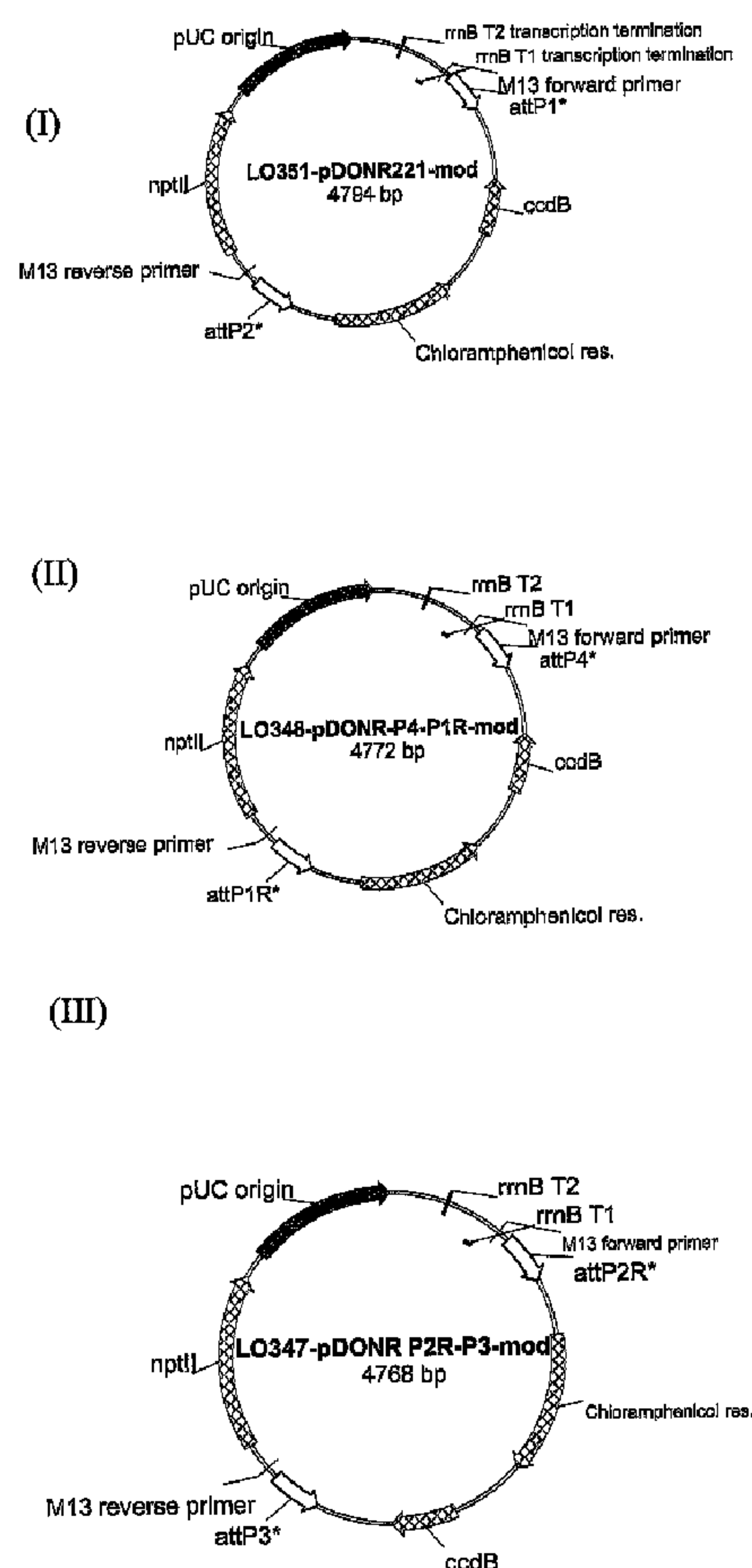




(86) Date de dépôt PCT/PCT Filing Date: 2005/04/28
(87) Date publication PCT/PCT Publication Date: 2005/11/17
(85) Entrée phase nationale/National Entry: 2006/10/16
(86) N° demande PCT/PCT Application No.: EP 2005/004545
(87) N° publication PCT/PCT Publication No.: 2005/108568
(30) Priorité/Priority: 2004/05/10 (EP04011041.3)

(51) Cl.Int./Int.Cl. *C12N 15/09* (2006.01),
C12N 5/10 (2006.01)
(71) Demandeur/Applicant:
BASF PLANT SCIENCE GMBH, DE
(72) Inventeur/Inventor:
LOYALL, LINDA PATRICIA, DE
(74) Agent: ROBIC

(54) Titre : PROCEDES D'ASSEMBLAGE DE PRODUITS D'EXPRESSION MULTIPLES
(54) Title: METHODS FOR ASSEMBLING MULTIPLE EXPRESSION CONSTRUCTS



(57) **Abrégé/Abstract:**

The present invention relates to methods for assembling constructs with multiple gene expression cassettes by directed recombinational cloning. DNA and vectors having engineered recombination sites are provided for use in a recombinational cloning

(57) **Abrégé(suite)/Abstract(continued):**

method that enables efficient and specific assembly of expression cassettes to one multiple expression construct in a single recombinational cloning step.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
17 November 2005 (17.11.2005)

PCT

(10) International Publication Number
WO 2005/108568 A1

(51) International Patent Classification⁷: C12N 15/09, 5/10

(74) Agent: BIEBERBACH, Andreas; c/o Basf Aktiengesellschaft, 67056 Ludwigshafen (DE).

(21) International Application Number:

PCT/EP2005/004545

(22) International Filing Date: 28 April 2005 (28.04.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
04011041.3 10 May 2004 (10.05.2004) EP

(71) Applicant (for all designated States except US): BASF PLANT SCIENCE GMBH [DE/DE]; 67056 Ludwigshafen (DE).

(72) Inventor; and

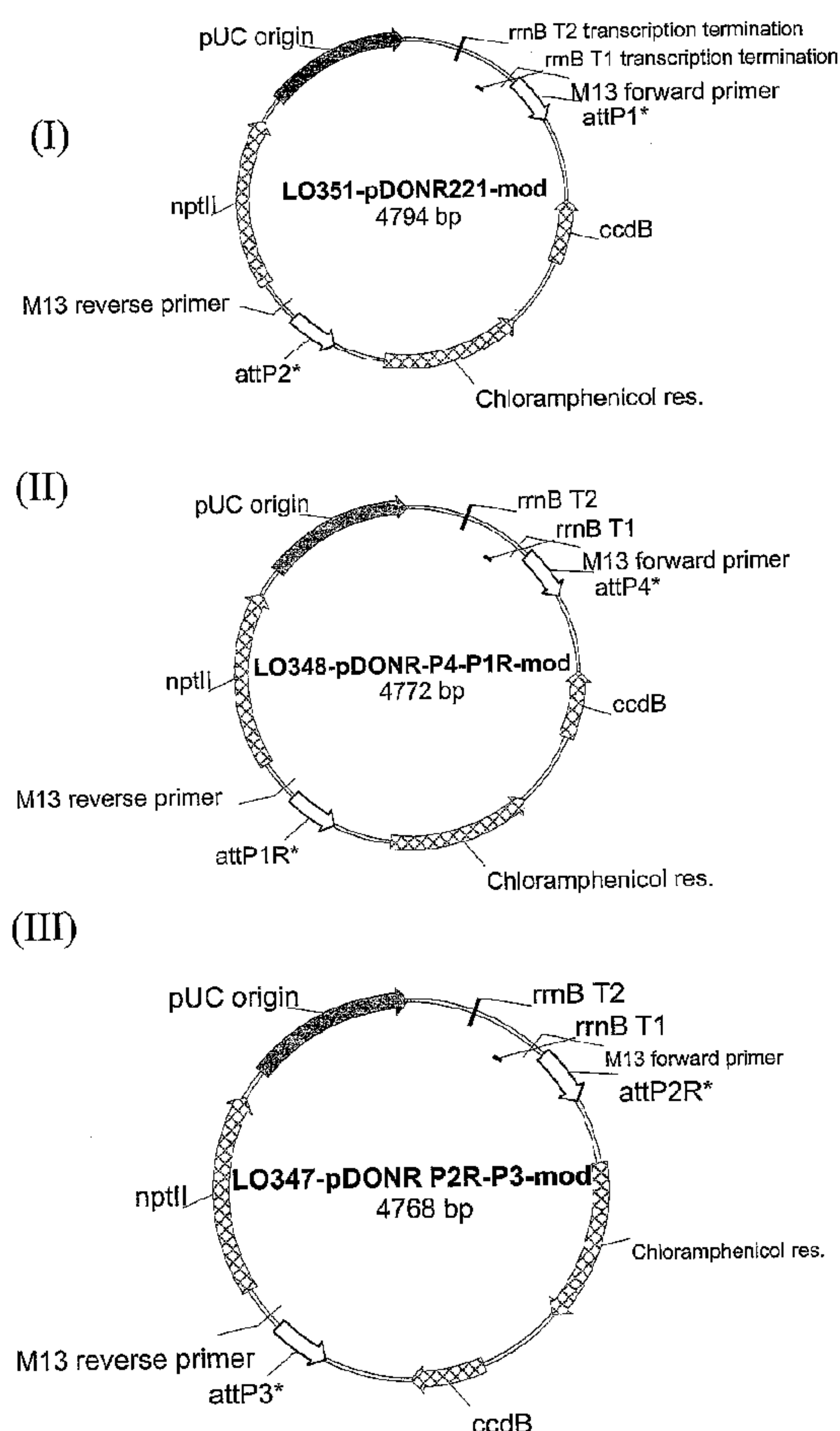
(75) Inventor/Applicant (for US only): LOYALL, Linda, Patricia [DE/DE]; Fred-Joachim-Schoeps-Str. 17, 68229 Mannheim (DE).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO,

[Continued on next page]

(54) Title: METHODS FOR ASSEMBLING MULTIPLE EXPRESSION CONSTRUCTS



(57) Abstract: The present invention relates to methods for assembling constructs with multiple gene expression cassettes by directed recombinational cloning. DNA and vectors having engineered recombination sites are provided for use in a recombinational cloning method that enables efficient and specific assembly of expression cassettes to one multiple expression construct in a single recombinational cloning step.

WO 2005/108568 A1

WO 2005/108568 A1



SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 78

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 78

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

Methods for assembling multiple expression constructs

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to methods for assembling constructs with multiple gene expression cassettes by directed recombinational cloning. DNA and vectors having engineered recombination sites are provided for use in a recombinational cloning method that enables efficient and specific assembly of expression cassettes to one multiple expression construct in a single recombinational cloning step.

2. Related Art

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 ligase 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)). Numerous recombination systems from various organisms have been described. See, e.g., Hoess et al. (1986) Nucleic Acids Res 14(6):2287; Abremski et al. (1986) J Biol Chem 261(1):391; Campbell (1992) J Bacteriol 174(23):7495; Qian et al. (1992) J Biol Chem 267(11):7794; Araki et al. (1992) J Mol Biol 225(1):25; Maeser and Kahnmann (1991) Mol Gen Genet 230:170-176; Esposito et al. (1997) Nucl Acids Res 25(18):3605. Many of these belong to the integrase family of recombinases (Argos et al. (1986) EMBO J. 5:433-440). Perhaps the best studied of these are the Integrase/att system from bacteriophage λ (Landy A (1993) Current Opinions in Genetics and Devel 3:699-707), the Cre/loxP system from bacteriophage P1 (Hoess and Abremski (1990) In Nucleic Acids and Molecular Biology, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the *Saccharomyces cerevisiae* 2 μ circle plasmid (Broach et al. (1982) Cell 29:227-234).

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. 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.

Boyd (Nucl. Acids Res. 21:817-821 (1993)) discloses a method to facilitate the cloning of blunt-ended DNA using conditions that encourage intermolecular 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.

Waterhouse et al. (Nucleic 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 transfect new *E. coli* cells.

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 transposons 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 (Lucklow et al. (1993) J. Virol. 67:4566-4579).

Devine and Boeke Nucl. 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 integrase 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.

US 5,888,732 is describing methods for recombinational cloning facilitating complex cloning procedures. The methods described therein are based on combining DNA fragments by employing recombinases instead using restriction endonucleases and ligases. By choice of certain recombination sites and their combination a practical applicability of recombinase in cloning to obtain predictable products was achieved. The methods described therein are employed to insert library derived cDNAs into a vector, to exchange inserts between various expression vectors, or to combine - for example - a promoter with a coding sequence.

The majority of experiments in biotechnology to date involve the manipulation of single genes. However, many important traits, as well as complex metabolic pathways, depend on interaction among a number of genes, and so genetic engineering had to proceed to the manipulation of polygenic traits, multiple traits, and multiple gene products. Introduction and expression of multiple transgenic expression cassettes is frequently required for both basic and applied studies. However, at present, multigene transformation is a time- and work intensive procedure due to limitations of the existing methods (for review see Francois IEJA et al. (2002) Plant Science 163:281-295; Halpin C et al. (2001) Plant Mol Biol 47:295-310). Current approaches include sexual crossing between plants carrying separate transgenes (Ma JK et al. (1995) Science 268:716-719; Bizily SP et al. (2000) Nat Biotechnol. 18:213-217), sequential retransformation (Lapierre C et al. (1999) Plant Physiol 119:153-163), co-transformation with multiple plasmids (Chen L et al. (1998) Nat Biotechnol 16:1060-1064; Ye X et al. (2000) Science 287:303-305; Hadi MZ et al. (1996) Plant Cell Rep. 15:500-505), use of internal ribosome entry sites (IRES) for combining several mRNA encoding sequences under control of one promoter (Urwin P et al. (2000) Plant J 24:583-589), or single plasmids on which several transgenes were linked by tedious cloning procedures (Goderis IJ et al. (2002) Plant Mol Biol 50:17-27; Dasgupta S et al. (1998) Plant J 16:107-116; Halpin C et al. (1999) Plant J 17: 453-459; Decosa B et al. (2001) Nat Biotechnol 19:71-74; De-Gray G et al. (2001) Plant Physiol 127:852-862; Thompson JM et al. (2002) In Vitro Cell Dev. Biol.-Plant 38: 537-542).

Lin L et al. (Proc Natl Acad Sci USA 2003, 100(10): 5962-5967) is describing a system for assembling up to ten expression cassettes based on subsequent action of site-

specific recombinases and vector back-bone removal by action of homing-
endonucleases. The system has the disadvantage that each expression cassette has
to be added to the transformation construct in a separate round of recombinase medi-
ated integration and subsequent vector backbone removal involving separate steps of
5 plasmid isolation and re-transformation. Although being an improvement and less error-
prone than a standard cloning approach, still numerous sequential operation steps are
necessary.

10 All of the above described technologies are very work- and time consuming. They are
error-prone and require careful verification of each cloning step by sequencing. Accord-
ingly, there is a long felt need to provide an alternative methods to allow for faster and
more efficient transformation of organisms with two or more expression cassettes.

SUMMARY OF THE INVENTION

15 The present invention provides nucleic acids, vectors and methods for obtaining DNA
constructs comprising at least two expression cassettes using recombination proteins
in vitro or in vivo. These methods are highly specific, rapid, and less labor intensive
than standard cloning techniques.

20 Accordingly, a first embodiment of this invention relates to a method for producing a
Multiple Expression Construct comprising at least two different expression cassettes
said method comprising combining in vitro or in vivo

25 i) one or more Insert Donor molecules, said Insert Donor molecules together compris-
ing at least two Inserts I(n), each Insert comprising at least one expression cassette,
said Inserts being flanked by two different recombination sites A(2n) and A(2n+1),
wherein n is an integer from 1 to m characterizing each Insert, and m is the total
number of different Inserts, and

30 ii) a Insert Acceptor molecule comprising two different recombination sites A1 and
A(2m+2), wherein m is the total number of different Inserts of step i)

35 wherein all recombination sites A(1) to A(2m+2) are different, and wherein a recombi-
nation side A(2i-1) for a specific i, said i being an integer from 1 to m+1, can recombine
with the recombination side A(2i) for the same i, but does not substantially recombine
with another recombination site, and

40 iii) at least one site specific recombination protein capable of recombining said recom-
bination sites in said Insert Donor molecules and said Insert Acceptor molecule,
and

incubating said combination under conditions sufficient to transfer all of said Inserts
into said Insert Acceptor molecule, thereby producing a Multiple Expression Construct.

45 The resulting Multiple Expression Construct molecules may optionally be selected or
isolated away from other molecules such as unreacted Insert Acceptor molecules or
Insert Donor molecules or other unintended by-products.

The Insert Donor DNA molecule may further comprise a DNA segment encoding for at least one marker selected from the group consisting of a cloning site, a restriction site, a promoter, an operon, an origin of replication, a functional DNA, an antisense RNA, a PCR fragment, a protein and a protein fragment. Preferably, at least one marker may be comprised within at least one Insert. Because of the high efficiency of the recombinational cloning procedure leading to an unambiguous result with high efficiency, the marker does not necessarily have to be a Selection Marker which would allow for isolation and selection of the resulting Multiple Expression Construct. .

10 In an preferred embodiment the Insert Acceptor molecule comprises a DNA segment flanked by said two different recombination sites A1 and A(2m+2) which is going to be replaced during the recombination process. Preferably, the Insert Acceptor molecule comprise at least one selectable marker. More preferred, the Insert Acceptor molecule further comprises (a) a toxic gene and (b) a Selectable Marker, wherein said toxic gene and said Selectable Marker are on different DNA segments, the DNA segments being separated from each other by at least two recombination sites. Preferably, the toxic gene is deleted from the Insert Acceptor molecule in consequence of the recombinational process.

20 The Selectable Marker, which may be comprised in the Insert, Insert Donor, Insert Acceptor or Vector Donor molecule the may comprise at least one DNA segment selected from the group consisting of:

- (a) a DNA segment that encodes a product that provides resistance in a recipient cell or organism against otherwise toxic compounds ("Negative Selection Marker");
- 25 (b) a DNA segment that encodes a product that is toxic in a recipient cell or organism ("Counter Selection Marker");
- (c) a DNA segment that encodes a product conferring to the recipient cell or organism a growth or proliferation advantage ("Positive Selection Marker");
- (d) a DNA segment that encodes a product that can be identified;
- 30 (e) a DNA segment that encodes a product that inhibits a cell function in a recipient cell;
- (f) a DNA segment that inhibits the activity of any of the DNA segments of (a)-(e) above;
- (g) a DNA segment that binds a product that modifies a substrate;
- 35 (h) a DNA segment that encodes a specific nucleotide recognition sequence which can be recognized by a protein, an RNA, DNA or chemical,
- (i) a DNA segment that, when deleted, directly or indirectly confers sensitivity to cell killing by particular compounds within a recipient cell;
- (j) a DNA segment that encodes a product that suppresses the activity of a gene product in a recipient cell;
- 40 (k) a DNA segment that encodes a product that is otherwise lacking in a recipient cell, and;
- (l) a DNA segment that can be used to isolate or identify a desired molecule.

45 Said Selectable Marker may preferably comprise at least one marker selected from the group consisting of an antibiotic resistance gene, a herbicide resistance gene, a tRNA

gene, an auxotrophic marker, a toxic gene, a phenotypic marker, an antisense oligonucleotide, a restriction endonuclease, a restriction endonuclease cleavage site, an enzyme cleavage site, a protein binding site, and a sequence complementary to a PCR primer sequence.

5

In an preferred embodiment the Insert Acceptor molecule is a Vector Donor molecule, preferably selected from prokaryotic and/or eukaryotic vectors. 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 prokaryotic vectors, eukaryotic vectors or vectors which may shuttle between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred eukaryotic vectors comprise vectors, which replicate in yeast cells, plant cells, fish cells, eukaryotic cells, mammalian cells, or insect cells. Preferred prokaryotic vectors comprise vectors which replicate in gram negative and/or gram positive bacteria, more preferably vectors which replicate in bacteria of the genus *Escherichia*, *Salmonella*, *Bacillus*, *Streptomyces*, *Agrobacterium*, *Rhizobium*, or *Pseudomonas*. Most preferred are vectors which replicates in both *E. coli* and *Agrobacterium*. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate and yeast cells, plant cells, mammalian cells (particularly human cells), fungal cells, insect cells, fish cells and the like. 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 prokaryotic and/or eukaryotic systems depending on the particular vector.

25 The Insert Donor DNA molecule, the Insert Acceptor DNA molecule, and/or the Vector Donor DNA molecule may be comprised of a circular or a circular DNA molecule, respectively. The Insert Donor molecules may comprise a vector or a DNA segment produced by amplification.

30 The method of the invention may further comprise the step of selecting the Multiple Expression Construct comprising all of said Inserts of said Insert Donor molecules.

Various recombination sites and corresponding recombination proteins may be employed in the method of the invention. Preferably, the recombination sites are selected from the group consisting of loxP, attB, attP, attL, and attR. More preferably, said recombination site comprises a DNA sequence selected from the group consisting of:

- 35 (a) RKYCWGCTTTYKTRTACNAASTSGB (m-att) (SEQ ID NO:1);
 (b) AGCCWGCTTTYKTRTACNAACTSGB (m-attB) (SEQ ID NO:2);
 (c) GTTCAGCTTTCKTRTACNAACTSGB (m-attR) (SEQ ID NO:3);
 40 (d) AGCCWGCTTTCKTRTACNAAGTSGB (m-attL) (SEQ ID NO:4);
 (e) GTTCAGCTTTYKTRTACNAAGTSGB (m-attP1) (SEQ ID NO:5);

and 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, C, or G or T/U; S=C or G; and B=C, G or T/U.

45

Most preferably, the recombination site comprises a DNA sequence selected from the group consisting of:

	(a)	AGCCTGCTTTTTTTGTACAAACTTGT	(attB1)	(SEQ ID NO:6);
	(b)	AGCCTGCTTTCTTGTACAAACTTGT	(attB2)	(SEQ ID NO:7);
	(c)	ACCCAGCTTTCTTGTACAAACTTGT	(attB3)	(SEQ ID NO:8);
	(d)	G TTCAGCTTTTTTTGTACAAACTTGT	(attR1)	(SEQ ID NO:9);
5	(e)	G TTCAGCTTTCTTGTACAAACTTGT	(attR2)	(SEQ ID NO:10);
	(f)	G TTCAGCTTTCTTGTACAAAGTTGG	(attR3)	(SEQ ID NO:11);
	(g)	AGCCTGCTTTTTTTGTACAAAGTTGG	(attL1)	(SEQ ID NO:12);
	(h)	AGCCTGCTTTCTTGTACAAAGTTGG	(attL2)	(SEQ ID NO:13);
	(i)	ACCCAGCTTTCTTGTACAAAGTTGG	(attL3)	(SEQ ID NO:14);
10	(j)	G TTCAGCTTTTTTTGTACAAAGTTGG	(attP1)	(SEQ ID NO:15);
	(k)	G TTCAGCTTTCTTGTACAAAGTTGG	(attP2,P3)	(SEQ ID NO:16);
	(l)	AGCCTGCTTTTTTTGTACAAACTTGC	(attB1*)	(SEQ ID NO:52);
	(m)	ACCCAGCTTTCTTGTACAAAGTGGC	(attB2*)	(SEQ ID NO:53);
	(N)	CAACTTTATTATACATAGTTG	(attB3*)	(SEQ ID NO:54)
15	(O)	CAACTTTTCTATACAAAGTTG	(attB4*)	(SEQ ID NO:55)
	(p)	G TTCAACTTTTTTTGTACAAACTTGC	(attR1*)	(SEQ ID NO:56)
	(q)	TTCAACTTTCTTGTACAAAGTGGG	(attR2*)	(SEQ ID NO:57)
	(r)	G TTCAACTTTATTATACATAGTTGA	(attR3*)	(SEQ ID NO:58)
	(s)	G TTCAACTTTTCTATACAAAGTTGA	(attR4*)	(SEQ ID NO:59)
20	(t)	AGCCTGCTTTTTTTGTACAAAGTTGG	(attL1*)	(SEQ ID NO:60)
	(u)	ACCCAGCTTTCTTGTACAAAGTTGG	(attL2*)	(SEQ ID NO:61)
	(v)	GGCAACTTTATTATACAAAGTTGG	(attL3*)	(SEQ ID NO:62)
	(w)	ACCCAAC TTTTCTATACAAAGTTGG	(attL4*)	(SEQ ID NO:63)
	(x)	G TTCAACTTTTTTTGTACAAAGTTGG	(att P1*)	(SEQ ID NO:64)
25	(y)	G TTCAGCTTTCTTGTACAAAGTTGG	(attP2*)	(SEQ ID NO:65)
	(z)	G TTCAACTTTATTATACAAAGTTGG	(att P3*)	(SEQ ID NO:66)
	(aa)	G TTCAACTTTTCTATACAAAGTTGG	(attP4*)	(SEQ ID NO:67)

and a corresponding or complementary DNA or RNA sequence.

- 30 Preferably, said recombination proteins are selected from the group consisting of Int, Cre, Flp, and Res.

The expression cassette comprised in the Inserts of the Insert Donor Molecules consists of a nucleic acid sequence of interest operably linked to a promoter sequence and
 35 – optionally – addition regulatory sequences such as for example, transcription terminator sequences, enhancer etc. The promoter is selected according to the target organism in which expression from the resulting Multiple Expression Construct is intended. For example, in case expression in plants is intended promoters are used which have transcriptional activity in plants. The nucleic acid sequence of interest is to
 40 be understood in the broad sense (as defined above). The expression cassette may result in transcription of untranslatable or translatable RNA. Untranslatable RNA may include for example antisense or double-stranded RNA, which may result in gene si-

lencing of the corresponding endogenous genes thereby conferring a valuable trait. Translatable RNA may result in production of protein thereby conferring a valuable trait.

5 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.

GENERAL DEFINITIONS

10 It is to be understood that this invention is not limited to the particular methodology, protocols, cell lines, plant species or genera, constructs, and reagents described as such. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims. 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.

20 It must be noted that as used herein and in the appended claims, the singular forms "a," "and," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a vector" is a reference to one or more vectors and includes equivalents thereof known to those skilled in the art, and so forth.

25 The term "about" is used herein to mean approximately, roughly, around, or in the region of. When the term "about" is used in conjunction with a numerical range, it modifies that range by extending the boundaries above and below the numerical values set forth. In general, the term "about" is used herein to modify a numerical value above and below the stated value by a variance of 20 percent up or down (higher or lower).

30 As used herein, the word "or" means any one member of a particular list and also includes any combination of members of that list.

35 The term "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, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives include, for example, [α S]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. 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.

45 The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers or hybrids thereof in either single- or double-stranded, sense or antisense form.

Unless otherwise indicated, a particular nucleic acid sequence also implicitly encom-

passes conservatively modified variants thereof (e. g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. The term "nucleic acid" is used interchangeably herein with "gene", "cDNA", "mRNA", "oligonucleotide," and "polynucleotide".

5

The phrase "nucleic acid sequence" as used herein refers to a consecutive list of abbreviations, letters, characters or words, which represent nucleotides. In one embodiment, a nucleic acid can be a "probe" which is a relatively short nucleic acid, usually less than 100 nucleotides in length. Often a nucleic acid probe is from about 50 nucleotides in length to about 10 nucleotides in length. A "target region" of a nucleic acid is a portion of a nucleic acid that is identified to be of interest. A "coding region" of a nucleic acid is the portion of the nucleic acid which is transcribed and translated in a sequence-specific manner to produce into a particular polypeptide or protein when placed under the control of appropriate regulatory sequences. The coding region is said to encode such a polypeptide or protein.

10

The term "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.

15

The term "antisense" is understood to mean a nucleic acid having a sequence complementary to a target sequence, for example a messenger RNA (mRNA) sequence the blocking of whose expression is sought to be initiated by hybridization with the target sequence.

20

The term "sense" is understood to mean a nucleic acid having a sequence which is homologous or identical to a target sequence, for example a sequence which binds to a protein transcription factor and which is involved in the expression of a given gene. According to a preferred embodiment, the nucleic acid comprises a gene of interest and elements allowing the expression of the said gene of interest.

25

The term "gene" refers to a coding region operably joined to appropriate regulatory sequences capable of regulating the expression of the polypeptide in some manner. A gene includes untranslated regulatory regions of DNA (e. g., promoters, enhancers, repressors, etc.) preceding (upstream) and following (downstream) the coding region (open reading frame, ORF) as well as, where applicable, intervening sequences (i.e., introns) between individual coding regions (i.e., exons).

30

As used herein the term "coding region" when used in reference to a structural gene refers to the nucleotide sequences which encode the amino acids found in the nascent polypeptide as a result of translation of a mRNA molecule. The coding region is bounded, in eukaryotes, on the 5'-side by the nucleotide triplet "ATG" which encodes the initiator methionine and on the 3'-side by one of the three triplets which specify stop codons (i.e., TAA, TAG, TGA). In addition to containing introns, genomic forms of a gene may also include sequences located on both the 5'- and 3'-end of the sequences which are present on the RNA transcript. These sequences are referred to as "flanking" sequences or regions (these flanking sequences are located 5' or 3' to the non-

35

40

45

translated sequences present on the mRNA transcript). The 5'-flanking region may contain regulatory sequences such as promoters and enhancers which control or influence the transcription of the gene. The 3'-flanking region may contain sequences which direct the termination of transcription, posttranscriptional cleavage and polyadenylation.

5

The term "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 to 100 "cycles" of denaturation and synthesis of a DNA molecule.

10
15

The terms "polypeptide", "peptide", "oligopeptide", "polypeptide", "gene product", "expression product" and "protein" are used interchangeably herein to refer to a polymer or oligomer of consecutive amino acid residues.

20 The term "isolated" as used herein means that a material has been removed from its original environment. For example, a naturally occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides can be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and would be isolated in that such a vector or composition is not part of its original environment.

Preferably, the term "isolated" when used in relation to a nucleic acid, as in "an isolated nucleic acid sequence" refers to a nucleic acid sequence that is identified and separated from at least one contaminant nucleic acid with which it is ordinarily associated in its natural source. Isolated nucleic acid is nucleic acid present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated nucleic acids are nucleic acids such as DNA and RNA which are found in the state they exist in nature. For example, a given DNA sequence (e.g., a gene) is found on the host cell chromosome in proximity to neighboring genes; RNA sequences, such as a specific mRNA sequence encoding a specific protein, are found in the cell as a mixture with numerous other mRNAs which encode a multitude of proteins. For example, an isolated nucleic acid sequence encoding for a specific trait includes, by way of example, such nucleic acid sequences in cells which ordinarily contain said nucleic acid sequence, wherein said nucleic acid sequence is in a chromosomal or extrachromosomal location different from that of natural cells, or is otherwise flanked by a different nucleic acid sequence than that found in nature. The isolated nucleic acid sequence may be present in single-stranded or double-stranded form. When an isolated nucleic acid sequence is to be utilized to express a protein, the nucleic acid sequence will contain at a minimum at least a portion of the sense or coding strand (i.e., the nucleic acid sequence may be single-stranded). Alternatively, it may contain both the sense and anti-sense strands (i.e., the nucleic acid sequence may be double-stranded).

30
35
40
45

As used herein, the term "purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated. An "isolated nucleic acid sequence" is therefore a purified nucleic acid sequence. "Substantially purified" molecules are at least 60% free, preferably at least 75% free, and
5 more preferably at least 90% free from other components with which they are naturally associated.

As used herein, the terms "complementary" or "complementarity" are used in reference to nucleotide sequences related by the base-pairing rules. For example, the sequence
10 5'-AGT-3' is complementary to the sequence 5'-ACT-3'. Complementarity can be "partial" or "total." "Partial" complementarity is where one or more nucleic acid bases is not matched according to the base pairing rules. "Total" or "complete" complementarity between nucleic acids is where each and every nucleic acid base is matched with another base under the base pairing rules. The degree of complementarity between nu-
15 cleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands.

A "complement" of a nucleic acid sequence as used herein refers to a nucleotide sequence whose nucleic acids show total complementarity to the nucleic acids of the nu-
20 cleic acid sequence.

The term "wild-type", "natural" or of "natural origin" means with respect to an organism, polypeptide, or nucleic acid sequence, that said organism is naturally occurring or available in at least one naturally occurring organism which is not changed, mutated, or
25 otherwise manipulated by man.

The term "transgenic" or "recombinant" when used in reference to a cell refers to a cell which contains a transgene, or whose genome has been altered by the introduction of a transgene. The term "transgenic" when used in reference to a tissue or to a plant
30 refers to a tissue or plant, respectively, which comprises one or more cells that contain a transgene, or whose genome has been altered by the introduction of a transgene. Transgenic cells, tissues and plants may be produced by several methods including the introduction of a "transgene" comprising nucleic acid (usually DNA) into a target cell or integration of the transgene into a chromosome of a target cell by way of human inter-
35 vention, such as by the methods described herein.

The term "transgene" as used herein refers to any nucleic acid sequence which is introduced into the genome of a cell by experimental manipulations. A transgene may be an "endogenous DNA sequence," or a "heterologous DNA sequence" (*i.e.*, "foreign DNA"). The term "endogenous DNA sequence" refers to a nucleotide sequence which is naturally found in the cell into which it is introduced so long as it does not contain some modification (*e.g.*, a point mutation, the presence of a selectable marker gene, *etc.*) relative to the naturally-occurring sequence. The term "heterologous DNA sequence" refers to a nucleotide sequence which is ligated to, or is manipulated to be-
40 come ligated to, a nucleic acid sequence to which it is not ligated in nature, or to which it is ligated at a different location in nature. Heterologous DNA is not endogenous to the cell into which it is introduced, but has been obtained from another cell. Heterologous DNA also includes an endogenous DNA sequence which contains some modification.

Generally, although not necessarily, heterologous DNA encodes RNA and proteins that are not normally produced by the cell into which it is expressed. Examples of heterologous DNA include reporter genes, transcriptional and translational regulatory sequences, selectable marker proteins (e.g., proteins which confer drug resistance), etc.

5 Preferably, the term "transgenic" or "recombinant" with respect to a regulatory sequence (e.g., a promoter of the invention) means that said regulatory sequence is covalently joined and adjacent to a nucleic acid to which it is not adjacent in its natural environment.

10 The term "foreign gene" refers to any nucleic acid (e.g., gene sequence) which is introduced into the genome of a cell by experimental manipulations and may include gene sequences found in that cell so long as the introduced gene contains some modification (e.g., a point mutation, the presence of a selectable marker gene, etc.) relative to the naturally-occurring gene.

15 "Recombinant polypeptides" or "recombinant proteins" refer to polypeptides or proteins produced by recombinant DNA techniques, i.e., produced from cells transformed by an exogenous recombinant DNA construct encoding the desired polypeptide or protein. Recombinant nucleic acids and polypeptide may also comprise molecules which as

20 such does not exist in nature but are modified, changed, mutated or otherwise manipulated by man.

The terms "heterologous nucleic acid sequence" or "heterologous DNA" are used interchangeably to refer to a nucleotide sequence which is ligated to a nucleic acid sequence to which it is not ligated in nature, or to which it is ligated at a different location

25 in nature. Heterologous DNA is not endogenous to the cell into which it is introduced, but has been obtained from another cell. Generally, although not necessarily, such heterologous DNA encodes RNA and proteins that are not normally produced by the cell into which it is expressed.

