

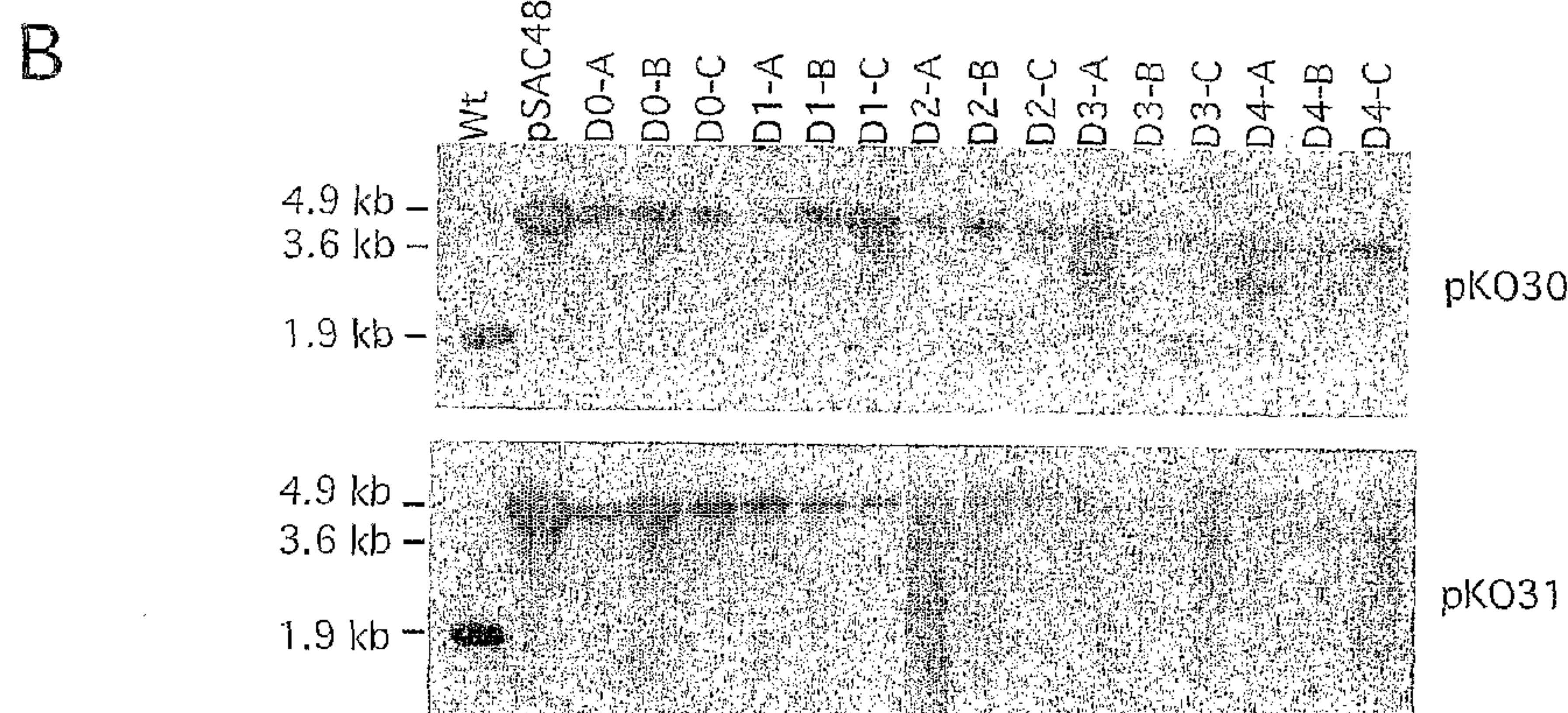
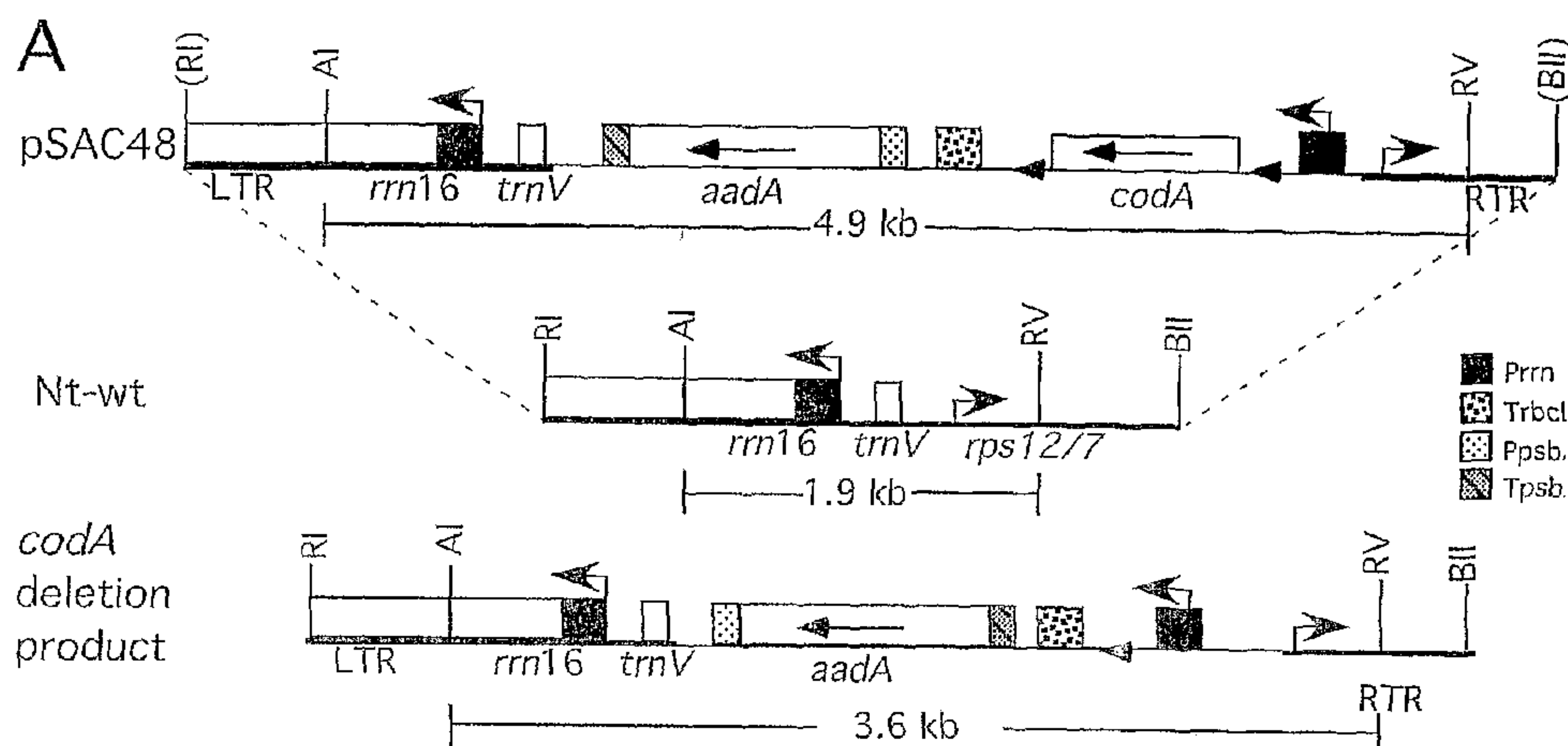


(86) Date de dépôt PCT/PCT Filing Date: 2004/03/03
(87) Date publication PCT/PCT Publication Date: 2004/09/16
(85) Entrée phase nationale/National Entry: 2005/09/01
(86) N° demande PCT/PCT Application No.: US 2004/006492
(87) N° publication PCT/PCT Publication No.: 2004/078935
(30) Priorités/Priorities: 2003/03/03 (60/451,779) US;
2004/01/08 (60/535,069) US

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 15/82, C12N 15/52
(71) Demandeur/Applicant:
RUTGERS, THE STATE UNIVERSITY OF NEW
JERSEY, US
(72) Inventeurs/Inventors:
MALIGA, PAL, US;
LUTZ, KERRY, US
(74) Agent: BERESKIN & PARR

(54) Titre : ELIMINATION DE SEQUENCES DE PLASTIDES A L'AIDE DE RECOMBINASES SPECIFIQUES DE SITES, A
EXPRESSION TRANSITOIRE

(54) Title: REMOVAL OF PLASTID SEQUENCES BY TRANSIENTLY EXPRESSED SITE-SPECIFIC RECOMBINASES



(57) Abrégé/Abstract:

Compositions and methods for manipulating the plastid genome of higher plants are provided.

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

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
16 September 2004 (16.09.2004)

PCT

(10) International Publication Number
WO 2004/078935 A2

(51) International Patent Classification⁷: **C12N**

(21) International Application Number:
PCT/US2004/006492

(22) International Filing Date: 3 March 2004 (03.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/451,779 3 March 2003 (03.03.2003) US
60/535,069 8 January 2004 (08.01.2004) US

(74) Agents: **RIGAUT, Kathleen, D.** et al.; Dann, Dorfman, Herrell and Skillman, 1601 Market Street, Suite 2400, Philadelphia, PA 19103-2307 (US).

(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, 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, 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, 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, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY** [US/US]; Old Queens, Somerset Street, New Brunswick, NJ 08903 (US).

(72) Inventors; and

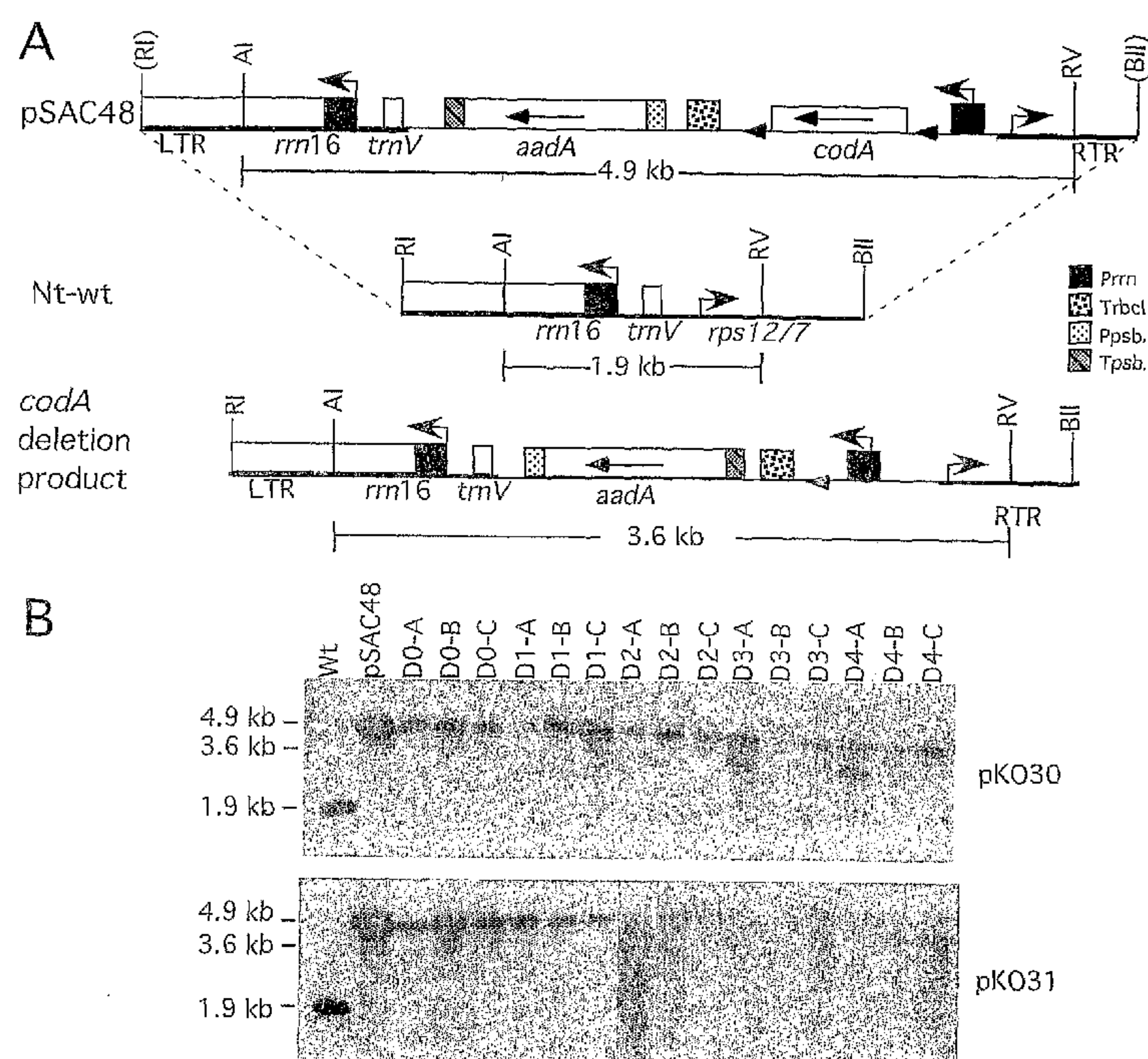
(75) Inventors/Applicants (*for US only*): **MALIGA, Pal** [US/US]; 72 Yorktown Road, East Brunswick, NJ 08816 (US). **LUTZ, Kerry** [US/US]; 100 Canal View Drive, Lawrenceville, NJ 08648 (US).

