381] Further analysis of the reaction products was performed using two different restriction enzymes, Avai and PvuII, each of which cleaves twice (but at different locations) within the predicted recombinant product, once within the PCR product sequence and once within the sequence contributed by pEZC7102. Both of these enzymes should cleave the intact pEZC7102 recombination partner plasmid at two sites, to give fragments easily distinguished from those of the expected recombinant products.

[0382] The two restriction enzyme digestes yielded the expected sizes of fragments (2373 and 430 bp for Avai; 2473 and 330 bp for PvuII) from the colonies generated from the recombination reactions with primer sets A and B, as well as for the three colonies from primer set C that displayed the expected size of supercoiled DNA. For the other five colonies from primer set C that yielded larger SC DNA, however, the PvuII digest revealed fragments of approximate size to those predicted from a digestion of pEZC7102, whereas the Avai digest revealed only a single fragment, approximately the size of linearized pEZC7102 (4161 bp).

[0383] Analysis

[0384] These results indicate that PCR products generated from templates containing a gene flanked by attB sites can serve as efficient substrates for the reverse recombination reaction. The addition of even short DNA sequences to the ends of the attB1 and attB3 sites or core regions (e.g., 5 bp and 8 bp, respectively, in primer set B) stimulated this reaction by 100 fold or more. Surprisingly, reverse recombination reactions with PCR products containing additional sequence beyond the attB sites appeared in these reactions to be more efficient recombination partners than the supercoiled positive control plasmid, pEZC7501.

[0385] All the recombination products were generated faithfully. A low level of background colonies emerged from the relatively inefficient recombination reactions with primer set C, which lacked additional sequence beyond the 25 bp attB sites. This background appeared to be due to a largely intact pEZC7102 (which encodes kanamycin resistance) lacking an active ccdB death gene, allowing it to survive. Consistent with this interpretation is that one of the two restriction sites for Avai in this plasmid was also altered. One of the Avai sites is present within the ccdB region of pEZC7102. It is likely therefore that the alteration of this site was secondary to mutational inactivation of the ccdB gene.

Example 11

Further Cloning of PCR Fragments

[0386] Two sets of 6 primers for preparing PCR products from the plasmid pBR322 as template were used. One set (Table 11) anneals to sequences flanking the TetR gene, including the TetR promoter. The other set (Table 12) anneals to sequences flanking the AmpR gene, including its promoter. The “tet” and “amp” primers used contain no attB sequences, only sequences inherent to the pBR322 plasmid; the “attB” primers contain, in addition to the pBR322 sequences, the 25 bp of attB 1 or attB3 sequences; the “attB4” primers contain the pBR322-specific sequences, plus the 25 bp of attB 1 or attB3 sequences, each with four Gs at the outer end. (Here “outer” refers to the end of the attB sequence not adjacent to the template-specific primer sequence.)

[0387] Preparation of pBR322 Template

[0388] To improve the efficiency of the PCR reaction, the supercoiled pBR322 DNA was linearized by incubating 3.5 μg of supercoiled (SC) pBR322 DNA in a 200 μl reaction with 15 units of the restriction enzyme NdeI and final concentration of 50 mM Tris-HCl, pH8.0, 10 mM MgCl2, and 50 mM NaCl, for one hour at 37°C.

[0389] The digested pBR322 DNA was extracted once with phenol, isoamyl alcohol, and chloroform, twice with isobutanol, and precipitated by adding two volumes of ethanol plus 0.15M sodium acetate. The precipitate was washed once with 100% ethanol, dried, then dissolved in TE buffer. Recovery of DNA, quantitated on a 1% agarose gel in TAE buffer, 0.5 μg/ml ethidium bromide, was estimated as greater than 80%.

### Table 11

<table>
<thead>
<tr>
<th>tet Primer</th>
<th>Primer Sequence</th>
<th>SEQ ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>tet-L</td>
<td>AAT TCT CAT GTT TGA CAG GCT TTT</td>
<td>48</td>
</tr>
<tr>
<td>tet-R</td>
<td>CCA TGG ATA TCT TCT GCC AAG</td>
<td>49</td>
</tr>
<tr>
<td>attB1-tetL</td>
<td>ACAA TTT CAAA AGCA GGA CAG 50</td>
<td></td>
</tr>
<tr>
<td>attB3-tetR</td>
<td>ACCAG CTT CAAA AGCA CAG 51</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 11-continued