30 The terms "organism", "host", "target organism" or "host organism" are referring to any prokaryotic or eukaryotic organism that can be a recipient of the recombinational cloning Multiple Expression Construct. A "host," as the term is used herein, includes prokaryotic or eukaryotic organisms that can be genetically engineered. For examples of

35 such hosts, see Maniatis T, Fritsch EF and Sambrook J (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor (NY)). Included are entire organisms but also organs, parts, cells, cultures, propagatable material derived therefrom. Preferred are microorganisms, non-human animal and plant organisms. Preferred microorganisms are bacteria, yeasts, algae or fungi.

40 Preferred bacteria are bacteria of the genus *Escherichia*, *Corynebacterium*, *Bacillus*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Clostridium*, *Proionibacterium*, *Butyrivibrio*, *Eubacterium*, *Lactobacillus*, *Phaeodactylum*, *Colpidium*, *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodium* or *cyanobacteria*, for example of the genus

45 *Synechocystis*. Especially preferred are microorganisms which are capable of infecting plants and thus of transferring the constructs according to the invention. Preferred microorganisms are those from the genus *Agrobacterium* and, in particular, the species *Agrobacterium tumefaciens*.

12

Preferred yeasts are *Candida*, *Saccharomyces*, *Hansenula* or *Pichia*. Preferred fungi are *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Fusarium*, *Beauveria* or other fungi. Plant organisms are furthermore, for the purposes of the invention, other organisms which are capable of photosynthetic activity such as, for example, algae or cyanobacteria, and also mosses. Preferred algae are green algae such as, for example, algae of the genus *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* or *Dunaliella*.

Preferred eukaryotic cells and organism comprise plant cells and organisms, animal cells, and non-human animal organism, including eukaryotic microorganism such as yeast, algae, or fungi.

"Non- human animal organisms" includes but is not limited to non-human vertebrates and invertebrates. Preferred are fish species, non-human mammals such as cow, horse, sheep, goat, mouse, rat or pig, birds such as chicken or goose. Preferred animal cells comprise for example CHO, COS, HEK293 cells. Invertebrate organisms include for example nematodes and insects. Insect cells include for example *Drosophila* S2 and *Spodoptera* Sf9 or Sf21 cells.

Preferred nematodes are those which are capable to invade plant, animal or human organism. Preferred nematodes include for example nematodes of the genus *Ancylostoma*, *Ascaridia*, *Ascaris*, *Bunostomum*, *Caenorhabditis*, *Capillaria*, *Chabertia*, *Cooperia*, *Dictyocaulus*, *Haemonchus*, *Heterakis*, *Nematodirus*, *Oesophagostomum*, *Ostertagia*, *Oxyuris*, *Parascaris*, *Strongylus*, *Toxascaris*, *Trichuris*, *Trichostrongylus*, *Tfhchonema*, *Toxocara* or *Uncinaria*. Especially preferred are plant parasitic nematodes such as *Bursaphelenchus*, *Criconemella*, *Ditylenchus*, *Globodera*, *Helicotylenchus*, *Heterodera*, *Longidorus*, *Meloidogyne*, *Nacobbus*, *Paratylenchus*, *Pratylenchus*, *Radopholus*, *Rotelynchus*, *Tylenchus* or *Xiphinema*. Preferred insects comprise those of the genus *Coleoptera*, *Diptera*, *Lepidoptera*, and *Homoptera*.

Preferred fungi are *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Fusarium*, *Beauveria* or other fungi described in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) on page 15, table 6. Especially preferred is the filamentic Hemiascomycete *Ashbya gosypii*.

Preferred yeasts are *Candida*, *Saccharomyces*, *Hansenula* or *Pichia*, especially preferred are *Saccharomyces cerevisiae* and *Pichia pastoris* (ATCC Accession No. 201178).

The term "plant" or "plant organism" as used herein refers to a plurality of plant cells which are largely differentiated into a structure that is present at any stage of a plant's development. Such structures include one or more plant organs including, but are not limited to, fruit, shoot, stem, leaf, flower petal, etc. Host or target organisms which are preferred as transgenic organisms are especially plants. Included within the scope of the invention are all genera and species of higher and lower plants of the plant kingdom. Included are furthermore the mature plants, seeds, shoots and seedlings and parts, propagation material and cultures derived therefrom, for example cell cultures. The term "mature plants" is understood as meaning plants at any developmental stage

beyond the seedling. The term "seedling" is understood as meaning a young, immature plant in an early developmental stage.

Annual, biennial, monocotyledonous and dicotyledonous plants are preferred host or-
 5 ganisms for the generation of transgenic plants. The expression of genes is further-
 more advantageous in all ornamental plants, useful or ornamental trees, flowers, cut
 flowers, shrubs or lawns. Plants which may be mentioned by way of example but not by
 limitation are angiosperms, bryophytes such as, for example, *Hepaticae* (liverworts)
 and *Musci* (mosses); *Pteridophytes* such as ferns, horsetail and club mosses; gymno-
 10 sperms such as conifers, cycads, ginkgo and *Gnetatae*; algae such as *Chlorophyceae*,
Phaeophyceae, *Rhodophyceae*, *Myxophyceae*, *Xanthophyceae*, *Bacillariophyceae*
 (diatoms) and *Euglenophyceae*.

Preferred are plants which are used for food or feed purpose such as the families of the
 15 *Leguminosae* such as pea, alfalfa and soya; *Gramineae* such as rice, maize, wheat,
 barley, sorghum, millet, rye, triticale, or oats; the family of the *Umbelliferae*, especially
 the genus *Daucus*, very especially the species *carota* (carrot) and *Apium*, very
 especially the species *Graveolens dulce* (celery) and many others; the family of the
Solanaceae, especially the genus *Lycopersicon*, very especially the species
 20 *esculentum* (tomato) and the genus *Solanum*, very especially the species *tuberosum*
 (potato) and *melongena* (egg plant), and many others (such as tobacco); and the
 genus *Capsicum*, very especially the species *annuum* (peppers) and many others; the
 family of the *Leguminosae*, especially the genus *Glycine*, very especially the species
max (soybean), alfalfa, pea, lucerne, beans or peanut and many others; and the family
 25 of the *Cruciferae* (*Brassicaceae*), especially the genus *Brassica*, very especially the
 species *napus* (oil seed rape), *campestris* (beet), *oleracea* cv Tastie (cabbage),
oleracea cv Snowball Y (cauliflower) and *oleracea* cv Emperor (broccoli); and of the
 genus *Arabidopsis*, very especially the species *thaliana* and many others; the family of
 the *Compositae*, especially the genus *Lactuca*, very especially the species *sativa*
 30 (lettuce) and many others; the family of the *Asteraceae* such as sunflower, *Tagetes*,
 lettuce or *Calendula* and many other; the family of the *Cucurbitaceae* such as melon,
 pumpkin/squash or zucchini, and linseed. Further preferred are cotton, sugar cane,
 hemp, flax, chillies, and the various tree, nut and wine species. Very especially
 preferred are *Arabidopsis thaliana*, *Nicotiana tabacum*, *Tagetes erecta*, *Calendula*
 35 *officinalis*, *Gycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, *Pisum sativum*,
Phaseolus vulgaris, *Hordium vulgare*, *Brassica napus*.

The term "cell" refers to a single cell. The term "cells" refers to a population of cells.
 The population may be a pure population comprising one cell type. Likewise, the popu-
 40 lation may comprise more than one cell type. In the present invention, there is no limit
 on the number of cell types that a cell population may comprise. The cells may be syn-
 chronize or not synchronized, preferably the cells are synchronized.

The term "organ" with respect to a plant (or "plant organ") means parts of a plant and
 45 may include (but shall not limited to) for example roots, fruits, shoots, stem, leaves,
 anthers, sepals, petals, pollen, seeds, etc.

The term "tissue" with respect to a plant (or "plant tissue") means arrangement of mul-

multiple plant cells including differentiated and undifferentiated tissues of plants. Plant tissues may constitute part of a plant organ (e.g., the epidermis of a plant leaf) but may also constitute tumor tissues and various types of cells in culture (e.g., single cells, protoplasts, embryos, calli, protocorm-like bodies, etc.). Plant tissue may be *in planta*,
5 in organ culture, tissue culture, or cell culture.

The term "chromosomal DNA" or "chromosomal DNA-sequence" is to be understood as the genomic DNA of the cellular nucleus independent from the cell cycle status. Chromosomal DNA might therefore be organized in chromosomes or chromatids, they
10 might be condensed or uncoiled. An insertion into the chromosomal DNA can be demonstrated and analyzed by various methods known in the art like e.g., polymerase chain reaction (PCR) analysis, Southern blot analysis, fluorescence *in situ* hybridization (FISH), and *in situ* PCR.

15 The term "structural gene" as used herein is intended to mean a DNA sequence that is transcribed into mRNA which is then translated into a sequence of amino acids characteristic of a specific polypeptide.

The term "nucleotide sequence of interest" refers to any nucleotide sequence, the manipulation of which may be deemed desirable for any reason (e.g., confer improved qualities), by one of ordinary skill in the art. Such nucleotide sequences include, but are not limited to, coding sequences of structural genes (e.g., reporter genes, selection
20 marker genes, oncogenes, drug resistance genes, growth factors, etc.), and non-coding regulatory sequences which do not encode an mRNA or protein product, (e.g., promoter sequence, polyadenylation sequence, termination sequence, enhancer sequence, etc.).
25

The term "expression" refers to the biosynthesis of a gene product. For example, in the case of a structural gene, expression involves transcription of the structural gene into
30 mRNA and - optionally - the subsequent translation of mRNA into one or more polypeptides.

The term "expression cassette" or "expression construct" as used herein is intended to mean the combination of any nucleic acid sequence to be expressed in operable linkage with a promoter sequence and - optionally - additional elements (like e.g., terminator and/or polyadenylation sequences) which facilitate expression of said nucleic acid
35 sequence.

The term "promoter," "promoter element," or "promoter sequence" as used herein, refers to a DNA sequence which when ligated to a nucleotide sequence of interest is capable of controlling the transcription of the nucleotide sequence of interest into mRNA. A promoter is typically, though not necessarily, located 5' (*i.e.*, upstream) of a nucleotide sequence of interest (e.g., proximal to the transcriptional start site of a structural gene) whose transcription into mRNA it controls, and provides a site for specific binding
40 by RNA polymerase and other transcription factors for initiation of transcription. 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,
45

though it can vary under the influence of general metabolic conditions.

Promoters may be tissue specific or cell specific. The term "tissue specific" as it applies to a promoter refers to a promoter that is capable of directing selective expression of a nucleotide sequence of interest to a specific type of tissue (e.g., petals) in the relative absence of expression of the same nucleotide sequence of interest in a different type of tissue (e.g., roots). Tissue specificity of a promoter may be evaluated by, for example, operably linking a reporter gene to the promoter sequence to generate a reporter construct, introducing the reporter construct into the genome of a plant such that the reporter construct is integrated into every tissue of the resulting transgenic plant, and detecting the expression of the reporter gene (e.g., detecting mRNA, protein, or the activity of a protein encoded by the reporter gene) in different tissues of the transgenic plant. The detection of a greater level of expression of the reporter gene in one or more tissues relative to the level of expression of the reporter gene in other tissues shows that the promoter is specific for the tissues in which greater levels of expression are detected. The term "cell type specific" as applied to a promoter refers to a promoter which is capable of directing selective expression of a nucleotide sequence of interest in a specific type of cell in the relative absence of expression of the same nucleotide sequence of interest in a different type of cell within the same tissue. The term "cell type specific" when applied to a promoter also means a promoter capable of promoting selective expression of a nucleotide sequence of interest in a region within a single tissue. Cell type specificity of a promoter may be assessed using methods well known in the art, e.g., GUS activity staining or immunohistochemical staining. Briefly, tissue sections are embedded in paraffin, and paraffin sections are reacted with a primary antibody which is specific for the polypeptide product encoded by the nucleotide sequence of interest whose expression is controlled by the promoter. A labeled (e.g., peroxidase conjugated) secondary antibody which is specific for the primary antibody is allowed to bind to the sectioned tissue and specific binding detected (e.g., with avidin/biotin) by microscopy.

Promoters may be constitutive or regulatable. The term "constitutive" when made in reference to a promoter means that the promoter is capable of directing transcription of an operably linked nucleic acid sequence in the absence of a stimulus (e.g., heat shock, chemicals, light, etc.). Typically, constitutive promoters are capable of directing expression of a transgene in substantially any cell and any tissue. In contrast, a "regulatable" promoter is one which is capable of directing a level of transcription of an operably linked nucleic acid sequence in the presence of a stimulus (e.g., heat shock, chemicals, light, etc.) which is different from the level of transcription of the operably linked nucleic acid sequence in the absence of the stimulus.

The term "operable linkage" or "operably linked" is to be understood as meaning, for example, the sequential arrangement of a regulatory element (e.g. a promoter) with a nucleic acid sequence to be expressed and, if appropriate, further regulatory elements (such as e.g., a terminator) in such a way that each of the regulatory elements can fulfill its intended function to allow, modify, facilitate or otherwise influence expression of said nucleic acid sequence. The expression may result depending on the arrangement of the nucleic acid sequences in relation to sense or antisense RNA. To this end, direct linkage in the chemical sense is not necessarily required. Genetic control sequences

such as, for example, enhancer sequences, can also exert their function on the target sequence from positions which are further away, or indeed from other DNA molecules. Preferred arrangements are those in which the nucleic acid sequence to be expressed recombinantly is positioned behind the sequence acting as promoter, so that the two sequences are linked covalently to each other. The distance between the promoter sequence and the nucleic acid sequence to be expressed recombinantly is preferably less than 200 base pairs, especially preferably less than 100 base pairs, very especially preferably less than 50 base pairs. Operable linkage, and an expression construct, can be generated by means of customary recombination and cloning techniques as described (e.g., in Maniatis T, Fritsch EF and Sambrook J (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Silhavy TJ, Berman ML and Enquist LW (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Ausubel FM et al. (1987) *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley Interscience; Gelvin et al. (Eds) (1990) *Plant Molecular Biology Manual*; Kluwer Academic Publisher, Dordrecht, The Netherlands). However, further sequences which, for example, act as a linker with specific cleavage sites for restriction enzymes, or as a signal peptide, may also be positioned between the two sequences. The insertion of sequences may also lead to the expression of fusion proteins. Preferably, the expression construct, consisting of a linkage of promoter and nucleic acid sequence to be expressed, can exist in a vector-integrated form and be inserted into a plant genome, for example by transformation.

The term "transformation" as used herein refers to the introduction of genetic material (e.g., a transgene) into a cell. Transformation of a cell may be stable or transient. The term "transient transformation" or "transiently transformed" refers to the introduction of one or more transgenes into a cell in the absence of integration of the transgene into the host cell's genome. Transient transformation may be detected by, for example, enzyme-linked immunosorbent assay (ELISA) which detects the presence of a polypeptide encoded by one or more of the transgenes. Alternatively, transient transformation may be detected by detecting the activity of the protein (e.g., β -glucuronidase) encoded by the transgene (e.g., the uid A gene) as demonstrated herein [e.g., histochemical assay of GUS enzyme activity by staining with X-gluc which gives a blue precipitate in the presence of the GUS enzyme; and a chemiluminescent assay of GUS enzyme activity using the GUS-Light kit (Tropix)]. The term "transient transformant" refers to a cell which has transiently incorporated one or more transgenes.

In contrast, the term "stable transformation" or "stably transformed" refers to the introduction and integration of one or more transgenes into the genome of a cell, preferably resulting in chromosomal integration and stable heritability through meiosis. Stable transformation of a cell may be detected by Southern blot hybridization of genomic DNA of the cell with nucleic acid sequences which are capable of binding to one or more of the transgenes. Alternatively, stable transformation of a cell may also be detected by the polymerase chain reaction of genomic DNA of the cell to amplify transgene sequences. The term "stable transformant" refers to a cell which has stably integrated one or more transgenes into the genomic DNA. Thus, a stable transformant is distinguished from a transient transformant in that, whereas genomic DNA from the stable transformant contains one or more transgenes, genomic DNA from the transient

transformant does not contain a transgene. Transformation also includes introduction of genetic material into plant cells in the form of plant viral vectors involving epichromosomal replication and gene expression which may exhibit variable properties with respect to meiotic stability. Stable transformation also includes introduction of genetic material into cells in the form of viral vectors involving epichromosomal replication and gene expression which may exhibit variable properties with respect to meiotic stability.

Cloning and transformation techniques for manipulation of ciliates and algae are well known in the art WO 98/01572; Falciatore et al. (1999) Marine Biotechnology 1(3):239-251; Dunahay et al. (1995) J Phycol 31:10004-1012).

Principally speaking transformation techniques suitable for plant cells or organisms (as described below) can also be employed for animal or yeast organism and cells. Preferred are direct transformation techniques such as calcium phosphate or liposome mediated transformation, or electroporation.

The terms "infecting" and "infection" with a bacterium refer to co-incubation of a target biological sample, (e.g., cell, tissue, etc.) with the bacterium under conditions such that nucleic acid sequences contained within the bacterium are introduced into one or more cells of the target biological sample.

The term "*Agrobacterium*" refers to a soil-borne, Gram-negative, rod-shaped phytopathogenic bacterium which causes crown gall. The term "*Agrobacterium*" includes, but is not limited to, the strains *Agrobacterium tumefaciens*, (which typically causes crown gall in infected plants), and *Agrobacterium rhizogenes* (which causes hairy root disease in infected host plants). Infection of a plant cell with *Agrobacterium* generally results in the production of opines (e.g., nopaline, agropine, octopine etc.) by the infected cell. Thus, *Agrobacterium* strains which cause production of nopaline (e.g., strain LBA4301, C58, A208) are referred to as "nopaline-type" *Agrobacteria*; *Agrobacterium* strains which cause production of octopine (e.g., strain LBA4404, Ach5, B6) are referred to as "octopine-type" *Agrobacteria*; and *Agrobacterium* strains which cause production of agropine (e.g., strain EHA105, EHA101, A281) are referred to as "agropine-type" *Agrobacteria*.

The terms "bombarding," "bombardment," and "biolistic bombardment" refer to the process of accelerating particles towards a target biological sample (e.g., cell, tissue, etc.) to effect wounding of the cell membrane of a cell in the target biological sample and/or entry of the particles into the target biological sample. Methods for biolistic bombardment are known in the art (e.g., US 5,584,807, the contents of which are herein incorporated by reference), and are commercially available (e.g., the helium gas-driven microprojectile accelerator (PDS-1000/He) (BioRad)).

The terms "homology" or "identity" when used in relation to nucleic acids refers to a degree of complementarity. Homology or identity between two nucleic acids is understood as meaning the identity of the nucleic acid sequence over in each case the entire length of the sequence, which is calculated by comparison with the aid of the program algorithm GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA) with the parameters being set as follows:

Gap Weight: 12

Length Weight: 4

Average Match: 2,912

Average Mismatch:-2,003

5

Alternatively, a partially complementary sequence is understood to be one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid and is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe (*i.e.*, an oligonucleotide which is capable of hybridizing to another oligonucleotide of interest) will compete for and inhibit the binding (*i.e.*, the hybridization) of a completely homologous sequence to a target under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (*i.e.*, selective) interaction. The absence of non-specific binding may be tested by the use of a second target which lacks even a partial degree of complementarity (*e.g.*, less than about 30% identity); in the absence of non-specific binding the probe will not hybridize to the second non-complementary target.

25

When used in reference to a double-stranded nucleic acid sequence such as a cDNA or genomic clone, the term "substantially homologous" refers to any probe which can hybridize to either or both strands of the double-stranded nucleic acid sequence under conditions of low stringency as described *infra*. When used in reference to a single-stranded nucleic acid sequence, the term "substantially homologous" refers to any probe which can hybridize to the single-stranded nucleic acid sequence under conditions of low stringency as described *infra*.

30

The terms "hybridization" and "hybridizing" as used herein includes "any process by which a strand of nucleic acid joins with a complementary strand through base pairing." (Coombs (1994) Dictionary of Biotechnology, Stockton Press, New York N.Y). Hybridization and the strength of hybridization (*i.e.*, the strength of the association between the nucleic acids) is impacted by such factors as the degree of complementarity between the nucleic acids, stringency of the conditions involved, the T_m of the formed hybrid, and the G:C ratio within the nucleic acids.

35

As used herein, the term " T_m " is used in reference to the "melting temperature." The melting temperature is the temperature at which a population of double-stranded nucleic acid molecules becomes half dissociated into single strands. The equation for calculating the T_m of nucleic acids is well known in the art. As indicated by standard references, a simple estimate of the T_m value may be calculated by the equation: $T_m = 81.5 + 0.41(\% \text{ G+C})$, when a nucleic acid is in aqueous solution at 1 M NaCl [see *e.g.*, Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985)]. Other references include more sophisticated computations which take structural as well as sequence characteristics into account for the calculation of T_m .

40

45

Low stringency conditions when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 68°C. in a solution consisting of 5x SSPE (43.8 g/L NaCl, 6.9 g/L NaH₂PO₄·H₂O and 1.85 g/L EDTA, pH adjusted to 7.4 with NaOH), 1% SDS, 5x Denhardt's reagent [50x Denhardt's contains the following
5 per 500 mL: 5 g Ficoll (Type 400, Pharmacia), 5 g BSA (Fraction V; Sigma)] and 100 µg/mL denatured salmon sperm DNA followed by washing in a solution comprising 0.2x SSPE, and 0.1% SDS at room temperature when a DNA probe of about 100 to about 1000 nucleotides in length is employed.

10 High stringency conditions when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 68° C. in a solution consisting of 5x SSPE, 1% SDS, 5x Denhardt's reagent and 100 µg/mL denatured salmon sperm DNA followed by washing in a solution comprising 0.1x SSPE, and 0.1% SDS at 68° C. when a probe of about 100 to about 1000 nucleotides in length is employed.

15 The term "equivalent" when made in reference to a hybridization condition as it relates to a hybridization condition of interest means that the hybridization condition and the hybridization condition of interest result in hybridization of nucleic acid sequences which have the same range of percent (%) homology. For example, if a hybridization
20 condition of interest results in hybridization of a first nucleic acid sequence with other nucleic acid sequences that have from 80% to 90% homology to the first nucleic acid sequence, then another hybridization condition is said to be equivalent to the hybridization condition of interest if this other hybridization condition also results in hybridization of the first nucleic acid sequence with the other nucleic acid sequences that have from
25 80% to 90% homology to the first nucleic acid sequence.

When used in reference to nucleic acid hybridization the art knows well that numerous equivalent conditions may be employed to comprise either low or high stringency conditions; factors such as the length and nature (DNA, RNA, base composition) of the
30 probe and nature of the target (DNA, RNA, base composition, present in solution or immobilized, etc.) and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate, polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency hybridization different from, but equivalent to, the above-listed conditions. Those skilled in the art know that whereas higher stringencies may be preferred
35 to reduce or eliminate non-specific binding between the nucleotide sequence of interest and other nucleic acid sequences, lower stringencies may be preferred to detect a larger number of nucleic acid sequences having different homologies to the nucleotide sequence of interest.

40 The term "byproduct" is referring to a daughter molecule lacking one or more or all of the expression cassettes which are desired to be subcloned.

45 The term "cointegrate" is referring to at least one recombination intermediate DNA molecule of the present invention that contains both parental (starting) DNA molecules (Insert Donor and Insert Acceptor). It will usually be circular. In some embodiments it can be linear.

The term "Insert" or "Inserts" as used within the context of this invention is referring a nucleic acid sequence (preferably a DNA sequence) flanked by recombination sites. Such Insert may comprise one or more expression cassettes.

- 5 The term "Insert Donor" is referring to is one of the two classes of 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 se-
10 quence outside of the recombination signals.

The term "Multiple Expression Construct" is referring to a desired daughter molecule comprising the Inserts of the Insert Donor molecules which is produced after the recombination events during the recombinational cloning process (see FIG. 6, 7).

- 15 The term "recognition sequence" refers to a 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 recog-
20 nition 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 Sauer B (1994) Current Opinion in Biotechnology 5:521-527; figure 1). Other examples of recognition sequences are the attB, attP, attL, and attR sequences which are recognized by the recombinase enzyme λ
25 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 (IHF), FIS and excisionase (Xis). See Landy, Current Opinion in Biotechnology 3:699-707 (1993). Such
30 sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. When such engineered sites lack the P1 or H1 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.

- 35 The term "recombinase" is referring to an enzyme which catalyzes the exchange of DNA segments at specific recombination sites.

- 40 The term "recombinational cloning" is referring to 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, by action of a site-specific recombinase.

- 45 The term "Recombination proteins" refers to polypeptide including excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites (Landy (1993) Current Opinion in Biotechnology 3:699-707).

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

The term "selectable marker" is referring to a DNA segment that allows one to select for or against a molecule or a cell that contains it, of ten 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:

(a) a DNA segment that encodes a product that provides resistance in a recipient cell or organism against otherwise toxic compounds ("Negative Selection Marker"); (e.g., antibiotics). Negative Selection Markers confer a resistance to a biocidal compound such as a metabolic inhibitor (e.g., 2-deoxyglucose-6-phosphate, WO 98/45456), antibiotics (e.g., kanamycin, G 418, bleomycin or hygromycin) or herbicides (e.g., phosphinothricin or glyphosate). Especially preferred Negative Selection Markers are those which confer resistance to herbicides. Examples which may be mentioned are:

- Phosphinothricin acetyltransferases (PAT; also named Bialophos[®] resistance; bar; de Block *et al.* (1987) EMBO J 6:2513-2518; EP 0 333 033; US 4,975,374)
- 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) conferring resistance to Glyphosate[®] (N-(phosphonomethyl)glycine) (Shah *et al.* (1986) Science 233: 478)
- Glyphosate[®] degrading enzymes (Glyphosate[®] oxidoreductase; gox),
- Dalapon[®] inactivating dehalogenases (deh)
- sulfonylurea- and imidazolinone-inactivating acetolactate synthases (for example mutated ALS variants with, for example, the S4 and/or Hra mutation)
- Bromoxynil[®] degrading nitrilases (bxn)
- Kanamycin- or G418- resistance genes (NPTII; NPTI) coding e.g., for neomycin phosphotransferases (Fraley *et al.* (1983) Proc Natl Acad Sci USA 80:4803)
- 2-Desoxyglucose-6-phosphate phosphatase (DOG^R1-Gene product; WO 98/45456; EP 0 807 836) conferring resistance against 2-desoxyglucose (Randez-Gil *et al.* (1995) Yeast 11:1233-1240).
- hygromycin phosphotransferase (HPT), which mediates resistance to hygromycin (Vanden Elzen *et al.* (1985) Plant Mol Biol. 5:299).
- dihydrofolate reductase (Eichholtz *et al.* (1987) Somatic Cell and Molecular Genetics 13:67-76)

Additional negative selectable marker genes of bacterial origin that confer resistance to antibiotics include the aadA gene, which confers resistance to the antibiotic spectinomycin, gentamycin acetyl transferase, streptomycin phosphotransferase (SPT), aminoglycoside-3-adenyl transferase and the bleomycin resistance determinant (Hayford *et al.*, Plant Physiol. 86:1216 (1988); Jones *et al.* (1987) Mol Gen Genet 210:86; Svab *et al.* (1990) Plant Mol. Biol. 14:197; Hille *et al.* (1986) Plant Mol. Biol. 7:171).

Especially preferred are negative selection marker which confer resistance against the toxic effects imposed by D-amino acids like e.g., D-alanine and D-serine (WO 03/060133). Especially preferred as negative selection marker in this contest are

the daol gene (EC: 1.4. 3.3 : GenBank Acc.-No.: U60066) from the yeast *Rhodotorula gracilis* (*Rhodospiridium toruloides*) and the *E. coli* gene *dsdA* (D-serine dehydratase (D-serine deaminase) [EC: 4.3. 1.18; GenBank Acc.-No.: J01603).

5 Selection Marker suitable in prokaryotic or non-plant eukaryotic systems can also be based on the Selection Markers described above for plants (beside that expression cassettes are based on other host-specific promoters). For mammal cells preferred are resistance against neomycin (G418), hygromycin, Bleomycin, Zeocin Gatignol et al. (1987) Mol. Gen. Genet. 207:342; Drocourt et al. (1990) Nucl. Acids Res. 18:4009), puromycin (see, for example, Kaufman (1990) Meth. Enzymol. 10 185:487; Kaufman (1990) Meth. Enzymol. 185:537). Corresponding selectable marker genes are known in the art (see, for example, Srivastava and Schlessinger, Gene 103:53 (1991); Romanos et al., "Expression of Cloned Genes in Yeast," in DNA Cloning 2: Expression Systems, 2.sup.nd Edition, pages 123-167 (IRL Press 1995); Markie, Methods Mol. Biol. 54:359 (1996); Pfeifer et al., Gene 188:183 15 (1997); Tucker and Burke, Gene 199:25 (1997); Hashida-Okado et al., FEBS Letters 425:117 (1998). For prokaryotes preferred are resistance against Ampicillin, Kanamycin, Specomycin, or Tetracyclin. Selectable marker genes can be cloned or synthesized using published nucleotide sequences, or marker genes can be obtained commercially.

20

(b) a DNA segment that encodes a product that is toxic in a recipient cell or organism ("Counter Selection Marker"). Counter Selection Marker are especially suitable to select organisms with defined deleted sequences comprising said marker (Koprek T et al. (1999) Plant J 19(6): 719-726). Examples for negative selection marker 25 comprise thymidin kinases (TK), cytosine deaminases (Gleave AP et al. (1999) Plant Mol Biol. 40(2):223-35; Perera RJ et al. (1993) Plant Mol. Biol 23(4): 793-799; Stougaard J; (1993) Plant J 3:755-761), cytochrom P450 proteins (Koprek T et al. (1999) Plant J 19(6): 719-726), haloalkan dehalogenases (Naested H (1999) Plant J 18:571-576), *iaaH* gene products (Sundaresan *et al.* (1995) Gene Develop 9: 1797-1810), cytosine deaminase *codA* (Schlaman HRM and Hooykaas PJJ (1997) Plant J 11:1377-1385), or *tms2* gene products (Fedoroff NV & Smith DL (1993) Plant J 3:273- 289).

(c) a DNA segment that encodes a product conferring to the recipient cell or organism a growth or proliferation advantage ("Positive Selection Marker"). Genes like 35 isopentenyltransferase from *Agrobacterium tumefaciens* (strain:PO22; Genbank Acc.-No.: AB025109) may – as a key enzyme of the cytokinin biosynthesis – facilitate regeneration of transformed plants (e.g., by selection on cytokinin-free medium). Corresponding selection methods are described (Ebinuma *et al.* (2000) Proc Natl Acad Sci USA 94:2117-2121; Ebinuma *et al.* (2000) Selection of Marker-free transgenic plants using the oncogenes (*ipt*, *rol A*, *B*, *C*) of *Agrobacterium* as 40 selectable markers, In Molecular Biology of Woody Plants. Kluwer Academic Publishers). Additional Positive Selection Markers, which confer a growth advantage to a transformed plant in comparison with a non-transformed one, are described e.g., in EP-A 0 601 092. Growth stimulation selection markers may include (but shall not 45 be limited to) β -glucuronidase (in combination with e.g., a cytokinin glucuronide), mannose-6-phosphate isomerase (in combination with mannose), UDP-galactose-

4-epimerase (in combination with e.g., galactose), wherein mannose-6-phosphate isomerase in combination with mannose is especially preferred.

- 5 (d) a DNA segment that encodes a product that can be readily identified ("Reporter Genes or Proteins"; e.g., phenotypic markers such as β -galactosidase, green fluorescent protein (GFP), and cell surface proteins). Reporter genes encode readily quantifiable proteins and, via their color or enzyme activity, make possible an assessment of the transformation efficacy, the site of expression or the time of expression. Very especially preferred in this context are genes encoding reporter proteins (Schenborn E, Groskreutz D. (1999) *Mol Biotechnol* 13(1):29-44) such as the
- 10 green fluorescent protein (GFP) (Sheen et al. (1995) *Plant J* 8(5):777-784; Haseloff et al. (1997) *Proc Natl Acad Sci USA* 94(6):2122-2127; Reichel et al. (1996) *Proc Natl Acad Sci USA* 93(12):5888-5893; Tian et al. (1997) *Plant Cell Rep* 16:267-271; WO 97/41228; Chui WL et al. (1996) *Curr Biol* 6:325-330; Leffel SM et al. (1997) *Biotechniques* 23(5):912-8), chloramphenicol transferase, a luciferase (Ow et al. (1986) *Science* 234:856-859; Millar et al. (1992) *Plant Mol Biol Rep* 10:324-414), the aequorin gene (Prasher et al. (1985) *Biochem Biophys Res Commun* 126(3):1259-1268), β -galactosidase, R locus gene (encoding a protein which regulates the production of anthocyanin pigments (red coloring) in plant tissue and thus makes possible the direct analysis of the promoter activity without addition of
- 20 further auxiliary substances or chromogenic substrates (Dellaporta et al. (1988) In: *Chromosome Structure and Function: Impact of New Concepts*, 18th Stadler Genetics Symposium, 11:263-282; Ludwig et al. (1990) *Science* 247:449), with β -glucuronidase (GUS) being very especially preferred (Jefferson (1987b) *Plant Mol. Bio. Rep.*, 5:387-405, Jefferson et al. (1987a) *EMBO J.*, 6:3901-3907). β -glucuronidase (GUS) expression is detected by a blue color on incubation of the tissue with 5-bromo-4-chloro-3-indolyl- β -D-glucuronic acid, bacterial luciferase (LUX) expression is detected by light emission; firefly luciferase (LUC) expression is detected by light emission after incubation with luciferin; and galactosidase expression is detected by a bright blue color after the tissue is stained with 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside. Reporter genes may also be used as
- 30 scorable markers as alternatives to antibiotic resistance markers. Such markers are used to detect the presence or to measure the level of expression of the transferred gene. The use of scorable markers in plants to identify or tag genetically modified cells works well only when efficiency of modification of the cell is high.
- 35 (e) a DNA segment that encodes a product that inhibits a cell function in a recipient cell;
- (f) a DNA segment that inhibits the activity of any of the DNA segments of (a)-(e) above;
- 40 (g) a DNA segment that binds a product that modifies a substrate (e.g. restriction endonucleases);
- (h) a DNA segment that encodes a specific nucleotide recognition sequence which can be recognized by a protein, an RNA, DNA or chemical,
- (i) a DNA segment that, when deleted or absent, directly or indirectly confers resistance or sensitivity to cell killing by particular compounds within a recipient cell;
- 45 (j) a DNA segment that encodes a product that suppresses the activity of a gene product in a recipient cell;

- (k) a DNA segment that encodes a product that is otherwise lacking in a recipient cell (e.g, tRNA genes, auxotrophic markers), and;
- (l) a DNA segment that can be used to isolate or identify a desired molecule (e.g., specific protein binding sites).