Published:

— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) Title: REMOVAL OF PLASTID SEQUENCES BY TRANSIENTLY EXPRESSED SITE-SPECIFIC RECOMBINASES



(57) Abstract: Compositions and methods for manipulating the plastid genome of higher plants are provided.

WO 2004/078935 A2

WO 2004/078935 A2



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.

**METHODS AND COMPOSITIONS FOR
REMOVAL OF PLASTID SEQUENCES BY TRANSIENTLY EXPRESSED
SITE-SPECIFIC RECOMBINASES**

By Pal Maliga & Kerry Lutz

5

This application claims priority to US Provisional Applications 60/451,779 and 60/535,069 filed March 3, 2003 and January 8, 2004 respectively. Each of the foregoing applications is incorporated by reference
10 herein.

Pursuant to 35 U.S.C. §202(c) it is acknowledged that the U.S. Government has certain rights in the invention described, which was made in part with funds
15 from the National Science Foundation Eukaryotic Genetics Program, Grant Number, MCB-0319958.

FIELD OF THE INVENTION

This invention relates to the fields of molecular
20 biology and transgenic plants. More specifically, the invention provides nucleic acid constructs and methods of use thereof for removing sequences of interest using transiently expressed site-specific recombinases.

25 **BACKGROUND OF THE INVENTION**

Several publications and patent documents are cited throughout this application in order to better describe the state of the art to which this invention pertains. Each of these citations is incorporated by reference
30 herein.

Engineering the plastid genome (ptDNA) rather than the nuclear genome gives rise to high protein expression levels and facilitates transgene containment. Plastid transformation involves targeted insertion of the
35 transforming DNA into the plastid genome (ptDNA) by

homologous recombination and amplification of the rare, transformed copies by selection for antibiotic resistance encoded in the vector (Bock, 2001; Staub, 2002; Maliga, 2004). Selection for vector-encoded
5 antibiotic resistance genes is essential to obtain uniform transformation of the 1000 to 10000 ptDNA copies present in a higher plant. Following selection, continued expression of the selectable marker gene places an unnecessary metabolic burden on the plant,
10 interferes with the need to use the same marker gene for multistep engineering and reduces consumer acceptance of the transgeneic plant. Thus, once transformation of ptDNA is accomplished, it is often desirable to eliminate the marker gene (Maliga, 2004).

15 Accordingly, it is an object of the invention to provide nucleic acid constructs and methods of use thereof for selective removal of predetermined sequences from the plastid genome.

20 SUMMARY OF THE INVENTION

In accordance with the present invention, compositions and methods are provided which facilitate the manipulation of the plastid genomes of higher plants. The methods of the invention may be employed to
25 remove predetermined target sequences from the plastid genome, such as selectable marker genes following successful isolation of transformed progeny. Advantageously, the present inventors have developed methods and compositions wherein both the heterologous
30 nucleic acids (encoding a selectable marker gene for example) and the recombinase are eliminated from the transplastomic plant cell, from which progeny plants producing a protein of interest can be regenerated.

In one embodiment of the invention, the method entails providing a transplastomic plant cell comprising plastids having heterologous nucleic acid sequence(s) flanked by excision sites and a nucleic acid sequence encoding at least one foreign gene of interest which is not flanked by excision sites. The plant cell is then contacted with a DNA construct which comprises a promoter operably linked to a nucleic acid encoding a protein having excision activity and optionally further comprises a sequence encoding a plastid transit peptide sequence and a selectable marker gene flanked by plant specific 5' and 3' regulatory regions under conditions where the DNA construct enters said cell and the proteins encoded thereby are transiently expressed for a suitable time period. The proteins catalyze the excision of the heterologous sequence from the plastids in the plant cell, thereby removing the heterologous sequence. Because the protein having excising activity is only transiently expressed, expression of this protein will also be lost. The method further entails identifying those transplastomic cells which lack both both the heterologous nucleic acid sequence and the protein having excising activity. In yet another embodiment, a progeny plant is generated from the plant cell following removal of the heterologous sequences.

Proteins having excision activity for use the practice of the present invention include, without limitation, CRE, flippase, resolvase, FLP, SSVI-encoded integrase, phiC31 integrase and transposases. Table I provides additional recombinase encoding sequences and excision sites recognized thereby.

In preferred embodiments, the heterologous nucleic acid encodes a selectable marker gene which confers

resistance to a selection agent. Selection agents suitable for use in the present invention include spectinomycin, kanamycin, hygromycin, streptomycin, phophinotricin, basta, glyphosate, and bromxynil.

5 The constructs of the invention can be introduced into plant cells using a variety of methods, which include, without limitation, Agroinfiltration, PEG fusion, biolistic delivery, CaPO₄-mediated transfection, and electroporation. Preferably, agroinfiltration is
10 used and the construct is transiently expressed in the nucleus.

 The methods of the present invention can be used to generate tranplastomic plant species which include for example, tobacco, rice, potato, maize, soybean, oil seed
15 rape, cotton and wheat.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows a schematic diagram of plastid
20 vector pSAC48 for CRE-mediated *codA* deletion. Shown are the plastid targeting region of vector pSAC48, the targeted region of the plastid genome (Nt-wt) and *codA* deletion derivative obtained by CRE-mediated
recombination via *lox* sites. Positions of plastid genes
25 *rrn16*, *trnV* and *rps12/7* (Shinozaki et al., 1986), *aadA* and *codA* transgenes, *lox*-sites and relevant restriction sites (AI, *ApaI*; BII, *BglII*; RI, *EcoRI*; RV, *EcoRV*) are marked.

Figure 1B shows DNA gel blot analysis tracking the
30 *codA* deletion the first four days after Agroinfiltration. Total leaf cellular DNA was digested with the *ApaI* and *EcoRV* restriction endonucleases and probed with the targeting region (*rrn16*, 1.9-kb *ApaI*-

EcoRV fragment). Fragments detected by the probing are marked in Figure 1A.

Figure 2A shows DNA gel blot analysis testing CRE-mediated *codA* deletion in plants regenerated from Agrobioinfiltrated leaves. Total leaf cellular DNA was digested with the *ApaI* and *EcoRV* restriction endonucleases and probed with the targeting region (*rrn16*, 1.9-kb *ApaI-EcoRV* fragment). Fragments detected by the probing are marked in Figure 1A.

Figure 2B. PCR amplification to probe for *Cre* in DNA samples which were analyzed for *codA* deletion as shown in Figure 2A.

Figure 3A is a schematic diagram showing the insertion of *bar* (dark arrow) and *aadA* (light arrow) genes into the plastid genome (ptDNA) by homologous recombination via left (L) and right (R) targeting sequences to obtain T1-ptDNA. Triangles represent *lox* sites.

Figure 3B depicts maps of ptDNA, plastid vectors pPAL1, pMHB10 and pMHB11, the T1-ptDNA of Nt-pMHB10 and Nt-pMHB11 plants, and of the T2-ptDNA after marker gene excision. Transcripts from vectors pMHB10 and pMHB11 with 5' and 3' stem-loop structures are shown in red above the maps. Abbreviations: *rrn16*, *trnV* and 3'-*rps12/rps7* are plastid genes; RBS, ribosome binding site; PL, promoter-leader cassette; T, transcription terminator. Map distances and restriction enzyme recognition sites are marked.

Figure 3C shows DNA gel blot analysis to confirm T1-ptDNA structure in Nt-pMHB10 and Nt-pMHB11 transformed plants. Total leaf cellular DNA was digested

with *EcoRI* and probed with the *StuI/ApaI* plastid DNA fragment (P1 probe, heavy line, Fig. 3B). The position and size of hybridizing fragments is marked in Figure 3B.

5

Figure 4A shows that expression of *Ebar* confers herbicide resistance in a leaf assay of wt, Nt-pMHB10 and Nt-pMHB11 plants grown in the presence 4mg/L of PPT.

Figure 4B depicts wild-type (wt), Nt-pMHB10 and Nt-pMHB11 plants one month after spraying with Liberty (2%; AgrEvo, Wilmington, DE).

Figure 5A shows excision of floxed *aadA* (light arrow) by transiently expressed CRE. Cre introduced into the plant nucleus on an *Agrobacterium* T-DNA is expressed in the cytoplasm, enters chloroplasts and excises floxed *aadA* from T1-ptDNA to yield T2-ptDNA. See Description of Figure 3A.

Figure 5B shows the results of short-term testing for excision. Vacuum infiltration of leaf disks (green squares) in *Agrobacterium* solution was followed by culture of leaf disks (1-4 days; D1-D4) on RMOP shoot regeneration medium (Svab and Maliga, 1993). Probing of *EcoRI*-digested total cellular DNA isolated from leaf disks with *StuI/ApaI* ptDNA fragment confirmed *aadA* excision: T1-ptDNA yields a 3.6-kb, T2-ptDNA a 2.8-kb diagnostic fragment (**Fig. 3B**). *Agrobacterium* was killed on day 3 by transferring leaf sections onto carbenicillin (500mg/L) medium.

Figure 5C shows the identification of T1 and/or T2 ptDNA in regenerated shoots by Southern blot analysis (for details see Fig. 5B). Information on Cre and the T-DNA plant marker gentamycin resistance ($Gent^r$) in the

nucleus is listed below. Abbreviations: +, present; -, absent; nd, not determined.

Figure 6 depicts INT-mediated excision of an atted
5 marker gene (upper) and the marker-free plastid genome
(lower).

Figure 7 shows directly oriented *attP* and *attB*
sequences with restriction sites in pBluescriptKSII+
10 plasmid derivative pK095.

Figure 8 provides schematic drawings of progenitor
plasmids for construction of plastid transformation
vectors with atted marker genes. Such vectors targeting
15 insertions at the *trnV*-3'-*rps12* intergenic region can be
obtained by combining sequences from plastid vector
pPRV1 described in US Patents 5,073,675, plasmid pK095
and the marker genes cited in this application.

20

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions:

The following definitions are provided to aid in
understanding the subject matter regarded as the
25 invention.

Heteroplastomic refers to the presence of a mixed
population of different plastid genomes within a single
plastid or in a population of plastids contained in
plant cells or tissues.