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence</th>
<th>SEQ ID No:</th>
</tr>
</thead>
<tbody>
<tr>
<td>attB1 + 4-tetL</td>
<td>GGGG ACAAG TTTGTA CAAAA AGC-</td>
<td>52</td>
</tr>
<tr>
<td>attB1 + 4-tetL</td>
<td>AGGCT AAT TCT CAT GTC TGA CAG CTTATC</td>
<td></td>
</tr>
<tr>
<td>attB3 + 4-tetR</td>
<td>GGGG ACCAC TTTGTA CAAAA AGC-</td>
<td>53</td>
</tr>
<tr>
<td>attB3 + 4-tetR</td>
<td>TGGGT CGA TGG ATA TCT CTT CTT CTT GGC AAG</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 12

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence</th>
<th>SEQ ID No:</th>
</tr>
</thead>
<tbody>
<tr>
<td>amp-L</td>
<td>AAT ACA TTC AAA TAT GTA TCC GC</td>
<td>54</td>
</tr>
<tr>
<td>amp-R</td>
<td>TTA CCA ATG CTT AAT CAG TGA G</td>
<td></td>
</tr>
<tr>
<td>attB1-4-ampL</td>
<td>ACAA TTTGTA CAAAA AGC-</td>
<td>56</td>
</tr>
<tr>
<td>attB1-4-ampL</td>
<td>AGGCT AAT ACA TTC AAA TAT GTA TCC GC</td>
<td></td>
</tr>
<tr>
<td>attB3-4-ampR</td>
<td>ACCAC TTTGTA CAAAA AGC-</td>
<td>57</td>
</tr>
<tr>
<td>attB3-4-ampR</td>
<td>GGGT TTA CCA ATG CTT AAT CAG TGA G</td>
<td></td>
</tr>
</tbody>
</table>

[0390] PCR Amplification of Tet and Amp Gene Sequences

[0392] Six PCR reactions were performed, in 100 μl, consisting of 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM dNTPs, 2 ng linearized pBR322, 2.5 units of Taq DNA polymerase (GIBCO-BRL), and 0.5 μM of each pair of PCR primers listed in Tables 3 and 4. The reactions were first heated to 94°C for 3 min, then subjected to 25 cycles of 94°C for 45 seconds, 55°C for 30 seconds, and 72°C for 1 min. Based on 1% agarose gel analysis, all the reactions generated products of the expected size, in reasonable yields.

[0393] Purification of PCR Products

[0394] The products from duplicate reactions were pooled; extracted with an equal volume of phenol, isomyl alcohol, and chloroform; extracted twice with an equal volume of isobutanol; and precipitated with two volumes of ethanol, as above. The six precipitates were washed once with 100% ethanol, dried and dissolved in 100 μl TE. 1 μl aliquots were taken for gel analysis of the product before PEG precipitation.

[0395] To each tube was added 50 μl of 30% PEG 8000, 30 mM MgCl₂. The solution was mixed well and centrifuged at 13,000 rpm for 10 min at room temperature. The supernatant was carefully removed, and the precipitate dissolved in 100 μl TE. Recovery was quantitated on a 1% agarose and estimated to be over 90%. The gel analysis also revealed that nucleic acid products smaller than about 300 nucleotides had been effectively removed by the PEG precipitation step.

[0396] Recombination Reactions

[0397] Seven recombination reactions were performed, each in a total volume of 8 μl, containing 1 integrase buffer, 40 ng pEZC7102 (FIG. 11A), and 2 μl recombinase mixture (see above; Example 10). Each of the reactions also contained approximately 40 ng of one of the six above PCR products or, as a positive control, 40 ng of pEZC7501 (FIG. 11B). The amp and tet PCR products with attB sites at their termini are shown in FIGS. 12A and 12B. The reactions were incubated at 25°C for 45 min, at 70°C for 10 min, then held at 0-5°C for 1-2 hours until to transform E. coli.