5

The term "site-specific recombinase" as used herein is referring to 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 reseal the cleaved strands of nucleic acid. See Sauer, B., Current Opinions in Biotechnology 5:521-527 (1994). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of specificity for both partners. The strand exchange mechanism involves the cleavage and rejoining of specific DNA sequences in the absence of DNA synthesis (Landy, A. (1989) Ann. Rev. Biochem. 58:913-949).

10
15

The term "subcloning vector" is referring to a cloning vector comprising a circular or linear nucleic acid molecule which includes preferably an appropriate replicon. In the present invention, the subcloning vector 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. The subcloning vector can also contain a Selectable marker (preferably DNA).

20

The term vector is referring to 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 (US 5,334,575, entirely incorporated herein by reference), TA Cloning™ brand PCR cloning (Invitrogen Corp., Carlsbad, Calif.), 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.

25
30
35
40

The term "Insert Acceptor" is referring to 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 backbone (e.g. a vector backbone) which is to become part of the desired Multiple Expression Construct. Preferably, the Insert Acceptor is a vector molecule constituting a Vector Donor. The Vector Donor comprises a subcloning vector (or it can be called the cloning vector if the Insert Donor does not already contain a cloning vec-

45

tor) and a segment flanked by recombination sites (see FIG. 6, 7). Both the segment flanked by the recombination sites and the outside of these can contain elements that contribute to selection for the desired Multiple Expression Construct molecule, as described above for selection schemes. The recombination 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.

The term "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 1000 bases, 1 to about 500 bases, and 1 to about 100 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.

The term "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 polymerases 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.

The term "adapter" refers to 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 molecule as well as other nucleic acid molecules described herein. When using portions of recombination sites, the missing portion may be provided by the Insert Donor molecule. Such adapters may be added at any location within a circular or 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. The synthesis of adapters (e.g., by oligonucleotide synthesis, annealing procedures, and or PCR) is a standard technique well known to the person skilled in the art. 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.

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 arts.

20

DETAILED DESCRIPTION OF THE INVENTION

Accordingly, a first embodiment of this invention relates to a method for producing an Multiple Expression Construct comprising at least two different expression cassettes said method comprising combining in vitro or in vivo

25

i) one or more Insert Donor molecules, said Insert Donor molecules together comprising at least two Inserts I(n), each Insert comprising at least one expression cassette, said Inserts being flanked by two different recombination sites A(2n) and A(2n+1), wherein n is an integer from 1 to m characterizing each Insert, and m is the total number of different Inserts, and

30

ii) a Insert Acceptor molecule IA comprising two different recombination sites A1 and A(2m+2), wherein m is the total number of different Inserts of step i)

wherein all recombination sites A(1) to A(2m+2) are different, and wherein a recombination side A(2i-1) for a specific i, said i being an integer from 1 to m+1, can recombine with the recombination side A(2i) for the same i, but does not substantially recombine with another recombination site, and

35

iii) at least one site specific recombination protein capable of recombining said recombination sites in said Insert Donor molecules and said Insert Acceptor molecule, and

40

incubating said combination under conditions sufficient to transfer all of said Inserts into said Insert Acceptor molecule, thereby producing a Multiple Expression Construct.

45

By choice of certain recombination sites, from which each allows only for recombination with one other recombination site, directed assembling of multiple expression cassettes

to an unambiguous product (Multiple Expression Construct) is achieved. The improved specificity, speed and/or yields of the present invention facilitates clustering of multiple expression cassettes in one DNA molecule (e.g., one expression vector).

5 The Insert Donor molecules used in accordance with the invention comprise two or more recombination sites which allow the Inserts (comprising the expression cassette for the nucleic acid segment of interest) of the Insert Donor molecules to be transferred or moved into one or more Insert Acceptor molecules (e.g., Vector Donor molecules) in accordance with the invention. The Insert Donor molecules of the invention may be
10 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
15 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 US 5,539,082 and 5,482,836). Particularly preferred nucleic acid molecules are DNA molecules (single stranded or double stranded). Additionally, the
20 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).

30 1. Recombination Proteins

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

35 Cre: A protein from bacteriophage P1 (Abremski 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
40 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 cofactor, as is well known in the art. The DNA substrates can be either linear or
45 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.

Integrase: A protein from bacteriophage lambda that mediates the integration of the lambda genome into the *E. coli* chromosome. The bacteriophage λ 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, IHF and FIS) determine the direction of recombination.

Integrative recombination involves the Int and IHF 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, HF, 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, intermolecular recombination between molecules containing attL and attR, in the presence of Int, IHF 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 recombination proteins, one can direct excisive or integrative recombination, as reverse reactions of each other.

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., *Ann. 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.

Integrase 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 integrase and IHF works *in vitro*, with simple buffer containing spermidine. Integrase can be obtained as described by Nash, H. A., *Methods of Enzymology* 100:210-216 (1983). IHF can be obtained as described by Filutowicz, M., et al., *Gene* 147:149-150 (1994).

A preferred ready-to-use mixture of lambda integrase with its corresponding co-factors can be obtained from Invitrogen Inc. (GatewayTM LR ClonaseTM Plus enzyme). GatewayTM LR ClonaseTM Plus enzyme mix contains the bacteriophage lambda recombination proteins Integrase (Int) and Excisionase (Xis), and the *E. coli*-encoded protein Integration Host Factor (IHF) (Landy, A. (1989) *Ann. Rev. Biochem.* 58, 913). GatewayTM LR ClonaseTM Plus enzyme mix promotes *in vitro* recombination between attL- and

attR-flanked regions on entry clones and destination vectors to generate *attB*-containing expression clones.

5 Numerous recombination systems from various organisms can also be used, based on the teaching and guidance provided herein. See, e.g., Hoess et al., *Nucleic Acids Research* 14(6):2287 (1986); Abremski et al., *J. Biol. Chem.* 261(1):391 (1986); Campbell, J. *Bacteriol.* 174(23):7495 (1992); Qian et al., *J. Biol. Chem.* 267(11):7794 (1992); Araki et al., *J. Mol. Biol.* 225(1):25 (1992)). Many of these belong to the integrase family of recombinases (Argos et al. *EMBO J.* 5:433-440 (1986)). Perhaps the best studied
10 of these are the Integrase/*att* system from bacteriophage λ (Landy, A. (1993) *Current Opinions in Genetics and Devel.* 3:699-707), the Cre/*loxP* system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the *Saccharomyces cerevisiae* 2 μ circle plasmid (Broach et al. *Cell*
15 29:227-234 (1982)).

Members of a second family of site-specific recombinases, the resolvase family (e.g., $\gamma\delta$, 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., in-
20 versions and excisions) and can require host-encoded factors. Mutants have been isolated that relieve some of the requirements for host factors (Maeser and Kahnmann (1991) *Mol. Gen. Genet.* 230:170-176), as well as some of the constraints of intramolecular recombination.

25 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.

30 While Cre and Int are described in detail for reasons of example, many related recombinase systems exist and their application to the described 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
35 the present invention, as site-specific recombination proteins encoded by, for example bacteriophage lambda, 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 C terminus is particularly well conserved in all the proteins and is homologous to a region near
40 the C terminus of the yeast 2 μ plasmid Fip 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
45 rejoining. See, e.g., Argos, P. et al., *EMBO J.* 5:433-40 (1986).

The recombinases of some transposons, such as those of conjugative transposons (e.g., Tn916) (Scott and Churchward. 1995. *Ann Rev Microbiol* 49:367; Taylor and

Churchward, 1997. *J Bacteriol* 179:1837) belong to the integrase family of recombinases and in some cases show strong preferences for specific integration sites (Ike et al 1992. *J Bacteriol* 174:1801; Trieu-Cuot et al, 1993. *Mol. Microbiol* 8:179).

5 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
10 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.

15 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 transposi-
20 tion, 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).

An unrelated family of recombinases, the transposases, have also been used to transfer
25 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 repli-
30 cons (Lucklow et al. 1993. *J. Virol* 67:4566-4579).

A related element, the integron, are also translocatable-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
35 integrase (IntI1) from In2 is common to all integrons in this class and mediates recombination between two 59-bp elements or between a 59-bp element and an attI site that can lead to insertion into a recipient integron. The integrase also catalyzes excisive recombination. (Hall, 1997. *Ciba Found Symp* 207:192; Francia et al., 1997. *J Bacteriol* 179:4419).

40 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
45 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 (Ma-

tsuura 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 sub-cloning methods.

5

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.

10

2.1 Recombination Sites – General

The above recombinase proteins and corresponding recombination sites are suitable for use in recombination cloning according to the present invention. However, wild-type recombination sites may contain sequences that reduce the efficiency or specificity of recombination reactions, which is required for unambiguous assembly of the Multiple Expression Construct molecules by the methods of the present invention. Furthermore, multiple stop codons in attB, attR, attP, attL and loxP recombination sites occur in multiple reading frames on both strands, so translation efficiencies are reduced, e.g., where the coding sequence must cross the recombination sites, (only one reading frame is available on each strand of loxP and attB sites) or impossible (in attP, attR or attL).

20

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 (e.g., att1, att2, 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 subcloning according to the present invention.

25

30

Mutations can therefore be introduced into recombination sites for enhancing site specific recombination. Such mutations include, but are not limited to recombination sites recognized by the same proteins but differing in base sequence such that they react largely or exclusively only with their homologous partners. Parallel use of such mutated recombination sites allows multiple (parallel) reactions to be contemplated. 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 DNA constructs containing recombination sites attR1 and attR3; attL1 and attR2; attL2 and attL3, respectively.

35

40

There are well known procedures for introducing specific mutations into nucleic acid sequences. A number of these are described in Ausubel FM 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.

45

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:

- 5 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;
- 10 2. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) directly of the desired core sequence;
3. By mutagenesis (site-specific, PCR, random, spontaneous, etc) of parental DNA sequences, which are recombined to generate a desired core sequence;
4. By reverse transcription of an RNA encoding the desired core sequence; and
- 15 5. By de novo synthesis (chemical synthesis) of a sequence having the desired base changes.

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.

35 2.2 Recombination Sites - Choice of for Assembly of the Multiple Expression Construct

For the method of the invention one or more Insert Donor molecules are provided, which all together comprise at least two Inserts $I(n)$. Each Insert comprises at least one expression cassette. Each Insert is flanked by two different recombination sites $A(2n)$ and $A(2n+1)$, wherein n is an integer from 1 to m characterizing each Insert, and m is the total number of different Inserts. For example Insert I_1 is flanked by recombination sites A_2 and A_3 , Insert I_2 is flanked by recombination sites A_4 and A_5 , and so on.

Further at least one Insert Acceptor molecule IA is provided, comprising two different recombination sites A_1 and $A(2m+2)$, wherein m is the total number of different Inserts. For example, if two different Inserts (I_1 , and I_2) are employed the Insert Acceptor is comprising the different recombination sites A_1 and A_6 .

Now, all of the recombination sites A(1) to A(2m+2) are different. A recombination site A(2i-1) for a specific i, said i being an integer from 1 to m+1, can recombine with the recombination side A(2i) for the same i, but does not substantially recombine with another recombination site. In the above mentioned case for two Inserts (I1 with recombination sites A2 and A3, and I2 with recombination sites A4 and A5) and one Insert Acceptor (with recombination sites A1 and A6), A1 can recombine only with A2, A3 can recombine only with A4, and A5 can recombine only with A6.

The feature, that a recombination site A(2i-1) recombines only with the recombination site A(2i), but does not substantially recombine with another recombination site, is intended to mean that the recombination frequency or velocity between A(2i-1) and A(2i) is at least two times, preferably, 5 times, more preferably 10 times, most preferably 100 times higher than the recombination frequency or velocity of A(2i-1) with any other recombination site beside A(2i) or of A(2i) with any other recombination site beside A(2i-1). The recombination frequency or velocity can – for example – be assessed by the specificity of the recombination reaction, i.e. by the kind and number of unintended byproducts.

By incubating the combination of Insert Donor(s) and Insert Acceptors(s) with at least one site specific recombination protein capable of recombining the recombination sites in said Insert Donor molecules and said Insert Acceptor molecule, the all Inserts are transferred into said Insert Acceptor molecule, thereby producing a Multiple Expression Construct. Only when all Inserts are transferred a circular, replicable Insert Acceptor (e.g., expression vector) can be obtained and the DNA segment formerly between A1 and A(2m+2) is properly replaced.

The following examples are given to facilitate understanding of the choice and combination of the recombination sites and their ability to recombine among each other:

Number of Insert Acceptor (IA)	Number of Inserts (I)	Recombination Sites in IA	Recombination Sites in Inserts	Recombination capability
1	2 (I1, I2)	A1, A6	I1: A2, A3 I2: A4, A5	A1/A2 A3/A4 A5/A6
1	3 (I1, I2, I3)	A1, A8	I1: A2, A3 I2: A4, A5 I3: A6, A7	A1/A2 A3/A4 A5/A6 A7/A8
1	4 (I1, I2, I3, I4)	A1, A10	I1: A2, A3 I2: A4, A5 I3: A6, A7 I4: A8, A9	A1/A2 A3/A4 A5/A6 A7/A8 A9/A10

Table 1: Example for combinations of one Insert Acceptor (Vector Donor) molecule and 2, 3, or 4 Inserts, respectively. Preferably, each Insert is comprised in a separate Insert Donor molecule. Recombination capability is indicating that the two specified recombination sites (e.g., A1 and A2 = A1/A2) can recombine, but that none of these sites can recombine with other recombination sites.

The recombination sites A1 to A(2m+2) can be easily obtained by methods as described e.g., in US 5,888,732, hereby incorporated entirely by reference. Such recombination sites comprises a core region having at least one engineered mutation. Beside

modifying recombination specificity, mutations of the recombination sites may further confer enhancement of recombination, said enhancement selected from the group consisting of substantially (i) favoring integration; (ii) favoring recombination; (ii) relieving the requirement for host factors; (iii) increasing the efficiency of Cointegrate DNA or
 5 Multiple Expression Construct DNA formation; and (iv) increasing the specificity of said Cointegrate DNA or Multiple Expression Construct DNA formation. The nucleic acid molecule constituting the recombination site preferably comprises at least one recombination site derived from attB, attP, attL or attR, such as attR' or attP'. More preferably the att site is selected from att1, att2, or att3, as described herein.

10

In a preferred embodiment, the core region of the recombination site comprises a DNA sequence selected from the group consisting of:

- | | | | | |
|----|-----|------------------------------|-----------|-----------------|
| | (a) | RKYCWGCTTTYKTRTACNAASTSGB | (m-att) | (SEQ ID NO:1); |
| | (b) | AGCCWGCTTTYKTRTACNAACTSGB | (m-attB) | (SEQ ID NO:2); |
| 15 | (c) | G TTCAGCTTTCKTRTACNAACTSGB | (m-attR) | (SEQ ID NO:3); |
| | (d) | AGCCWGCTTTCKTRTACNAAGTSGB | (m-attL) | (SEQ ID NO:4); |
| | (e) | G TTCAGCTTTYKTRTACNAAGTSGB | (m-attP1) | (SEQ ID NO:5); |
| | (f) | RBYCW GCTTTYTTRTACWAA STKGD | (n-att) | (SEQ ID NO:39); |
| | (g) | ASCCW GCTTTYTTRTACWAA STKGW | (n-attB) | (SEQ ID NO:40); |
| 20 | (h) | ASCCW GCTTTYTTRTACWAA GTTGG | (n-attL) | (SEQ ID NO:41); |
| | (i) | G TTCA GCTTTYTTRTACWAA STKGW | (n-attR) | (SEQ ID NO:42); |
| | (j) | G TTCA GCTTTYTTRTACWAA GTTGG | (n-attP) | (SEQ ID NO:43); |

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
 25 T/U, as presented in 37 C.F.R. §1.822 ("Symbols and format to be used for nucleotide and/or amino acid sequence data"), which is entirely incorporated herein by reference, wherein the core region does not contain a stop codon in one or more reading frames.

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

- | | | | | |
|----|-----|----------------------------|------------|-----------------|
| | (a) | AGCCTGCTTTTTTGTACAAACTTGT | (attB1) | (SEQ ID NO:6); |
| | (b) | AGCCTGCTTTCTTGTACAAACTTGT | (attB2) | (SEQ ID NO:7); |
| | (c) | ACCCAGCTTTCTTGTACAAAGTGGT | (attB3) | (SEQ ID NO:8); |
| | (d) | G TTCAGCTTTTTTGTACAAACTTGT | (attR1) | (SEQ ID NO:9); |
| 35 | (e) | G TTCAGCTTTCTTGTACAAACTTGT | (attR2) | (SEQ ID NO:10); |
| | (f) | G TTCAGCTTTCTTGTACAAAGTGGT | (attR3) | (SEQ ID NO:11); |
| | (g) | AGCCTGCTTTTTTGTACAAAGTTGG | (attL1) | (SEQ ID NO:12); |
| | (h) | AGCCTGCTTTCTTGTACAAAGTTGG | (attL2) | (SEQ ID NO:13); |
| | (i) | ACCCAGCTTTCTTGTACAAAGTTGG | (attL3) | (SEQ ID NO:14); |
| 40 | (j) | G TTCAGCTTTTTTGTACAAAGTTGG | (attP1) | (SEQ ID NO:15); |
| | (k) | G TTCAGCTTTCTTGTACAAAGTTGG | (attP2,P3) | (SEQ ID NO:16); |
| | (l) | AGCCTGCTTTTTTGTACAAACTTGC | (attB1*) | (SEQ ID NO:52); |
| | (m) | ACCCAGCTTTCTTGTACAAAGTGGC | (attB2*) | (SEQ ID NO:53); |

- (N) CAACTTTATTATACATAGTTG (attB3*) (SEQ ID NO:54)
 (O) CAACTTTTCTATACAAAGTTG (attB4*) (SEQ ID NO:55)
 (p) GTTCAACTTTTTTTGTACAAACTTGC (attR1*) (SEQ ID NO:56)
 (q) TTCAACTTTTCTTGTACAAAGTGGG (attR2*) (SEQ ID NO:57)
 5 (r) GTTCAACTTTATTATACATAGTTGA (attR3*) (SEQ ID NO:58)
 (s) GTTCAACTTTTCTATACAAAGTTGA (attR4*) (SEQ ID NO:59)
 (t) AGCCTGCTTTTTTTGTACAAAGTTGG (attL1*) (SEQ ID NO:60)
 (u) ACCCAGCTTTTCTTGTACAAAGTTGG (attL2*) (SEQ ID NO:61)
 (v) GGCAACTTTATTATACAAAGTTGG (attL3*) (SEQ ID NO:62)
 10 (w) ACCCAACTTTTCTATACAAAGTTGG (attL4*) (SEQ ID NO:63)
 (x) GTTCAACTTTTTTTGTACAAAGTTGG (att P1*) (SEQ ID NO:64)
 (y) GTTCAGCTTTTCTTGTACAAAGTTGG (attP2*) (SEQ ID NO:65)
 (z) GTTCAACTTTATTATACAAAGTTGG (att P3*) (SEQ ID NO:66)
 (aa) GTTCAACTTTTCTATACAAAGTTGG (attP4*) (SEQ ID NO:67)
 15 or a corresponding or complementary DNA or RNA sequence.

Clearly, various types and permutations of the above specified core regions are considered to be suitable for carrying out the method of the invention and the number of sequences constituting suitable recombination sites of the present invention is virtually
 20 unlimited, 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. Preferably, a recombination site comprises 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
 25 which hybridizes under stringent conditions thereto, as known in the art.

3. The Insert Donor Molecule and the Inserts

The Insert Donor Molecule may comprise one or more Inserts. In an preferred embodiment a Insert Donor Molecule comprises one Insert. The Insert Donor Molecule
 30 may be comprised of a linear or a circular DNA molecule. The Insert Donor molecules may comprise a vector (preferably a circular plasmid vector) or a DNA segment produced by amplification. For vectors any cloning or expression vector may be used. Various examples are known in the art and exemplified below (see "Vector Donor"). Insert Donor Molecules based on cloning vectors have the advantage that they can be
 35 replicated with low error rates avoiding unintended mutations. In case of a circular Insert Donor Molecule said molecule may be supercoiled or relaxed, preferably supercoiled.

However, in another embodiment of the present invention the Insert Donor Molecule
 40 can be a linear DNA sequence. Such sequence can be obtained, for example, by adding adapters comprising the respective recombination sites to a expression cassette e.g., by a polymerase chain reaction (PCR) mediated amplification process. The adapters can be added by using oligonucleotide primers consisting of the recombination site and a sequence characterizing the 5'- or 3'- end of the target expression cassette.
 45

In an preferred embodiment of the invention, the Insert Donor DNA molecule may further comprise a DNA segment encoding for at least one marker selected from the group consisting of a cloning site, a restriction site, a promoter, an operon, an origin of replication, a functional DNA, an antisense RNA, a PCR fragment, a protein and a protein fragment. Preferably, at least one marker may be comprised within at least one Insert. More preferably, said marker is a Selection Marker (preferably a Negative selection Marker) or a Reporter Gene or Protein. Preferred Selection Markers are those which allow for selection of the resulting Multiple Expression Construct (which by incorporation of the Insert comprises said Selection Marker). Said Selectable Marker may preferably comprise at least one marker selected from the group consisting of an antibiotic resistance gene, a herbicide resistance gene, a tRNA gene, an auxotrophic marker, a toxic gene, a phenotypic marker, an antisense oligonucleotide, a restriction endonuclease, a restriction endonuclease cleavage site, an enzyme cleavage site, a protein binding site, and a sequence complementary to a PCR primer sequence. Suitable Selection Marker may comprise at least one DNA segment selected from the group consisting of:

- (a) a DNA segment that encodes a product that provides resistance in a recipient cell or organism against otherwise toxic compounds ("Negative Selection Marker");
- (b) a DNA segment that encodes a product that is toxic in a recipient cell or organism ("Counter Selection Marker");
- (c) a DNA segment that encodes a product conferring to the recipient cell or organism a growth or proliferation advantage ("Positive Selection Marker");
- (d) a DNA segment that encodes a product that can be identified;
- (e) a DNA segment that encodes a product that inhibits a cell function in a recipient cell;
- (f) a DNA segment that inhibits the activity of any of the DNA segments of (a)-(e) above;
- (g) a DNA segment that binds a product that modifies a substrate;
- (h) a DNA segment that encodes a specific nucleotide recognition sequence which can be recognized by a protein, an RNA, DNA or chemical,
- (i) a DNA segment that, when deleted, directly or indirectly confers sensitivity to cell killing by particular compounds within a recipient cell;
- (j) a DNA segment that encodes a product that suppresses the activity of a gene product in a recipient cell;
- (k) a DNA segment that encodes a product that is otherwise lacking in a recipient cell, and;
- (l) a DNA segment that can be used to isolate or identify a desired molecule.

Specific examples are given above.

The method of the invention may further comprise the step of selecting the Multiple Expression Construct comprising all of said Inserts of said Insert Donor molecules.

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.

5 **4. The Insert Acceptor Molecule**

In an preferred embodiment the Insert Acceptor molecule comprises a DNA segment flanked by said two different recombination sites A1 and A(2m+2) which is going to be replaced during the recombination process.

10 The Insert Acceptor Molecule may be comprised of a linear or a circular DNA molecule. The Insert Acceptor molecules may comprise a vector. In this case the Insert Acceptor Molecule is becoming a Vector Donor Molecule ("preferably a circular plasmid vector). For vectors any cloning or expression vector may be used. Various examples are known in the art and exemplified below. Insert Acceptor Molecules based on cloning
15 vectors have the advantage that they can be replicated with low error rates avoiding unintended mutations. It is preferred that the Vector Donor Molecule is a vector which can be employed for transformation of the selected target organism with the Multiple Expression Construct. For example, for plant cells or organisms an Agrobacterium binary vector can be used as base for a Vector Donor Molecule (see below for details). In
20 case of a circular Insert Acceptor Molecule (Vector Donor Molecule) said molecule may be supercoiled or relaxed, preferably relaxed.

However, in another embodiment of the present invention the Insert Acceptor Molecule can be a linear DNA sequence. Such sequence can be obtained, for example, by add-
25 ing adapters comprising the respective recombination sites to a expression cassette e.g., by a polymerase chain reaction (PCR) mediated amplification process. The adapters can be added by using oligonucleotide primers consisting of the recombination site and a sequence characterizing the 5'- or 3'- end of the target expression cassette.

30 In another preferred embodiment of the invention the Insert Acceptor Molecule is a genomic DNA molecule, for example a chromosomal or plastidic DNA molecule. Thereby, the Multiple Expression Construct can be assembled directly into the genome of the host organism (e.g., by co-transformation of the Insert Donor molecules). Cells or
35 organisms comprising suitable recombination site in the genomic DNA (thereby providing a Insert Acceptor Molecule) can be generated – for example - in a precedent step by standard transformation techniques inserting into the genomic DNA a DNA construct comprising said recombination sites thereby providing a master cell or organism in which various Multiple Expression Construct can be assembled.

40 Preferably, the Insert Acceptor or Vector Donor Molecule comprise at least one selectable marker. More preferred, the Insert Acceptor or Vector Donor Molecule further comprises (a) a toxic gene and (b) a Selectable Marker, wherein said toxic gene and said Selectable Marker are on different DNA segments, the DNA segments being separated from each other by at least two recombination sites. Preferably, the toxic gene is
45 deleted from the Insert Acceptor or Vector Donor Molecule in consequence of the recombinational process.

Said Selectable Marker may preferably comprise at least one marker selected from the group consisting of an antibiotic resistance gene, a herbicide resistance gene, a tRNA gene, an auxotrophic marker, a toxic gene, a phenotypic marker, an antisense oligonucleotide, a restriction endonuclease, a restriction endonuclease cleavage site, an enzyme cleavage site, a protein binding site, and a sequence complementary to a PCR primer sequence.

The Multiple Expression Construct according to the invention can advantageously be introduced into cells, preferably into plant cells, using vectors. In an advantageous embodiment, the Multiple Expression Construct is a vector or is comprised in a vector or inserted into a vector, preferably selected from prokaryotic and/or eukaryotic vectors. Thus, in an preferred embodiment of the invention the Insert Acceptor is a Vector Donor, providing the relevant vector sequences. In one embodiment, the methods of the invention involve transformation of organism or cells (e.g. plants or plant cells) with a transgenic expression vector comprising the Multiple Expression Construct of the invention (as described above). As used herein, the terms "vector" and "vehicle" are used interchangeably in reference to nucleic acid molecules that transfer DNA segment(s) from one cell to another. The term "expression vector" as used herein refers to a recombinant DNA molecule containing a desired coding sequence and appropriate nucleic acid sequences necessary for the expression of the operably linked coding sequence in a particular host organism.

The methods of the invention are not limited to the expression vectors disclosed herein. Any expression vector which is capable of introducing a nucleic acid sequence of interest into a plant cell is contemplated to be within the scope of this invention. Typically, expression vectors comprise the transgenic expression cassette of the invention in combination with elements which allow cloning of the vector into a bacterial or phage host. The vector preferably, though not necessarily, contains an origin of replication which is functional in a broad range of prokaryotic hosts. A selectable marker is generally, but not necessarily, included to allow selection of cells bearing the desired vector.

Examples of vectors may be plasmids, cosmids, phages, viruses or *Agrobacteria*. More specific examples are given below for the individual transformation technologies.

Preferred are those vectors which make possible a stable integration of the expression construct into the host genome. In the case of injection or electroporation of DNA into cells (e.g., plant cells), the plasmid used need not meet any particular requirements. Simple plasmids such as those of the pUC series can be used. If intact plants are to be regenerated from the transformed cells, it is necessary for an additional selectable marker gene to be present on the plasmid. A variety of possible plasmid vectors are available for the introduction of foreign genes into plants, and these plasmid vectors contain, as a rule, a replication origin for multiplication in *E.coli* and a marker gene for the selection of transformed bacteria. Examples are pBR322, pUC series, M13mp series, pACYC184 and the like.

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 prokaryotic vectors, eukaryotic vectors or vectors which may shuttle

- between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred eukaryotic vectors comprise vectors, which replicate in yeast cells, plant cells, fish cells, eukaryotic cells, mammalian cells, or insect cells. Preferred prokaryotic vectors comprise vectors which replicate in gram negative and/or gram positive bacteria, more preferably vectors which replicate in bacteria of the genus *Escherichia*, *Salmonella*, *Bacillus*, *Streptomyces*, *Agrobacterium*, *Rhizobium*, or *Pseudomonas*. Most preferred are vectors which replicates in both *E. coli* and *Agrobacterium*. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate and yeast cells, plant cells, mammalian cells (particularly human cells), fungal cells, insect cells, fish cells and the like. 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 prokaryotic and/or eukaryotic systems depending on the particular vector.
- 15 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, Invitrogen, Vector Laboratories Inc., InVitrogen, Promega, Novagen, NEB, Clontech, Boehringer Mannheim, 20 Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, PerkinElmer, 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 25 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.
- 30 Other vectors of interest include viral origin vectors (M13 vectors, bacterial phage 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 (pCDM8).
- 35 Particular vectors of interest include prokaryotic expression vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Inc.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZZ18, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Life Technologies, Inc.) and variants and derivatives thereof Vector donors can also be made from 40 eukaryotic expression vectors such as pFastBac, pFastBacHT, pFastBacDUAL, pSFV, and pTet-Splice (Life Technologies, Inc.), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, pMC1neo, and 45 pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A, B, and C, pVL1392, pBsueBacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Inc.) and variants or derivatives thereof.

Other vectors of particular interest include pUC18, pUC19, pBlueScript, pSPORT, cosmids, phagemids, YAC's (yeast artificial chromosomes), BAC's (bacterial artificial chromosomes), P1 (E. coli phage), pQE70, pQE60, pQE9 (quagan), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (Invitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SPORT1 (Life Technologies, Inc.) and variants or derivatives thereof

Additional vectors of interest include pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2, pcDNA3.1/His, pcDNA3.1(-)/Myc-His, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ α , pGAPZ, pGAPZ α , pBlueBac4.5, pBlueBacHis2, pMelBac, pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZErO1.1, pZErO-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSe, SV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP 10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen; λ ExCell, λ gt11, pTrc99A, pKK223-3, pGEX-1 λ T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZZ18, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTAg, pET-32LIC, pET-30LIC, pBAC-2cp LIC, pBACgus-2cp LIC, pT7Blue-2 LIC, pT7Blue-2, λ SCREEN-1, λ BlueSTAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta-Hyg, and Selecta Vecta-Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1, pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP, pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPuv, pGFP, p6xHis-GFP, pSEAP2-Basic, pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p β gal-Basic, p β gal-Control, p β gal-Promoter, p β gal-Enhancer, pCMV β , pTet-Off, pTet-On, pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX, pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo, pYEX4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6, pTriplEx, λ gt10, λ gt11, pWE15, and $\square$ TriplEx from Clontech; Lambda ZAP II, pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4, pBD-GAL4 Cam, pSurfsript, Lambda FIX II, Lambda DASH, Lambda EMBL3, Lambda EMBL4, SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n, pCAL-c, pCAL-kc, pET-3abcd, pET-11'abcd, pSPUTK, pESP-1, pCMVLacI, pOPRSVI/MCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMC1neo, pMC1neo Poly A, pOG44, pOG45, pFRT β GAL, pNEO β GAL, pRS403, pRS404, pRS405, pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene.

Two-hybrid and reverse two-hybrid vectors of particular interest include pPC86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pAct, pACT2, pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4, pLexA, pBD-GAL4,

pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202, pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof.

5 Preferred vectors for expression in *E. coli* are pQE70, pQE60 und pQE-9 (QIAGEN, Inc.); pBluescript vektors, phagescript vektors, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene Cloning Systems, Inc.); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia Biotech, Inc.).

10 Preferred vectors for expression in eukaryotic, animals systems comprise pWLNE0, pSV2CAT, pOG44, pXT1 and pSG (Stratagene Inc.); pSVK3, pBPV, pMSG und pSVL (Pharmacia Biotech, Inc.). Examples for inducible vectors are pTet-tTak, pTet-Splice, pcDNA4/TO, pcDNA4/TO/LacZ, pcDNA6/TR, pcDNA4/TO/Myc-His/LacZ, pcDNA4/TO/Myc-His A, pcDNA4/TO/Myc-His B, pcDNA4/TO/Myc-His C, pVgRXR (Invitrogen, Inc.) or the pMAM-Serie (Clontech, Inc.; GenBank Accession No.: U02443).

15 Preferred vectors for the expression in yeast comprise for example pYES2, pYD1, pTEFI/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, PHIL-D2, PHIL-SI, pPIC3SK, pPIC9K, and PA0815 (Invitrogen, Inc.).

20 Preferred vector for plant transformation are described herein below and preferably comprise vectors for *Agrobacterium* mediated transformation. *Agrobacterium tumefaciens* and *A. rhizogenes* are plant-pathogenic soil bacteria, which genetically transform plant cells. The Ti and Ri plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of the plant (Kado (1991) Crit Rev
25 Plant Sci 10:1). Vectors of the invention may be based on the *Agrobacterium* Ti- or Ri-plasmid and may thereby utilize a natural system of DNA transfer into the plant genome.

30 As part of this highly developed parasitism *Agrobacterium* transfers a defined part of its genomic information (the T-DNA; flanked by about 25 bp repeats, named left and right border) into the chromosomal DNA of the plant cell (Zupan et al. (2000) Plant J 23(1):11-28). By combined action of the so-called *vir* genes (part of the original Ti-plasmids) said DNA-transfer is mediated. For utilization of this natural system, Ti-plasmids were developed which lack the original tumor inducing genes ("disarmed vectors"). In a further improvement, the so called "binary vector systems", the T-DNA was
35 physically separated from the other functional elements of the Ti-plasmid (e.g., the *vir* genes), by being incorporated into a shuttle vector, which allowed easier handling (EP-A 120 516; US 4.940.838). These binary vectors comprise (beside the disarmed T-DNA with its border sequences), prokaryotic sequences for replication both in *Agrobacterium* and *E. coli*. It is an advantage of *Agrobacterium*-mediated transformation that in
40 general only the DNA flanked by the borders is transferred into the genome and that preferentially only one copy is inserted. Descriptions of *Agrobacterium* vector systems and methods for *Agrobacterium*-mediated gene transfer are known in the art (Miki et al. (1993) "Procedures for Introducing Foreign DNA into Plants" in METHODS IN PLANT
45 MOLECULAR BIOLOGY AND BIOTECHNOLOGY; pp.67-88; Gruber et al. (1993) "Vectors for Plant Transformation," in METHODS IN PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY; pp.89-119; Moloney et al. (1989) Plant Cell Reports 8: 238-242). The use of T-DNA for the transformation of plant cells has been studied and de-

scribed intensively (EP 120516; Hoekema (1985) In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasterdam, Chapter V; Fraley 1985; and An et al. (1985) EMBO J 4:277-287). Various binary vectors are known, some of which are commercially available such as, for example, pBIN19 (Clontech Laboratories, Inc. U.S.A.).