30 Homoplastomic refers to a pure population of
plastid genomes, either within a plastid or within a
population contained in plant cells and tissues.
Homoplastomic plastids, cells or tissues are genetically

stable because they contain only one type of plastid genome. Hence, they remain homoplastomic even after the selection pressure has been removed, and selfed progeny are also homoplastomic. For purposes of the present
5 invention, heteroplastomic populations of genomes that are functionally homoplastomic (i.e., contain only minor populations of wild-type DNA or transformed genomes with sequence variations) may be referred to herein as "functionally homoplastomic" or "substantially
10 homoplastomic." These types of cells or tissues can be readily purified to a homoplastomic state by continued selection.

Plastome refers to the genome of a plastid.

Transplastome refers to a transformed plastid
15 genome.

Transformation of plastids refers to the stable integration of transforming DNA into the plastid genome that is transmitted to the seed progeny of plants containing the transformed plastids. Alternatively,
20 transformation may also include the introduction and transient expression of heterologous DNA into the plastid or nuclear genomes.

Selectable marker gene refers to a gene that upon expression confers a phenotype by which successfully
25 transformed plastids or cells or tissues carrying the transformed plastid can be identified.

Transforming DNA refers to homologous DNA, or heterologous DNA flanked by homologous DNA, which when introduced into plastids becomes part of the plastid
30 genome by homologous recombination.

An alternative type of transforming DNA refers to a DNA which contains recombination site sequences for a site-specific recombinase or integrase. Insertion of

this type of DNA is not dependent of the degree of
homology between the transforming DNA and the plastid to
be transformed but rather is catalyzed by the action of
the recombinase or integrase on the first and second
5 recombination sites.

Operably linked refers to two different regions or
two separate genes spliced together in a construct such
that both regions will function to promote gene
expression and/or protein translation.

10 "Nucleic acid" or a "nucleic acid molecule" as used
herein refers to any DNA or RNA molecule, either single
or double stranded and, if single stranded, the molecule
of its complementary sequence in either linear or
circular form. In discussing nucleic acid molecules, a
15 sequence or structure of a particular nucleic acid
molecule may be described herein according to the normal
convention of providing the sequence in the 5' to 3'
direction. With reference to nucleic acids of the
invention, the term "isolated nucleic acid" is sometimes
20 used. This term, when applied to DNA, refers to a DNA
molecule that is separated from sequences with which it
is immediately contiguous in the naturally occurring
genome of the organism in which it originated. For
example, an "isolated nucleic acid" may comprise a DNA
25 molecule inserted into a vector, such as a plasmid or
virus vector, or integrated into the genomic DNA of a
prokaryotic or eukaryotic cell or host organism.

When applied to RNA, the term "isolated nucleic
acid" refers primarily to an RNA molecule encoded by an
30 isolated DNA molecule as defined above. Alternatively,
the term may refer to an RNA molecule that has been
sufficiently separated from other nucleic acids with
which it would be associated in its natural state (i.e.,

in cells or tissues). An isolated nucleic acid (either DNA or RNA) may further represent a molecule produced directly by biological or synthetic means and separated from other components present during its production.

5 The term "functional" as used herein implies that the nucleic or amino acid sequence is functional for the recited assay or purpose.

 The phrase "consisting essentially of" when referring to a particular nucleotide or amino acid
10 means a sequence having the properties of a given SEQ ID No:. For example, when used in reference to an amino acid sequence, the phrase includes the sequence per se and molecular modifications that would not affect the basic and novel characteristics of the sequence.

15 A "vector" is a replicon, such as a plasmid, cosmid, bacmid, phage or virus, to which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication of the attached sequence or element. Vectors, expression cassettes and
20 methods suitable for the generation of transplastomic plants are described in US Patents 6,624,296, 6,472,586, 6,388,168, 6,376,744, 6,297,054, 5,877,402, and 5,451,513, by Maliga et al., the disclosures of which are incorporated by reference herein.

25 An "expression operon" refers to a nucleic acid segment that may possess transcriptional and translational control sequences, such as promoters, enhancers, translational start signals (e.g., ATG or AUG codons), polyadenylation signals, terminators, and the
30 like, and which facilitate the production of a polypeptide coding sequence in a host cell or organism. Such expression signals may be combined such that

production of said polypeptide occurs transiently or is produced stably over the life of the cell.

The term "oligonucleotide," as used herein refers to primers and probes of the present invention, and is
5 defined as a nucleic acid molecule comprised of two or more ribo- or deoxyribonucleotides, preferably more than three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide.

10 The term "probe" as used herein refers to an oligonucleotide, polynucleotide or nucleic acid, either RNA or DNA, whether occurring naturally as in a purified restriction enzyme digest or produced synthetically, which is capable of annealing with or specifically
15 hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be either single-stranded or double-stranded. The exact length of the probe will depend upon many factors, including temperature, source of probe and use of the method.

20 The term "primer" as used herein refers to an oligonucleotide, either RNA or DNA, either single-stranded or double-stranded, either derived from a biological system, generated by restriction enzyme digestion, or produced synthetically which, when placed
25 in the proper environment, is able to functionally act as an initiator of template-dependent nucleic acid synthesis. When presented with an appropriate nucleic acid template, suitable nucleoside triphosphate precursors of nucleic acids, a polymerase enzyme,
30 suitable cofactors and conditions such as a suitable temperature and pH, the primer may be extended at its 3' terminus by the addition of nucleotides by the action of a polymerase or similar activity to yield an primer

extension product. The primer may vary in length depending on the particular conditions and requirement of the application.

Amino acid residues described herein are preferred
5 to be in the "L" isomeric form. However, residues in the "D" isomeric form may be substituted for any L-amino acid residue, provided the desired properties of the polypeptide are retained.

All amino-acid residue sequences represented herein
10 conform to the conventional left-to-right amino-terminus to carboxy-terminus orientation

The term "tag," "tag sequence" or "protein tag" refers to a chemical moiety, either a nucleotide, oligonucleotide, polynucleotide or an amino acid,
15 peptide or protein or other chemical, that when added to another sequence, provides additional utility or confers useful properties, particularly in the detection or isolation, to that sequence. Thus, for example, a homopolymer nucleic acid sequence or a nucleic acid
20 sequence complementary to a capture oligonucleotide may be added to a primer or probe sequence to facilitate the subsequent isolation of an extension product or hybridized product. In the case of protein tags, histidine residues (e.g., 4 to 8 consecutive histidine
25 residues) may be added to either the amino- or carboxy-terminus of a protein to facilitate protein isolation by chelating metal chromatography. Alternatively, amino acid sequences, peptides, proteins or fusion partners representing epitopes or binding
30 determinants reactive with specific antibody molecules or other molecules (e.g., flag epitope, c-myc epitope, transmembrane epitope of the influenza A virus hemagglutinin protein, protein A, cellulose binding

domain, calmodulin binding protein, maltose binding protein, chitin binding domain, glutathione S-transferase, and the like) may be added to proteins to facilitate protein isolation by procedures such as

5 affinity or immunoaffinity chromatography. Chemical tag moieties include such molecules as biotin, which may be added to either nucleic acids or proteins and facilitates isolation or detection by interaction with avidin reagents, and the like. Numerous other tag

10 moieties are known to, and can be envisioned by, the trained artisan, and are contemplated to be within the scope of this definition.

As used herein, the terms "reporter," "reporter system", "reporter gene," or "reporter gene product"

15 shall mean an operative genetic system in which a nucleic acid comprises a gene that encodes a product that when expressed produces a reporter signal that is a readily measurable, e.g., by biological assay, immunoassay, radioimmunoassay, or by colorimetric,

20 fluorogenic, chemiluminescent or other methods. The nucleic acid may be either RNA or DNA, linear or circular, single or double stranded, antisense or sense polarity, and is operatively linked to the necessary control elements for the expression of the reporter gene

25 product. The required control elements will vary according to the nature of the reporter system and whether the reporter gene is in the form of DNA or RNA, but may include, but not be limited to, such elements as promoters, enhancers, translational control sequences,

30 poly A addition signals, transcriptional termination signals and the like. The terms "transform", "transfect", "transduce", shall refer to any method or means by which a nucleic acid is introduced into a cell

or host organism and may be used interchangeably to convey the same meaning. Such methods include, but are not limited to, transfection, electroporation, microinjection, PEG-fusion, biolistic bombardment and
5 the like.

"Agroinfiltration" refers to Agrobacterium mediated DNA transfer. Specifically, this process involves vacuum treatment of leaf segments in an Agrobacterium suspension and a subsequent release of vacuum which
10 facilitates entry of bacterium cells into the inter-cellular space.

"T-DNA" refers to the transferred-region of the Ti (tumor-inducing) plasmid of Agrobacterium tumefaciens. Ti plasmids are natural gene transfer systems for the
15 introduction of heterologous nucleic acids into the nucleus of higher plants. Binary Agrobacterium vectors such pBIN20 and pPZP22 (GenBank Accession Number 10463) are known in the art.

A "plastid transit peptide" is a sequence which,
20 when linked to the N-terminus of a protein, directs transport of the protein from the cytoplasm to the plastid.

A "clone" or "clonal cell population" is a population of cells derived from a single cell or common
25 ancestor by mitosis.

A "cell line" is a clone of a primary cell or cell population that is capable of stable growth *in vitro* for many generations.

30 II. Removal of predetermined target sequences from the plastids of transgenic plants

The plastid genetic system is highly polyploid, as there are as many as 10,000 genome copies in a leaf

cell. Plastid transformation is facilitated by selective amplification of the rare, transformed copies.

This present invention addresses the reverse process, removal of the marker gene. This has been accomplished
5 in the past by employing a site-specific recombinase, which has been stably incorporated in the nuclear genome. Since the recombinase gene is present in the nucleus and is constitutively expressed, there is ample time for the excision of all marker gene copies to take
10 place. An alternative approach is described herein where excision function is obtained from transient expression of the site-specific recombinase. Excision of the marker genes by a transiently expressed recombinase in the absence of any selection resulted in the successful
15 elimination of all copies from the plastid genome within a very short period of time. Thus, the method has the advantage of also eliminating the recombinase without resorting to segregation of seed progeny for this purpose.

20 Efficient protocols for marker gene excision rely on a CRE site-specific recombinase which excises marker genes flanked by directly-oriented, 34-bp *loxP* sites (floxed) (Corneille et al., 2001; Hajdukiewicz et al., 2001). CRE is fused with a plastid targeting sequence at
25 its N-terminus, and is expressed from a stably integrated nuclear gene. After excision, Cre is removed from the nucleus by segregation in the seed progeny.

Providing CRE from a nuclear Cre gene facilitates efficient removal of the selectable marker gene, however
30 stable expression of the CRE gene in progeny plants is undesirable. In accordance with the present invention, *Agrobacterium tumefaciens*-mediated transient gene expression or Agroinfiltration has been employed to

provide, high, yet transient levels, of nuclear gene expression of recombinases, such as CRE, from a transiently introduced T-DNA region. It has been previously reported that histochemical staining after
5 Agroinfiltration showed transgene-expressing sectors comprising up to 90% of the leaf area (Kapila et al., 1997). Additionally, this system is suitable for the delivery of multiple proteins carried on multiple Agrobacterium binary vectors (for example (Goodin et
10 al., 2002)) and protein yields may be enhanced by suppression of RNA silencing (Johansen and Carrington, 2001). Finally, it is known that the T-region rarely integrates in the nucleus.

In accordance with the present invention, a P1
15 bacteriophage CRE-*loxP* site-specific recombination system is also provided which is suitable for efficient elimination of marker genes from the plastid genome by transiently expressing CRE from an introduced T-DNA region shown in Examples I and II. However, the utility
20 of the system is not limited to the CRE-*loxP* system. Example III describes excision of marker genes by the ϕ C31 phage integrase via directly oriented *attP* and *attB* sites. There are many additional site-specific recombinases, which would be equally suitable for this
25 application (Smith and Thorpe, 2002). See Table I. Advantageously, transient CRE expression significantly shortens the time required for selectable marker gene excision and enables applications in crops in which variety preservation is incompatible with seed
30 propagation.