[0398] E. coli Transformation with Recombination Reaction Products

[0399] 1 μl of each of the recombination reactions was transformed into Max Efficiency DH5α in a 50 μl transformation reaction, and following expression in SOC medium, 1/100 (100 μl) and 1/400 (400 μl) of each reaction were plated on culture plates containing 50 μg/ml kanamycin. The plates were incubated overnight and colonies were counted. The number of colonies obtained from each set of duplicate reactions are displayed in Table 13:

<table>
<thead>
<tr>
<th>Reaction</th>
<th>No. Colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>tet 100 (100 μl)</td>
<td>6, 10</td>
</tr>
<tr>
<td>tet 400 (400 μl)</td>
<td>27, 32</td>
</tr>
<tr>
<td>attB-tet 100</td>
<td>9, 6</td>
</tr>
<tr>
<td>attB-tet 400</td>
<td>27, 26</td>
</tr>
<tr>
<td>attB + 4-tet 100</td>
<td>&gt;2000, &gt;2000</td>
</tr>
<tr>
<td>amp 100</td>
<td>7, 13</td>
</tr>
<tr>
<td>amp 400</td>
<td>59, 65</td>
</tr>
<tr>
<td>attB-amp 100</td>
<td>18, 22</td>
</tr>
<tr>
<td>attB-amp 400</td>
<td>66, 66</td>
</tr>
<tr>
<td>attB + 4-amp 100</td>
<td>2020, 2040</td>
</tr>
<tr>
<td>amp + 4-amp 400</td>
<td>&gt;5000, &gt;5000</td>
</tr>
<tr>
<td>pEZC7501</td>
<td>320, 340</td>
</tr>
<tr>
<td>pEZC7501</td>
<td>1188, 1400</td>
</tr>
</tbody>
</table>

[0400] Analysis of the Colonies Obtained

[0401] As a rapid phenotypic screen, 10 of the colonies from the tet EZC reactions and 33 of the colonies from the attB4-tet EZC reactions were streaked onto an LB culture plate containing tetracycline (15 μg/ml). As a control for the potency of the tetracycline, 3 colonies of pUC19-transformed cells, lacking a TetR gene, were also streaked onto the plate. All colonies from the attB4-tet EZC reactions grew well; colonies from the tet EZC reactions grew only very slightly, and the pUC19 colonies grew not at all.

[0402] Analogous results were obtained by streaking colonies from the amp PCR reactions on culture plates containing ampicillin (100 μg/ml). All 21 colonies generated from the attB4-amp recombination reactions grew well, whereas only one of 13 colonies from the attB-amp reactions grew in the presence of ampicillin. No growth was seen with any of the 15 colonies from the recombination reaction with amp PCR products.

[0403] To characterize plasmid DNA, eight colonies generated from the six EZC reactions with PCR products were
picked into LB broth containing 50 μg/ml kanamycin and grown overnight at 37°C. Miniprep DNA was prepared from 0.9 ml of each culture, and the size of the supercoiled DNA was analyzed on a 1% agarose gel in TAE buffer containing 0.5 μg/ml ethidium bromide. The results are displayed in Table 14. The predicted structures of the recombination products are shown in FIG. 12C and 12D.

<table>
<thead>
<tr>
<th>TABLE 14</th>
<th>Recombination Reactions</th>
<th>DNA (supercoiled)</th>
<th>Predicted Size (bp)</th>
<th>Number with Predicted Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>tet</td>
<td>SC</td>
<td>3386</td>
<td>0/8</td>
<td></td>
</tr>
<tr>
<td>attB-tet</td>
<td>SC</td>
<td>3386</td>
<td>1/8</td>
<td></td>
</tr>
<tr>
<td>attB + 4-tet</td>
<td>SC</td>
<td>3386</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td>attB-amp</td>
<td>SC</td>
<td>2931</td>
<td>0/8</td>
<td></td>
</tr>
<tr>
<td>attB+4-amp</td>
<td>SC</td>
<td>2931</td>
<td>3/8</td>
<td></td>
</tr>
<tr>
<td>attB+4-amp</td>
<td>Pst</td>
<td>429, 2502</td>
<td>3/3</td>
<td></td>
</tr>
</tbody>
</table>

[0404] Analysis

[0405] These results, based on the amplification of two different gene sequences, tet and amp, within the plasmid pBR322 clearly demonstrate that PCR products generated using primers containing the 25 bp attB1 and attB3 recombination sequence serve as highly efficient substrates for the recombination reaction. Addition of a short sequence to the outside of each 25 bp attB site stimulates the recombination reaction by over 100 fold, as also observed in the experiments of Example 10. Also similar to Example 10, the efficiency of the recombination reactions using linear PCR products with attB sites exceeded the efficiency obtained with the positive control SC DNA plasmid, pEZC7501.

[0406] Further, a high percentage of the reaction products are as predicted, since all 33 colonies tested from the attB+4-tet reactions displayed functional tetracycline resistance, and all 21 of the colonies from the attB+4-amp reactions displayed ampicillin resistance. All 16 of the miniprep DNAs, examined from the recombination reactions of either attB+4-tet or attB+4-amp PCR products with pEZC7102, generated supercoiled DNA and restriction digest fragments of the correct sizes.