Hence, for *Agrobacterium*-mediated transformation the Multiple Expression Construct is integrated into specific plasmids, either into a shuttle or intermediate vector, or into a binary vector. If a Ti or Ri plasmid is to be used for the transformation, at least the right border, but in most cases the right and left border, of the Ti or Ri plasmid T-DNA is linked to the transgenic expression construct to be introduced in the form of a flanking region. Binary vectors are preferably used. Binary vectors are capable of replication both in *E.coli* and in *Agrobacterium*. They may comprise a selection marker gene and a linker or polylinker (for insertion of e.g. the expression construct to be transferred) flanked by the right and left T-DNA border sequence. They can be transferred directly into *Agrobacterium* (Holsters et al. (1978) Mol Gen Genet 163:181-187). The selection marker gene permits the selection of transformed *Agrobacterium* and is, for example, the *nptII* gene, which confers resistance to kanamycin. The *Agrobacterium* which acts as host organism in this case should already contain a plasmid with the *vir* region. The latter is required for transferring the T-DNA to the plant cell. An *Agrobacterium* transformed in this way can be used for transforming plant cells. The use of T-DNA for transforming plant cells has been studied and described intensively (EP 120 516; Hoekema (1985) In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasterdam, Chapter V; An et al. (1985) EMBO J 4:277-287; see also below).

Common binary vectors are based on "broad host range"-plasmids like pRK252 (Bevan et al. (1984) Nucl Acid Res 12,8711-8720) or pTJS75 (Watson et al. (1985) EMBO J 4(2):277- 284) derived from the P-type plasmid RK2. Most of these vectors are derivatives of pBIN19 (Bevan et al. (1984) Nucl Acid Res 12,8711-8720). Various binary vectors are known, some of which are commercially available such as, for example, pBI101.2 or pBIN19 (Clontech Laboratories, Inc. USA). Additional vectors were improved with regard to size and handling (e.g. pPZP; Hajdukiewicz et al. (1994) Plant Mol Biol 25:989-994). Improved vector systems are described also in WO 02/00900.

In a preferred embodiment, *Agrobacterium* strains for use in the practice of the invention include octopine strains, e.g., LBA4404 or agropine strains, e.g., EHA101 or EHA105. Suitable strains of *A. tumefaciens* for DNA transfer are for example EHA101pEHA101 (Hood et al. (1986) J Bacteriol 168:1291-1301), EHA105[pEHA105] (Li et al. (1992) Plant Mol Biol 20:1037-1048), LBA4404[pAL4404] (Hoekema et al. (1983) Nature 303:179-181), C58C1[pMP90] (Koncz & Schell (1986) Mol Gen Genet 204:383-396), and C58C1[pGV2260] (Deblaere et al. (1985) Nucl Acids Res 13:4777-4788). Other suitable strains are *Agrobacterium tumefaciens* C58, a nopaline strain. Other suitable strains are *A. tumefaciens* C58C1 (Van Laerebeke et al. (1974) Nature 252,169-170), A136 (Watson et al. (1975) J. Bacteriol 123, 255-264) or LBA4011 (Klapwijk et al. (1980) J. Bacteriol., 141,128-136). In a preferred embodiment, the *Agrobacterium* strain used to transform the plant tissue pre-cultured with the plant phenolic compound contains a L,L-succinamopine type Ti-plasmid, preferably disarmed, such as pEHA101. In another preferred embodiment, the *Agrobacterium* strain used to

transform the plant tissue pre-cultured with the plant phenolic compound contains an octopine-type Ti-plasmid, preferably disarmed, such as pAL4404. Generally, when using octopine-type Ti-plasmids or helper plasmids, it is preferred that the *virF* gene be deleted or inactivated. In a preferred embodiment, the *Agrobacterium* strain used to

5 transform the plant tissue pre-cultured with the plant phenolic compound such as acetosyringone. The method of the invention can also be used in combination with particular *Agrobacterium* strains, to further increase the transformation efficiency, such as *Agrobacterium* strains wherein the *vir* gene expression and/or induction thereof is altered due to the presence of mutant or chimeric *virA* or *virG* genes (e.g. Hansen et al.

10 (1994) Proc. Natl. Acad. Sci. USA 91:7603-7607; Chen and Winans (1991) J. Bacteriol. 173: 1139-1144; Scheeren-Groot et al. (1994) J. Bacteriol 176: 6418-6426).

A binary vector or any other vector can be modified by common DNA recombination techniques, multiplied in *E. coli*, and introduced into *Agrobacterium* by e.g., electroporation or other transformation techniques (Mozo & Hooykaas, Plant Mol. Biol. 16

15 (1991), 917-918). *Agrobacterium* is grown and used as described in the art. The vector comprising *Agrobacterium* strain may, for example, be grown for 3 days on YP medium (5 g/L yeast extract, 10 g/L peptone, 5 g/L Nail, 15 g/L agar, pH 6.8) supplemented with the appropriate antibiotic (e.g., 50 mg/L spectinomycin). Bacteria are collected with a

20 loop from the solid medium and resuspended. For the purpose of this invention, *Agrobacterium* compatible vectors are provided by inserting site-specific recombination sites as described – for example – in the Examples.

After constructing a vector (e.g., a Vector Donor Molecule or a Insert Donor Molecule

25 or a Multiple Expression Construct constituting a vector), the vector can be propagated in a host cell to synthesize nucleic acid molecules for the generation of a nucleic acid polymer. Vectors, often referred to as "shuttle vectors," are capable of replicating in at least two unrelated expression systems. To facilitate such replication, the vector should include at least two origins of replication, one effective in each replication system. Typi-

30 cally, shuttle vectors are capable of replicating in a eukaryotic system and a prokaryotic system. This enables detection of protein expression in eukaryotic hosts, the "expression cell type," and the amplification of the vector in the prokaryotic hosts, the "amplification cell type." As an illustration, one origin of replication can be derived from SV40, while another origin of replication can be derived from pBR322. Those of skill in the art

35 know of numerous suitable origins of replication.

After constructing a vector (e.g., a Vector Donor Molecule or a Insert Donor Molecule or a Multiple Expression Construct constituting a vector), the vector is typically propagated in a host cell. Vector propagation is conveniently carried out in a prokaryotic host

40 cell, such as *E. coli* or *Bacillus subtilis*. Suitable strains of *E. coli* include BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pLysE, DB2, DB3.1, DH1, DH4I, DH5, DH5I, DH5IF, DH5IMCR, DH10B, DH10B/p3, DH11S, C600, HB101, JM101, JM105, JM109, JM110, K38, RR1, Y1088, Y1089, CSH18, ER1451, and ER1647 (see, for example, Brown (ed.), Molecular Biology Labfax (Academic Press 1991)). Suitable strains of *Bacillus*

45 *subtilis* include BR151, YB886, MI119, MI120, and B170 (see, for example, Hardy, "Bacillus Cloning Methods," in DNA Cloning: A Practical Approach, Glover (ed.) (IRL Press 1985)). Standard techniques for propagating vectors in prokaryotic hosts are well-known to those of skill in the art (see, for example, Ausubel et al. (eds.), Short Pro-

protocols in Molecular Biology, 3rd Edition (John Wiley & Sons 1995) ["Ausubel 1995"]; Wu et al., Methods in Gene Biotechnology (CRC Press, Inc. 1997)).

5. Recombination Schemes – Selection of the Multiple Expression Construct

5 One general scheme for an in vitro or in vivo method of the invention is shown in FIG. 6 and 7, where the Insert Donor and the Insert Acceptor (e.g., a Vector Donor) can be either circular or linear DNA, but is shown as linear for the Insert Donor and circular for the Insert Acceptor (Vector Donor). The Multiple Expression Construct is introduced (or assembled) into the Insert Acceptor or Vector Donor via the recombinational cloning
10 procedure of the invention.

The resulting Multiple Expression Construct molecules may optionally be selected or isolated away from other molecules such as unreacted Insert Acceptor molecules or Insert Donor molecules or other unintended by-products. Such selection is preferably
15 realized in vivo (e.g., in *E.coli*) employing one or more Selection Marker comprised in the Inserts or the Vector Donor, which in consequence of the recombinational cloning procedure become introduced or deleted from the final Multiple Expression Construct.

20 In an preferred embodiment of the invention, the resulting construct (e.g., a plasmid) may first be introduced into *E.coli* for further selection and amplification. Correctly transformed *E.coli* are selected and grown, and the recombinant construct is obtained by methods known to the skilled worker. Restriction analysis and sequencing can be used for verifying the cloning step.

25 The fragment flanked by the recombination sites in the Insert Acceptor or Vector Donor Molecule (which may comprise a repression cassette or a Counter Selection Marker) is exchanged for the cluster of the Inserts comprised in the Insert Donor. The method of the invention allows the Inserts to be transferred into any number of vectors. The Inserts may be transferred to a particular Vector or may be transferred to a number of
30 vectors in one step. Additionally, the cluster of Inserts in the resulting Multiple Expression Construct may be transferred from said resulting Multiple Expression Construct to any number of vectors sequentially.

35 In accordance with the invention, it is desirable to select for the resulting Multiple Expression Construct against other molecules (including unreacted parental molecules, by-products or intermediate Cointegrate(s)). For this purpose, in an preferred embodiment at least one of the Inserts and/or the region flanked by the recombination sites in the Insert Acceptor comprises at least one Selection Marker, expression signals, origins of replication, or specialized functions for detecting, selecting, expressing, map-
40 ping or sequencing DNA.

A variety of other selection schemes can be used that are known in the art which may be employed to select an Multiple Expression Construct of the invention. Depending upon individual preferences and needs (e.g., target host species), a number of different
45 types of selection schemes can be used. 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 encodes 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.

5 The term "selection scheme" is referring to any method which allows selection, enrichment, or identification of a desired Multiple Expression Construct or Multiple Expression Construct(s) from a mixture containing the Insert Donor, Vector Donor, any intermediates (e.g. a Cointegrate), and/or Byproducts. Various methods are known to the person skilled in the art and – for example – described in US 5,888,732. A preferred requirement is that the selection scheme results in selection of or enrichment for only one or
10 more desired Multiple Expression Constructs. 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.

15 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. 6 and 7. The first, exemplified herein with a Negative Selection Marker (SN) and a Counter Selection Marker (e.g., a toxic gene product; SC), selects for Multiple Expression Construct molecules comprising the Inserts and lacking the segment flanked in the parental Insert
20 Acceptor by the recombination sites. 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".)

25 Examples of such toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (e.g., DpnI), 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 ϕ X174 or bacteriophage T4; antibiotic sensitivity genes such as rpsL, antimicrobial
30 sensitivity genes such as pheS, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-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.

35 In the second form, one or more of the Inserts are comprising an additional expression cassette for a negative selection marker (e.g., which confers resistance against a antibiotic, herbicide, or other biozide) or a positive selection marker (which confers a growth advantage). Additional examples include but are not limited to:

- 40 (i) Generation of new primer sites for PCR (e.g., juxtaposition of two DNA sequences that were not previously juxtaposed);
- (ii) Inclusion of a DNA sequence acted upon by a restriction endonuclease or other DNA modifying enzyme, chemical, ribozyme, etc.;
- 45 (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;

- (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;
 - (vi) 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;
 - (vii) Selection of the desired product by size (e.g., fractionation) or other physical property of the molecule(s); and
 - (viii) Inclusion of a DNA sequence required for a specific modification (e.g., methylation) that allows its identification.
- After formation of the Multiple Expression Construct and Byproduct in the method of the present invention, the selection step can be carried out either in vitro or in vivo depending upon the particular selection scheme which has been optionally devised in the particular recombinational cloning procedure. For example, an in vitro method of selection can be devised for the Insert Donor and Insert Acceptor (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.

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 Multiple Expression Construct. 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 Multiple Expression Construct molecule is amplified.

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 Insert Acceptor (Vector Donor) molecule, and a drug marker controlled by that repressor on the other segment of the same plasmid. Another is to put a killer gene in the region flanked by recombination sites in the Insert Acceptor molecule (FIG. 6 and 7). Of course a way must exist for growing such a Insert Acceptor molecule (e.g., a plasmid vector), 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 DpnI, which will not cleave unless its recognition sequence GATC is methylated. Many popular common E. coli strains methylate GATC sequences, but there are mutants in which cloned DpnI 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 gene encoding the corresponding methylase gene provides protected host for use in the invention. Similarly, the ccdB protein is a potent poison of DNA gy-

rase, efficiently trapping gyrase molecules in a cleavable 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 *ccdB*. 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, Ill. 61604 USA as deposit number NRRL B-21852.

10

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 Multiple Expression Construct in eukaryotic cells.

15

6. Preferred Expression Cassettes of the Inserts

Preferably, an Insert of the invention comprises at least one expression cassette, which allows for expression of a nucleic acid of interest and may – for example – facilitate expression of selection marker gene, trait genes, antisense RNA or double-stranded RNA. Preferably said expression cassettes comprise a promoter sequence functional in the targeted host cell or organism (preferably plant cells) operatively linked to a nucleic acid sequence which – upon expression – confers an advantageous phenotype to the so cell or organism. The person skilled in the art is aware of numerous sequences which may be utilized in this context. For plants sequences may be employed e.g. to increase quality of food and feed, to produce chemicals, fine chemicals or pharmaceuticals (e.g., vitamins, oils, carbohydrates; Dunwell JM (2000) J Exp Bot 51 Spec No:487-96), conferring resistance to herbicides, or conferring male sterility. Furthermore, growth, yield, and resistance against abiotic and biotic stress factors (like e.g., fungi, viruses or insects) may be enhanced. Advantageous properties may be conferred either by overexpressing proteins or by decreasing expression of endogenous proteins by e.g., expressing a corresponding antisense (Sheehy et al. (1988) Proc Natl Acad Sci USA 85: 8805-8809; US 4,801,340; Mol et al. (1990) FEBS Lett 268(2):427-430) or double-stranded RNA (Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; Waterhouse et al. (1998) Proc Natl Acad Sci USA 95:13959-64; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364).

25

30

35

40

45

It is an advantage of the current invention that by choice of the recombination sites not only an unambiguous Multiple Expression Construct is derived (comprising all of the intended expression cassettes), but that also the orientation of the individual expression cassettes in said Multiple Expression Construct is pre-determined and can be arranged as needed. This is of advantage to avoid unintended effects (like e.g., gene silencing) which may occur for certain orientations of multiple expression cassettes.

6.1 Preferred Promoter

To express a gene, a nucleic acid molecule of interest to be expressed must be operably linked to regulatory sequences that control transcriptional expression and then, introduced into a host cell. In addition to transcriptional regulatory sequences, such as
 5 promoters and enhancers, expression vectors can include transcriptional and translational regulatory sequences (see above for details under "Definitions").

As an illustration, the transcriptional and translational regulatory signals suitable for a mammalian host may be derived from viral sources, such as adenovirus, bovine papilloma virus, simian virus, or the like, in which the regulatory signals are associated with
 10 a particular gene that has a high level of expression. Suitable transcriptional and translational regulatory sequences also can be obtained from mammalian genes, such as actin, collagen, myosin, and metallothionein genes.

Suitable transcriptional regulatory sequences include a promoter region sufficient to direct the initiation of RNA synthesis. Illustrative eukaryotic promoters include the promoter of the mouse metallothionein I gene (Hamer et al., J. Molec. Appl. Genet. 1:273 (1982)), the TK promoter of Herpes virus (McKnight, Cell 31:355 (1982)), the SV40 early promoter (Benoist et al., Nature 290:304 (1981)), the Rous sarcoma virus promoter (Gorman et al., Proc. Nat'l Acad. Sci. USA 79:6777 (1982)), the cytomegalovirus promoter (Foecking et al., Gene 45:101 (1980)), and the mouse mammary tumor virus promoter (see, generally, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in Protein Engineering: Principles and Practice, Cleland et al. (eds.), pages 163-181 (John Wiley & Sons, Inc. 1996)). Alternatively, a prokaryotic promoter,
 20 such as the bacteriophage T3 RNA polymerase promoter, can be used to control expression of the gene of interest in mammalian cells if the prokaryotic promoter is regulated by a eukaryotic promoter (Zhou et al., Mol. Cell. Biol. 10:4529 (1990), and Kaufman et al., Nucl. Acids Res. 19:4485 (1991)).

For expression in plants, plant-specific promoters are preferred. The term "plant-specific promoter" is understood as meaning, in principle, any promoter which is capable of governing the expression of genes, in particular foreign genes, in plants or plant parts, plant cells, plant tissues or plant cultures. In this context, expression can be, for example, constitutive, inducible or development-dependent. The following are preferred:
 35

a) Constitutive promoters
 "Constitutive" promoters refers to those promoters which ensure expression in a large number of, preferably all, tissues over a substantial period of plant development, preferably at all times during plant development. A plant promoter or promoter originating from a plant virus is especially preferably used. The promoter of the CaMV (cauliflower mosaic virus) 35S transcript (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228) or the 19S CaMV promoter (US 5,352,605; WO
 40 84/02913) are especially preferred. Another suitable constitutive promoter is the rice actin promoter (McElroy et al., Plant Cell 2: 163171 (1990)), Rubisco small subunit (SSU) promoter (US 4,962,028), the legumin B promoter (GenBank Acc. No. X03677), the promoter of the nopaline synthase from Agrobacterium, the TR dual promoter, the
 45

OCS (octopine synthase) promoter from *Agrobacterium*, the ubiquitin promoter (Holtorf S et al. (1995) *Plant Mol Biol* 29:637-649), the ubiquitin 1 promoter (Christensen et al. (1989) *Plant Mol. Biol.* 12: 619-632, Christensen et al. (1992) *Plant Mol Biol* 18:675-689; Bruce et al. (1989) *Proc Natl Acad Sci USA* 86:9692-9696), the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (US 5,683,439), the promoters of the vacuolar ATPase subunits, the pEMU promoter (Last DI et al. (1991) *Theor. Appl. Genet.* 81, 581-588); the MAS promoter (Velten J et al. (1984) *EMBO J.* 3(12): 2723-2730) and maize H3 histone promoter (Lepetit et al., *Mol. Gen. Genet.* 231: 276-285 (1992); Atanassova et al. (1992) *Plant J* 2(3): 291-300), the promoter of the *Arabidopsis thaliana* nitrilase-1 gene (GenBank Acc. No.: U38846, nucleotides 3862 to 5325 or else 5342) or the promoter of a proline-rich protein from wheat (WO 91/13991), and further promoters of genes whose constitutive expression in plants.

b) Tissue-specific or tissue-preferred promoters

Furthermore preferred are promoters with specificities for seeds, such as, for example, the phaseolin promoter (US 5,504,200; Bustos et al. (1989) *Plant Cell* 1(9):839-53, Murai et al., *Science* 23: 476-482 (1983); Sengupta-Gopalan et al., *Proc. Nat'l Acad. Sci. USA* 82: 3320-3324 (1985)), the promoter of the 2S albumin gene (Joseffson LG et al. (1987) *J Biol Chem* 262:12196-12201), the legumine promoter (Shirsat A et al. (1989) *Mol Gen Genet* 215(2):326-331), the USP (unknown seed protein) promoter (Bäumlein et al. (1991a) *Mol Gen Genet* 225(3):459-467), the napin gene promoter (US 5,608,152; Stalberg K et al. (1996) *Planta* 199:515-519), the promoter of the sucrose binding proteins (WO 00/26388) or the legumin B4 promoter (LeB4; Bäumlein H et al. (1991b) *Mol Gen Genet* 225:121-128, 1992), the *Arabidopsis* oleosin promoter (WO 98/45461), and the *Brassica* Bce4 promoter (WO 91/13980). Further preferred are a leaf-specific and light-induced promoter such as that from cab or Rubisco (Simpson et al. (1985) *EMBO J* 4:2723-2729; Timko et al. (1985) *Nature* 318: 579-582); an anther-specific promoter such as that from LAT52 (Twell et al. (1989b) *Mol Gen Genet* 217:240-245); a pollen-specific promoter such as that from Zml3 (Guerrero et al. (1993) *Mol Gen Genet* 224:161-168); and a microspore-preferred promoter such as that from apg (Twell et al. (1983) *Sex. Plant Reprod.* 6: 217-224).

c) Chemically inducible promoters

The expression cassettes may also contain a chemically inducible promoter (review article: Gatz et al. (1997) *Annu Rev Plant Physiol Plant Mol Biol* 48:89-108), by means of which the expression of the exogenous gene in the plant can be controlled at a particular point in time. Such promoters such as, for example, the PRP1 promoter (Ward et al. 1993), a salicylic acid-inducible promoter (WO 95/19443), a benzenesulfonamide-inducible promoter (EP 0 388 186), a tetracyclin-inducible promoter (Gatz et al. (1991) *Mol Gen Genetics* 227:229-237; Gatz et al. (1992) *Plant J* 2:397-404), an abscisic acid-inducible promoter EP 0 335 528) or an ethanol-cyclohexanone-inducible promoter (WO 93/21334) can likewise be used. Also suitable is the promoter of the glutathione-S transferase isoform II gene (GST-II-27), which can be activated by exogenously applied safeners such as, for example, N,N-diallyl-2,2-dichloroacetamide (WO 93/01294) and which is operable in a large number of tissues of both monocots and dicots. Further exemplary inducible promoters that can be utilized in the instant invention include that from the ACE1 system which responds to copper (Mett et al. *PNAS* 90: 4567-4571 (1993)); or the In2 promoter from maize which responds to benzenesulfonamide herbi-

cide safeners (Hershey et al. (1991) *Mol Gen Genetics* 227:229-237; Gatz et al. (1994) *Mol Gen Genetics* 243:32-38). A promoter that responds to an inducing agent to which plants do not normally respond can be utilized. An exemplary inducible promoter is the inducible promoter from a steroid hormone gene, the transcriptional activity of which is induced by a glucocorticosteroid hormone (Schena et al. (1991) *Proc Nat'l Acad Sci USA* 88:10421).

Particularly preferred are constitutive promoters. Furthermore, further promoters may be linked operably to the nucleic acid sequence to be expressed, which promoters make possible the expression in further plant tissues or in other organisms, such as, for example, *E. coli* bacteria. Suitable plant promoters are, in principle, all of the above-described promoters.

The genetic component and/or the expression cassette may comprise further genetic control sequences in addition to a promoter. The term "genetic control sequences" is to be understood in the broad sense and refers to all those sequences which have an effect on the materialization or the function of the expression cassette according to the invention. For example, genetic control sequences modify the transcription and translation in prokaryotic or eukaryotic organisms. Preferably, the expression cassettes according to the invention encompass a promoter functional in plants 5'-upstream of the nucleic acid sequence in question to be expressed recombinantly, and 3'-downstream a terminator sequence as additional genetic control sequence and, if appropriate, further customary regulatory elements, in each case linked operably to the nucleic acid sequence to be expressed recombinantly.

Genetic control sequences furthermore also encompass the 5'-untranslated regions, introns or noncoding 3'-region of genes, such as, for example, the actin-1 intron, or the Adh1-S introns 1, 2 and 6 (general reference: *The Maize Handbook*, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994)). It has been demonstrated that they may play a significant role in the regulation of gene expression. Thus, it has been demonstrated that 5'-untranslated sequences can enhance the transient expression of heterologous genes. Examples of translation enhancers which may be mentioned are the tobacco mosaic virus 5'-leader sequence (Gallie et al. (1987) *Nucl Acids Res* 15:8693-8711) and the like. Furthermore, they may promote tissue specificity (Rouster J et al. (1998) *Plant J* 15:435-440).

The expression cassette may advantageously comprise one or more enhancer sequences, linked operably to the promoter, which make possible an increased recombinant expression of the nucleic acid sequence. Additional advantageous sequences, such as further regulatory elements or terminators, may also be inserted at the 3'-end of the nucleic acid sequences to be expressed recombinantly.

The expression cassette can also include a transcription termination sequence, and optionally, a polyadenylation signal sequence. Polyadenylation signals which are suitable as control sequences are plant polyadenylation signals, preferably those which essentially correspond to T-DNA polyadenylation signals from *Agrobacterium tumefaciens*, in particular the OCS (octopine synthase) terminator and the NOS (nopaline synthase) terminator. An expression vector need not contain transcription termination and

polyadenylation signal sequences, because these elements can be provided by the cloned gene or gene fragment.

5 The genetic component and/or expression cassette of the invention may comprise further functional elements. The term functional element is to be understood in the broad sense and refers to all those elements which have an effect on the generation, amplification or function of the genetic component, expression cassettes or recombinant organisms according to the invention. Functional elements may include for example (but shall not be limited to) selectable marker genes (including negative, positive, and
10 counter selection marker, see below for details), reporter genes, and

1) Origins of replication, which ensure amplification of the expression cassettes or vectors according to the invention in, for example, *E. coli*. Examples which may be mentioned are ORI (origin of DNA replication), the pBR322 ori or the P15A ori
15 (Maniatis T, Fritsch EF and Sambrook J (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor (NY)). Additional examples for replication systems functional in *E. coli*, are ColE1, pSC101, pACYC184, or the like. In addition to or in place of the *E. coli* replication system, a broad host range replication system may be employed, such as the replication systems of the P-1 Incompatibility plasmids; e.g., pRK290. These plasmids are particularly effective with armed and disarmed Ti-plasmids for transfer of T-DNA to the
20 plant species host. A expression vector can also include an SV40 origin. This element can be used for episomal replication and rescue in cell lines expressing SV40 large T antigen.

25 2) Elements which are necessary for *Agrobacterium*-mediated plant transformation, such as, for example, the right and/or – optionally - left border of the T-DNA or the *vir* region.

30 3) Cloning Sites: The cloning site can preferably be a multicloning site. Any multicloning site can be used, and many are commercially available.

4) S/MAR (scaffold/matrix attachment regions). Matrix attachment regions (MARs) are operationally defined as DNA elements that bind specifically to the nuclear matrix (nuclear scaffold proteins) in vitro and are proposed to mediate the attachment
35 of chromatin to the nuclear scaffold in vivo. It is possible, that they also mediate binding of chromatin to the nuclear matrix in vivo and alter the topology of the genome in interphase nuclei. When MARs are positioned on either side of a transgene their presence usually results in higher and more stable expression in transgenic organisms (especially plants) or cell lines, most likely by minimizing gene silencing (for review: Allen GC et al. (2000) *Plant Mol Biol* 43(2-3):361-76). Various S/MARS sequences and their effect on gene expression are described (Sidorenko L et al. (2003) *Transgenic Research*. 12(2):137-54; Allen CG et al. (1996) *Plant Cell* 8(5), 899-913; Villemure JF et al. (2001) *J. Mol. Biol.* 312, 963-974; Mlynarova, L.
40 et al. (2002) *Genetics* 160, 727-40). S/MAR elements may be preferably employed to reduce gene silencing (Mlynarova L et al. (2003) *Plant Cell*. 15(9):2203-17), which may occur for certain orientations of expression cassettes in a Multiple Expression Construct. An example for an S/MAR being the chicken lysozyme A

element (Stief et al., 1989, Nature 341: 343).

5. Sequences which further modify transcription, translation, and/or transport of an expressed protein. For example the expressed protein may be a chimeric protein comprising a secretory signal sequence. The secretory signal sequence is operably linked to a gene of interest such that the two sequences are joined in the correct reading frame and positioned to direct the newly synthesized polypeptide of interest into the secretory pathway of the host cell. Secretory signal sequences are commonly positioned 5' to the nucleotide sequence encoding the amino acid sequence of interest, although certain secretory signal sequences may be positioned elsewhere in the nucleotide sequence of interest (US 5,037,743, US 5,143,830). Expression vectors can also comprise nucleotide sequences that encode a peptide tag to aid the purification of the polypeptide of interest. Peptide tags that are useful for isolating recombinant polypeptides include polyHistidine tags (which have an affinity for nickel-chelating resin), c-myc tags, calmodulin binding protein (isolated with calmodulin affinity chromatography), substance P, the RYIRS tag (which binds with anti-RYIRS antibodies), the Glu-Glu tag, and the FLAG tag (which binds with anti-FLAG antibodies). See, for example, Luo et al., Arch. Biochem. Biophys. 329:215 (1996), Morganti et al., Biotechnol. Appl. Biochem. 23:67 (1996), and Zheng et al., Gene 186:55 (1997). Nucleic acid molecules encoding such peptide tags are available, for example, from Sigma-Aldrich Corporation (St. Louis, Mo.).

All of the compositions and methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the composition, methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference for all purposes.

6.2 Preferred Nucleic Acids of Interest

As noted above, a wide variety of nucleotide sequences may be carried by the Multiple Expression Constructs of the present invention. The methods of the present invention can preferably be used to obtain transgenic cells and organism with valuable traits, which require expression of two or more nucleic acids of interest. The nucleic acid of interest can – for example - be used to suppression an endogenous gene (e.g., by expression of an antisense or double stranded RNA) or to express or over-express a protein.

Preferably, proteins are expressed having value in industry, therapeutics, diagnostics, or research. Illustrative proteins include antibodies and antibody fragments, receptors, immunomodulators, hormones, and the like. For example, an expression cassette can

include a nucleic acid molecule that encodes a pharmaceutically active molecule, such as Factor VIIa, proinsulin, insulin, follicle stimulating hormone, tissue type plasminogen activator, tumor necrosis factor, interleukins (e.g., interleukin-1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, and IL-21), colony stimulating factors (e.g., granulocyte-colony stimulating factor, and granulocyte macrophage-colony stimulating factor), interferons (e.g., interferons- α , - β , - γ , - ω , - δ , - τ , and - ϵ), a stem cell growth factor, erythropoietin, lymphokines, toxins (e.g., ricin, abrin, Pseudomonas exotoxin A, herpes simplex virus thymidine kinase, E. coli guanine phosphoribosyl transferase, Shigella toxin, tritin, antiviral protein, pokeweed, gelonin, diphtheria toxin), prodrugs, prodrug-activating proteins, antigens which stimulate an immune response, ribozymes, and proteins which assist or inhibit an immune response, as well as antisense sequences (or sense sequences for "antisense applications"), and thrombopoietin. Additional examples of a protein of interest include an antibody, an antibody fragment, an anti-idiotypic antibody (or, fragment thereof), a chimeric antibody, a humanized antibody, an antibody fusion protein, and the like. Preferably, the method of the invention can be employed to express proteins which comprise two or more different subunits. The Multiple Expression Construct of the invention can be employed to simultaneously introduce all expression cassettes for all subunits into the host cell in a single step.

Sequences which encode the above-described proteins may be readily obtained from a variety of sources, including for example, depositories such as the American Type Culture Collection (ATCC, Rockville, Md.), or from commercial sources such as British Biotechnology Limited (Cowley, Oxford, England). Sequences which encode the above-described proteins may also be synthesized, for example, on an Applied Biosystems Inc. DNA synthesizer (e.g., APB DNA synthesizer model 392 (Foster City, Calif.)).

Preferably, the transgenic expression construct of the invention to be inserted into the genome of the target plant comprises at least one expression construct, which may – for example – facilitate expression of selection markers, trait genes, antisense RNA or double-stranded RNA. Preferably said expression constructs comprise a promoter sequence functional in plant cells (either – and preferably – a promoter of the invention or another suitable promoter as for example described above operatively linked to a nucleic acid sequence which – upon expression – confers an advantageous phenotype to the so transformed plant. The person skilled in the art is aware of numerous sequences which may be utilized in this context, e.g. to increase quality of food and feed, to produce chemicals, fine chemicals or pharmaceuticals (e.g., vitamins, oils, carbohydrates; Dunwell JM (2000) J Exp Bot 51 Spec No:487-96), conferring resistance to herbicides, or conferring male sterility. Furthermore, growth, yield, and resistance against abiotic and biotic stress factors (like e.g., fungi, viruses, nematodes, or insects) may be enhanced. Advantageous properties may be conferred either by overexpressing proteins or by decreasing expression of endogenous proteins by e.g., expressing a corresponding antisense (Sheehy et al. (1988) Proc Natl Acad Sci USA 85: 8805-8809; US 4,801,340; Mol 1990) or double-stranded RNA (Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; Waterhouse et al. (1998) Proc Natl Acad Sci USA 95:13959-64; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). Nucleic acids of interest may encode for the following (but shall not be limited to):

1. Selection markers and Reporter Genes

Selection markers are useful to select and separate successfully transformed or homologous recombined cells. Various Negative Selection Marker (conferring resistance
5 against toxic or biocidal compounds), Positive Selection Marker (conferring growth or proliferation advantages), or Counter Selection Marker (conferring toxic effects preferably in combination with otherwise non-toxic compounds) are known in the art and described above (see "Definitions"). Furthermore, various reporter genes and Proteins
10 are known and described above (see "Definitions"), which allow for readily identification of the transformed cell or organisms preferably by visual identification.

2. Trait Genes

The skilled worker is familiar with a multiplicity of nucleic acids of interest (or proteins encoded thereby) whose transgenic expression is advantageous. The skilled worker is
15 furthermore familiar with a multiplicity of genes by whose repression or silencing by means of expression of a corresponding antisense or double stranded RNA advantageous effects may also be achieved. The following may be mentioned by way of example, but not by way of limitation, as advantageous effects:

- 20 - Obtaining a resistance to abiotic stresses (high and low temperatures, drought, increased humidity, environmental toxins, UV radiation)
- Obtaining a resistance to biotic stresses (pathogens, viruses, insects and diseases)
- Obtaining resistance against phytotoxic substances or herbicides
- Improving the growth rate or the yield.

25

The following may be mentioned by way of example but not by way of limitation as nucleic acid sequences or polypeptides which can be used for these applications:

1. Improved protection of the plant embryo against abiotic stresses such as drought,
30 high or low temperatures, for example by overexpressing the antifreeze polypeptides from *Myoxocephalus scorpius* (WO 00/00512), *Myoxocephalus octodecemspinosus*, the *Arabidopsis thaliana* transcription activator CBF1, glutamate dehydrogenases (WO 97/12983, WO 98/11240), a late embryogenesis gene (LEA), for example from barley (WO 97/13843), calcium-dependent protein kinase genes
35 (WO 98/26045), calcineurins (WO 99/05902), farnesyl transferases (WO 99/06580, Pei 1998), ferritin (Deak 1999), oxalate oxidase (WO 99/04013; Dunwell (1998) Biotechnology and Genetic Engineering Reviews 15:1-32), DREB1A factor (dehydration response element B 1A; Kasuga 1999), mannitol or trehalose synthesis genes, such as trehalose-phosphate synthase or trehalose-phosphate phosphatase (WO
40 97/42326), or by inhibiting genes such as the trehalase gene (WO 97/50561). Especially preferred nucleic acids are those which encode the transcriptional activator CBF1 from *Arabidopsis thaliana* (GenBank Acc. No.: U77378) or the *Myoxocephalus octodecemspinosus* antifreeze protein (GenBank Acc. No.: AF306348), or functional equivalents of these.