The following examples are provided to illustrate certain embodiments of the invention. They are not intended to limit the invention in any way.

5

EXAMPLE I
REMOVAL OF CODA FROM THE PLASTID GENOME BY
AGROINFILTRATION

10 The previously described CRE-*loxP* test system has two components: a plastid tester strain carrying a cytosine deaminase (*codA*) transgene flanked by *lox* sites (floxed) conferring sensitivity to 5-fluorocytosine (Nt-pSAC48 plant line) and a nuclear CRE line carrying a
15 nuclear-encoded, plastid-targeted CRE (Nt-Cre1-100 and Nt-Cre2-300). Both the plastid tester (no CRE activity) and the nuclear CRE line (no *lox* sequence) were genetically stable. However, *codA* was eliminated at a very fast rate when the plastid-targeted CRE was
20 introduced into the plastid tester strain by transformation or crossing. The gene for the nuclear-encoded CRE was subsequently separated from the transformed plastids by segregation in the seed progeny. This approach, i.e., removal of plastid sequences by CRE
25 expressed from a stably integrated gene has been described in PCT/US00/25930 and in Corneille et al., 2001. This PCT application also provides T-DNA constructs for expressing recombinases such as CRE. PCT/US02/09537 also provides T-DNA constructs for
30 expressing integrase.

To avoid the need to eliminate CRE expression via segregation in seed progeny, marker gene elimination by transient expression of CRE was assessed. Leaf segments of Nt-pSAC48 plants (1 cm²) were placed in vacuum in an

Agrobacterium suspension in which T-DNA transfer has been induced with acetosyringone as described (Kapila et al., 1997). Agrobacterium containing pK030 (P2'long:TP5:cre:Tnos) or pK031 (P2'long:TP22:cre:Tnos) (Corneille et al., 2001) were inoculated into 100 ml YEB medium with 100 µg/mL spectinomycin and grown overnight at 27°C. 1 ml from the above culture was inoculated into fresh YEB medium containing 10 µM MES 2-(N-morpholino)ethanesulfonic acid, pH adjusted to 5.6, 20 µM acetosyringone, and 100 µg/mL spectinomycin, and grown overnight at 27°C. The culture was centrifuged and resuspended in MMA medium to a final OD₆₆₀=2.4 as described in Kapila et al., 1997. After incubation at room temperature for 1 hour, vacuum infiltration was performed. Leaves were cut into small pieces (1cm²) and mixed with the Agrobacterium culture. The mixture was placed under continuous vacuum of 2 Torr for 20 minutes while shaking gently. Upon releasing the vacuum, the Agrobacterium suspension infiltrated the leaves.

For two days, the leaves were incubated on an antibiotic-free medium to allow T-DNA transfer by Agrobacterium into the tobacco nuclei. After two days, the leaves were transferred onto RMOP medium containing carbenicillin (500 mg/L) to kill Agrobacterium. Notably, no selection for the antibiotic resistance marker gene carried on the T-DNA was employed. Removal of *codA* from the plastid genome was tested in the infiltrated leaf segments during the first four days after Agroinfection, and in plants regenerated from the leaves.

Initially, elimination of *codA* was assessed during the first four days after Agroinfiltration. Three leaf pieces were collected daily and DNA was extracted to test for *codA* excision using the plastid-targeting

region as the probe (Figure 1). DNA gel blot analysis revealed that by day 3 *codA* was excised in ~50% of the plastid genome copies in three of the six samples; and by day 4 in four of the six leaf segments. Lack of
5 excision in a small proportion of the day 4 samples is likely due to a lack in the uniformity of Agrobacterium infiltration. Proteins introduced on the Agrobacterium T-DNA are expected to be expressed highly only for a few days. It was important, therefore, that
10 deletion of *codA* occurred in 50% of the genome copies in this time frame suggesting that a significant fraction of regenerated plants will have the *codA* deleted.

Subsequently, *codA* excision in regenerated plants was tested. For this, shoots from the leaf sections were
15 regenerated in the absence of selection for Agrobacterium-mediated T-DNA transfer by antibiotic resistance. Excision of *codA* was tested by DNA gel blot analysis of total cellular DNA isolated from leaves. To ensure independent, clonal origin of the plants, the
20 infiltrated leaf samples (25 times 1 cm² each) were further subdivided (about 9 segments each), and only one regenerated plant was studied per piece.

Excision of *codA* was studied in 20 independently regenerated plants (Figure 2). Excision of the target
25 sequence was seen in six plants: three were homoplasmic for the deletion event (48-30-2B, 48-30-3B, 48-31-2B); one had a minor fraction of unexcised copies (48-31-1A) and two were heteroplasmic, containing genomes with *codA* deletion and *codA-aadA-trnV*
30 deletion (48-31-4A, 48-31-5A). PCR analyses were carried out to test for Cre integration. Cre was present in four of the six samples in which *codA* was excised. Two clones (48-31-1A and 48-31-2B), i.e., 10% of the regenerated

plants lacked an integrated *Cre* gene as well as all or most of the *codA* sequences used for selection of transformed plastids.

5

EXAMPLE II

TRANSIENTLY EXPRESSED CRE YIELDS HERBICIDE RESISTANT TRANSPLASTOMIC TOBACCO PLANTS FREE OF SELECTABLE MARKERS

10 Construction of transplastomic herbicide resistant tobacco plants utilizing transient expression of recombinase is described in the present example. To test the efficiency of plastid marker gene excision with transiently expressed CRE, additional vectors were
15 constructed which encode an herbicide resistance (*bar*) and a spectinomycin resistance (*aadA*) gene expressed from a single promoter by cloning a promoter-*bar* fragment into a promoter-less *aadA/lox* vector (pPAL1; Figure 3B). The *bar* gene confers herbicide
20 (phosphinothricin or PPT) resistance when the majority of ptDNA copies have been transformed, but is not suitable for direct selection of transplastomic clones when present in only a few of the 1,000 to 10,000 genome copies in a cell (Lutz et al., 2001; Ye et al., 2003).
25 Linking *bar* to spectinomycin resistance (*aadA*) ensured efficient selection of transplastomic clones. The selectable *aadA* gene was floxed to facilitate its excision by the CRE. The *bar* genes encoded in vectors pMHB10 and pMHB11 differ in their leader sequence.
30 Expression of the *bar* gene in plasmid pMHB11 is dependent on RNA editing (*Ebar*) (Chaudhuri and Maliga, 1996), a post-transcriptional event that converts an ACG codon (encoding a threonine) to an AUG translation

initiation codon. In contrast, *bar* expression in plasmid pMHB10 is independent of RNA editing (Figure 3B).

Plasmid pPAL1 was obtained by replacing the EcoRI-HindIII region in plastid vector pPRV100B (Zoubenko et al., 1994) with a nucleic acid of SEQ ID NO: 1. Plasmid pMHB10 was obtained by replacing the EcoRI-HindIII region in plastid vector pPRV100B (Zoubenko et al., 1994) with a nucleic acid of SEQ ID NO: 2. Plasmid pMHB11 was obtained by replacing the EcoRI-HindIII region in plastid vector pPRV100B (Zoubenko et al., 1994) with a nucleic acid of SEQ ID NO: 3.

Plasmids pMHB10 and pMHB11 were introduced into plastids by biolistic bombardment (Svab and Maliga, 1993), followed by integration of *bar* and *aadA* into the plastid genome, a process that yielded T1 transplastomes (T1-ptDNA; Figure 3A). Transformed shoots were identified by spectinomycin resistance (500 mg/L) on a selective RMOP shoot regeneration medium (Svab and Maliga, 1993) and integration of the transforming DNA was verified by Southern blot analysis of total leaf cellular DNA (Figure 3C). Plants were designated by plasmid name, a serial number and letters, for example Nt-pMHB10-23AH.

T1-ptDNA plants were tested for herbicide resistance in cell culture and in the greenhouse. Leaf segments of both the Nt-pMHB10 and Nt-pMHB11 plants proliferated on a culture medium containing 4 mg/L phosphinothricin (PPT), the active compound of the herbicide Liberty, whereas the wild-type tobacco leaf segments bleached and died (Figure 4A). Herbicide resistance was tested in the greenhouse by spraying (2% Liberty). The wild type plants died within two weeks, while the transgenic Nt-pMHB10 and Nt-pMHB11 plants

flowered and set seed (Figure 4B). The *bar* genes encode phosphinothricin acetyl transferase (PAT), an enzyme that detoxifies PPT. Acetylation of PPT was confirmed in leaf extracts of plants carrying the *bar* and *Ebar* genes.

5 We also tested editing of the chimeric mRNA and found that >80% of the mRNAs had the ACG codon converted to an AUG translation initiation codon (data not shown). However, the editing-dependent *Ebar* gene (pMHB11) was not expressed in *E. coli*, as expected (Lutz et al.,

10 2001) (data not shown), because prokaryotes lack the capacity for mRNA editing (Smith et al., 1997). Since no mRNA sequence is edited in the plant nucleus (Smith et al., 1997), *Ebar* mRNA will not confer herbicide resistance if the gene accidentally escapes to the plant

15 nucleus (Huang et al., 2003; Stegemann et al., 2003). However, this *Ebar* will be expressed in the plastids of tobacco, and other higher plant species which share the capacity to edit the initiation codon of *psbL* mRNA (Kudla et al., 1992; Chateigner-Boutin and Hanson, 2002), the

20 source of the edited segment (pSC4) (Chaudhuri et al., 1995). The editing-based herbicide-resistance gene described here is a useful alternative to intein-based protein splicing (Chin et al., 2003) and also provides an effective means of transgene containment.

25 According to the current protocols, the floxed marker gene is excised by CRE expressed from an integrated nuclear gene (Corneille et al., 2001; Hajdukiewicz et al., 2001). When tobacco plants with T1-ptDNA were obtained, we tested the feasibility of

30 excising the floxed *aadA* marker gene from the plastid genome with a transiently expressed CRE (Figure 5A). Based on the data presented in Example I, we anticipated that transient CRE expression would yield enough CRE to

excise a significant fraction of the marker genes from the ~10,000 T1-ptDNA copies. Agroinfiltration has been previously described and is suitable for high-level nuclear gene expression from a transiently introduced T-DNA region. Agroinfiltration involves vacuum treatment of leaf segments in an *Agrobacterium* suspension, and subsequent release of the vacuum to facilitate entry of bacterium cells into the inter-cellular space (Kapila et al., 1997). Infiltration was performed with *Agrobacterium* strain EHA101 (Hood et al., 1986) carrying the pK031 binary vector encoding a plastid-targeted CRE and gentamycin resistance as the plant marker (Corneille et al., 2003). Protein expressed from the T-DNA accumulates in two to four days after agroinfiltration (Johansen and Carrington, 2001; Voinnet et al., 2003). Therefore, we tested marker gene excision in T1-ptDNA leaf segments up to four days after agroinfiltration. Marker gene excision was observed as early as three days after agroinfiltration. By day four, a significant fraction of plastid genome copies lacked the marker gene (Figure 5B).

Encouraged by the levels of marker gene excision in the short-term experiment, agroinfiltration was used in efforts to obtain marker-free transplastomic plants. Infiltration of Nt-pMHB10 and Nt-pMHB11 leaf sections was carried out with *Agrobacterium* EHA101 carrying vector pK031 as described above. Two days after infiltration the leaf segments were placed on shoot regeneration medium, which contained carbenicillin (500 mg/L), to kill the *Agrobacterium* cells. Plants were then regenerated in the absence of selection for T-DNA transfer (gentamycin resistance). A month later the regenerated plants were individually tested for marker

gene excision by Southern blot analysis (Figure 5C). Marker-excised (T2) plastid genomes were detected in 19 of the 61 tested plants (11 of 27 Nt-pMHB10; 8 of 34 Nt-pMHB11). The nuclear Cre gene was absent in 7 clones (3
5 Nt-pMHB10; 4 Nt-pMHB11), as determined by PCR analysis (Figure 5C). One of these (Nt-pMHB11-12AK) carries the unselected gentamycin resistance gene (data not shown).