Example 12

Use of Topoisomerase to Stimulate Recombination

[0407] The stimulation of the recombination reaction by making one or the parental plasmids linear was not expected. If the stimulation resulted from relief of some conformation constraint arising during the two recombination reactions (formation of the Coincriste and resolution to the two daughter molecules), then unwinding of the plasmids with a topoisomerase might also be stimulatory when one or both parental plasmids were circular.

[0408] The Insert Donor was pEZC2901 (FIG. 7A), and the Vector Donor was pEZC3101 (FIG. 7B). A portion of pEZC3101 was linearized with Mlu I. 20 ng of pEZC2901 and/or pEZC3101 were used in each 10 2 reaction (29 ng Int, 2.9 ng Xis, 5.4 ng HIF in 50 mM Tris HCl pH about 7.8, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 0.375 mg/ml BSA, 0.25 mM EDTA, 2% glycerol). Topoisomerase I (from calf Thymus; Life Technologies, Inc.) was diluted from 15 units/μl to the concentrations indicated in Table 15 in 1×EZC buffer.

<table>
<thead>
<tr>
<th>TABLE 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Linear</td>
</tr>
<tr>
<td>Circular</td>
</tr>
</tbody>
</table>

[0409] These reactions were assembled in the following order: buffer; TE; DNAs; Clonase; Topoisomerase. The reactions were incubated at 22°-28° for 45 minutes, then at 70° for 5 minutes. 1 μl aliquots were transformed into UltraMax DH5α competent E. coli (Life Technologies, Inc.). Following expression, aliquots were plated on 100 μg/ml kanamycin and incubated at 30° for 48 hours. Results: see Table 16.

<table>
<thead>
<tr>
<th>TABLE 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction #</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Linear</td>
</tr>
<tr>
<td>Circular</td>
</tr>
</tbody>
</table>

[0410] Analysis

[0411] Linearizing the Vector Donor increased the number of colonies about 10 fold (reaction 2 vs. reaction 7). Addition of 0.5 to 5 units of topoisomerase I to reactions containing circular Insert Donor and linear Vector Donor had little or no effect on the number of colonies (reaction 2 compared to reactions 3, 4, and 5; maximum 1.4 fold). In contrast, if both parental plasmids were circular (reaction 7-10), the addition of topoisomerase stimulated the number of colonies 5 to 9 fold. Thus addition of topoisomerase I to reactions in which both parental plasmids were circular stimulated the recombination reactions nearly as much as linearizing the Vector Donor parent. Topoisomerase I was active when used in combination with the three recombination proteins, in recombination buffer. The addition of topoisomerase I to the recombination reaction releases the necessity to linearize the Vector Donor to achieve stimulation of the recombination reactions.
[0412] Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

[0413] All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

```
SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 60
<210> SEQ ID NO 1
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<221> NAME/KEY: OTHER
<222> LOCATION: 18
<223> OTHER INFORMATION: "n" may be any nucleotide
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

rkycwctttt yktrtaanaa stsgb

<210> SEQ ID NO 2
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<221> NAME/KEY: OTHER
<222> LOCATION: 18
<223> OTHER INFORMATION: "n" may be any nucleotide
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

agccwctttt yktrtaanaa ctsgb

<210> SEQ ID NO 3
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<221> NAME/KEY: OTHER
<222> LOCATION: 18
<223> OTHER INFORMATION: "n" may be any nucleotide
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

gttcagcttt ckttrtaanaa ctsgb

<210> SEQ ID NO 4
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
```
This page contains a continuation of the data provided in the previous page. It includes details on sequences with tags for names, keys, locations, and other information.