45

2. Obtaining resistance for example against fungi, insects, nematodes and diseases by the targeted secretion or concentration of specific metabolites or proteins in the em-

bryol epidermis. Examples which may be mentioned are glucosinolates (defence against herbivores), chitinases or glucanases and other enzymes which destroy the cell wall of parasites, ribosome-inactivating proteins (RIPs) and other proteins of the plant's resistance and stress response as are induced upon wounding or microbial attack of plants or chemically by, for example, salicylic acid, jasmonic acid or ethylene; lysozymes from nonplant sources such as, for example, T4 lysozyme or lysozyme from a variety of mammals, insecticidal proteins such as *Bacillus thuringiensis* endotoxin, α -amylase inhibitor or protease inhibitors (cowpea trypsin inhibitor), glucanases, lectins such as phytohemagglutinin, snowdrops lectin, wheat-germ agglutinin, RNAses or ribozymes. Nucleic acids which are especially preferred are those which encode the *Trichoderma harzianum* chit42 endochitinase (GenBank Acc. No.: S78423) or the *Sorghum bicolor* N-hydroxylating multifunctional cytochrome P-450 (CYP79) proteins (GenBank Acc. No.: U32624), or functional equivalents of these.

The transgenic expression constructs of the invention can be employed for suppressing or reducing expression of endogenous target genes by "gene silencing". Preferred genes or proteins whose suppression brings about an advantageous phenotype are known to the skilled worker. Examples may include but are not limited to down-regulation of the β -subunit of *Arabidopsis* G protein for increasing root mass (Ullah *et al.* (2003) Plant Cell 15:393-409), inactivating cyclic nucleotide-gated ion channel (CNGC) for improving disease resistance (WO 2001007596), and down-regulation of 4-coumarate-CoA ligase (4CL) gene for altering lignin and cellulose contents (US 2002138870).

Gene silencing can be realized by antisense or double-stranded RNA or by co-suppression (sense-suppression). An "antisense" nucleic acid is firstly understood as meaning a nucleic acid sequence which is fully or partially complementary to at least part of the "sense" strand of said target protein. The skilled worker knows that he can use alternative cDNA or the corresponding gene as starting template for suitable antisense constructs. The "antisense" nucleic acid is preferably complementary to the coding region of the target protein or part thereof. However, the "antisense" nucleic acid may also be complementary to the non-coding region or part thereof. Starting from the sequence information on a target protein, an antisense nucleic acid can be designed in the manner with which the skilled worker is familiar, taking into consideration Watson's and Crick's rules of base pairing. An antisense nucleic acid can be complementary to the entire or part of the nucleic acid sequence of a target protein.

Likewise encompassed is the use of the above-described sequences in sense orientation, which, as is known to the skilled worker, can lead to co-suppression (sense-suppression). It has been demonstrated that expression of sense can reduce or switch off expression of same, analogously to what has been described for antisense approaches (Goring *et al.* (1991) Proc. Natl Acad. Sci. USA 88:1770-1774; Smith *et al.* (1990) Mol. Gen. Genet. 224:447-481; Napoli *et al.* (1990) Plant Cell 2:279-289; Van der Krol *et al.* (1990) Plant Cell 2:291-99). In this context, the construct introduced may represent the gene to be reduced fully or only in part. The possibility of translation is not necessary.

Especially preferred is the use of gene regulation methods by means of double-stranded RNAi ("double-stranded RNA interference"). Such methods are known to the person skilled in the art (e.g., Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; 5 WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). The processes and methods described in the references stated are expressly referred to.

Furthermore, artificial transcription factors (e.g. of the zinc finger protein type; Beerli *et al.* (2000) Proc Natl Acad Sci USA 97(4):1495-500) can be expressed under control of 10 a promoter of the invention to modulate expression of specific endogenous genes. These factors attach to the regulatory regions of the endogenous genes to be expressed or to be repressed and, depending on the design of the factor, bring about expression or repression of the endogenous gene.

15 In an preferred embodiment of the method of the invention to assemble Multiple Expression Constructs can be employed for the purpose of complex pathway engineering or to produce complex products requiring two or more enzymatic reactions. Example may include but shall not be limited to:

20 a) Synthesis of polyunsaturated fatty acids (PUFA):

b) Increasing oil content:

25 c) Synthesis of vitamins (e.g., tocopherol) or carotinoids (e.g., asthaxanthin): For synthesis of astaxanthins various combinations of at least two sequences selected from the group consisting of ketolases, β -cyclases, β -hydroxylases, β -hydroxylase suppressing sequences, ϵ -cyclases, and ϵ -cyclases suppressing sequences. More preferably, the sequences are selected from the group of consisting of sequences coding for HMG-CoA-reductases, (E)-4-hydroxy-3-methyl-but-2-enyl-diphosphate-reductases, 30 1-deoxy-D-xylose-5-phosphate-synthases, 1-deoxy-D-Xylose-5-phosphate-reductoisomerases, isopentenyl-diphosphate- Δ -isomerases, geranyl-diphosphate-synthases, farnesyl-diphosphat-synthase, geranyl-geranyl-diphosphate-synthases, phytoen-synthases, phytoen-desaturases, Zeta-Carotin-desaturases, crtISO proteins, FtsZ proteins, and MinD proteins.

35 Altering the content of a substance in a biological system is an aim in pathway manipulation. In most cases the biosynthesis of one substance or a substance class depends on more than one precursor. For example the synthesis of substances summarized in the term vitamin E (Tocopherols and Tocotrienols) depends on the 40 synthesis of homogentisate and phytylpyrophosphate. Phytylpyrophosphate is a product of the isoprenoid pathway and homogentisate from the shikimate pathway. Both products are necessary for the final steps of the tocopherol biosynthesis, which involves methylation and cyclisation steps. All these enzymatic steps and the gene products involved are possible bottle necks in manipulating the pathway. To 45 increase the endproduct Vitamin E it is necessary to find the limiting enzymatic activities. These enzymatic activities should be increased in a correct order. To fulfill this aim many combinations have to be tested, to find the most promising gene combination.

In an preferred embodiment of the invention the expression cassettes introduced into the plant by said Multiple Expression Construct of the invention are able to confer to said transgenic plant at least one phenotype selected from the group consisting of increased nutritional value, increased oil content, increased starch content, increased protein content, increased vitamin content, increased carotinoid content, increased pathogen resistance, modified oil or starch composition, and increased stress tolerance. The term "increased" is intended to mean a quantity of a compound or quality (e.g., oil) of a property (e.g., stress tolerance) which is higher than the same property in the same plant variety which is lacking the expression cassettes. Preferably the increase is at least 10%, preferably at least 50%, more preferably at least 100%, most preferably at least 500%. The term "modified" is intended to mean a change in quality or quantity, preferably in composition of a complex mixture. With respect to oil composition, "modified" means preferably a higher content of unsaturated and/or polyunsaturated fatty acids. With respect to starch composition, "modified" mean preferably a change in the amylopectin to amylose ratio.

Thus another embodiment of the present invention related to a Multiple Expression Construct obtained by the method(s) of the invention. This Multiple Expression Construct is comprising at least two Inserts I(n), wherein n is an integer from 1 to m characterizing each Insert, and m is the total number of different Inserts. Each Insert comprising at least one expression cassette. Each Insert is flanked by a sequence resulting from the recombination of a recombination side A(2i-1) with the recombination side A(2i) for the same i, wherein said recombination site are the same as defined above for the individual Inserts comprised in the Insert Donor Molecules. For example recombination of recombination site attR1* with attL1* results in attB1*, attR2* with attL2* results in attB2*, attR3* with attL3* results in attB3*, attR4* with attL4* results in attB4* (for sequence specifications see above).

30 7. Target Organism and Transformation Techniques

Another subject matter of the invention relates to transgenic cells or non-human organisms transformed with at least one Multiple Expression Construct of the invention, and to cells, cell cultures, tissues, organs (e.g., leaves, roots and the like in the case of plant organisms), or propagation material derived from such organisms.

35 Preferably, the transgenic cell or non-human organism is selected from the group comprising prokaryotic and eukaryotic cells or organism (as defined and specified above). Most preferably, said transgenic cell or organism is selected from the group comprising of plant cells or organism (as defined and specified above).

40 The generation of a transformed organism or a transformed cell requires introducing the DNA in question into the host cell in question. A multiplicity of methods is available for this procedure, which is termed transformation (see also Keown (1990) Methods in Enzymology 185:527-537). For example, the DNA can be introduced directly by microinjection or by bombardment via DNA-coated microparticles. Also, the cell can be permeabilized chemically, for example using polyethylene glycol, so that the DNA can enter the cell by diffusion. The DNA can also be introduced by protoplast fusion with other

DNA-containing units such as minicells, cells, lysosomes or liposomes. Another suitable method of introducing DNA is electroporation, where the cells are permeabilized reversibly by an electrical pulse.

- 5 As described above the host cell or organism can be any prokaryotic or eukaryotic organism. Preferred are mammalian cells, non-human mammalian organism, plant cells and plant organisms as defined above.

10 The Multiple Expression Construct of the invention is preferably introduced into a eukaryotic cell. It may be preferably inserted into the genome (e.g., plastids or chromosomal DNA) but may also exist extra-chromosomal or epichromosomal. Preferred eukaryotic cells are mammalian cell, fungal cell, plant cell, insect cell, avian cell, and the like. Examples of suitable mammalian host cells include African green monkey kidney cells (Vero; ATCC CRL 1587), human embryonic kidney cells (293-HEK; ATCC
15 CRL 1573), baby hamster kidney cells (BHK-21, BHK-570, ATCC CRL 8544, ATCC CRL 10314), canine kidney cells (MDCK; ATCC CCL 34), Chinese hamster ovary cells (CHO-K1; ATCC CCL61; CHO DG44 (Chasin et al., Som. Cell. Molec. Genet. 12:555, 1986)), rat pituitary cells (GH1; ATCC CCL82), HeLa S3 cells (ATCC CCL2.2), rat hepatoma cells (H-4-II-E; ATCC CRL 1548) SV40-transformed monkey kidney cells
20 (COS-1; ATCC CRL 1650) and murine embryonic cells (NIH-3T3; ATCC CRL 1658).

25 An Multiple Expression Construct can be introduced into host cells using a variety of standard techniques including calcium phosphate transfection, liposome-mediated transfection, microprojectile-mediated delivery, electroporation, and the like. Transfected cells can be selected and propagated to provide recombinant host cells that comprise the gene of interest stably integrated in the host cell genome.

30 The Multiple Expression Construct may be a baculovirus expression vector to be employed in a baculovirus system. The baculovirus system provides an efficient means to introduce cloned genes of interest into insect cells. Suitable expression vectors are based upon the Autographa californica multiple nuclear polyhedrosis virus (AcMNPV), and contain well-known promoters such as Drosophila heat shock protein (hsp) 70 promoter, Autographa californica nuclear polyhedrosis virus immediate-early gene promoter (ie-1) and the delayed early 39K promoter, baculovirus p10 promoter, and the
35 Drosophila metallothionein promoter. A second method of making recombinant baculovirus utilizes a transposon-based system described by Luckow (Luckow, et al., J. Virol. 67:4566 (1993)). This system, which utilizes transfer vectors, is sold in the BAC-to-BAC kit (Life Technologies, Rockville, Md.). This system utilizes a transfer vector, PFASTBAC (Life Technologies) containing a Tn7 transposon to move the DNA encoding the polypeptide of interest into a baculovirus genome maintained in *E. coli* as a large plasmid called a "bacemid." See, Hill-Perkins and Possee, J. Gen. Virol. 71:971 (1990), Bonning, et al., J. Gen. Virol. 75:1551 (1994), and Chazenbalk, and Rapoport, J. Biol. Chem. 270:1543 (1995). In addition, transfer vectors can include an in-frame
40 fusion with DNA encoding an epitope tag at the C- or N-terminus of the expressed polypeptide, for example, a Glu-Glu epitope tag (Grussenmeyer et al., Proc. Nat'l Acad. Sci. 82:7952 (1985)). Using a technique known in the art, a transfer vector containing a gene of interest is transformed into *E. coli*, and screened for bacmids, which contain an interrupted lacZ gene indicative of recombinant baculovirus. The bacmid DNA contain-

ing the recombinant baculovirus genome is then isolated using common techniques. The recombinant virus or bacinid is used to transfect host cells. Suitable insect host cells include cell lines derived from IPLB-Sf-21, a *Spodoptera frugiperda* pupal ovarian cell line, such as Sf9 (ATCC CRL 1711), Sf21AE, and Sf21 (Invitrogen Corporation; 5 San Diego, Calif.), as well as *Drosophila* Schneider-2 cells, and the HIGH FIVEO cell line (Invitrogen) derived from *Trichoplusia ni* (U.S. Pat. No. 5,300,435). Commercially available serum-free media can be used to grow and to maintain the cells. Suitable media are Sf900 II.TM. (Life Technologies) or ESF 921.TM. (Expression Systems) for the Sf9 cells; and Ex-cello405.TM. (JRH Biosciences, Lenexa, Kans.) or Express Fi- 10 veO.TM. (Life Technologies) for the *T. ni* cells. When recombinant virus is used, the cells are typically grown up from an inoculation density of approximately 2-5.times.10.sup.5 cells to a density of 1-2.times.10.sup.6 cells at which time a recombinant viral stock is added at a multiplicity of infection of 0.1 to 10, more typically near 3. Established techniques for the baculovirus systems are provided by Bailey et al., "Ma- 15 nipulation of Baculovirus Vectors," in *Methods in Molecular Biology*, Volume 7: Gene Transfer and Expression Protocols, Murray (ed.), pages 147-168 (The Humana Press, Inc. 1991), by Patel et al., "The baculovirus expression system," in *DNA Cloning 2: Expression Systems*, 2nd Edition, Glover et al. (eds.), pages 205-244 (Oxford University Press 1995), by Ausubel (1995) at pages 16-37 to 16-57, by Richardson (ed.), 20 *Baculovirus Expression Protocols* (The Humana Press, Inc. 1995), and by Lucknow, "Insect Cell Expression Technology," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 183-218 (John Wiley & Sons, Inc. 1996).

Fungal cells, including yeast cells, can also be used as host cells for transformation 25 with the Multiple Expression Construct of the invention. Yeast species of particular interest in this regard include *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Pichia methanolica*. Suitable promoters for expression in yeast include promoters from GAL1 (galactose), PGK (phosphoglycerate kinase), ADH (alcohol dehydrogenase), AOX1 (alcohol oxidase), HIS4 (histidinol dehydrogenase), and the like. Many yeast cloning 30 vectors have been designed and are readily available to be employed e.g., as basic vector to derive Insert Acceptor (Vector Donor) Molecules. These vectors include Ylp-based vectors, such as Ylp5, YRp vectors, such as YRp17, YEp vectors such as YEp13 and YCp vectors, such as YCp19. Methods for transforming *S. cerevisiae* cells with exogenous DNA and producing recombinant polypeptides therefrom are disclosed 35 by, for example, US 4,599,311, US 4,931,373, US 4,870,008, US 5,037,743, and US 4,845,075. Transformed cells are selected by phenotype determined by the selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient (e.g., leucine). An illustrative vector system for use in *Saccharomyces cerevisiae* is the POT1 vector system (US 4,931,373), which allows transformed cells to be 40 selected by growth in glucose-containing media. Additional suitable promoters and terminators for use in yeast include those from glycolytic enzyme genes (see, e.g., US 4,599,311, US 4,615,974, and US 4,977,092) and alcohol dehydrogenase genes. See also US 4,990,446, 5,063,154, 5,139,936, and 4,661,454.

45 Transformation systems for other yeasts, including *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces fragilis*, *Ustilago maydis*, *Pichia pastoris*, *Pichia methanolica*, *Pichia guilliermondii* and *Candida maltosa* are known in the art. See, for example, Gleeson et al., *J. Gen. Microbiol.* 132:3459 (1986),

and US 4,882,279. *Aspergillus* cells may be utilized according to the methods of McKnight et al. (US 4,935,349). Methods for transforming *Acremonium chrysogenum* are disclosed (US 5,162,228). Methods for transforming *Neurospora* are disclosed (US 4,486,533).

5

For example, the use of *Pichia methanolica* as host for the production of recombinant proteins is disclosed (US 5,716,808, US 5,736,383, Raymond et al., *Yeast* 14:11-23 (1998), WO 97/17450, WO 97/17451, WO 98/02536, and WO 98/02565). DNA molecules for use in transforming *P. methanolica* will commonly be prepared as double-stranded, circular plasmids, which are preferably linearized prior to transformation. For polypeptide production in *P. methanolica*, the promoter and terminator in the plasmid can be that of a *P. methanolica* gene, such as a *P. methanolica* alcohol utilization gene (AUG1 or AUG2). Other useful promoters include those of the dihydroxyacetone synthase (DHAS), formate dehydrogenase (FMD), and catalase (CAT) genes. To facilitate integration of the DNA into the host chromosome, it is preferred to have the entire expression segment of the plasmid flanked at both ends by host DNA sequences. For large-scale, industrial processes where it is desirable to minimize the use of methanol host cells can be used in which both methanol utilization genes (AUG1 and AUG2) are deleted. For production of secreted proteins, host cells can be used that are deficient in vacuolar protease genes (PEP4 and PRB1). Electroporation is used to facilitate the introduction of a plasmid containing DNA encoding a polypeptide of interest into *P. methanolica* cells. *P. methanolica* cells can be transformed by electroporation using an exponentially decaying, pulsed electric field having a field strength of from 2.5 to 4.5 kV/cm, preferably about 3.75 kV/cm, and a time constant (t) of from 1 to 40 milliseconds, most preferably about 20 milliseconds.

10

15

Standard methods for introducing nucleic acid molecules into bacterial, yeast, insect, mammalian, and plant cells are provided, for example, by Ausubel (1995). General methods for expressing and recovering foreign protein produced by a mammalian cell system are provided by, for example, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 163 (Wiley-Liss, Inc. 1996). Established methods for isolating recombinant proteins from a baculovirus system are described by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc. 1995).

30

35

Especially preferred in transfer of the Multiple Expression Construct into plant cells, tissues and/or organism. Methods for introduction of a transgenic expression construct or vector into plant tissue may include but are not limited to, e.g., electroinjection (Nan et al. (1995) In "Biotechnology in Agriculture and Forestry," Ed. Y. P. S. Bajaj, Springer-Verlag Berlin Heidelberg, Vol 34:145-155; Griesbach (1992) *Hort. Science* 27:620); fusion with liposomes, lysosomes, cells, minicells or other fusible lipid-surfaced bodies (Fraley et al. (1982) *Proc. Natl. Acad. Sci. USA* 79:1859-1863); polyethylene glycol (Krens et al. (1982) *Nature* 296:72-74); chemicals that increase free DNA uptake; transformation using virus, and the like. Furthermore, the biolistic method with the gene gun, electroporation, incubation of dry embryos in DNA-containing solution, and micro-injection may be employed.

40

45

Protoplast based methods can be employed (e.g., for rice), where DNA is delivered to the protoplasts through liposomes, PEG, or electroporation (Shimamoto *et al.* (1989) Nature 338:274-276; Datta *et al.* (1990b) Bio/Technology 8:736-740). Transformation by electroporation involves the application of short, high-voltage electric fields to create
5 "pores" in the cell membrane through which DNA is taken-up. These methods are – for example - used to produce stably transformed monocotyledonous plants (Paszkowski *et al.* (1984) EMBO J 3:2717-2722; Shillito *et al.* (1985) Bio/Technology, 3:1099-1103; Fromm *et al.* (1986) Nature 319:791-793) especially from rice (Shimamoto *et al.* (1989) Nature 338:274-276; Datta *et al.* (1990b) Bio/Technology 8:736-740; Hayakawa *et al.*
10 (1992) Proc Natl Acad Sci USA 89:9865-9869).

Particle bombardment or "biolistics" is a widely used method for the transformation of plants, especially monocotyledonous plants. In the "biolistics" (microprojectile-mediated DNA delivery) method microprojectile particles are coated with DNA and accelerated
15 by a mechanical device to a speed high enough to penetrate the plant cell wall and nucleus (WO 91/02071). The foreign DNA gets incorporated into the host DNA and results in a transformed cell. There are many variations on the "biolistics" method (Sanford (1990) Physiologia Plantarum 79:206-209; Fromm *et al.* (1990) Bio/Technology 8:833-839; Christou *et al.* (1988) Plant Physiol 87:671-674; Sautter *et al.* (1991)
20 Bio/Technology, 9:1080-1085). The method has been used to produce stably transformed monocotyledonous plants including rice, maize, wheat, barley, and oats (Christou *et al.* (1991) Bio/Technology 9:957-962; Gordon-Kamm *et al.* (1990) Plant Cell 2:603-618; Vasil *et al.* (1992) Bio/Technology, 10:667-674, Wan & Lemaux (1994) Plant Physiol. 104:3748).

25 In addition to these "direct" transformation techniques, transformation can also be effected by bacterial infection by means of *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. These strains contain a plasmid (Ti or Ri plasmid) which is transferred to the plant following *Agrobacterium* infection. Part of this plasmid, termed T-DNA (transferred DNA), is integrated into the genome of the plant cell (see above for description
30 of vectors). To transfer the DNA to the plant cell, plant explants are cocultured with a transgenic *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. Starting from infected plant material (for example leaf, root or stem sections, but also protoplasts or suspensions of plant cells), intact plants can be generated using a suitable medium
35 which may contain, for example, antibiotics or biocides for selecting transformed cells. The plants obtained can then be screened for the presence of the DNA introduced, in this case the expression construct according to the invention. As soon as the DNA has integrated into the host genome, the genotype in question is, as a rule, stable and the insertion in question is also found in the subsequent generations. As a rule, the ex-
40 pression construct integrated contains a selection marker which imparts a resistance to a biocide (for example a herbicide) or an antibiotic such as kanamycin, G 418, bleomycin, hygromycin or phosphinotricin and the like to the transformed plant. The selection marker permits the selection of transformed cells from untransformed cells (McCormick 1986). The plants obtained can be cultured and hybridized in the customary fashion.
45 Two or more generations should be grown in order to ensure that the genomic integration is stable and hereditary.

The abovementioned methods are described (for example, in Jenes B et al.(1993) Techniques for Gene Transfer, in: Recombinant Plants, Vol. 1, Engineering and Utilization, edited by SD Kung and R Wu, Academic Press, pp. 128-143; and in Potrykus (1991) Ann Rev Plant Physiol Plant Mol Biol 42:205-225).

5

One of skill in the art knows that the efficiency of transformation by *Agrobacterium* may be enhanced by using a number of methods known in the art. For example, the inclusion of a natural wound response molecule such as acetosyringone (AS) to the *Agrobacterium* culture has been shown to enhance transformation efficiency with *Agrobacterium tumefaciens* (Shahla et al. (1987) Plant Molec. Biol. 8:291-298). Alternatively, transformation efficiency may be enhanced by wounding the target tissue to be transformed. Wounding of plant tissue may be achieved, for example, by punching, maceration, bombardment with microprojectiles, etc. (see, e.g., Bidney et al. (1992) Plant Molec. Biol. 18:301-313).

15

A number of other methods have been reported for the transformation of plants (especially monocotyledonous plants) including, for example, the "pollen tube method" (WO 93/18168; Luo & Wu (1988) Plant Mol. Biol. Rep. 6:165-174), macro-injection of DNA into floral tillers (Du et al. (1989) Genet Manip Plants 5:8-12; de la Pena et al. (1987) Nature 325:274-276), injection of *Agrobacterium* into developing caryopses (WO 00/63398), and tissue incubation of seeds in DNA solutions (Topfer et al. (1989) Plant Cell 1:133-139). Direct injection of exogenous DNA into the fertilized plant ovule at the onset of embryogenesis was disclosed in WO 94/00583. WO 97/48814 disclosed a process for producing stably transformed fertile wheat and a system of transforming wheat via *Agrobacterium* based on freshly isolated or pre-cultured immature embryos, embryogenic callus and suspension cells.

25

It may be desirable to target the nucleic acid sequence of interest to a particular locus on the plant genome. Site-directed integration of the nucleic acid sequence of interest into the plant cell genome may be achieved by, for example, homologous recombination using *Agrobacterium*-derived sequences. Generally, plant cells are incubated with a strain of *Agrobacterium* which contains a targeting vector in which sequences that are homologous to a DNA sequence inside the target locus are flanked by *Agrobacterium* transfer-DNA (T-DNA) sequences, as previously described (US 5,501,967, the entire contents of which are herein incorporated by reference). One of skill in the art knows that homologous recombination may be achieved using targeting vectors which contain sequences that are homologous to any part of the targeted plant gene, whether belonging to the regulatory elements of the gene, or the coding regions of the gene. Homologous recombination may be achieved at any region of a plant gene so long as the nucleic acid sequence of regions flanking the site to be targeted is known.

30

35

40

Where homologous recombination is desired, the targeting vector used may be of the replacement- or insertion-type (US 5,501,967; supra). Replacement-type vectors generally contain two regions which are homologous with the targeted genomic sequence and which flank a heterologous nucleic acid sequence, e.g., a selectable marker gene sequence. Replacement-type vectors result in the insertion of the selectable marker gene which thereby disrupts the targeted gene. Insertion-type vectors contain a single region of homology with the targeted gene and result in the insertion of the entire tar-

45

geting vector into the targeted gene.

Transformed cells, *i.e.* those which contain the introduced DNA integrated into the DNA of the host cell, can be selected from untransformed cells if a selectable marker is part of the introduced DNA. A selection marker gene may confer positive or negative selection.

A positive selection marker gene may be used in constructs for random integration and site-directed integration. Positive selection marker genes include antibiotic resistance genes, and herbicide resistance genes and the like. Transformed cells which express such a marker gene are capable of surviving in the presence of concentrations of the antibiotic or herbicide in question which kill an untransformed wild type. Examples are the bar gene, which imparts resistance to the herbicide phosphinotricin (bialaphos; Vasil et al. (1992) *Bio/Technology*, 10:667-674; Weeks et al. (1993) *Plant Physiol* 102:1077-1084; Rathore et al. (1993) *Plant Mol Biol* 21(5):871-884), the *nptII* gene, which imparts resistance to kanamycin, the *hpt* gene, which imparts resistance to hygromycin, or the EPSP gene, which imparts resistance to the herbicide glyphosate, geneticin (G-418) (aminoglycoside) (Nehra et al. (1994) *Plant J.* 5:285-297), glyphosate (Della-Cioppa et al. (1987) *Bio/Technology* 5:579-584) and the ALS gene (chlorsulphuron resistance). Further preferred selectable and screenable marker genes are disclosed above.

A negative selection marker gene may also be included in the constructs. The use of one or more negative selection marker genes in combination with a positive selection marker gene is preferred in constructs used for homologous recombination. Negative selection marker genes are generally placed outside the regions involved in the homologous recombination event. The negative selection marker gene serves to provide a disadvantage (preferably lethality) to cells that have integrated these genes into their genome in an expressible manner. Cells in which the targeting vectors for homologous recombination are randomly integrated in the genome will be harmed or killed due to the presence of the negative selection marker gene. Where a positive selection marker gene is included in the construct, only those cells having the positive selection marker gene integrated in their genome will survive. The choice of the negative selection marker gene is not critical to the invention as long as it encodes a functional polypeptide in the transformed plant cell. The negative selection gene may for instance be chosen from the aux-2 gene from the Ti-plasmid of *Agrobacterium*, the tk-gene from SV40, cytochrome P450 from *Streptomyces griseolus*, the *Adh* gene from Maize or *Arabidopsis*, etc. Any gene encoding an enzyme capable of converting a substance which is otherwise harmless to plant cells into a substance which is harmful to plant cells may be used. Further preferred negative selection markers are disclosed above.

However, insertion of an expression cassette or a vector into the chromosomal DNA can also be demonstrated and analyzed by various other methods (not based on selection marker) known in the art like including, but not limited to, restriction mapping of the genomic DNA, PCR-analysis, DNA-DNA hybridization, DNA-RNA hybridization, DNA sequence analysis and the like. More specifically such methods may include e.g., PCR analysis, Southern blot analysis, fluorescence *in situ* hybridization (FISH), and *in situ* PCR.

As soon as a transformed plant cell has been generated, an intact plant can be obtained using methods known to the skilled worker. Accordingly, the present invention provides transgenic plants. The transgenic plants of the invention are not limited to
5 plants in which each and every cell expresses the nucleic acid sequence of interest under the control of the promoter sequences provided herein. Included within the scope of this invention is any plant which contains at least one cell which expresses the nucleic acid sequence of interest (e.g., chimeric plants). It is preferred, though not necessary, that the transgenic plant comprises the nucleic acid sequence of interest in more
10 than one cell, and more preferably in one or more tissue.

Once transgenic plant tissue which contains an expression vector has been obtained, transgenic plants may be regenerated from this transgenic plant tissue using methods known in the art. The term "regeneration" as used herein, means growing a whole plant
15 from a plant cell, a group of plant cells, a plant part or a plant piece (e.g., from a protoplast, callus, protocorm-like body, or tissue part).

Species from the following examples of genera of plants may be regenerated from transformed protoplasts: *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*,
20 *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Hyoscyamus*, *Lycopersicon*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Pisum*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and
25 *Datura*.

For regeneration of transgenic plants from transgenic protoplasts, a suspension of transformed protoplasts or a petri plate containing transformed explants is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, somatic embryo formation can be induced in the callus
30 tissue. These somatic embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and plant hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. These three
35 variables may be empirically controlled to result in reproducible regeneration.

Plants may also be regenerated from cultured cells or tissues. Dicotyledonous plants which have been shown capable of regeneration from transformed individual cells to
40 obtain transgenic whole plants include, for example, apple (*Malus pumila*), blackberry (*Rubus*), Blackberry/raspberry hybrid (*Rubus*), red raspberry (*Rubus*), carrot (*Daucus carota*), cauliflower (*Brassica oleracea*), celery (*Apium graveolens*), cucumber (*Cucumis sativus*), eggplant (*Solanum melongena*), lettuce (*Lactuca sativa*), potato (*Solanum tuberosum*), rape (*Brassica napus*), wild soybean (*Glycine canescens*), strawberry
45 (*Fragaria ananassa*), tomato (*Lycopersicon esculentum*), walnut (*Juglans regia*), melon (*Cucumis melo*), grape (*Vitis vinifera*), and mango (*Mangifera indica*). Monocotyledonous plants which have been shown capable of regeneration from transformed individual cells to obtain transgenic whole plants include, for example, rice (*Oryza sativa*), rye

(*Secale cereale*), and maize (*Zea mays*).

In addition, regeneration of whole plants from cells (not necessarily transformed) has also been observed in: apricot (*Prunus armeniaca*), asparagus (*Asparagus officinalis*),
5 banana (hybrid *Musa*), bean (*Phaseolus vulgaris*), cherry (hybrid *Prunus*), grape (*Vitis vinifera*), mango (*Mangifera indica*), melon (*Cucumis melo*), ochra (*Abelmoschus esculentus*), onion (hybrid *Allium*), orange (*Citrus sinensis*), papaya (*Carrica papaya*), peach (*Prunus persica*), plum (*Prunus domestica*), pear (*Pyrus communis*), pineapple (*Ananas comosus*), watermelon (*Citrullus vulgaris*), and wheat (*Triticum aestivum*).

10 The regenerated plants are transferred to standard soil conditions and cultivated in a conventional manner. After the expression vector is stably incorporated into regenerated transgenic plants, it can be transferred to other plants by vegetative propagation or by sexual crossing. For example, in vegetatively propagated crops, the mature
15 transgenic plants are propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. In seed propagated crops, the mature transgenic plants are self crossed to produce a homozygous inbred plant which is capable of passing the transgene to its progeny by Mendelian inheritance. The inbred plant produces seed containing the nucleic acid sequence of interest. These seeds can
20 be grown to produce plants that would produce the selected phenotype. The inbred plants can also be used to develop new hybrids by crossing the inbred plant with another inbred plant to produce a hybrid.

Confirmation of the transgenic nature of the cells, tissues, and plants may be performed by PCR analysis, antibiotic or herbicide resistance, enzymatic analysis and/or
25 Southern blots to verify transformation. Progeny of the regenerated plants may be obtained and analyzed to verify whether the transgenes are heritable. Heritability of the transgene is further confirmation of the stable transformation of the transgene in the plant. The resulting plants can be bred in the customary fashion. Two or more generations should be grown in order to ensure that the genomic integration is stable and hereditary. Corresponding methods are described, (Jenes B et al.(1993) Techniques for
30 Gene Transfer, in: Recombinant Plants, Vol. 1, Engineering and Utilization, edited by SD Kung and R Wu, Academic Press, pp. 128-143; Potrykus (1991) Ann Rev Plant Physiol Plant Mol Biol 42:205-225).

35 Also in accordance with the invention are cells, cell cultures, tissues, parts, organs,—such as, for example, roots, leaves and the like in the case of transgenic plant organisms — derived from the above-described transgenic organisms, and transgenic propagation material such as seeds or fruits.

40 Another embodiment of the invention related to a method of producing food or feed products, pharmaceuticals, or chemicals, said method comprising providing a transgenic cell or organism comprising the Multiple Expression Construct of the Invention, growing said cell or organism, and — optionally — isolating said food or feed
45 product, pharmaceutical, or chemical.

Genetically modified plants according to the invention which can be consumed by humans or animals can also be used as food or feedstuffs, for example directly or following processes known *per se*.

5 A further subject matter of the invention relates to the use of the above-described transgenic organisms according to the invention and the cells, cell cultures, parts, tissues, organs— such as, for example, roots, leaves and the like in the case of transgenic plant organisms – derived from them, and transgenic propagation material such as seeds or fruits, for the production of foods or feedstuffs, pharmaceuticals or fine chemicals.

10 Preferred is furthermore a method for the recombinant production of pharmaceuticals or fine chemicals in host organisms, where a host organism is transformed with one of the above-described expression constructs, and this expression construct contains one or more structural genes which encode the desired fine chemical or catalyze the biosynthesis of the desired fine chemical, the transformed host organism is cultured, and the desired fine chemical is isolated from the culture medium. This process can be
15 used widely for fine chemicals such as enzymes, vitamins, amino acids, sugars, fatty acids, natural and synthetic flavorings, aroma substances and colorants. Especially preferred is the production of tocopherols and tocotrienols, carotenoids, oils, polyunsaturated fatty acids *etc.* Culturing the transformed host organisms, and isolation from the host organisms or the culture medium, is performed by methods known to the
20 skilled worker. The production of pharmaceuticals such as, for example, antibodies, vaccines, enzymes or pharmaceutically active proteins is described (Hood & Jilka (1999) *Curr Opin Biotechnol.* 10(4):382-6; Ma & Vine (1999) *Curr Top Microbiol. Immunol.* 236:275-92; Russel (1999) *Current Topics in Microbiology and Immunology* 240:119-138; Cramer *et al.* (1999) *Current Topics in Microbiology and Immunology* 240:95-118; Gavilondo & Larrick (2000) *Biotechniques* 29(1):128-138; Holliger & Bohlen (1999) *Cancer & Metastasis Reviews* 18(4):411-419).

30 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
35 the invention.