Overall, excision of *aadA* by the transiently expressed nuclear CRE yielded marker-free transplastomic
10 plants in ~10% (6 out of 61) of the regenerated plants, in which *Agrobacterium* T-DNA delivery was not followed by T-DNA integration. This number is comparable to the ~12% excision frequency (13 out of 57) of a floxed plastid *codA* gene observed in regenerated shoots after
15 agroinfiltration (See Example I) and is much higher than excision efficiency of a nuclear gene (0.25%) by a transiently expressed CRE in the absence of direct selection for the excision event (Gleave et al., 1999). For an overview of marker gene excision from the nuclear
20 genome, see Hare and Chua, 2002 and Ow, 2002.

Presently four systems are available for plastid marker gene elimination. One approach relies on the loop-out of the marker gene through directly repeated sequences (Iamtham and Day, 2000). A second approach
25 involves co-transformation with two independently targeted plastid transgenes (Ye et al., 2003). Both approaches are difficult to control since transformation and marker gene elimination occur simultaneously. The third approach is based on restoration of the green
30 phenotype in pigment deficient mutants and depends on the availability of knockout mutants of plastid photosynthetic genes (Klaus et al., 2004). The most efficient, generally applicable approach excises floxed

marker genes by CRE expressed from an integrated nuclear gene(Corneille et al., 2001; Hajdukiewicz et al., 2001). Plastid marker gene excision by transiently expressed CRE significantly accelerates the production of marker-free transplastomic plants as it eliminates the requirement to segregate out the nuclear Cre in the seed progeny. Until the present invention, CRE-mediated excision was not applicable to vegetatively propagated species such as potato, apple, and poplar, as variety preservation is incompatible with seed propagation in these highly heterozygous crops. Transient expression of CRE for marker gene removal now enables production of marker free transplastomic plants that belong to this group.

15

EXAMPLE III
EXCISION OF MARKER GENES BY A TRANSIENTLY EXPRESSED
PHIC31 PHAGE SITE-SPECIFIC INTEGRASE (INT)

20

In addition to CRE-*loxP*, there are many site-specific recombination systems, which are suitable for the excision of plastid marker genes (Smith and Thorpe, 2002). One of these is the *phiC31* phage integrase (INT), which mediates recombination between bacterial *attB* and phage *attP* sequences. The minimal sequences for *attB* and *attP* function were defined to be 34 and 39 bp, respectively (Groth et al., 2000). In mammalian cells longer sequences were more efficient. Therefore, in tobacco plastids, a 55-bp *attB* site was incorporated in the plastid genome to facilitate INT-mediated insertion of a vector carrying a 217-bp *attP* sequence (Lutz et al., 2004).

30

Molecular reactions carried out by the recombinase depend on the orientation and location (cis or trans) of target sequences. If the marker gene is flanked by directly orientated *attB* and *attP* target sites (atted),

5 INT mediates excision of the atted marker gene and leaves behind a recombinant *attB/attP* site. A general scheme for a plastid vector with an atted marker gene is shown in Figure 6. The atted marker gene can be excised by INT, which is stably integrated in the nucleus. Such

10 stably transformed lines can be obtained by transformation with *Agrobacterium* binary vector pKO117 (Lutz et al., 2004). INT can also be expressed transiently by Agroinfiltration using binary vector pKO117 as described in Example II. To create a suitable

15 plastid transformation vector containing an atted marker gene, *attP* and *attB* sequences from *E. coli* plasmid pKO95 can be used. The map of plasmid pKO95 is shown in Figure 7; the DNA sequence of the region between the *XhoI* and *SpeI* sites is included within SEQ ID NO: 4. Plastid

20 transformation vectors with an atted marker gene can be constructed using the progenitor plasmids shown in Figure 8. Marker genes can be inserted at the *Sca I* site located between the directly oriented *attP* and *attB* sequences. Suitable marker genes for insertion have been

25 described in US Patents 5,073,675 and pending patent application US99/17806. The multiple cloning sites are also present for cloning of passenger genes.

The following sequences were used in the methods of the

30 present invention:

SEQ ID NO: 1

GAATTCGAGC TCGGTACCCG GGGATCCGGT ACatctagat GTACCATAAC

TTCGTATAAT GTATGCTATA CGAAGTTATa gatcaattca gttgtaggga
 5 gggattttATG gGGGAAGCGG TGATCGCCGA AGTATCGACT CAACTATCAG
 AGGTAGTTGG CGTCATCGAG CGCCATCTCG AACCGACGTT GCTGGCCGTA
 CATTTGTACG GCTCCGCAGT GGATGGCGGC CTGAAGCCAC ACAGTGATAT
 10 TGATTTGCTG GTTACGGTGA CCGTAAGGCT TGATGAAACA ACGCGGCGAG
 CTTTGATCAA CGACCTTTTG GAAACTTCGG CTTCCCCTGG AGAGAGCGAG
 15 ATTCTCCGCG CTGTAGAAGT CACCATTTGTT GTGCACGACG ACATCATTC
 GTGGCGTTAT CCAGCTAAGC GCGAACTGCA ATTTGGAGAA TGGCAGCGCA
 ATGACATTCT TGCAGGTATC TTCGAGCCAG CCACGATCGA CATTGATCTG
 20 GCTATCTTGC TGACAAAAGC AAGAGAACAT AGCGTTGCCT TGGTAGGTCC
 AGCGGCGGAG GAACTCTTTG ATCCGGTTCC TGAACAGGAT CTATTTGAGG
 25 CGCTAAATGA AACCTTAACG CTATGGAAC TCGCCGCCGA CTGGGCTGGC
 GATGAGCGAA ATGTAGTGCT TACGTTGTCC CGCATTTGGT ACAGCGCAGT
 AACCGGCAAA ATCGCGCCGA AGGATGTGCG TGCCGACTGG GCAATGGAGC
 30 GCCTGCCGGC CCAGTATCAG CCCGTCATAC TTGAAGCTAG ACAGGCTTAT
 CTTGGACAAG AAGAAGATCG CTTGGCCTCG CGCGCAGATC AGTTGGAAGA
 35 ATTTGTCCAC TACGTGAAAG GCGAGATCAC CAAGGTAGTC GGCAAATAAT
 GTctagctag ataacttcgt ataatgtatg ctatacgaag ttatAGACAT
 TAGCAGATAA ATTAGCAGGA AATAAAGAAG GATAAGGAGA AAGAACTCAA
 40 GTAATTATCC TTCGTTCTCT TAATTGAATT GCAATTAAAC TCGGCCCAAT
 CTTTFACTAA AAGGATTGAG CCGAATACAA CAAAGATTCT ATTGCATATA
 45 TTTTGACTAA GTATATACTT ACCTAGATAT ACAAGATTTG AAATACAAAA
 TCTAGcaagc tt

SEQ ID NO: 2:

50 GAATTCGAGC TCGCTCCCC GCGTCGTTT AATGAGAATG GATAAGAGGC
 TCGTGGGATT GACGTGAGGG GGCAGGGATG GCTATATTTT TGGGAGAATT
 AACCGATCGA CGTGCaAGCG GACATTTATT TTaAATTCGA TAATTTTTC
 55 AAAAACATTT CGACATATTT ATTTATTTTA TTATTATGAG AATCAATCCT
 ACTACTTCTG GTTCTGGGGT TTCCACGgct AGCCCAGAAa GAaGaCCGGC

CGACATCCGC CGTGCCACCG AGGCGGACAT GCCGGCGGTC TGCACCATCG
 TCAACCACTA CATCGAGACA AGCACGGTCA ACTTCCGTAC CGAGCCGCAG
 5 GAACCGCAGG AGTGGACGGA CGACCTCGTC CGTCTGCGGG AGCGCTATCC
 CTGGCTCGTC GCCGAGGTGG ACGGCGAGGT CGCCGGCATC GCCTACGCGG
 10 GCCCCTGGAA GGCACGCAAC GCCTACGACT GGACGGCCGA GTCGACCGTG
 TACGTCTCCC CCCGCCACCA GCGGACGGGA CTGGGCTCCA CGCTCTACAC
 CCACCTGCTG AAGTCCCTGG AGGCACAGGG CTTCAAGAGC GTGGTCGCTG
 15 TCATCGGGCT GCCCAACGAC CCGAGCGTGC GCATGCACGA GGCCTCGGA
 TATGCCCCC GCGGCATGCT GCGGGCGGCC GGCTTCAAGC ACGGGAAGTG
 GCATGACGTG GGTTTCTGGC AGCTGGACTT CAGCCTGCCG GTACCGCCCC
 20 GTCCGGTCCT GCCCGTCACC GAGATCTGAT GAtctagatG TACCATAACT
 TCGTATAATG TATGCTATAC GAAGTTATag atcaattcag ttgtagggag
 25 ggatttATGg GGAAGCGGT GATCGCCGAA GTATCGACTC AACTATCAGA
 GGTAGTTGGC GTCATCGAGC GCCATCTCGA ACCGACGTTG CTGGCCGTAC
 ATTTGTACGG CTCCGCAGTG GATGGCGGCC TGAAGCCACA CAGTGATATT
 30 GATTTGCTGG TTACGGTGAC CGTAAGGCTT GATGAAACAA CGCGGCGAGC
 TTTGATCAAC GACCTTTTGG AAActTCGGC TTCCCCTGGA GAGAGCGAGA
 35 TTCTCCGCGC TGTAGAAGTC ACCATTGTTG TGCACGACGA CATCATTCGG
 TGGCGTTATC CAGCTAAGCG CGAACTGCAA TTTGGAGAAT GGCAGCGCAA
 TGACATTCTT GCAGGTATCT TCGAGCCAGC CACGATCGAC ATTGATCTGG
 40 CTATCTTGCT GACAAAAGCA AGAGAACATA GCGTTGCCTT GGTAGGTCCA
 GCGGCGGAGG AACTCTTTGA TCCGGTTCCT GAACAGGATC TATTTGAGGC
 45 GCTAAATGAA ACCTTAACGC TATGGAAGTC GCCGCCCAGC TGGGCTGGCG
 ATGAGCGAAA TG TAGTGCTT ACGTTGTCCC GCATTTGGTA CAGCGCAGTA
 ACCGGCAAAA TCGCGCCGAA GGATGTCGCT GCCGACTGGG CAATGGAGCG
 50 CCTGCCGGCC CAGTATCAGC CCGTCATACT TGAAGCTAGA CAGGCTTATC
 TTGGACAAGA AGAAGATCGC TTGGCCTCGC GCGCAGATCA GTTGGAAGAA
 55 TTTGTCCACT ACGTGAAAGG CGAGATCACC AAGGTAGTCG GCAAATAATG
 Tctagctaga taacttcgta taatgtatgc tatacgaagt tatAGACATT
 AGCAGATAAA TTAGCAGGAA ATAAAGAAGG ATAAGGAGAA AGAACTCAAG