**Sequence 1:**
```
agccwgcattt ctttracnaa gtsgb
```

**Sequence 2:**
```
gtccagcttt yktrtacnaa gtsgb
```

**Sequence 3:**
```
agcctgcattt ttgtacaaaa cttgt
```

**Sequence 4:**
```
agcctgcattt cttgtacaaaa cttgt
```

**Sequence 5:**
```
agcctgcattt cttgtacaaaa cttgt
```

**Sequence 6:**
```
agcctgcattt cttgtacaaaa cttgt
```

**Sequence 7:**
```
agcctgcattt cttgtacaaaa cttgt
```

**Sequence 8:**
```
agcctgcattt cttgtacaaaa cttgt
```

The sequences are tagged with information related to their origin and the type of data they represent. Additionally, there are descriptions of unknown organisms and recombination products.
-continued

<400> SEQUENCE: 15

gttctagcttt tttgtacaaa gttgg 25

<210> SEQ ID NO 16
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism; recombination products

<400> SEQUENCE: 16

gttctagcttt tttgtacaaa gttgg 25

<210> SEQ ID NO 17
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism; recombination products

<400> SEQUENCE: 17

ccacoacaaa gcggtcactg gaaactact ttaatcag 39

<210> SEQ ID NO 18
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism; recombination products

<400> SEQUENCE: 18

ccacoacact gagcgtcactg cgcagcact tcctactgt 39

<210> SEQ ID NO 19
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence; synthetic oligonucleotide

<400> SEQUENCE: 19

gcggtcactgcgtcactgcgtcactgcgtcactgcgtcactgcgtc 39

<210> SEQ ID NO 20
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence; synthetic oligonucleotide

<400> SEQUENCE: 20

caggttttccc gtcggttggc tattgtaactc 30

<210> SEQ ID NO 21
<211> LENGTH: 47
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 21

ggcgagattcagatacoca agacgcgaaa acctgtattt tcaggtt 47

<210> SEQ ID NO 22
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 22

cagcttttgc gtcgtagga ttagtatsct t 31

<210> SEQ ID NO 23
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 23

ggcaagattcagatacoca aagacgcgaaa acctgtattt tcaggtt 48

<210> SEQ ID NO 24
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 24

cagcttttgc gtcgtagga ttagtatsct t 32

<210> SEQ ID NO 25
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 25

cacggtttacg tggac 15

<210> SEQ ID NO 26
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 26

tgagttccac gtaacgcgtt cacacttatt a 31

<210> SEQ ID NO 27
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 27
uauuuucagg guatggagaa aaaaactcct ggtatacc

<210> SEQ ID NO 28
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 28
uuccacauu uacgccgccc cctgccacct ato

<210> SEQ ID NO 29
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 29
uauuuucagg guatgcctgt tctggasac cgg

<210> SEQ ID NO 30
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 30
uuccacauu uatttacgcc caaggggccc tttc

<210> SEQ ID NO 31
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 31
tcgttgaag cctgcttttt tatactaact tgaggaag cctgggtcga gcataagg

<210> SEQ ID NO 32
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 32
cgatactt cgttagcat acattattcg asgttatcg cccctgtga cataactcg
<table>
<thead>
<tr>
<th>SEQ ID NO</th>
<th>LENGTH</th>
<th>TYPE</th>
<th>ORGANISM</th>
<th>Feature</th>
<th>Other Information</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>20</td>
<td>DNA</td>
<td>Artificial Sequence</td>
<td></td>
<td>Description of Artificial Sequence: synthetic oligonucleotide</td>
<td>tcactagtgc gggcgcacca</td>
</tr>
<tr>
<td>34</td>
<td>20</td>
<td>DNA</td>
<td>Artificial Sequence</td>
<td></td>
<td>Description of Artificial Sequence: synthetic oligonucleotide</td>
<td>ggccgccag cggcccacac</td>
</tr>
<tr>
<td>35</td>
<td>21</td>
<td>DNA</td>
<td>Artificial Sequence</td>
<td></td>
<td>Description of Artificial Sequence: synthetic oligonucleotide</td>
<td>ggccgccag ttgtagaca</td>
</tr>
<tr>
<td>36</td>
<td>20</td>
<td>DNA</td>
<td>Artificial Sequence</td>
<td></td>
<td>Description of Artificial Sequence: synthetic oligonucleotide</td>
<td>cccgcgcgc cactttgtac</td>
</tr>
<tr>
<td>37</td>
<td>21</td>
<td>DNA</td>
<td>Artificial Sequence</td>
<td></td>
<td>Description of Artificial Sequence: synthetic oligonucleotide</td>
<td>acaagttgt acaaaaagc</td>
</tr>
<tr>
<td>38</td>
<td>21</td>
<td>DNA</td>
<td>Artificial Sequence</td>
<td></td>
<td>Description of Artificial Sequence: synthetic oligonucleotide</td>
<td>acacactttct acagagagc</td>
</tr>
</tbody>
</table>
<210> SEQ ID NO 39
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