SEQUENCES

- | | | |
|----|--------------------|--|
| | 1. SEQ ID NO:1 | Recombination site m-att
5'-RKYCWGCTTTYKTRTACNAASTSGB-3' |
| 5 | 2. SEQ ID NO:2 | Recombination site m-attB
5'-AGCCWGCTTTYKTRTACNAACTSGB-3' |
| | 3. SEQ ID NO:3 | Recombination site m-attR
5'-G TTCAGCTTCKTRTACNAACTSGB-3' |
| | 4. SEQ ID NO:4 | Recombination site m-attL
5'-AGCCWGCTTCKTRTACNAAGTSGB-3' |
| 10 | 5. SEQ ID NO:5 | Recombination site m-attP1
5'-G TTCAGCTTTYKTRTACNAAGTSGB-3' |
| | 6. SEQ ID NO:6 | Recombination site attB1
5'-AGCCTGCTTTTTGTACAAACTTGT-3' |
| 15 | 7. SEQ ID NO:7 | Recombination site attB2
5'-AGCCTGCTTCTTGTACAAACTTGT-3' |
| | 8. SEQ ID NO:8 | Recombination site attB3
5'-ACCCAGCTTCTTGTACAAACTTGT-3' |
| | 9. SEQ ID NO:9 | Recombination site attR1
5'-G TTCAGCTTTTTGTACAAACTTGT-3' |
| 20 | 10. SEQ ID NO:10 | Recombination site attR2
5'-G TTCAGCTTCTTGTACAAACTTGT-3' |
| | 11. SEQ ID NO:11 | Recombination site attR3
5'-G TTCAGCTTCTTGTACAAAGTTGG-3' |
| 25 | 12. SEQ ID NO:12 | Recombination site attL1
5'-AGCCTGCTTTTTGTACAAAGTTGG-3' |
| | 13. SEQ ID NO:13 | Recombination site attL2
5'-AGCCTGCTTCTTGTACAAAGTTGG-3' |
| | 14. SEQ ID NO:14 | Recombination site attL3
5'-ACCCAGCTTCTTGTACAAAGTTGG-3' |
| 30 | 15. SEQ ID NO:15 | Recombination site attP1
5'-G TTCAGCTTTTTGTACAAAGTTGG-3' |
| | 16. SEQ ID NO:16 | Recombination site attP2,P3
5'-G TTCAGCTTCTTGTACAAAGTTGG-3' |
| 35 | 17. SEQ ID NO: 17: | Phosphorylated oligonucleotide Loy344
5'-p-GTCGACCAGATCTGATATCTGCGGCCGCCTCGAGCATATG-3' |
| | 18. SEQ ID NO: 18: | Phosphorylated oligonucleotide Loy345
5'-p-GTCGACCAGATCTGATATCTGCGGCCGCCTCGAGCATATG-3' |
| | 19. SEQ ID NO: 19 | Double-stranded stuffer out of the E.Coli hisA-gene
5'-ccggtgcaggttggcggcggcgtgcgtagcgaaga-3' |

20. SEQ ID NO: 20 Oligonucleotide primer Loy 413-attB4*fwd-a
5'-GGGGACAACCTTTGTATAGAAAAGTTGGGTACCCGGGGATCCTCTA-3'
21. SEQ ID NO: 21 Oligonucleotide primer Loy 414 -attB1*rev-a
5'-GGGGACTGCTTTTTTTGTACAAACTTGCCATGATTACGCCAAGCTTGCA-3'
- 5 22. SEQ ID NO: 22 Oligonucleotide primer Loy447-attB4*fwd-a-inv
5'-GGGGACAACCTTTGTATAGAAAAGTTGCCATGATTACGCCAAGCTTGCA-3'
23. SEQ ID NO: 23 Oligonucleotide primer Loy448-attB1*rev-a-inv
5'-GGGGACTGCTTTTTTTGTACAAACTTGGGTACCCGGGGATCCTCTA -3'
- 10 24. SEQ ID NO: 24 Oligonucleotide primer Loy415-attB1*-fwd-b
5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGGTACCCGGGGATCCTCTA-3'
25. SEQ ID NO: 25 Oligonucleotide primer Loy416-attB2*rev-b
5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCCATGATTACGCCAAGC-
TTGCA-3'
- 15 26. SEQ ID NO: 26 Oligonucleotide primer Loy449-attB1*fwd-b-inv
5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGATTACGCCAAGC-
TTGCA-3'
27. SEQ ID NO: 27 Oligonucleotide primer Loy450-attB2*rev-b-inv
5'-GGGGACCACTTTGTACAAGAAAGCTGGGTGGTACCCGGGGATCCTCTA-3'
- 20 28. SEQ ID NO: 28 Oligonucleotide primer Loy417-attB2*fwd-c
5'-GGGGACAGCTTTCTTGTACAAAGTGGGGTACCCGGGGATCCTCTA-3'
29. SEQ ID NO: 29 Oligonucleotide primer Loy418-attB3*rev-c
5'-GGGGACAACCTTTGTATAATAAAGTTGCCATGATTACGCCAAGCTTGCA-3'
30. SEQ ID NO: 30 Oligonucleotide primer Loy451-attB2*fwd-c-inv
5'-GGGGACAGCTTTCTTGTACAAAGTGGCCATGATTACGCCAAGCTTGCA-3'
- 25 31. SEQ ID NO: 31 Oligonucleotide primer Loy452-attB3*rev-c-inv
5'-GGGGACAACCTTTGTATAATAAAGTTGGGTACCCGGGGATCCTCTA-3'
32. SEQ ID NO: 32 Insert Donor Molecule vector Lo391-pENTR-A1
33. SEQ ID NO: 33 Insert Donor Molecule vector Lo392-pENTR-B1
34. SEQ ID NO: 34 Insert Donor Molecule vector Lo393-pENTR-C1
- 30 35. SEQ ID NO: 35 Insert Donor Molecule vector Lo394-pENTR-A1-inv
36. SEQ ID NO: 36 Insert Donor Molecule vector Lo395-pENTR-B1-inv
37. SEQ ID NO: 37 Insert Donor Molecule vector Lo396-pENTR-C1-inv
38. SEQ ID NO: 38 Insert Donor Molecule vector Lo375-pENTR-A2
39. SEQ ID NO: 39 Insert Donor Molecule vector Lo376-pENTR-B2
- 35 40. SEQ ID NO: 40 Insert Donor Molecule vector Lo377-pENTR-C2
41. SEQ ID NO: 41 Insert Donor Molecule vector Lo397-pENTR-A2-inv
42. SEQ ID NO: 42 Insert Donor Molecule vector Lo398-pENTR-B2-inv
43. SEQ ID NO: 43 Insert Donor Molecule vector Lo399-pENTR-C2-inv
44. SEQ ID NO: 44 Insert Donor Molecule pENTR-A-USP::NAN::A7t

45. SEQ ID NO: 45 Insert Donor Molecule pENTR-B-L1-LuFad3-GUS-E9-L2
46. SEQ ID NO: 46 Insert Donor Molecule pENTR-C-R2-LeB4-700::GFP::LeB3t-L3
47. SEQ ID NO: 47 Oligonucleotide primer Loy277 attR4 EcoRV
5'-AAAAAAGATATCGATTACGCCAAGCTATCAACT-3'
- 5 48. SEQ ID NO: 48 Oligonucleotide primer Loy278 attR3 EcoRV
5'-AAAAAAGATATCCGGCCAGTGAATTATCAACT-3'
49. SEQ ID NO: 49 Insert Acceptor (Vector Donor) Molecule vector Lo338-pSUN2-GW-R4R3
- 10 50. SEQ ID NO: 50 Insert Acceptor (Vector Donor) Molecule vector Lo339-pSUN2-GW-R3R4
51. SEQ ID NO: 51 Multiple Expression Construct pSUN2-B4-USP-NAN-pa7-E9-GUS-FAD3-LeB4-GFP-LeB3
52. SEQ ID NO: 52 Recombination site attB1*
5'-AGCCTGCTTTTTTGTACAAACTTGC-3'
- 15 53. SEQ ID NO: 53 Recombination site attB2*
5'-ACCCAGCTTTCTTGTACAAAGTGGC
54. SEQ ID NO: 54 Recombination site attB3*
5'-CAACTTTATTATACATAGTTG-3'
- 20 55. SEQ ID NO: 55 Recombination site attB4*
5'-CAACTTTTCTATACAAAGTTG-3'
56. SEQ ID NO: 56 Recombination site attR1*
5'-GTTCAACTTTTTTGTACAAACTTGC-3'
57. SEQ ID NO: 57 Recombination site attR2*
5'-TTCAACTTTCTTGTACAAAGTGGG -3'
- 25 58. SEQ ID NO: 58 Recombination site attR3*
5'-GTTCAACTTTATTATACATAGTTGA -3'
59. SEQ ID NO: 59 Recombination site attR4*
5'-GTTCAACTTTTCTATACAAAGTTGA -3*
- 30 60. SEQ ID NO: 60 Recombination site attL1*
5'-AGCCTGCTTTTTTGTACAAAGTTGG -3'
61. SEQ ID NO: 61 Recombination site attL2*
5'-ACCCAGCTTTCTTGTACAAAGTTGG-3'
62. SEQ ID NO: 62 Recombination site attL3*
5'-GGCAACTTTATTATACAAAGTTGG -3'
- 35 63. SEQ ID NO: 63 Recombination site attL4*
5'-ACCCAAGTTTCTATACAAAGTTGG -3'
64. SEQ ID NO: 64 Recombination site attP1*
5'-GTTCAACTTTTTTGTACAAAGTTGG-3'
- 40 65. SEQ ID NO: 65 Recombination site attP2*
5'-GTTCAGCTTTCTTGTACAAAGTTGG-3'
66. SEQ ID NO: 66 Recombination site attP3*
5'-GTTCAACTTTATTATACAAAGTTGG-3'
67. SEQ ID NO: 67 Recombination site attP4*
5'-GTTCAACTTTTCTATACAAAGTTGG-3'

BRIEF DESCRIPTION OF THE DRAWINGS

Abbreviations in Figures have the following meaning:

	aadA	spectinomycin / streptomycin resistance gene
5	ccdB:	DNA gyrase inhibitor (counter selection marker)
	colE1:	origin of replication
	GUS:	β -glucuronidase (GUS) reporter gene
	LeB3'UT:	LeB-3'UT transcription terminator
	LB / RB:	Left (LB) and right (RB) border of Agrobacterium T-DNA
10	nan:	Nan reporter gene
	nosP:	nos promoter
	nosT:	nos transcription terminator
	nptII:	nptII kanamycin resistance gene
	pA7T:	35S transcription terminator pA7
15	rbcS E9 T:	rbcS (RUBISCO small subunit) E9 transcription terminator
	rrnBT1, rrnT2:	transcription terminator

Recombination sites are named starting with att (e.g., attP2*, attP1R*, attP3*, attP2R*, attL4*, attR1*, attL4*, attL1*, attL2*, attR2*, attL3*, attL4*, attR1* etc.)

20

Fig. 1A + B: Plasmid maps for modified pDONR vectors LO351-pDONR221-mod (I), LO348-pDONR-P4-P1R-mod (II), and LO347-pDON-P2R-P3-mod (III).

25

Fig. 2A + B: Plasmid maps for pENTR vectors LO391-pENTR-A1 (I), LO392-pENTR-B1 (II), and LO393-pENTR-C1 (III).

30

Fig. 3A + B: Plasmid maps for Insert Donor Molecule vectors pENTR A-L4-USP-NAN-pA7T-R1 (I), pENTR B-L1-E9-GUS-Fad3-L2 (II), and pENTR C-R2-LeB4-GFP-LeB3T-L3 (III).

35

Fig. 4: Plasmid maps for Insert Acceptor (Vector Donor) Molecule vectors Lo338-pSUN2-R4R3 (I), and Lo339-pSUN2-R3R4 (II).

Fig. 5: Plasmid maps for Multiple Expression Construct vector pSUN2-B4-USP-NAN-pA7-E9-GUS-Fad3-LeB4-GFP-LeB3.

40

Fig. 6: Reaction scheme for assembly of a Multiple Expression Cassette consisting of 2 expression cassettes (consisting of a nucleic acid on interest (N1, N2) under control of a promoter (p1, P2)). Two Insert Donor Molecules (I1 and I2) recombine with a Insert Acceptor Molecule (IA; Vector Donor = VD). The square, triangle, and circle are different sets of recombination sites (e.g., lox sites or att sites). Only a white "triangle" recombination site can recombine with a gray "triangle" site, a white "circu-

lar" site with a gray "circular" site and so on. The Insert Acceptor is preferably comprising a expression cassette for a counter selection maker (SC; under control of promoter P3). Recombination between the recombination sites leads to deletion of the counter selection marker. Only insertion of both expression cassettes is resulting a full deletion and recircularization of the Vector Donor Molecule.

Fig. 7: Reaction scheme for assembly of a Multiple Expression Cassette consisting of 3 expression cassettes (cassette 1,2,3). Insert Donor Molecules (pENTR1,2,3) recombined with Insert Acceptor Molecule (pSUN Gateway R4R3 = Vector Donor). Only recombination sites with the same end number on different molecules can recombine with each other (e.g. attR1 with attL1, attR2 with attL2 etc.). The Insert Acceptor is preferably comprising a ccdB expression cassette for a counter selection maker. Recombination between the recombination sites leads to deletion of the counter selection marker. Only insertion of all three expression cassettes is resulting a full deletion and recircularization of the Vector Donor Molecule.

Fig. 8: Scheme of the T-DNA region in the Multiple Expression Construct vector pSUN2-B4-USP-NAN-pA7-E9-GUS-Fad3-LeB4-GFP-LeB3. Expression cassettes are symbolized by the hexagons.

25 EXAMPLES

Chemicals

Unless indicated otherwise, chemicals and reagents in the Examples were obtained from Sigma Chemical Company (St. Louis, MO), restriction endonucleases were from New England Biolabs (Beverly, MA) or Roche (Indianapolis, IN), oligonucleotides were synthesized by MWG Biotech Inc. (High Point, NC), and other modifying enzymes or kits regarding biochemicals and molecular biological assays were from Clontech (Palo Alto, CA), Pharmacia Biotech (Piscataway, NJ), Promega Corporation (Madison, WI), or Stratagene (La Jolla, CA). Materials for cell culture media were obtained from Gibco/BRL (Gaithersburg, MD) or DIFCO (Detroit, MI). The cloning steps carried out for the purposes of the present invention, such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linking DNA fragments, transformation of *E. coli* cells, growing bacteria, multiplying phages and sequence analysis of recombinant DNA, are carried out as described by Sambrook (1989) (Sambrook *et al.* (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6). The sequencing of recombinant DNA molecules is carried out using ABI laser fluorescence DNA sequencer following the method of Sanger (Sanger *et al.* (1977) Proc Natl Acad Sci USA 74:5463-5467).

45 Buffers: 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. 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. One preferred buffer for P1 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. The buffer for other site-specific recombinases which are similar to lambda Int and P1 Cre are either known in the art or can be determined empirically by the skilled artisans, particularly in light of the above-described buffers.

10 **EXAMPLE 1: *Agrobacterium*-mediated transformation in dicotyledonous and monocotyledonous plants**

1.1: Transformation and regeneration of transgenic *Arabidopsis thaliana* (Columbia) plants

15 To generate transgenic *Arabidopsis* plants, *Agrobacterium tumefaciens* (strain C58C1 pGV2260) is transformed with various ptxA or SbHRGP3 promoter/GUS vector constructs. The agrobacterial strains are subsequently used to generate transgenic plants. To this end, a single transformed *Agrobacterium* colony is incubated overnight at 28°C in a 4 mL culture (medium: YEB medium with 50 µg/mL kanamycin and 25 µg/mL rifampicin). This culture is subsequently used to inoculate a 400 mL culture in the same medium, and this is incubated overnight (28°C, 220 rpm) and spun down (GSA rotor, 8,000 rpm, 20 min). The pellet is resuspended in infiltration medium (1/2 MS medium; 0.5 g/L MES, pH 5.8; 50 g/L sucrose). The suspension is introduced into a plant box (Duchefa), and 100 mL of SILWET L-77 (heptamethyltrisiloxan modified with polyalkylene oxide; Osi Specialties Inc., Cat. P030196) was added to a final concentration of 0.02%. In a desiccator, the plant box with 8 to 12 plants is exposed to a vacuum for 10 to 15 minutes, followed by spontaneous aeration. This is repeated twice or 3 times. Thereupon, all plants are planted into flowerpots with moist soil and grown under long-day conditions (daytime temperature 22 to 24°C, nighttime temperature 19°C; relative atmospheric humidity 65%). The seeds are harvested after 6 weeks.

As an alternative, transgenic *Arabidopsis* plants can be obtained by root transformation. White root shoots of plants with a maximum age of 8 weeks are used. To this end, plants which are kept under sterile conditions in 1 MS medium (1% sucrose; 100mg/L inositol; 1.0 mg/L thiamine; 0.5 mg/L pyridoxine; 0.5 mg/L nicotinic acid; 0.5 g MES, pH 5.7; 0.8 % agar) are used. Roots are grown on callus-inducing medium for 3 days (1x Gamborg's B5 medium; 2% glucose; 0.5 g/L mercaptoethanol; 0.8% agar; 0.5 mg/L 2,4-D (2,4-dichlorophenoxyacetic acid); 0.05 mg/L kinetin). Root sections 0.5 cm in length are transferred into 10 to 20 mL of liquid callus-inducing medium (composition as described above, but without agar supplementation), inoculated with 1 mL of the above-described overnight agrobacterial culture (grown at 28°C, 200 rpm in LB) and shaken for 2 minutes. After excess medium has been allowed to run off, the root explants are transferred to callus-inducing medium with agar, subsequently to callus-inducing liquid medium without agar (with 500 mg/L betabactyl, SmithKline Beecham Pharma GmbH, Munich), incubated with shaking and finally transferred to shoot-inducing medium (5 mg/L 2-isopentenyladenine phosphate; 0.15 mg/L indole-3-acetic acid; 50 mg/L kanamycin; 500 mg/L betabactyl). After 5 weeks, and after 1 or 2 medium changes, the small green shoots are transferred to germination medium (1 MS

medium; 1% sucrose; 100 mg/L inositol; 1.0 mg/L thiamine; 0.5 mg/L pyridoxine; 0.5 mg/L nicotinic acid; 0.5 g MES, pH 5.7; 0.8% agar) and regenerated into plants.

1.2: Transformation and regeneration of crop plants

5 The *Agrobacterium*-mediated plant transformation using standard transformation and regeneration techniques may also be carried out for the purposes of transforming crop plants (Gelvin & Schilperoort (1995) Plant Molecular Biology Manual, 2nd Edition, Dordrecht: Kluwer Academic Publ. ISBN 0-7923-2731-4; Glick & Thompson (1993) Methods in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, ISBN
10 0-8493-5164-2).

For example, oilseed rape can be transformed by cotyledon or hypocotyl transformation (Moloney *et al.* (1989) Plant Cell Reports 8: 238-242; De Block *et al.* (1989) Plant Physiol 91:694-701). The use of antibiotics for the selection of *Agrobacteria* and plants
15 depends on the binary vector and the *Agrobacterium* strain used for the transformation. The selection of oilseed rape is generally carried out using kanamycin as selectable plant marker.

The *Agrobacterium*-mediated gene transfer in linseed (*Linum usitatissimum*) can be carried out using for example a technique described by Mlynarova *et al.* (1994) Plant
20 Cell Report 13:282-285.

The transformation of soybean can be carried out using, for example, a technique described in EP-A1 0424 047 or in EP-A1 0397 687, US 5,376,543, US 5,169,770.
25

The transformation of maize or other monocotyledonous plants can be carried out using, for example, a technique described in US 5,591,616.

The transformation of plants using particle bombardment, polyethylene glycol-mediated DNA uptake or via the silicon carbonate fiber technique is described, for example, by
30 Freeling & Walbot (1993) "The maize handbook" ISBN 3-540-97826-7, Springer Verlag New York).

EXAMPLE 2: Generation of Insert Donor Molecules (pENTRs)

35 Aim of the here described pENTR-construction is to provide a Insert Donor molecules (shuttle system) for three independent expression cassettes as directed single insertions into a *Agrobacterium* binary vector as a Vector Donor (Insert Acceptor).

The pENTRs are generated by a BP recombination reaction of a PCR-product with specific attachment sites at the end into pDONRTM vectors of the GATEWAYTM Cloning system (Invitrogen). For this purpose, the pDONRTM vectors were modified by introducing a multiple cloning site grouped around a central NotI site. This grants the possibility
40 to introduce not only a cDNA into the NotI site but to clone the 5'element and the 3'element into the surrounding restriction sites.

45

2.1 Construction of a Multiple cloning site

First step is to build up the multiple cloning site (MCS) in a pUC19 vector. For this purpose the primer Loy344 was annealed to the primer Loy345 to create a double stranded adapter sequence.

5

Loy344 (SEQ ID NO: 17):

5'-p-GTCGACCAGATCTGATATCTGCGGCCGCCTCGAGCATATG-3'

Loy345 (SEQ ID NO: 18):

10 5'-p-GTCGACCAGATCTGATATCTGCGGCCGCCTCGAGCATATG-3'

The vector pUC19 is cleaved with XbaI and PstI, the overhangs converted to blunt ends with Pfu polymerase, and dephosphorylated. The phosphorylated adapter is ligated into the blunted vector resulting in the alternate orientations:

15

-UNE1 : KpnI SmaI BamHI NdeI XhoI NotI EcoRV BglII SalI SphI HindIII

-UNE2 : KpnI SmaI BamHI SalI BglII EcoRV NotI XhoI NdeI SphI HindIII

20 The resulting vectors are named Lo372-pUC-Polylinker-UNE1 and Lo373-pUC-Polylinker-UNE2, respectively .

2.2 Modification of the pDONRTM-vectors provided by Invitrogen

25 Cleavage sites for certain restriction endonucleases in the vector backbone outside the recombination region in Invitrogen vectors pDONRTM -P4P1R, pDONRTM 221 and pDONRTM P2RP3 were eliminated to facilitate subsequent cloning steps: In pDONRTM -P4P1R and pDONRTM P2RP3 the EcoRV and the NotI site were eliminated by cutting with these enzymes, blunting (fill-in) with Pfu polymerase, and religating the vectors - ending up with pDONR -P4P1R-mod (A) and pDONR-P2RP3-mod (C) respectively. In 30 pDONRTM 221 the EcoRV site in the backbone was eliminated by cleaving with said enzyme, dephosphorylation, and inserting a double-stranded stuffer out of the E.Coli hisA-gene (5'-ccggtgcaggttgccggcggcgtgcgtagcgaaga-3'; SEQ ID NO: 19). The resulting vector is named pDONR221-mod (C).

2.3 Generation of att-PCR-products

For amplification of the new MCS, PCR primers were designed which included recombinatorial attachment sites. The PCR reactions were carried out on Lo372-pUC-Polylinker-UNE1 and Lo373-pUC-Polylinker-UNE2 .

2.3.1 Generation of attB-PCR-A products:

Both Lo372-pUC-Polylinker-UNE1 and Lo373-pUC-Polylinker-UNE2 were separately amplified with

1) primer Loy 413-attB4*fwd-a and Loy 414 -attB1*rev-a

45 2) Loy447-attB4*fwd-a-inv and Loy448-attB1*rev-a-inv

The resulting two PCR products were recombined separately into the modified pDONR vector Lo347-pDONR-P2R-P3-mod (A) resulting in four pENTR vectors indicated in Table 1 below.

- 5 Loy 413-attB4*fwd-a (SEQ ID NO: 20):
 5'-GGGGACAACCTTTGTATAGAAAAGTTGGGTACCCGGGGATCCTCTA-3'
 Loy 414 -attB1*rev-a (SEQ ID NO: 21):
 5'-GGGGACTGCTTTTTTTGTACAAACTTGCCATGATTACGCCAAGCTTGCA-3'
 Loy447-attB4*fwd-a-inv (SEQ ID NO: 22)
 10 5'-GGGGACAACCTTTGTATAGAAAAGTTGGCATGATTACGCCAAGCTTGCA-3'
 Loy448-attB1*rev-a-inv (SEQ ID NO: 23):
 5'-GGGGACTGCTTTTTTTGTACAAACTTGGGTACCCGGGGATCCTCTA -3'

2.3.2 Generation of attB-PCR-B products:

- 15 Both Lo372-pUC-Polylinker-UNE1 and Lo373-pUC-Polylinker-UNE2 were separately amplified with
- 1) primer Loy415-attB1*-fwd-b and Loy416-attB2rev-b
 2) primer Loy449-attB1*fwd-b-inv and Loy450-attB2rev-b-inv

20

The resulting two PCR products were recombined separately into the modified pDONR vector Lo351-pDONR221-mod (B) resulting in four pENTR vectors indicated in Table 1 below.

- 25 Loy415-attB1*-fwd-b (SEQ ID NO: 24):
 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGGTACCCGGGGATCCTCTA-3'
 Loy416-attB2*rev-b (SEQ ID NO: 25):
 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCCATGATTACGCCAAGCTTGCA-3'
 primers Loy449-attB1*fwd-b-inv (SEQ ID NO: 26):
 30 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGATTACGCCAAGCTTGCA-3'
 Loy450-attB2*rev-b-inv (SEQ ID NO: 27):
 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTGGTACCCGGGGATCCTCTA-3'

2.3.3 Generation of attB-PCR-C products:

- 35 Both Lo372-pUC-Polylinker-UNE1 and Lo373-pUC-Polylinker-UNE2 were separately amplified with
- 1) primer Loy417-attB2*fwd-c and Loy418-attB3rev-c
 2) primer Loy451-attB2*fwd-c-inv and Loy452-attB3rev-c-inv

40

The resulting two PCR products were recombined separately into the modified pDONR vector Lo348-pDONR-P4P1R-mod (C) resulting in four pENTR vectors indicated in Table 1 below.

- 45 Loy417-attB2*fwd-c (SEQ ID NO: 28):
 5'-GGGGACAGCTTTCTTGTACAAAGTGGGGTACCCGGGGATCCTCTA-3'
 Loy418-attB3*rev-c (SEQ ID NO: 29):
 5'-GGGGACAACCTTTGTATAATAAAGTTGCCATGATTACGCCAAGCTTGCA-3'

Loy451-attB2*fwd-c-inv (SEQ ID NO: 30):

5'-GGGGACAGCTTTCTTGTACAAAGTGGCCATGATTACGCCAAGCTTGCA-3'

Loy452-attB3*rev-c-inv (SEQ ID NO: 31)

5'-GGGGACAACCTTTGTATAATAAAGTTGGGTACCCGGGGATCCTCTA-3'

5

2.4 Generation of a set of pENTRs with new MCS

To generate pENTR vectors the PCR-products obtained under 2.3 above (consisting of the multiple cloning site flanked by attachment sides for BP reaction with the modified pDONR vectors) are recombined in the respective modified pDONR vector as shown in the following Table 1. Recombination reactions were carried out according to the instructions of the manufacturer (Invitrogen).

10

pDONR-vector (Educt 1)	recombined with PCR-product (Educt 2)	resulting pENTR (Product)	site-setup of pENTR
Lo347-pDONR-P2R-P3-mod (A)	attb4-fwd-a-UNE1-attB1*-rev-a	Lo391-pENTR-A1 SEQ ID NO: 32	attL4* UNE2 attR1*
Lo347-pDONR-P2R-P3-mod (A)	attb4-fwd-a-inv-UNE1-attB1*-rev-a-inv	Lo394-pENTR-A1-inv SEQ ID NO: 35	attL4* UNE2-inv attR1*
Lo347-pDONR-P2R-P3-mod (A)	attb4-fwd-a-UNE2-attB1*-rev-a	Lo375-pENTR-A2 SEQ ID NO: 38	attL4* UNE1 attR1*
Lo347-pDONR-P2R-P3-mod (A)	attb4-fwd-a-inv-UNE2-attB1*-rev-a-inv	Lo397-pENTR-A2-inv SEQ ID NO: 41	attL4* UNE1-inv attR1*
Lo351-pDONR221-mod (B)	attB1-fwd-b-UNE2-attB2*-rev-b	Lo376-pENTR-B2 SEQ ID NO: 39	attL1*-UNE2-attL2*
Lo351-pDONR221-mod (B)	attB1-fwd-b-inv-UNE2-attB2*-rev-b-inv	Lo398-pENTR-B2-inv SEQ ID NO: 42	attL1*-UNE2-inv-attL2*
Lo351-pDONR221-mod (B)	attB1-fwd-b-UNE1-attB2*-rev-b	Lo392-pENTR-B1 SEQ ID NO: 33	attL1*-UNE1-attL2*
Lo351-pDONR221-mod (B)	attB1-fwd-b-inv-UNE1-attB2*-rev-b-inv	Lo395-pENTR-B1-inv SEQ ID NO: 36	attL1*-UNE1-inv-attL2*
Lo348-pDONR-P4P1R-mod (C)	attb2-fwd-c-UNE2-attB3*rev-c	Lo377-pENTR-C2 SEQ ID NO: 40	attR2*-UNE2-attL3*
Lo348-pDONR-P4P1R-mod (C)	attb2-fwd-c-inv-UNE2-attB3*rev-c-inv	Lo399-pENTR-C2-inv SEQ ID NO: 43	attR2*-UNE2-inv-attL3*
Lo348-pDONR-P4P1R-mod (C)	attb2-fwd-c-UNE1-attB3*rev-c	Lo393-pENTR-C1 SEQ ID NO: 34	attR2*-UNE1-attL3*
Lo348-pDONR-P4P1R-mod (C)	attb2-fwd-c-inv-UNE1-attB3*rev-c-inv	Lo396-pENTR-C1-inv SEQ ID NO: 37	attR2*-UNE1-inv-attL3*

Table: 1: Summery of the constrution of various pENTR vectors for insertion of expression cassettes. Modified pDONR vectors (Educt 1) are recombined with a multiple cloning site (UNE1 or 2, respectively) flanked with recombination sites (Educt 2) yielding in various pENTR vectors (Product) comprising the indicated recombination sites flanking the multiple cloning site.

15

3. Providing the Insert Donor Molecules

3.1 Cloning of pENTR-A-USP::NAN::A7t

A expression cassette consisting of the USP promoter of the unknown seed protein (*Vicia faba*), the NAN reporter (a codon optimised version of *nanH* of *Clostridium perfringens*, Kirby J et al. (2002) Plant J. 32(3):391-400; WO 03/052104), and the A7t terminator (Hirt H et al (1990) Curr Genet 17: 473-479; derivable from vector pCAM-BIA-1300; GenBank Acc.-No. AF234296) is inserted into Lo375-pENTR-A2 resulting in the vector pENTR-A-USP::NAN::A7t (SEQ ID NO: 44).

25

3.2 Cloning of pENTR-B-L1-LuFad3-GUS-E9-L2

An expression cassette consisting of the *Linum usitatissimum* promotor Fad3 (GenBank Acc.-No.: AX771967; Sequence 11 from WO 02/102970), the GUS (glucuronidase) gene, and the E9 terminator (terminator of small subunit of RUBP from pea; Genbank Acc.-No.: X00806) was inserted into Lo394-pENTR-A1-inv resulting in pENTR-B-L1-LuFad3-GUS-E9-L2 (SEQ ID NO: 45)..

3.3 Cloning of pENTR-C-R2-LeB4-700::GFP::LeB3t-L3

A expression cassette consisting of the *Vicia faba* LeB4-700 promoter, the *Aequorea victoria* gene for the green fluorescent protein GFP5er, and the *Vicia faba* LeguminB3 terminator was inserted into Lo391-pENTR-A1 resulting in pENTR-C-R2-LeB4-700::GFP::LeB3t-L3 (SEQ ID NO: 46).

4. Providing the Insert Acceptor (Vector Donor) Molecule

4.1 Cloning of a binary plant transformation pDEST, pSUN2-GW R4R3

The DNA fragment flanked the Gateway™ attachment sites R4 and R3 is amplified via PCR using the primers Lo277 and Lo278.

Lo277 attR4* EcoRV (SEQ ID NO: 47):

5'-AAAAAAGATATCGATTACGCCAAGCTATCAACT-3'

Lo278 attR3* EcoRV (SEQ ID NO: 48):

5'-AAAAAAGATATCCGGCCAGTGAATTATCAACT-3'

The DNA fragment is inserted into pTopo generating pTopo R4R3. The binary vector pSUN2 is opened with HindIII/EcoRI, blunt DNA ends are generated by a fill in reaction with Pfu polymerase. The DNA fragment containing the DNA region with the flanking attachment sites is isolated from pTOPO R4R3 by EcoRV digestion and ligated to the linearized pSUN2 vector fragment. The resulting destination vectors are named Lo338-pSUN2-GW-R4R3 (SEQ ID NO: 49) and Lo339-pSUN2-GW-R3R4 (inverse orientation of insert; SEQ ID NO: 50). These vectors are comprising the ccdB counter-selection marker and would cause toxic effects on standard E.coli strains (like DH5α). Therefore, these plasmids have to be transformed and propagated in E.coli strains like DB2 or DB3.1. DB2 or DB3.1 cells contain the gyrA462 mutation and are hence insensitive against the gyrase inhibitor ccdB.

5. LR reaction to create the binary Expression Vector pSUN2-GW-R3R4

The LR recombination reaction is carried out with the plasmids pENTR-C-R2-LeB4-700::GFP5er::LeB3t-attL3, pENTR-B-L1-E9-Gus-LuFad3-L2, pENTR-A-L4-USP::NAN::A7t-R1, pSUN2-GW-R3R4 employing LR Clonase™ Plus-Enzym (Invitrogen).

The recombinase reaction is carried out according to the manufacturers manual. Briefly, the enzyme and buffer are kept on - 80°C until preparation of the LR reactions. The plasmids used for LR reaction (pENTRs and pDEST) are diluted with TE-Puffer to the final concentration needed (20-25 fmol plasmid/LR reaction; a maximum

78

of 30 fmol/plasmid should not be exceeded, the total amount of DNA present in the LR reaction should be less than 250 ng DNA). The μg -amount corresponding to the molar amount to be employed can be calculated using the DNA (plasmid) size:

$$5 \quad C_n [\text{fmol}/\mu\text{l}] = \frac{C_m [\mu\text{g}/\mu\text{l}] \times 1000}{0,66 \times \text{plasmid size}[\text{kb}]}$$

The LR reaction is prepared follow order as shown in tabel

	Size [kb]	C_m [$\mu\text{g}/\mu\text{l}$]	dilution factor	final C_n [fmol/ μl]	LR-reactions [μL]		
					reaction	pos. control	neg. control
pSUN2 R4R3	9.2	1	1:7.5	~22	1	1	1
pENTR-A-L4- USP::NAN::A7t-R1	4.8	1	1:14.5	~22	1	0	0
Lo391-pENTR-A1	2.7	0.05	-	~22	0	1	0
pENTR-B-L1-E9- Gus-LuFad3-L2	6.3	1	1:11	~22	1	0	0
Lo392-pENTR-B2	2.7	0.05	-	~22		1	
pENTR-C-R2- LeB4- 700::GFP5er::LeB 3t-attL3	4.6	1	1:15	~22	1	0	0
Lo393-pENTR-C1	2.7	0.05	-	~22	0	1	0
5xLR-Puffer	-	-	-	-	4	4	4
TE	-	-	-	-	8	8	11
Summe					16 μl	16 μl	16 μl

- 10 The LR-Clonase™-Enzyme mix is placed on dry ice for transport. For use, the LR mix is thawed on ice for approx. 2 min, and gently vortexed for 2 seconds. 4 μl LR Clonase™ mix are added to each reaction tube, gently vortexed 2 x 2 sec., and incubated over night (16 hrs) at 25°C (small table incubator is sufficient). 2 μl proteinase K (2 $\mu\text{g}/\mu\text{l}$) are added to terminate the reaction, and incubated for 10 min at 37°C.