TAATTATCCT TCGTTCTCTT AATTGAATTG CAATTAACT CGGCCCAATC
 TTTTACTAAA AGGATTGAGC CGAATACAAC AAAGATTCTA TTGCATATAT
 5 TTTGACTAAG TATATACTTA CCTAGATATA CAAGATTTGA AATACAAAAT
 CTAGcaagct t
 10 **SEQ ID NO: 3:**
 GAATTCgagc tcggtaccca aaGCTCCCCC GCCGTCGTTC AATGAGAATG
 GATAAGAGGC TCGTGGGATT GACGTGAGGG GGCAGGGATG GCTATATTTT
 15 TGGGAGCGAA CTCCGGGCGA ATACGAAGCG CTTGGATACA GTTGTAGGGA
 GGGATCCATG gCTTTGGGAT CAATATCAGC AATGCAGTTC ATCCAACGAT
 20 AAACCTAATC CGAATTATAG AGCTACGACA CAATCAgctA GCCCAGAAaG
 AaGaCCGGCC GACATCCGCC GTGCCACCGA GGCGGACATG CCGGCGGTCT
 GCACCATCGT CAACCACTAC ATCGAGACAA GCACGGTCAA CTTCCGTACC
 25 GAGCCGCAGG AACCGCAGGA GTGGACGGAC GACCTCGTCC GTCTGCGGGA
 GCGCTATCCC TGGCTCGTCG CCGAGGTGGA CGGCGAGGTC GCCGGCATCG
 30 CCTACGCGGG CCCCTGGAAG GCACGCAACG CCTACGACTG GACGGCCGAG
 TCGACCGTGT ACGTCTCCCC CCGCCACCAG CGGACGGGAC TGGGCTCCAC
 GCTCTACACC CACCTGCTGA AGTCCCTGGA GGCACAGGGC TTCAAGAGCG
 35 TGGTCGCTGT CATCGGGCTG CCCAACGACC CGAGCGTGCG CATGCACGAG
 GCGCTCGGAT ATGCCCCCCG CGGCATGCTG CGGGCGGCCG GCTTCAAGCA
 40 CGGGAACCTGG CATGACGTGG GTTTCCTGGCA GCTGGACTTC AGCCTGCCGG
 TACCGCCCCG TCCGGTCCTG CCCGTCACCG AGATCTGATG AtctagatGT
 ACCATAACTT CGTATAATGT ATGCTATACG AAGTTATaga tcaattcagt
 45 tgtagggagg gatttATGgG GGAAGCGGTG ATCGCCGAAG TATCGACTCA
 ACTATCAGAG GTAGTTGGCG TCATCGAGCG CCATCTCGAA CCGACGTTGC
 50 TGGCCGTACA TTTGTACGGC TCCGCAGTGG ATGGCGGCCT GAAGCCACAC
 AGTGATATTG ATTTGCTGGT TACGGTGACC GTAAGGCTTG ATGAAACAAC
 GCGGCGAGCT TTGATCAACG ACCTTTTGGG AACTTCGGCT TCCCCTGGAG
 55 AGAGCGAGAT TCTCCGCGCT GTAGAAGTCA CCATTGTTGT GCACGACGAC
 ATCATTCCGT GCGGTTATCC AGCTAAGCGC GAACTGCAAT TTGGAGAATG

GCAGCGCAAT GACATTCTTG CAGGTATCTT CGAGCCAGCC ACGATCGACA
 TTGATCTGGC TATCTTGCTG ACAAAGCAA GAGAACATAG CGTTGCCTTG
 5 GTAGGTCCAG CGGCGGAGGA ACTCTTTGAT CCGGTTCTTG AACAGGATCT
 ATTTGAGGCG CTAAATGAAA CCTTAACGCT ATGGAACTCG CCGCCCGACT
 10 GGGCTGGCGA TGAGCGAAAT GTAGTGCTTA CGTTGTCCCG CATTTGGTAC
 AGCGCAGTAA CCGGCAAAAT CGCGCCGAAG GATGTCGCTG CCGACTGGGC
 AATGGAGCGC CTGCCGGCCC AGTATCAGCC CGTCATACTT GAAGCTAGAC
 15 AGGCTTATCT TGGACAAGAA GAAGATCGCT TGGCCTCGCG CGCAGATCAG
 TTGGAAGAAT TTGTCCACTA CGTGAAAGGC GAGATCACCA AGGTAGTCGG
 CAAATAATGT ctagctagat aacttcgtat aatgtatgct atacgaagtt
 20 atAGACATTA GCAGATAAAT TAGCAGGAAA TAAAGAAGGA TAAGGAGAAA
 GAACTCAAGT AATTATCCTT CGTTCTCTTA ATTGAATTGC AATTAACTC
 25 GGCCCAATCT TTTACTAAAA GGATTGAGCC GAATACAACA AAGATTCTAT
 TGCATATATT TTGACTAAGT ATATACTTAC CTAGATATAC AAGATTTGAA
 ATACAAAATC TAGcaagctt
 30 **SEQ ID NO: 4:**
 ctGAGCAAT CGCCCTGGGT GGGTTACACG ACGCCCCTCT ATGGCCCGTA
 CTGACGGACA CACCGAAGCC CCGGCGGCAA CCCTCAGCGG ATGCCCCGGG
 35 GCTTCACGTT TTCCAGGTC AGAAGCGGTT TTCGGGAGTA GTGCCCCAAC
 TGGGGTAACC TTTGAGTTCT CTCAGTTGGG GGCGTAGGGT CGCCGACATG
 40 ACACAAGGGG TTGTGACCGG Gaagcttgat atcgaattcc tgcAGCCGCG
 GTGCGGGTGC CAGGGCGTGC CCTTGGGCTC CCCGGGCGCG TACTCCActa
 45 gt

While the use of CRE and phi31 integrase recombinases
 are exemplified herein, other recombinases identified by
 GenBank accession numbers and their cognate excision
 50 site sequences are encompassed within the scope of the
 present invention and are set forth below in Table I.

Table 1.
Site-specific recombinases and the corresponding recognition site and references

Name	Rec. site	Recognition sequence	Ref.
XerC	dif	GGTGGCATAATGTATATTATGTTAAAT	(Cornet et al., 1997)
C37841			
XerD	dif	GGTGGCATAATGTATATTATGTTAAAT	(Cornet et al., 1997)
P21891			
Flp	FRT	GAAGTTCTATTCTCTAGAAAGTATAGGAACTTC	(Luetke and Sadowski, 1998)
J01347			
Hin	hixL	TTCTTGAAAAACCAAGGTTTTTGATA	(Glasgow et al., 1989)
NCC_00	hixR	TTATCAAAAAACCTTCCAAAAAGGAAAA	
3197			
R4	attB	AGTTGCCCATGACCATGCCGAAGCAGTGGTAGAAGGGCACCCGGCAGACAC	(Olivares et al., 2001)
D90361	attP	GCAATGTTCCCAAGCGATACCACTTGAAGCAGTGGTACTGCTGTGGGTACACTCTGCGGGTG	
P2	attB	ACCAGAGATAGGGCTTATGCATAAAAAATAAGCCCGTGTAAAGGAGATTACACAGGCTAAGGAGGTGGTTCCTGGTACA	(Yu et al., 1989)
AAD032	attP	TTTTATTTCAAATTATGTACGTAAAAATAAGCCCGTGTAAAGGAGATTAGGGTGTCAACAGTAGGGGCTTTTCAACGG	
97			
lambda	attB	TCCGTTGAAGCCCTGCTTTTATATACTAACTTGAGCGAAACG	(Ross et al., 1979)
AAA965	attP	CGTTTCTCGTTCAGCTTTTATATACTAAAGTTGGCATATATAAAAAAGCATTG	
62			
Mx8	attB	GAAAGGCCCCGGAACCTTGGGATTCGGGCCCTTCTTAGTGGCGAGGAGTACGGGACTTGAACCCGTGGCCCTCCGGCGTGAC	(Tojo et al., 1996)
AAC488	attP	AAAACCCCGAGCAAGTCCGAGAACTTGTGTGGGCGAGGGTGGCGAGGAGTACGGGACTTGAACCCGTACTCCTGGGCTCCTT	
95			
phiadh	attB	ATTACAGCTTAACCTGACCTAATGAAGAAAAATAAATTGTACACTTCTTAGGAGGAGAGTTTCGATCAGTCACCTATATCTAGTTCAAATT	(Raya et al., 1992)
JN0535	attP	CCCTATAGCACTATTGTCTACAGGGCTTTATTATTGCTCTACACTTCTTAGGAGGTTCAATGTGACGAAGTCACACCAATGTTGATATTAAA	

Comet, F., Hallet, B. and Sherratt, D.J.: Xer recombination in *Escherichia coli*. Site-specific DNA topoisomerase activity of the XerC and XerD recombinases. *The Journal of Biological Chemistry* 272 (1997) 21927-31.

Glasgow, A.C., Bruijs, M.F. and Simon, M.I.: DNA-binding properties of the Hin recombinase. *The Journal of Biological Chemistry* 264 (1989) 10072-82.

Luetke, K.H. and Sadowski, P.D.: DNA sequence determinant for Flp-induced DNA bending. *Molecular Microbiology* 29 (1998) 199-208.

Olivares, E.C., Hollis, R.P. and Calos, M.P.: Phage R4 integrase mediates site-specific integration in human cells. *Gene* 278 (2001) 167-76.

Raya, R.R., Fremaux, C., De_Antoni, G.L. and Klaenhammer, T.R.: Site-specific integration of the temperate bacteriophage phi adh into the *Lactobacillus gasseri* chromosome and molecular characterization of the phage (attP) and bacterial (attB) attachment sites. *Journal of Bacteriology* 174 (1992) 5584-92.

Ross, W., Landy, A., Kikuchi, Y. and Nash, H.: Interaction of int protein with specific sites on lambda att DNA. *Cell* 18 (1979) 297-307.

Tojo, N., Sanmiya, K., Sugawara, H., Inouye, S. and Komano, T.: Integration of bacteriophage Mx8 into the *Myxococcus xanthus* chromosome causes a structural alteration at the C-terminal region of the IntP protein. *Journal of Bacteriology* 178 (1996) 4004-11.

Yu, A., Bertani, L.E. and Haggard_Ljungquist, E.: Control of prophage integration and excision in bacteriophage P2: nucleotide sequences of the int gene and att sites. *Gene* 80 (1989) 1-11.

REFERENCES

- Bock, R. (2001) Transgenic plastids in basic research and plant biotechnology. *J. Mol. Biol.* 312, 425-438.
- Chateigner-Boutin, A.L. and Hanson, M.R. (2002) Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared *cis* elements. *Mol. Cell. Biol.* 22, 8448-8456.
- Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J.* 15, 5958-5964.
- Chaudhuri, S., Carrer, H. and Maliga, P. (1995) Site-specific factor involved in the editing of the *psbL* mRNA in tobacco plastids. *EMBO J.* 14, 2951-2957.
- Chin, H.H., Kim, G.D., Marin, I., Mersha, F., Evans, T.C., Chen, L., Xu, M.Q. and Pradhan, S. (2003) Protein trans-splicing in transgenic plant chloroplast: Reconstruction of herbicide resistance from split genes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4510-4515.
- Corneille, S., Lutz, K., Svab, Z. and Maliga, P. (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system. *Plant J.* 72, 171-178.
- Corneille, S., Lutz, K.A., Azhagiri, A.K. and Maliga, P. (2003) Identification of functional lox sites in the plastid genome. *Plant J.* 35, 753-762.
- Gleave, A.P., Mitra, D.S., Mudge, S.R. and Morris, B.A. (1999) Selectable marker-free transgenic plants without sexual crossing: transient expression of cre recombinase and use of a conditional lethal dominant gene. *Plant Mol. Biol.* 40, 223-235.
- Goodin, M.M., Dietzgen, R.G., Schichnes, D., Ruzin, S. and Jackson, A.O. (2002) pGD vectors: versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves. *Plant J.* 31, 375-383.
- Groth, A.C., Olivares, E.C., Thyagarajan, B. and Calos, M.P. (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5995-6000.
- Hajdukiewicz, P.T.J., Gilbertson, L. and Staub, J.M. (2001) Multiple pathways for Cre/lox-mediated recombination in plastids. *Plant J.* 27, 161-170.