<400> SEQUENCE: 39
rbycwgttt yttrracwas stkgd

<210> SEQ ID NO 40
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

<400> SEQUENCE: 40
ascwgytst yttrracwas stkgw

<210> SEQ ID NO 41
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

<400> SEQUENCE: 41
ascwgytst yttrracwas gtgg

<210> SEQ ID NO 42
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

<400> SEQUENCE: 42
gttoagttt yttrracwas stkgw

<210> SEQ ID NO 43
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

<400> SEQUENCE: 43
gttoagttt yttrracwas gtgg

<210> SEQ ID NO 44
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products

<400> SEQUENCE: 44
tcggacgaaa aaatagatt gaact
<210> SEQ ID NO 45
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products
<400> SEQUENCE: 45

tcggacgaaa aacacgtttt gaaca
<210> SEQ ID NO 46
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products
<400> SEQUENCE: 46

tcggacgaaa gacacgtttt gaaca
<210> SEQ ID NO 47
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown Organism: recombination products
<400> SEQUENCE: 47

tggtcgaaa gacacgtttt ccaaa
<210> SEQ ID NO 48
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide
<400> SEQUENCE: 48

aatattcatg ttgacagct tata
<210> SEQ ID NO 49
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide
<400> SEQUENCE: 49

cgatggatat ttgacagcaag
<210> SEQ ID NO 50
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide
<400> SEQUENCE: 50
acaaagtgtg tacagtaaac ggttaccacg ggttaaccg aacagtttgt acaaaaaagc aggctaatttc tcatgtttaa cagottatc 49

<210> SEQ ID NO 51
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 51
acaaagtgtg tacagtaaac ggttaccacg ggttaaccg aacagtttgt acaaaaaagc aggctaatttc tcatgtttaa cagottatc

<210> SEQ ID NO 52
<211> LENGTH: 53
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 52
gggtgtgatgtg tggagcttatg tagattgtct gccaag 46

<210> SEQ ID NO 53
<211> LENGTH: 50
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 53
gggtgtgtgtgtgtgatgtg tggagcttatg tagattgtct gccaag 53

<210> SEQ ID NO 54
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 54
aatatcatacaatgtataatc cgc 23

<210> SEQ ID NO 55
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: synthetic oligonucleotide

<400> SEQUENCE: 55
ttaccaatgc ttaatcagtg ag 22

<210> SEQ ID NO 56
<211> LENGTH: 48
<212> TYPE: DNA
What is claimed is:

1. An in vitro method of cloning an amplification product comprising:
   (a) obtaining an amplification product comprising a first recombination site and a second recombination site which do not recombine with each other; and
   (b) combining said amplification product in vitro with a vector comprising a third recombination site and a fourth recombination site which do not recombine with each other, under conditions such that recombination occurs between said first and third and said second and fourth recombination sites, thereby producing a product vector.

2. The method of claim 1, further comprising inserting said product vector into a host cell.

3. The method of claim 1, wherein said vector is an expression vector.

4. The method of claim 1, wherein said vector comprises at least one additional nucleic acid sequence selected from the group consisting of a selectable marker, a cloning site, a
restriction site, a promoter, an operon, an origin of replication, and a gene or partial gene.

5. The method of claim 1, wherein said vector comprises at least one origin of replication.

6. The method of claim 1, wherein said vector comprises at least one promoter.

7. The method of claim 1, wherein said vector comprises at least one selectable marker.

8. The method of claim 1, wherein said amplification product is linear.

9. The method of claim 1, wherein said first, second, third or fourth recombination sites are lox sites or functional mutants thereof.

10. The method of claim 9, wherein said lox sites are selected from the group consisting of loxP sites and loxP511 sites.

11. The method of claim 1, wherein said first, second, third or fourth recombination sites are att sites or functional mutants thereof.

12. The method of claim 11, wherein said att sites are selected from the group consisting of attB sites, attP sites, attL sites and attR sites.

13. The method of claim 1, wherein said first, second, third or fourth recombination sites are selected from the group consisting of a lox site, an att site, an FRT site, and functional mutants thereof.

14. The method of claim 1, wherein said amplification product and said vector are combined in the presence of at least one recombination protein.

15. The method of claim 14, wherein said recombination protein is Cre.

16. The method of claim 14, wherein said recombination protein is selected from the group consisting of Int, Xis and IHF.

17. The method of claim 1, wherein said amplification product is a polymerase chain reaction product.

18. The method of claim 17, wherein said polymerase chain reaction product is linear.

* * * * *