- 15 2 μl of the reaction mixture are used for transformation into E.coli DH5 α cells. The transformed cells are plated onto LB medium containing the appropriate selection marker. For pSUN2 derivatives the selection marker is streptomycin (100 $\mu\text{g}/\text{ml}$). Because DH5 α is a ccdB sensitive strain, only E.coli colonies will grow which comprise plasmids based on the Vector Donor Molecule but are lacking ccdB segment of said molecule in consequence of a recombination mediated replacement with the Inserts.
- 20

- For analysis of clones several colonies are picked for overnight culture. Plasmid DNA is prepared according to standard manuals. Correct assembly of the Multiple Expression Cassette is verified by thorough restriction analyses and sequencing of the clones. The resulting product pSUN2-B4-USP-NAN-pa7-E9-GUS-FAD3-LeB4-GFP-LeB3 (SEQ ID NO: 51).
- 25

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 78

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 78

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

What is claimed is:

1. A method for producing an Multiple Expression Construct comprising at least two different expression cassettes said method comprising combining in vitro or in vivo
 - i) one or more Insert Donor molecules, said Insert Donor molecules together comprising at least two Inserts I(n), each Insert comprising at least one expression cassette, said Inserts being flanked by two different recombination sites A(2n) and A(2n+1), wherein n is an integer from 1 to m characterizing each Insert, and m is the total number of different Inserts, and
 - ii) a Insert Acceptor molecule IA comprising two different recombination sites A1 and A(2m+2), wherein m is the total number of different Inserts of step i)

wherein all recombination sites A(1) to A(2m+2) are different, and wherein a recombination side A(2i-1) for a specific i, said i being an integer from 1 to m+1, can recombine with the recombination side A(2i) for the same i, but does not substantially recombine with another recombination site, and

 - iii) at least one site specific recombination protein capable of recombining said recombination sites in said Insert Donor molecules and said Insert Acceptor molecule, and

incubating said combination under conditions sufficient to transfer all of said Inserts into said Insert Acceptor molecule, thereby producing a Multiple Expression Construct.
2. The method as claimed in claim 1, wherein each Insert is comprised in a separate Insert Donor Molecule.
3. The method as claimed in claim 1 or 2, wherein at least one of said Insert Donor DNA Molecules and said Insert Acceptor Molecule is comprised of a circular DNA molecule.
4. The method as claimed in any of claim 1 to 3, wherein at least one of said Insert Donor DNA molecule and said Insert Acceptor Molecule is comprised of a linear DNA molecule.
5. The method of any of claim 1 to 4, wherein said Insert Acceptor Molecule and/or said Insert Donor Molecule and/or said Multiple Expression Construct comprise prokaryotic and/or eukaryotic vectors.
6. The method of claim 5, wherein said eukaryotic vectors comprise vectors which replicate in yeast cells, plant cells, fish cells, eukaryotic cells, mammalian cells, or insect cells.

7. The method of claim 5, wherein said prokaryotic vectors comprise vectors which replicate in gram negative or gram positive bacteria.
- 5 8. The method of claim 5 or 7, wherein said prokaryotic vectors comprise vectors which replicate in bacteria of the genus *Escherichia*, *Salmonella*, *Bacillus*, *Streptomyces*, *Agrobacterium*, or *Pseudomonas*.
- 10 9. The method of any of claim 5, 7, or 8, wherein said prokaryotic vector comprises a vector which replicates in both *E. coli* and *Agrobacterium*.
- 15 10. The method as claimed in any of claim 1 to 9, wherein said Insert Acceptor Molecule and/or said Insert Donor Molecule and/or said Multiple Expression Construct is a vector comprising at least one selectable marker.
- 20 11. The method as claimed in any of claim 1 to 10, wherein the Insert Acceptor Molecule further comprises (a) a toxic gene and (b) a selectable marker, wherein said toxic gene and said selectable marker are on different DNA segments, the DNA segments being separated from each other by at least two recombination sites.
- 25 12. The method of claim 10 or 11, wherein the selectable marker comprises at least one DNA segment selected from the group consisting of:
 - (a) a DNA segment that encodes a product that provides resistance in a recipient cell against otherwise toxic compounds;
 - (b) a DNA segment that encodes a product that is otherwise lacking in a recipient cell;
 - (c) a DNA segment that encodes a product that suppresses the activity of a gene product in a recipient cell;
 - (d) a DNA segment that encodes a product that can be identified;
 - 30 (e) a DNA segment that encodes a product that inhibits a cell function in a recipient cell;
 - (f) a DNA segment that inhibits the activity of any of the DNA segments of (a)-(e) above;
 - (g) a DNA segment that binds a product that modifies a substrate;
 - (h) a DNA segment that encodes a specific nucleotide recognition sequence which can be recognized by a protein, an RNA, DNA or chemical,
 - 35 (i) a DNA segment that, when deleted, directly or indirectly confers sensitivity to cell killing by particular compounds within a recipient cell;
 - (j) a DNA segment that encodes a product that is toxic in a recipient cell; and
 - (k) a DNA segment that can be used to isolate or identify a desired molecule.
- 40 13. The method of any of claim 10 to 12, wherein said selectable marker comprises at least one marker selected from the group consisting of an antibiotic resistance gene, a herbicide resistance gene, a tRNA gene, an auxotrophic marker, a toxic gene, a phenotypic marker, an antisense oligonucleotide, a restriction endonucle-

ase, a restriction endonuclease cleavage site, an enzyme cleavage site, a protein binding site, and a sequence complementary to a PCR primer sequence.

14. The method according to any of claim 1 to 13, further comprising the step of selecting the Multiple Expression Construct comprising all of said expression cassettes of said Inserts.
15. The method of any of claim 1 to 14, wherein said recombination sites are selected from the group consisting of loxP, attB, attP, attL, and attR.
16. The method of claim 15, wherein said recombination site comprises a DNA sequence selected from the group consisting of:

(a) RKYCWGCTTTYKTRTACNAASTSGB	(m-att)	(SEQ ID NO:1);
(b) AGCCWGCTTTYKTRTACNAACTSGB	(m-attB)	(SEQ ID NO:2);
(c) GTTCAGCTTTCKTRTACNAACTSGB	(m-attR)	(SEQ ID NO:3);
(d) AGCCWGCTTTCKTRTACNAAGTSGB	(m-attL)	(SEQ ID NO:4);
(e) GTTCAGCTTTYKTRTACNAAGTSGB	(m-attP1)	(SEQ ID NO:5);

 and 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, C, or G or T/U; S=C or G; and B=C, G or T/U.
17. The method of claim 15 or 16, wherein said DNA sequence comprises a sequence selected from the group consisting of:

(a) AGCCTGCTTTTTTGTACAAACTTGT	(attB1)	(SEQ ID NO:6);
(b) AGCCTGCTTTCTTGTACAAACTTGT	(attB2)	(SEQ ID NO:7);
(c) ACCCAGCTTTCTTGTACAAACTTGT	(attB3)	(SEQ ID NO:8);
(d) GTTCAGCTTTTTTGTACAAACTTGT	(attR1)	(SEQ ID NO:9);
(e) GTTCAGCTTTCTTGTACAAACTTGT	(attR2)	(SEQ ID NO:10);
(f) GTTCAGCTTTCTTGTACAAAGTTGG	(attR3)	(SEQ ID NO:11);
(g) AGCCTGCTTTTTTGTACAAAGTTGG	(attL1)	(SEQ ID NO:12);
(h) AGCCTGCTTTCTTGTACAAAGTTGG	(attL2)	(SEQ ID NO:13);
(i) ACCCAGCTTTCTTGTACAAAGTTGG	(attL3)	(SEQ ID NO:14);
(j) GTTCAGCTTTTTTGTACAAAGTTGG	(attP1)	(SEQ ID NO:15);
(k) GTTCAGCTTTCTTGTACAAAGTTGG	(attP2,P3)	(SEQ ID NO:16);
(l) AGCCTGCTTTTTTGTACAAACTTGC	(attB1*)	(SEQ ID NO:52);
(m) ACCCAGCTTTCTTGTACAAAGTGGC	(attB2*)	(SEQ ID NO:53);
(n) CAACTTTATTATACATAGTTG	(attB3*)	(SEQ ID NO:54)
(o) CAACTTTTCTATACAAAGTTG	(attB4*)	(SEQ ID NO:55)
(p) GTTCAACTTTTTTGTACAAACTTGC	(attR1*)	(SEQ ID NO:56)
(q) TTCAACTTTCTTGTACAAAGTGGG	(attR2*)	(SEQ ID NO:57)
(r) GTTCAACTTTATTATACATAGTTGA	(attR3*)	(SEQ ID NO:58)
(s) GTTCAACTTTTCTATACAAAGTTGA	(attR4*)	(SEQ ID NO:59)

82

- (t) AGCCTGCTTTTTTTGTACAAAGTTGG (attL1*) (SEQ ID NO:60)
 (u) ACCCAGCTTTCTTGTACAAAGTTGG (attL2*) (SEQ ID NO:61)
 (v) GGCAACTTTATTATACAAAGTTGG (attL3*) (SEQ ID NO:62)
 (w) ACCCAACTTTTCTATACAAAGTTGG (attL4*) (SEQ ID NO:63)
 5 (x) GTTCAACTTTTTTTGTACAAAGTTGG (att P1*) (SEQ ID NO:64)
 (y) GTTCAGCTTTCTTGTACAAAGTTGG (attP2*) (SEQ ID NO:65)
 (z) GTTCAACTTTATTATACAAAGTTGG (att P3*) (SEQ ID NO:66)
 (aa) GTTCAACTTTTCTATACAAAGTTGG (attP4*) (SEQ ID NO:67)

and a corresponding or complementary DNA or RNA sequence.

10

18. The method of any of claim 1 to 17, wherein said recombination proteins are selected from the group consisting of Int, IHF, Xis, Cre, Flp, and Res.
- 15 19. The method of any of claim 1 to 18, wherein said expression cassette comprised in said Insert in comprising at least a promoter sequences and a nucleic acid sequence of interest to be expressed operably linked to said promoter sequence,
- 20 20. The method of claim 19, wherein said promoter is selected from the group of promoters able to initiate transcription in eukaryotic or prokaryotic cells or organisms.
- 20 21. The method of claim 19 or 20, wherein said promoter is a promoter able to initiate transcription in plant cells.
- 25 22. The method of any of claim 19 to 21, wherein said nucleic acid of interest operably linked to said promoter is selected from the group of sequences encoding for
 a) sense, antisense, or double-stranded RNA,
 b) polypeptides or proteins.
- 30 23. The method of any of claim 19 to 21, wherein said expression cassettes are able to confer to a transgenic plant comprising said expression cassettes at least one phenotype selected from the group consisting of increased nutritional value, increased oil content, increased starch content, increased protein content, increased vitamin content, increased carotinoid content, increased pathogen resistance, modified oil composition, and increased stress tolerance.
- 35 24. A Multiple Expression Construct obtained by the method of any of Claim 1 to 23.
- 40 25. The Multiple Expression Construct of claim 24, wherein said Multiple Expression Construct is comprising at least two Inserts I(n), each Insert comprising at least one expression cassette, and wherein each Insert is flanked by a sequence resulting from the recombination of a recombination side A(2i-1) with the recombination side A(2i) for the same i, wherein said Inserts I(n) and said recombination sites are defined as in any of claim 1 to 19.

26. The Multiple Expression Construct of claim 24 or 25, wherein said Multiple Expression Construct is a vector as defined in any of claim 5 to 13.
- 5 27. A transgenic cell or non-human organism comprising a Multiple Expression Construct of any of Claim 24 to 26.
- 10 28. The transgenic cell or non-human organism of claim 27, wherein said cell or organism is selected from the group comprising prokaryotic and eukaryotic cells or organism.
29. The transgenic cell or non-human organism of claim 27 o 28, wherein said cell or organism is selected from the group comprising of plant cells or organism.
- 15 30. The method of producing food or feed products, pharmaceuticals, or chemicals, said method comprising providing a transgenic cell or organism of any of claim 27 to 29, growing said cell or organism, and – optionally – isolating said food or feed product, pharmaceutical, or chemical.

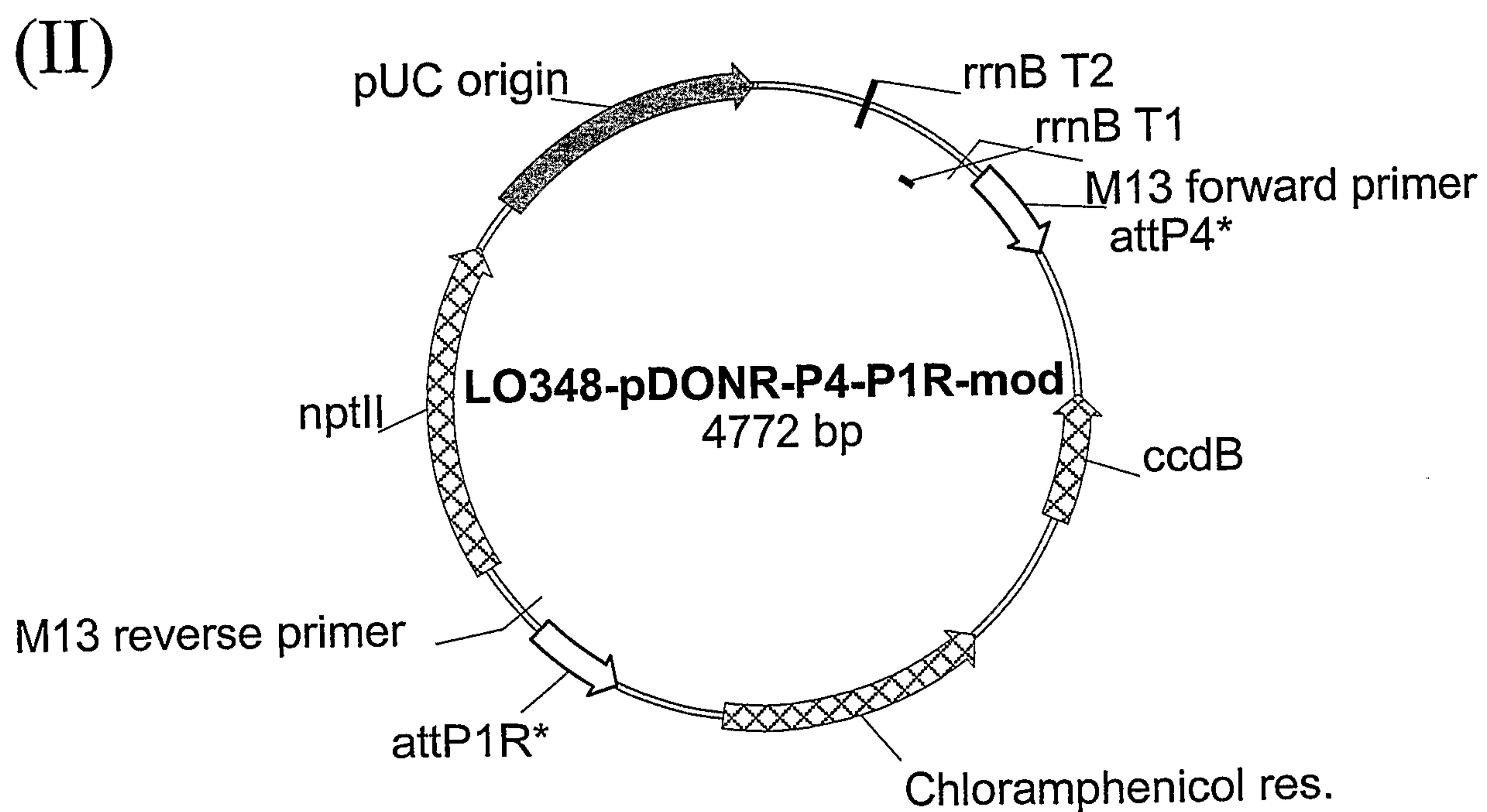
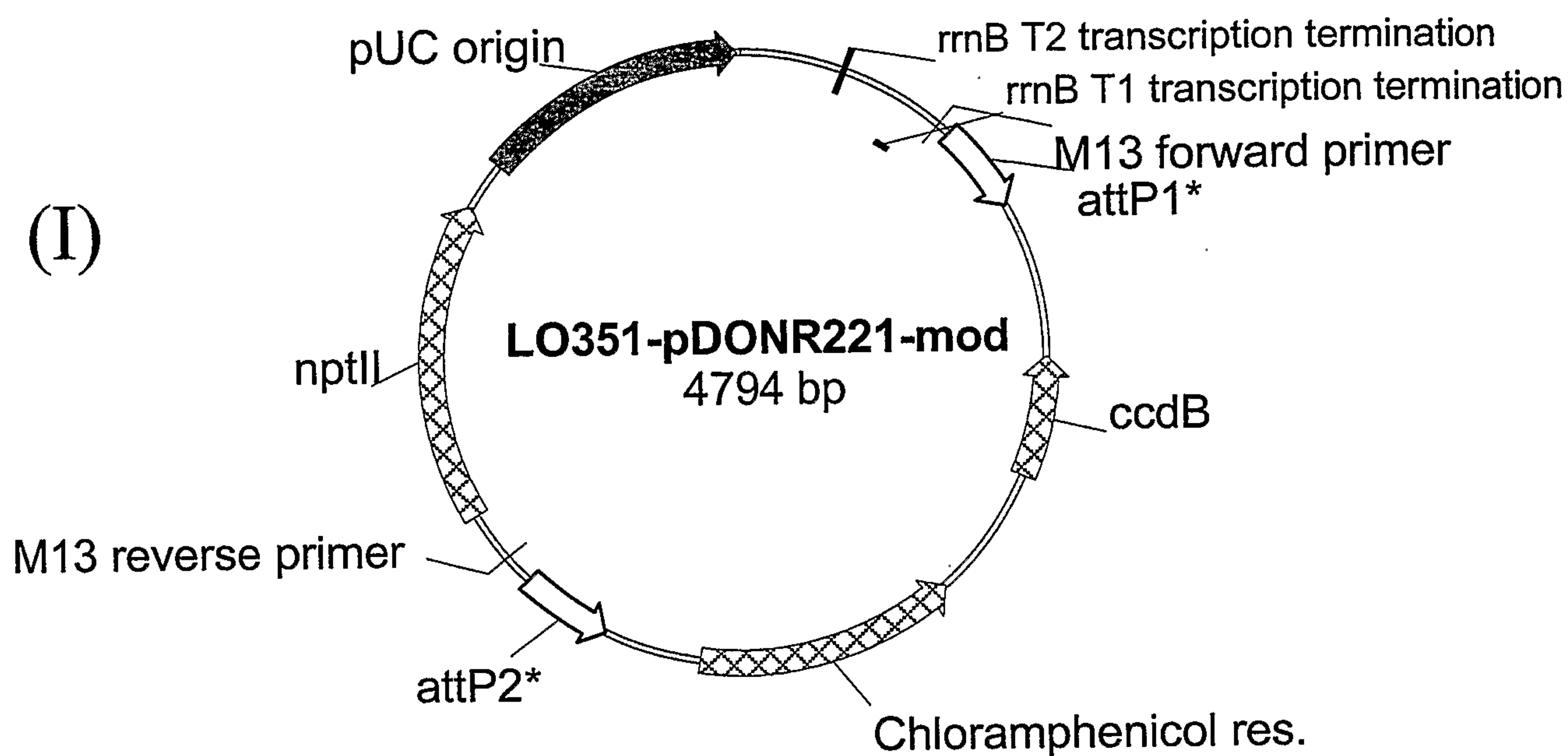


Fig. 1A

(III)

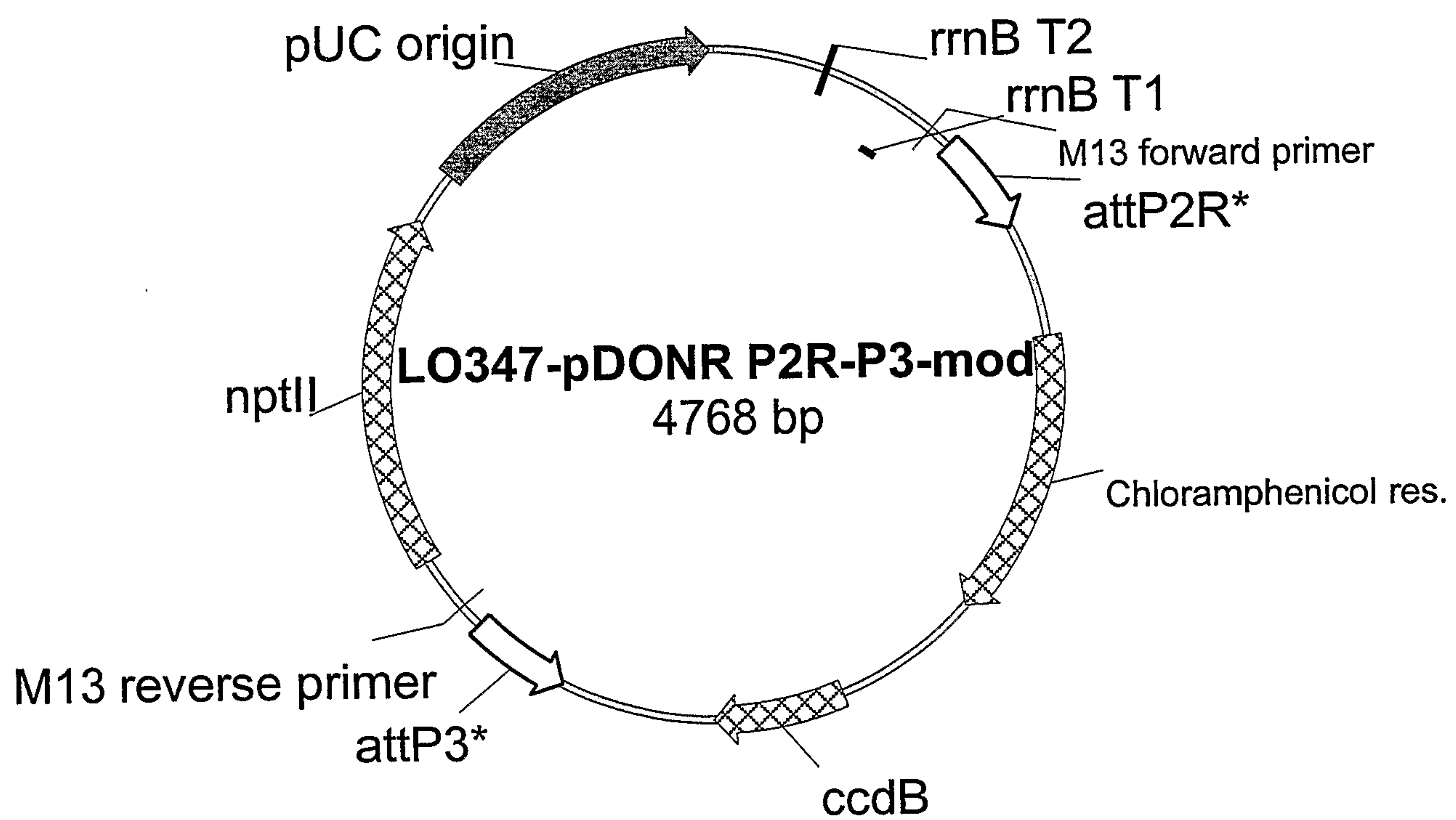


Fig. 1B

3/11

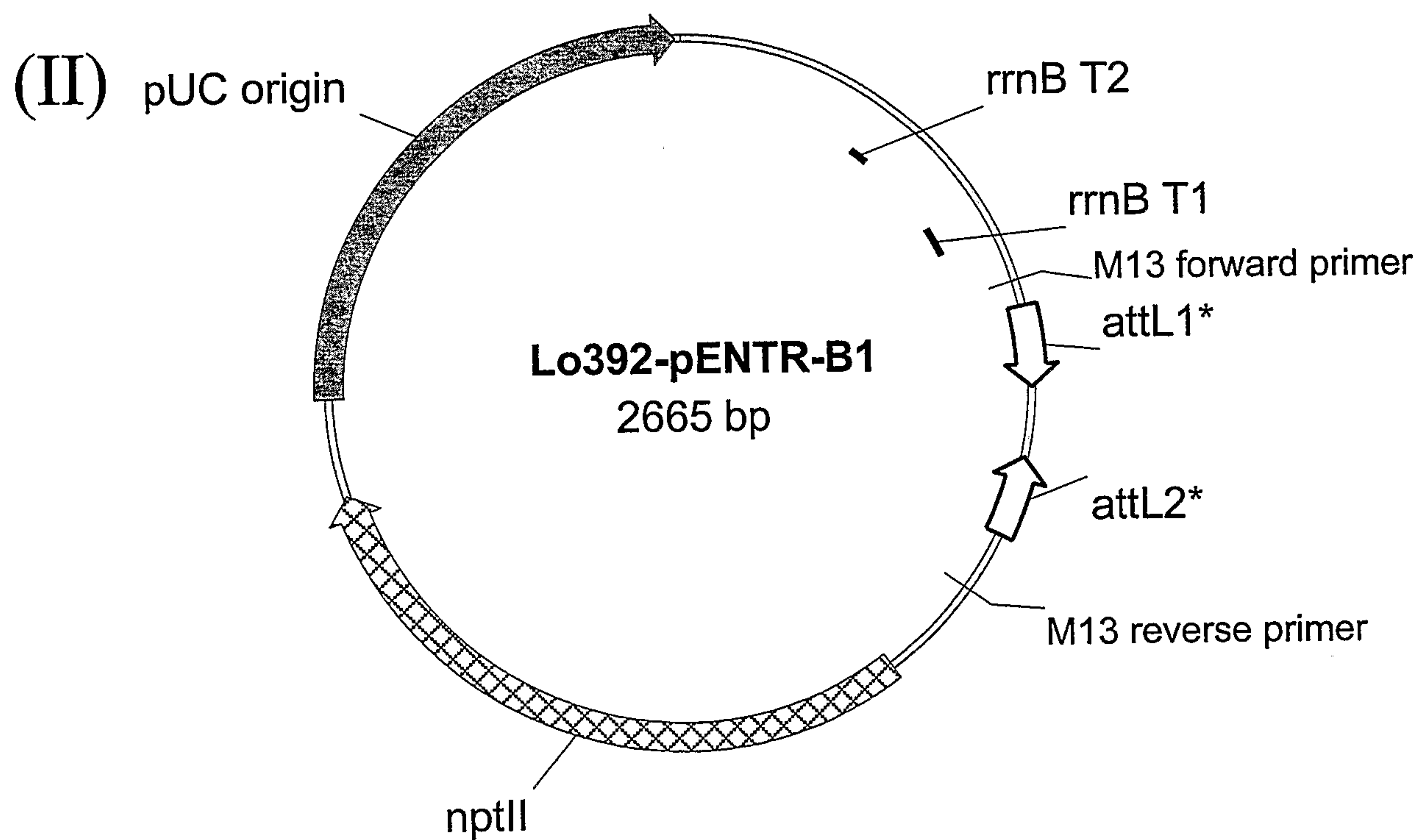
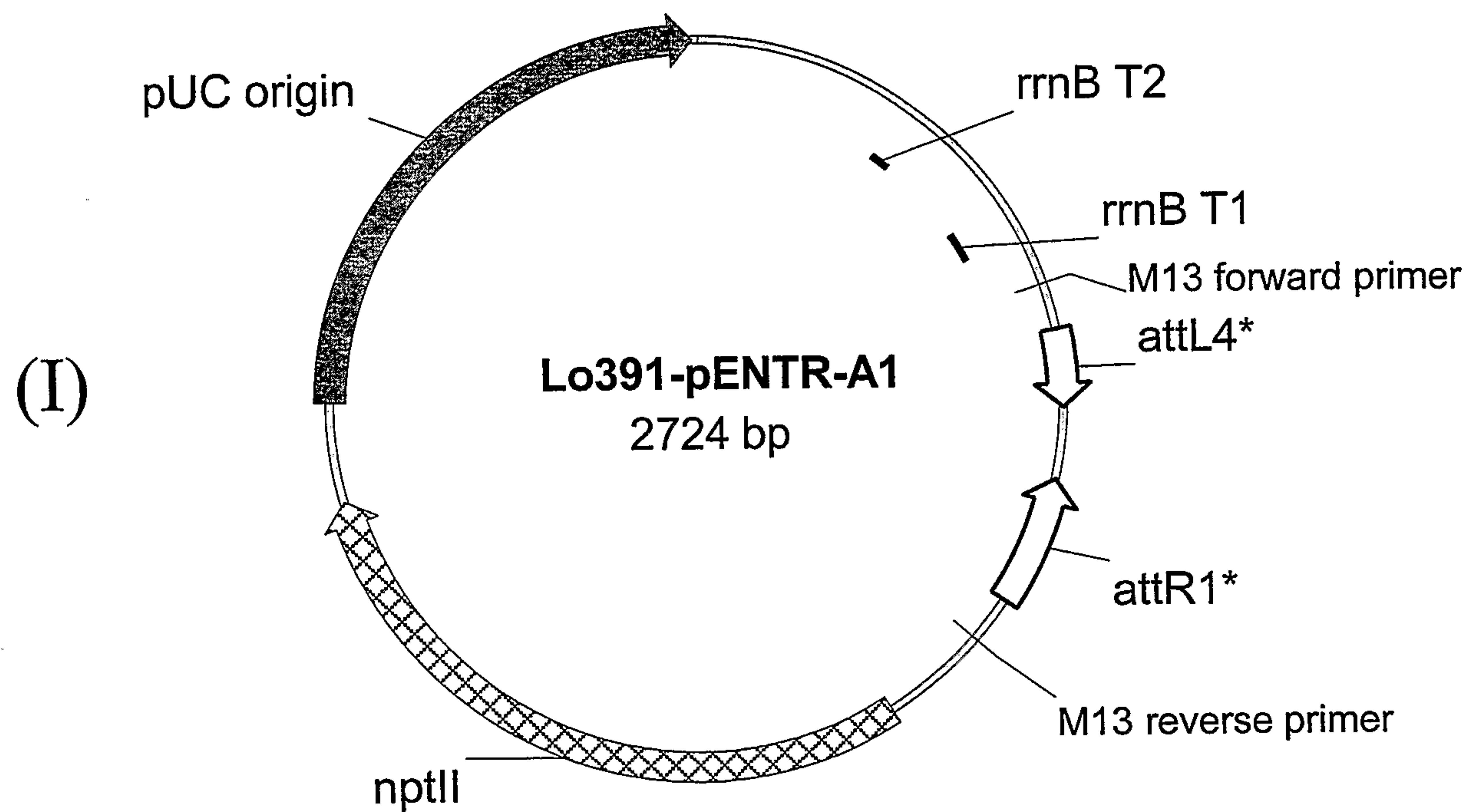


Fig. 2A

(III)