- Hare, P.D. and Chua, N.H. (2002) Excision of selectable marker genes from transgenic plants. *Nat. Biotechnol.* 20, 575-580.
- 5 Hood, E.E., Helmer, G.L., Fraley, R.T. and Chilton, M.D. (1986) The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. *J. Bacteriol.* 168, 1291-1301.
- 10 Huang, C.Y., Ayliffe, M.A. and Timmis, J.N. (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422, 72-76.
- Iamtham, S. and Day, A. (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. *Nat. Biotechnol.* 18, 1172-1176.
- 15 Johansen, L.K. and Carrington, J.C. (2001) Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiol.* 126, 930-938.
- 20 Kapila, J., De Rycke, R., Van Montagu, M. and Angenon, G. (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci.* 122, 101-108.
- Klaus, S.M.J., Huang, F.C., Golds, T.J. and Koop, H.-U. (2004) Generation of marker-free plastid transformants using a transiently cointegrated selection gene. *Nat. Biotechnol.* 22, 225-229.
- 25 Kudla, J., Igloi, G.L., Metzlauff, M., Hagemann, R. and Kössel, H. (1992) RNA editing in tobacco chloroplasts leads to the formation of a translatable *psbL* mRNA by a C to U substitution within the initiation codon. *EMBO J.* 11, 1099-1103.
- 30 Lutz, K., Corneille, S., Azhagiri, A.K., Svab, Z. and Maliga, P. (2004) A novel approach to plastid transformation utilizes the *phiC31* phage integrase. *Plant J.* in press.
- 35 Lutz, K.A., Knapp, J.E. and Maliga, P. (2001) Expression of *bar* in the plastid genome confers herbicide resistance. *Plant Physiol.* 125, 1585-1590.
- 40 Maliga, P. (2004) Plastid transformation in higher plants. *Annu. Rev. Plant Biol.* 55, 289-313.
- Ow, D.W. (2002) Recombinase-directed plant transformation for the post-genomic era. *Plant Mol. Biol.* 48, 183-200.
- 45 Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.-Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H.

- and Sugiura, M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* 5, 2043-2049.
- 5 Smith, H.C., Gott, J.M. and Hanson, M.R. (1997) A guide to RNA editing. *RNA* 3, 1105-1123.
- Smith, M.C.M. and Thorpe, H.M. (2002) Diversity in the serine recombinases. *Mol. Microbiol.* 44, 299-307.
- 10 Staub, J.M. (2002) Expression of recombinant proteins via the plastid genome. In *Handbook of industrial cell culture: mammalian, microbial and plant cells*. (Parekh, S.R. and Vinci, V.A., eds). Totowa, NJ: Humana Press Inc., pp. 261-280.
- 15 Stegemann, S., Hartmann, S., Ruf, S. and Bock, R. (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8828-8833.
- 20 Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc. Natl. Acad. Sci. U.S.A.* 90, 913-917.
- Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein from tomato bushy stunt virus. *Plant J.* 33, 949-956.
- 25 Ye, G.N., Colburn, S., Xu, C.W., Hajdukiewicz, P.T.J. and Staub, J.M. (2003) Persistence of unselected transgenic DNA during a plastid transformation and segregation approach to herbicide resistance. *Plant Physiol.* 133, 402-410.
- 30 Zoubenko, O.V., Allison, L.A., Svab, Z. and Maliga, P. (1994) Efficient targeting of foreign genes into the tobacco plastid genome. *Nucleic Acids Res.* 22, 3819-3824.
- 35

While certain of the preferred embodiments of the present invention have been described and specifically

40 exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention, as set forth in the following claims.

What is claimed is:

1. A site specific recombination method for removal of heterologous nucleic acid sequences from the plastid genome, said method comprising:

- 5 a) providing a transplastomic plant cell, said plastids of said cell comprising a heterologous nucleic acid sequence flanked by excision sites and a nucleic acid sequence encoding at least one foreign gene of interest which is not flanked by excision sites;
- 10 b) providing a DNA construct, said construct comprising a promoter operably linked to a nucleic acid encoding a protein having excision activity and optionally comprising a sequence encoding a plastid transit peptide sequence and a selectable marker gene
- 15 c) introducing the construct of step b) into the plant cell of step a) such that said DNA construct enters said cell and the proteins encoded thereby are transiently expressed for a suitable time period, said
- 20 proteins catalyzing the excision of the heterologous sequence from the plastids in the plant cell of step a), thereby removing said heterologous sequence; and
- 25 d) identifying those transplastomic cells which lack both said heterologous nucleic acid sequence and said protein having excising activity.

2. The method of claim 1, further comprising generating a plant from the transplastomic plant cell of step d).

30 3. The method of claim 1, wherein said heterologous nucleic acid encodes a selectable marker gene which confers resistance to a selection agent.

4. The method of claim 3, wherein expression of said selectable marker gene confers resistance to a selection agent selected from the group consisting of spectinomycin, kanamycin, hygromycin, streptomycin, 5 phophinotricin, basta, glyphosate, and bromxynil.
5. The method of claim 1, wherein said protein having excision activity is selected from the group consisting of CRE, and phi31 integrase and said excision sites are 10 selected from the group consisting of LOX and ATT.
6. The method of claim 1, wherein said recombinase and its cognate excision site are selected from the group listed in Table I.
- 15 7. The method of claim 1, wherein the construct of step b) is introduced into said cells via a method selected from the group consisting of Agrobacterium infiltration, PEG fusion, biolistic delivery, CaPO₄-mediated transfection, 20 and electroporation.
8. The method of claim 1, wherein said construct of step b) is introduced into said plant cell by agrobacterium infiltration and is expressed in the nucleus of said 25 plant cell.
9. The method of claim 1, wherein said plant cell is obtained from a plant selected from the group consisting of tobacco, rice, potato, oil seed rape, corn, wheat, 30
10. A DNA construct for use in the method of claim 1, said construct consisting essentially of a sequence

selected from the group consisting of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3 and SEQ ID NO:4.