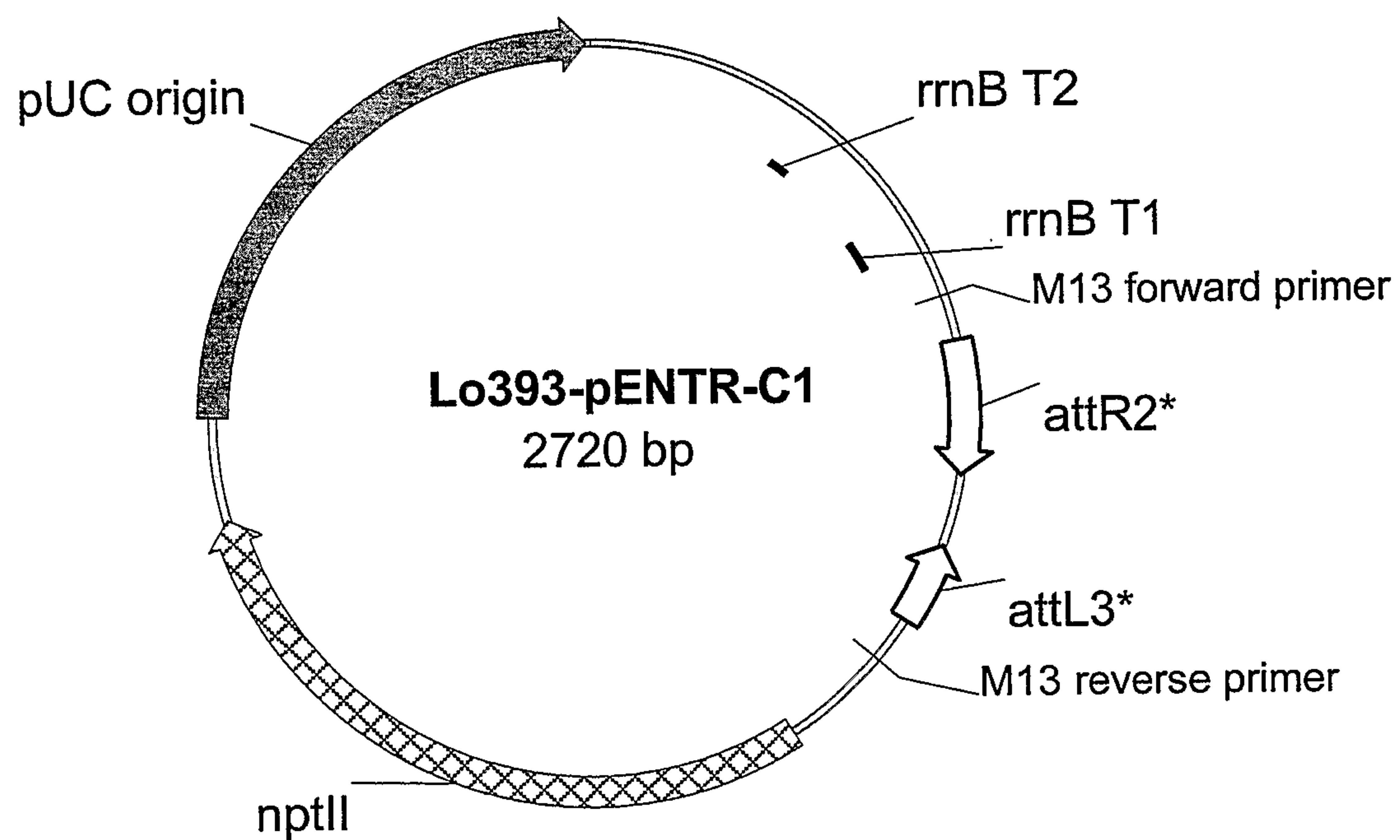


Fig. 2B

5/11

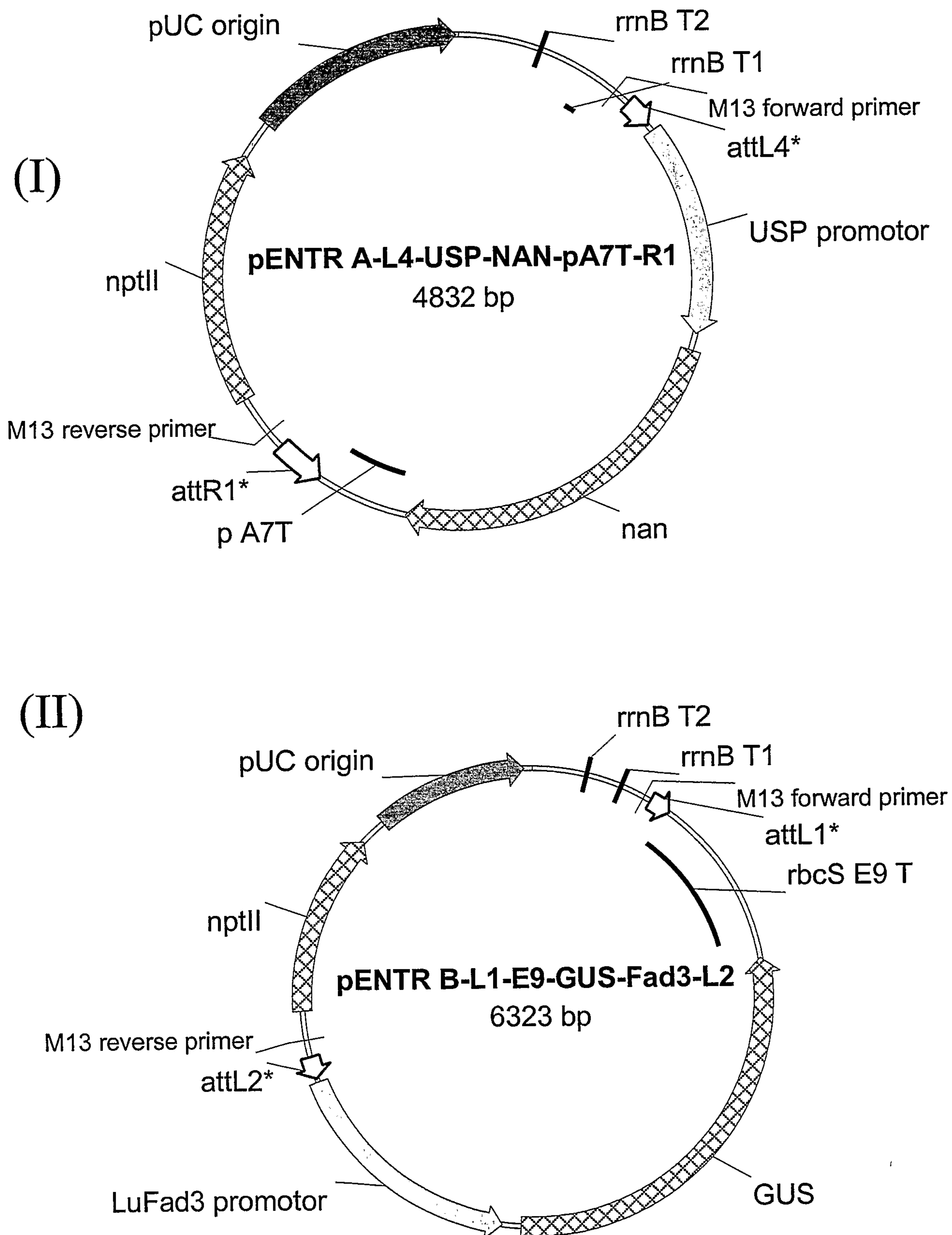
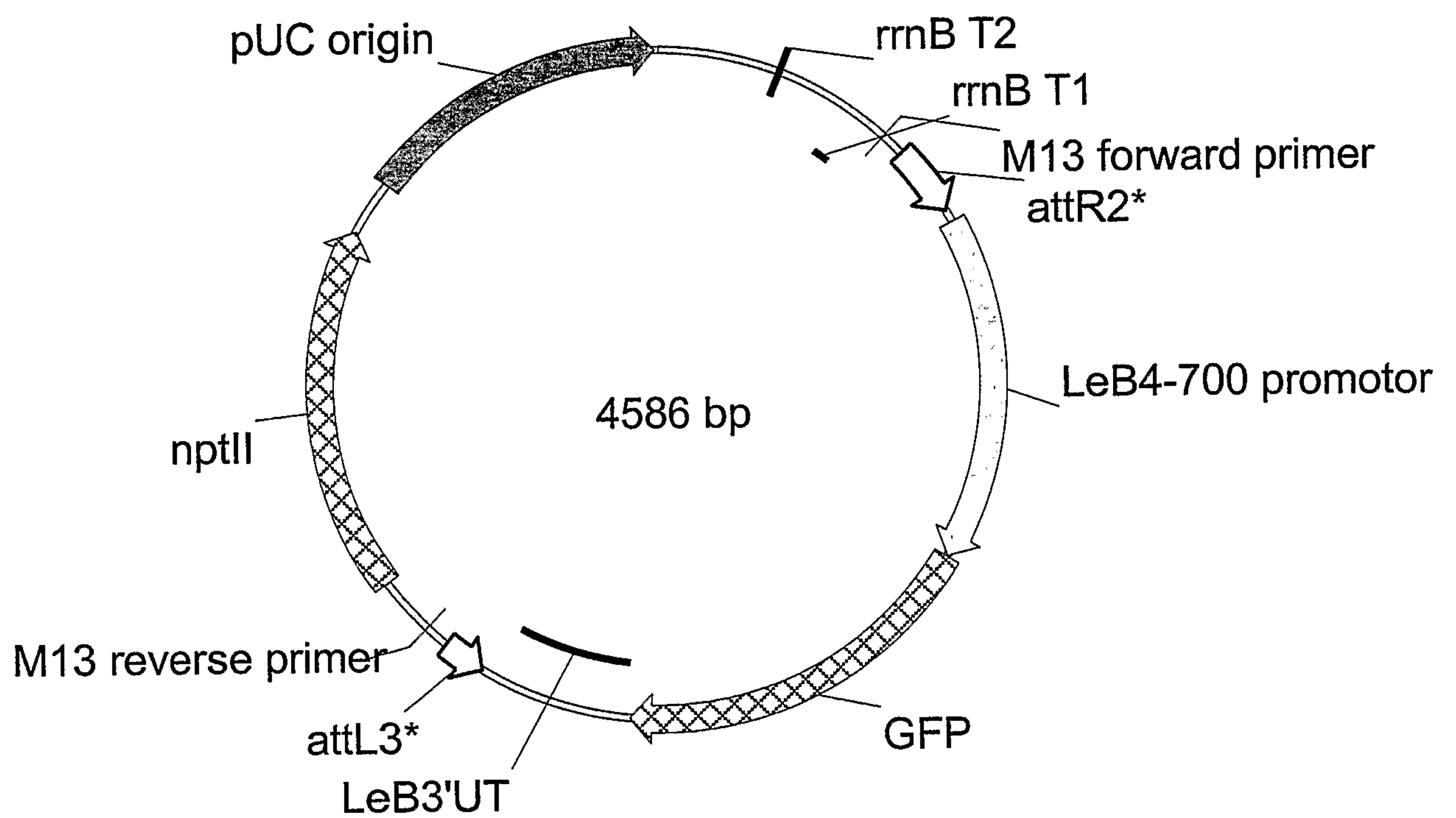


Fig. 3A

(III)

**pENTR C-R2-LeB4-GFP-LeB3T-L3****Fig. 3B**

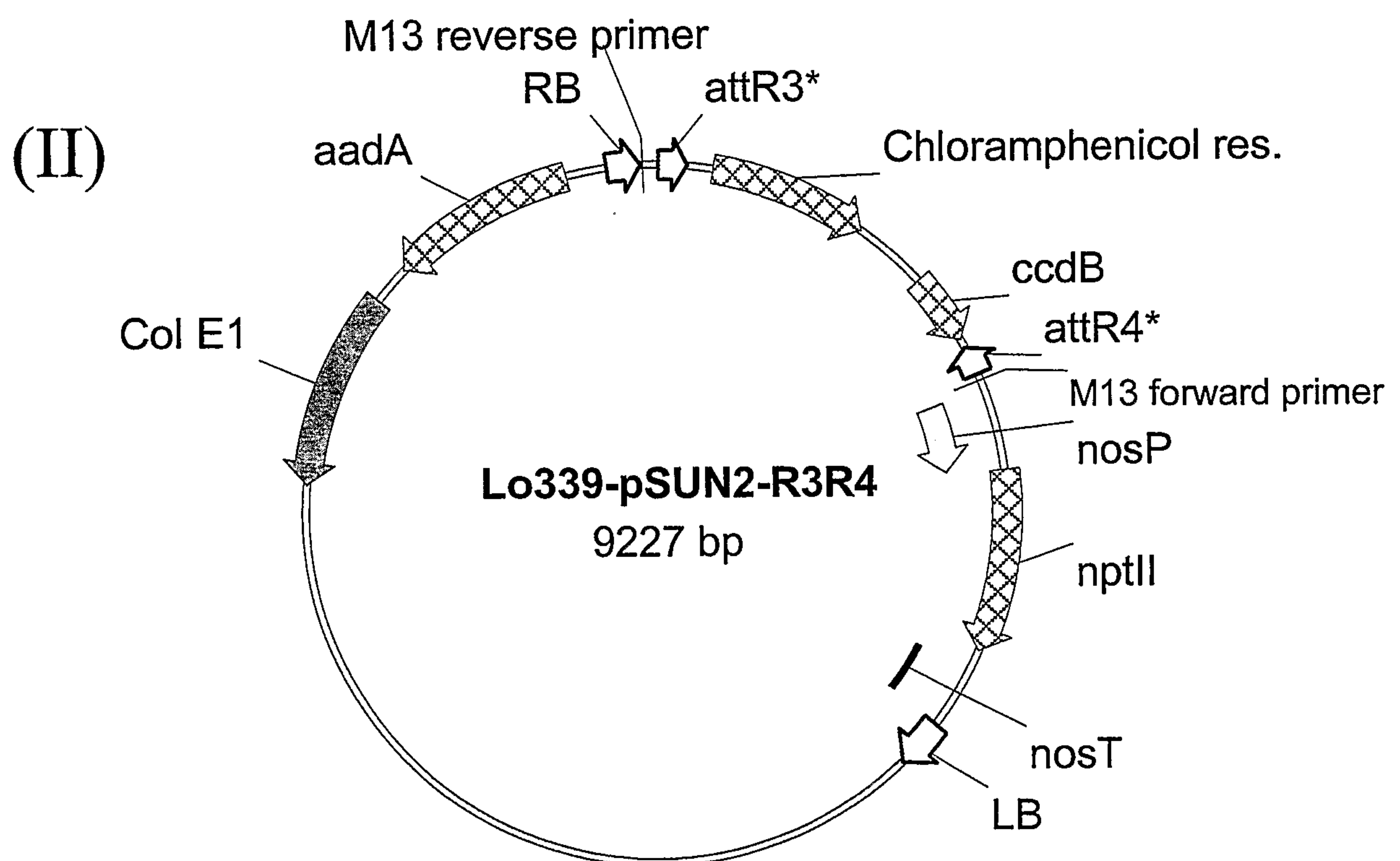
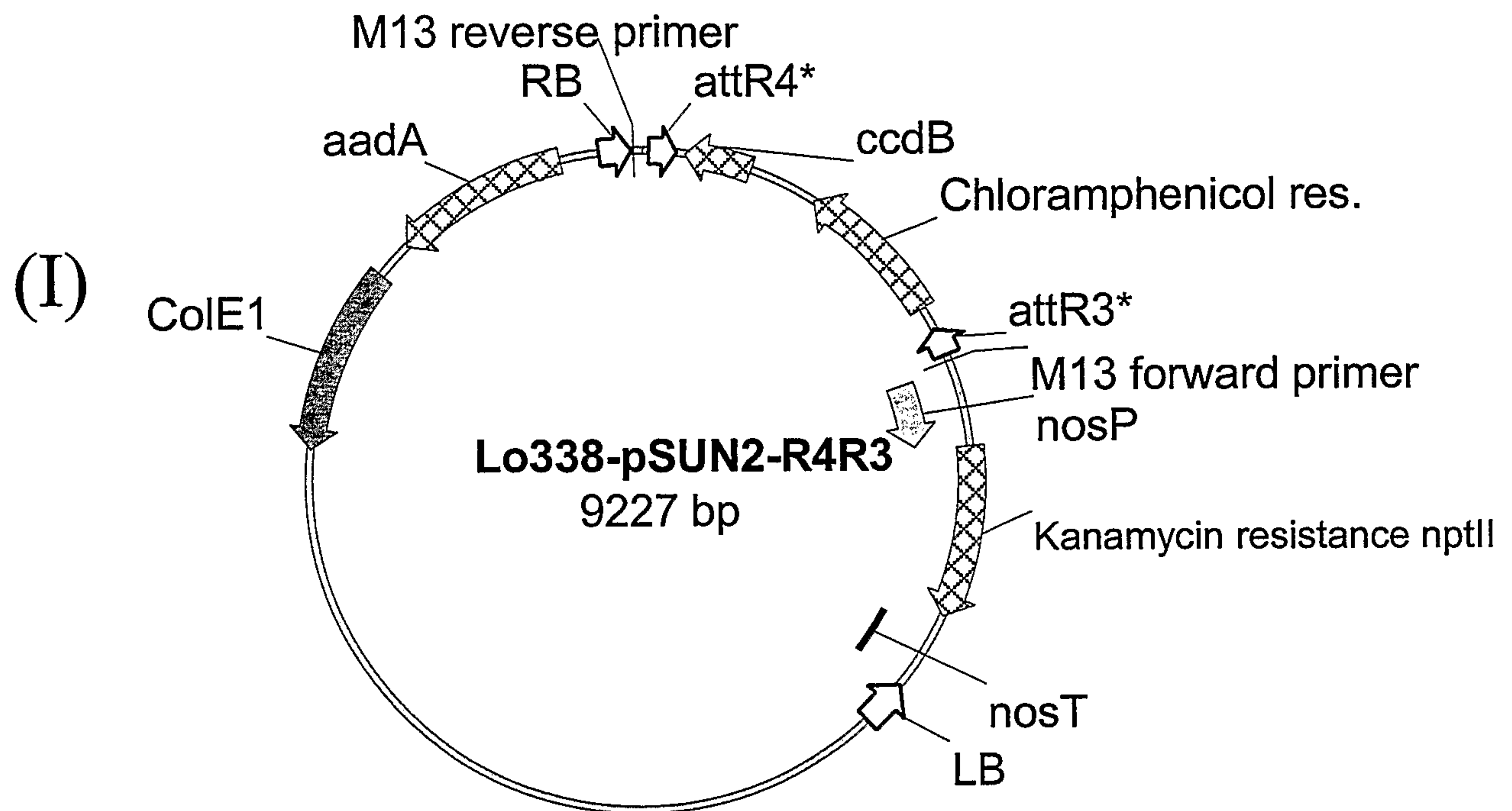
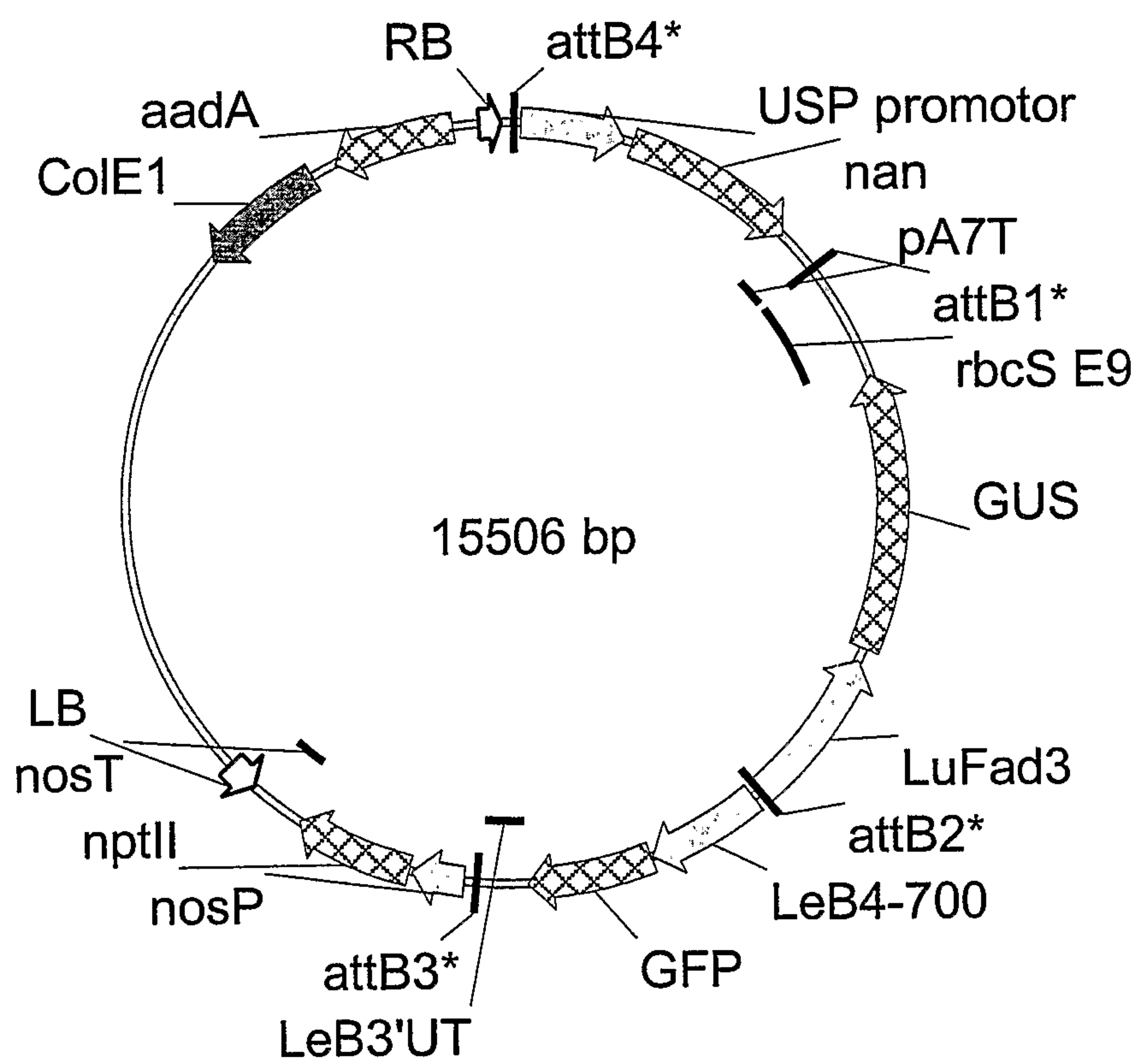


Fig. 4



pSUN2-B4-USP-NAN-pA7-E9-GUS-Fad3-LeB4-GFP-LeB3

Fig. 5

9/11

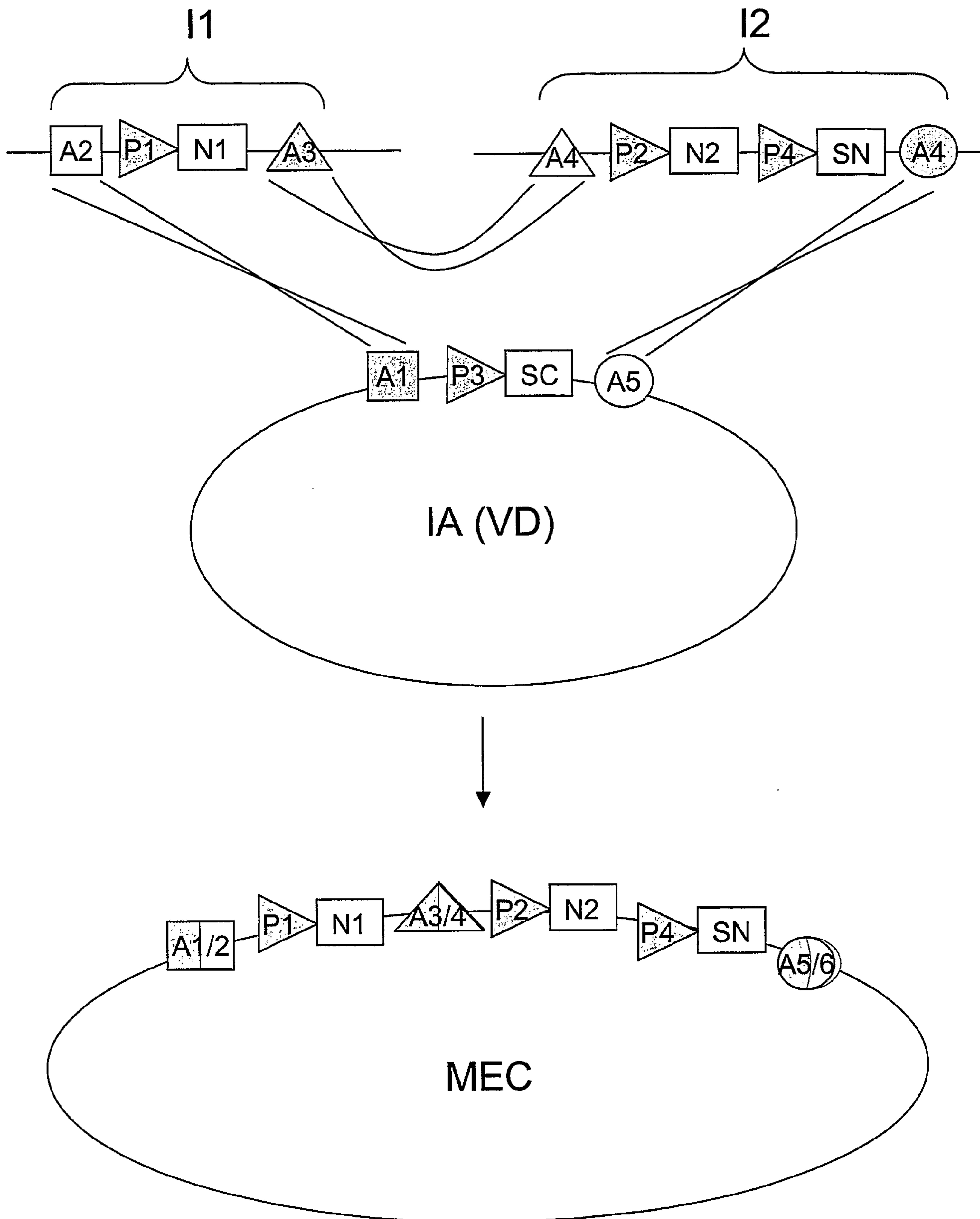


FIG. 6

10/11

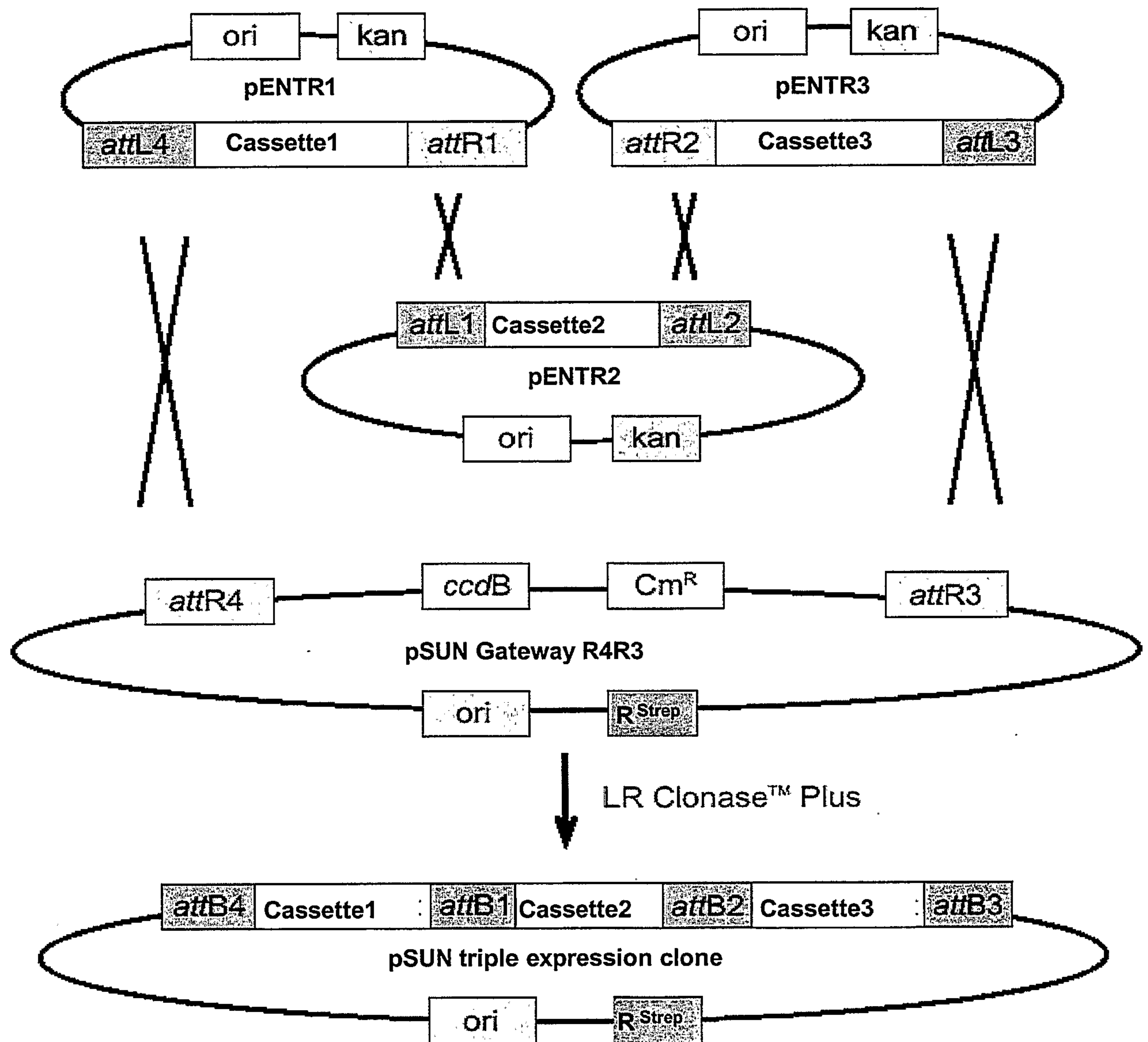


Fig. 7

11/11

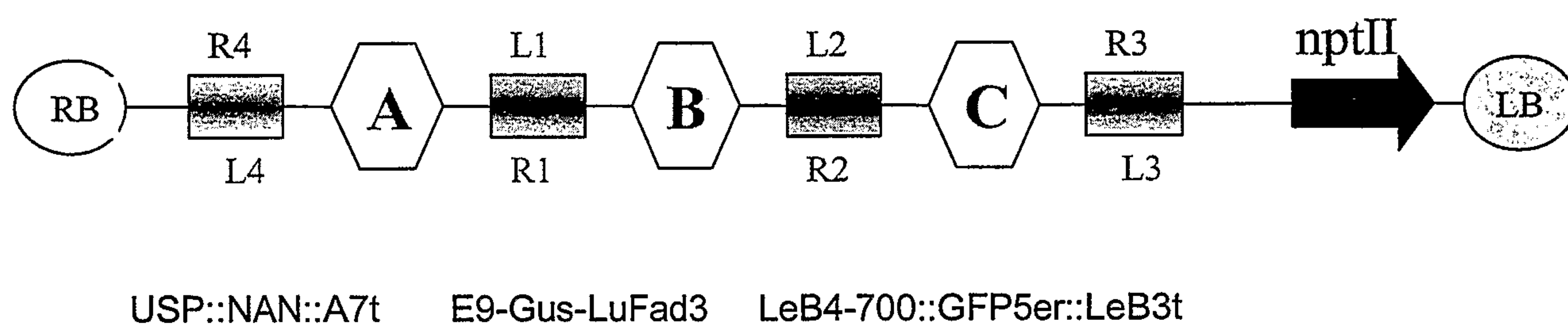
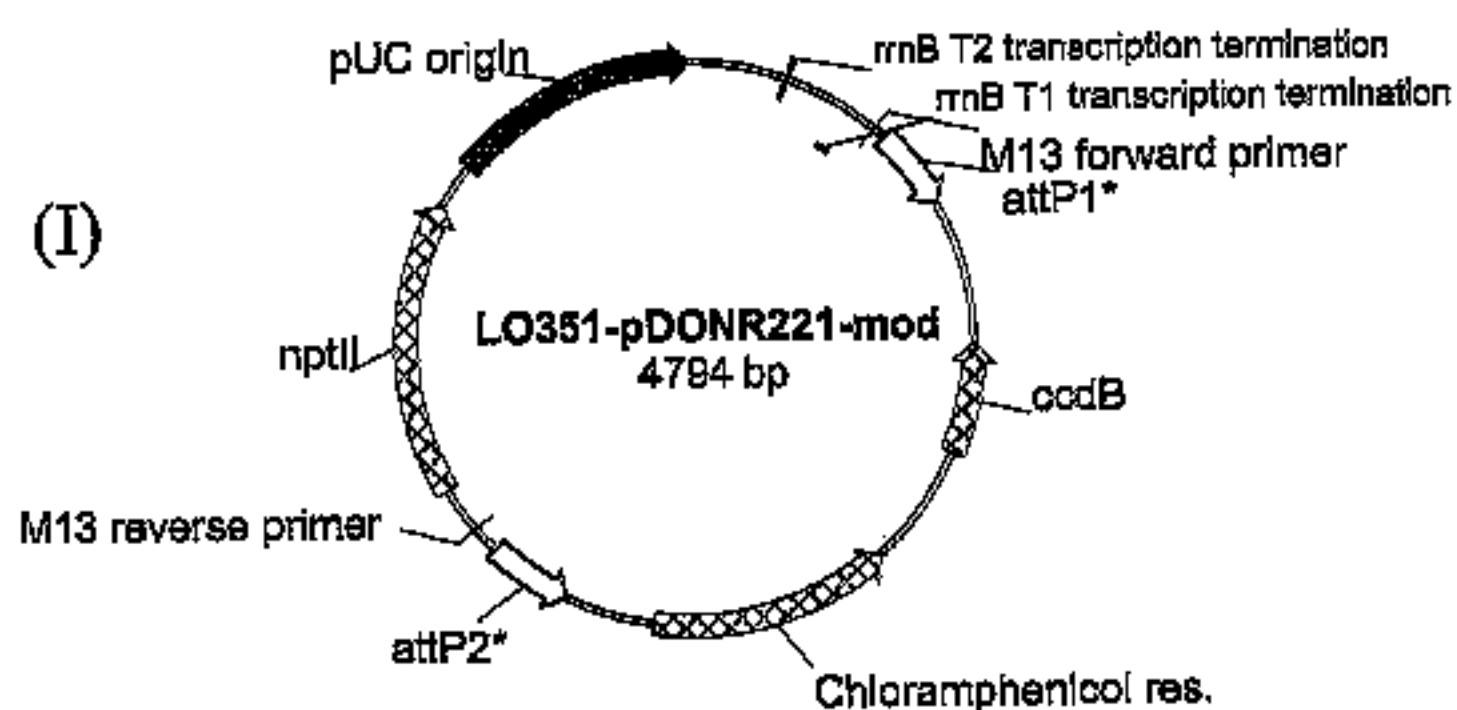
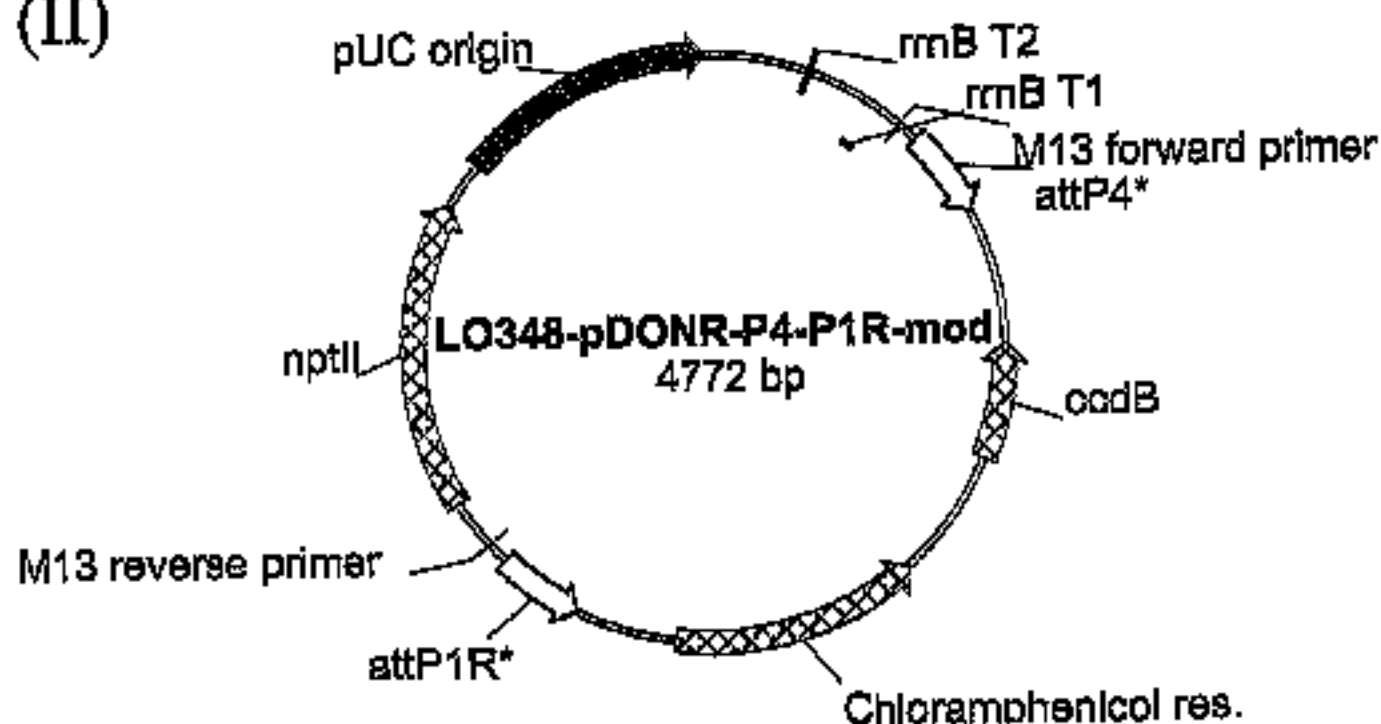


Fig. 8

(I)



(II)



(III)