1/8

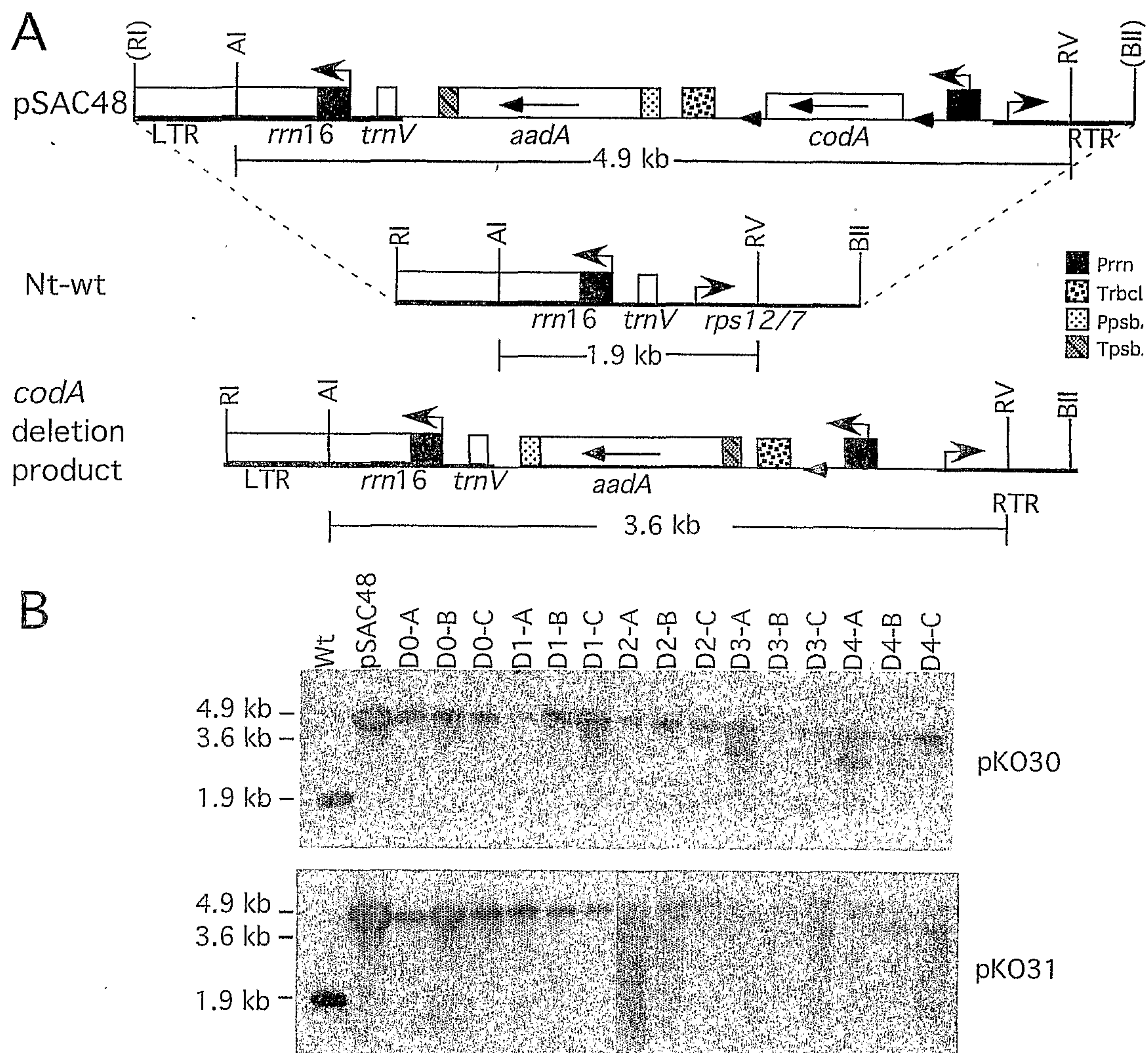


Figure 1

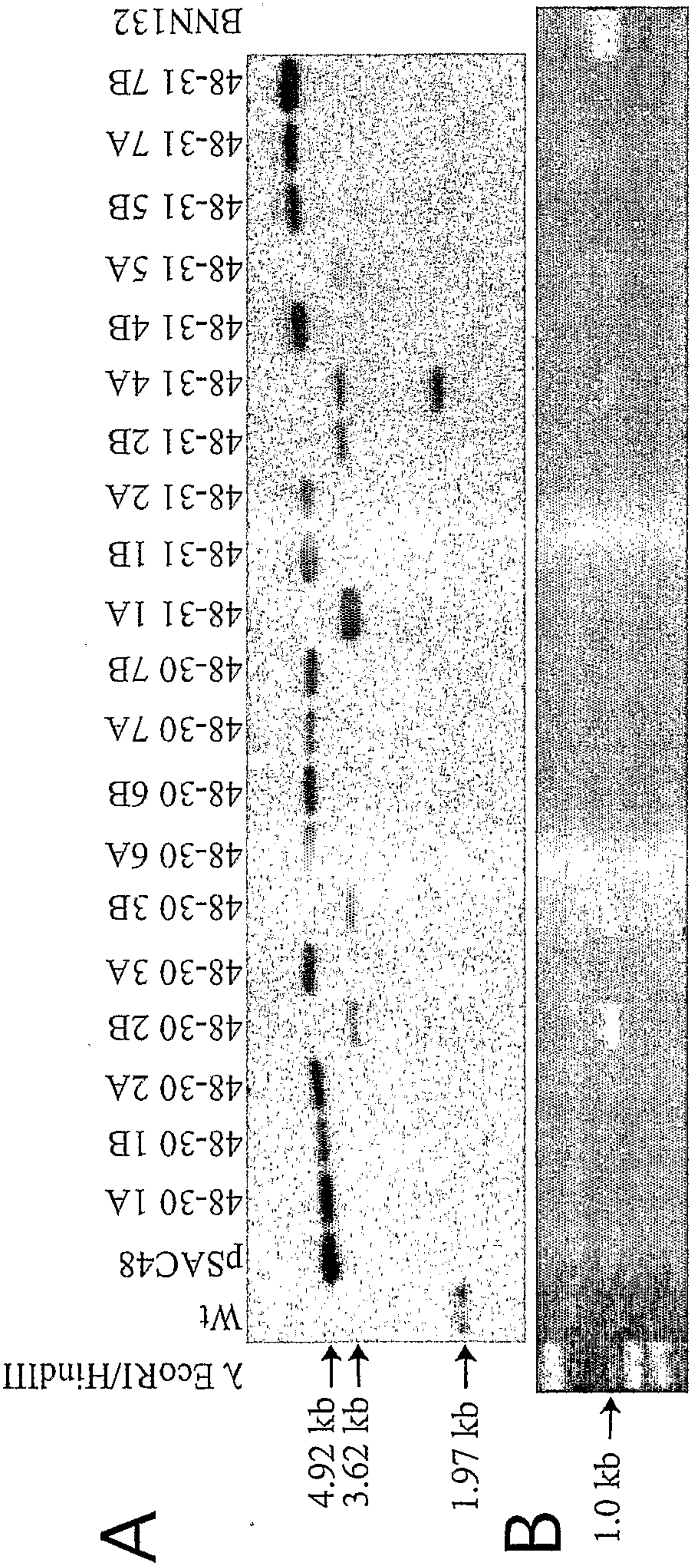


Figure 2

3/8

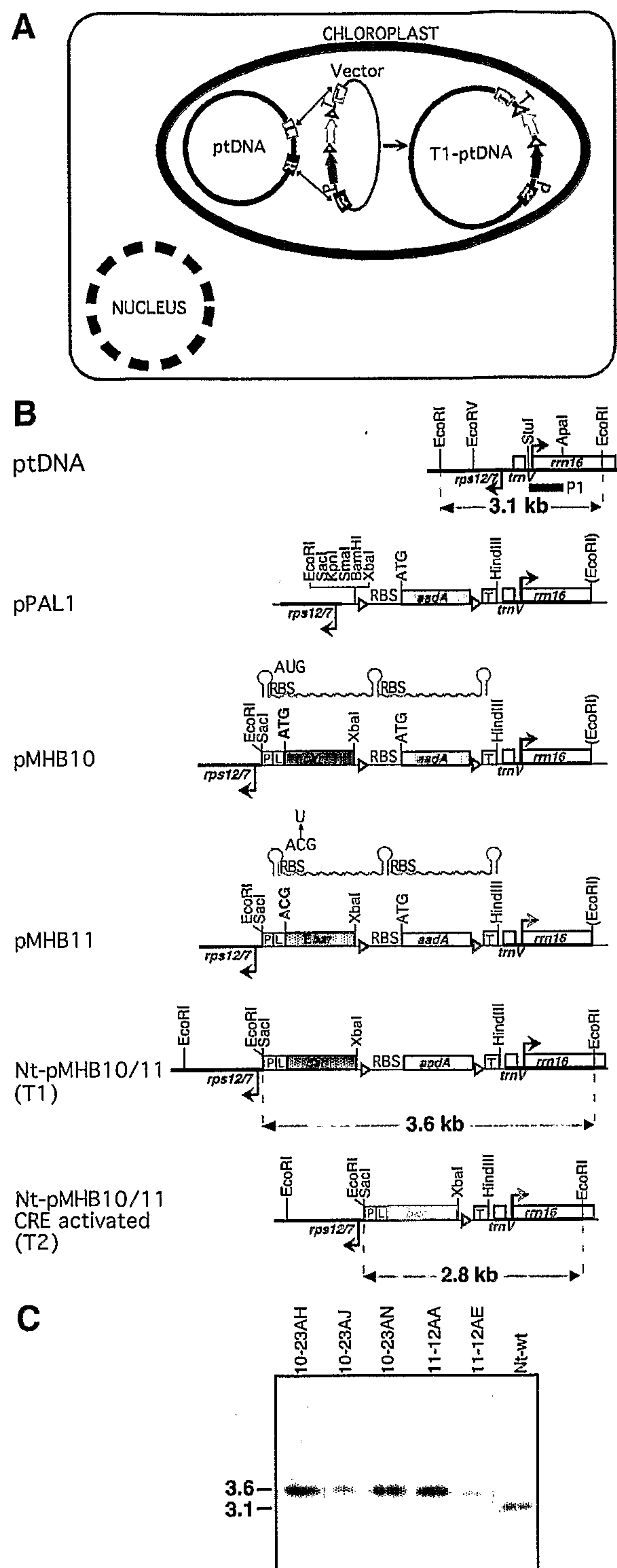


Figure 3

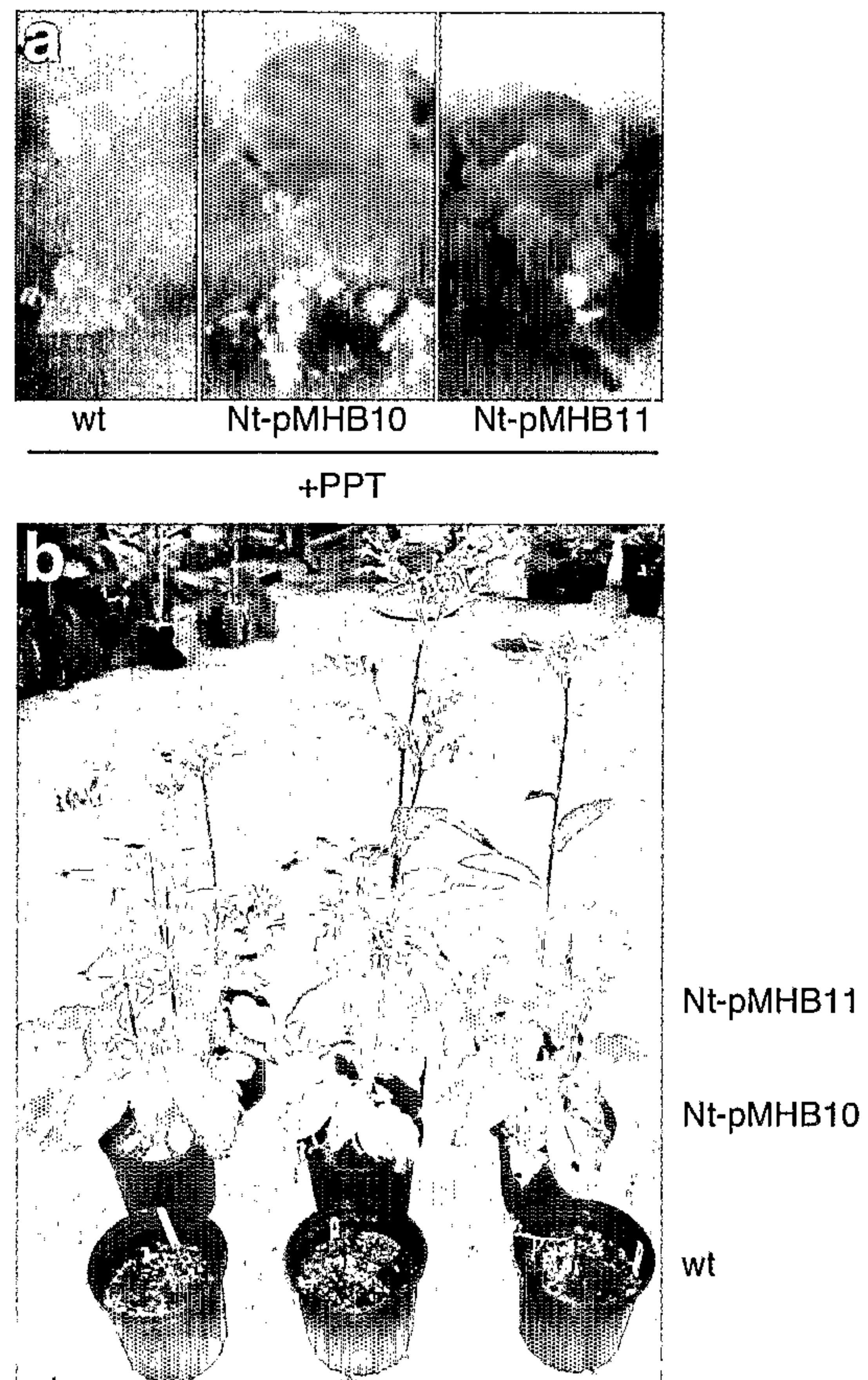


Figure 4

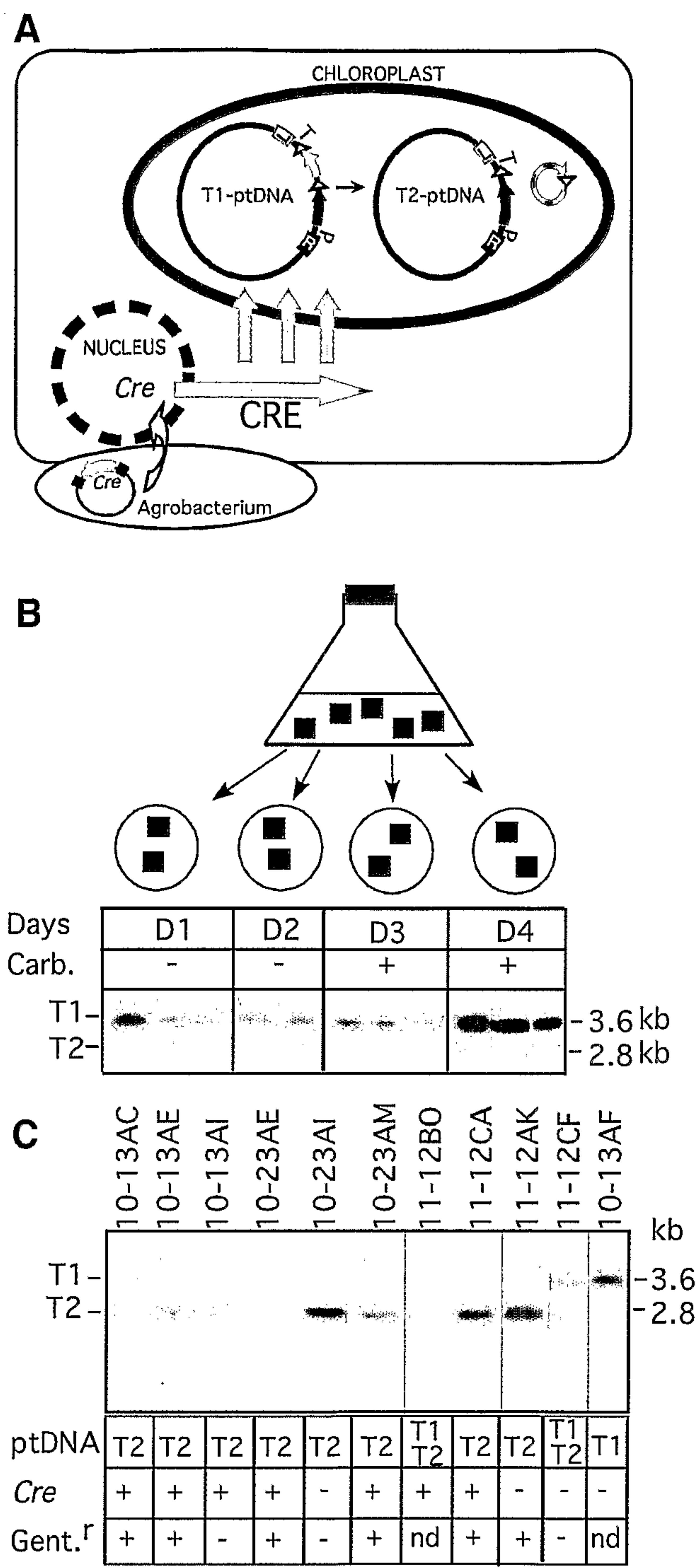


Figure 5

6/8

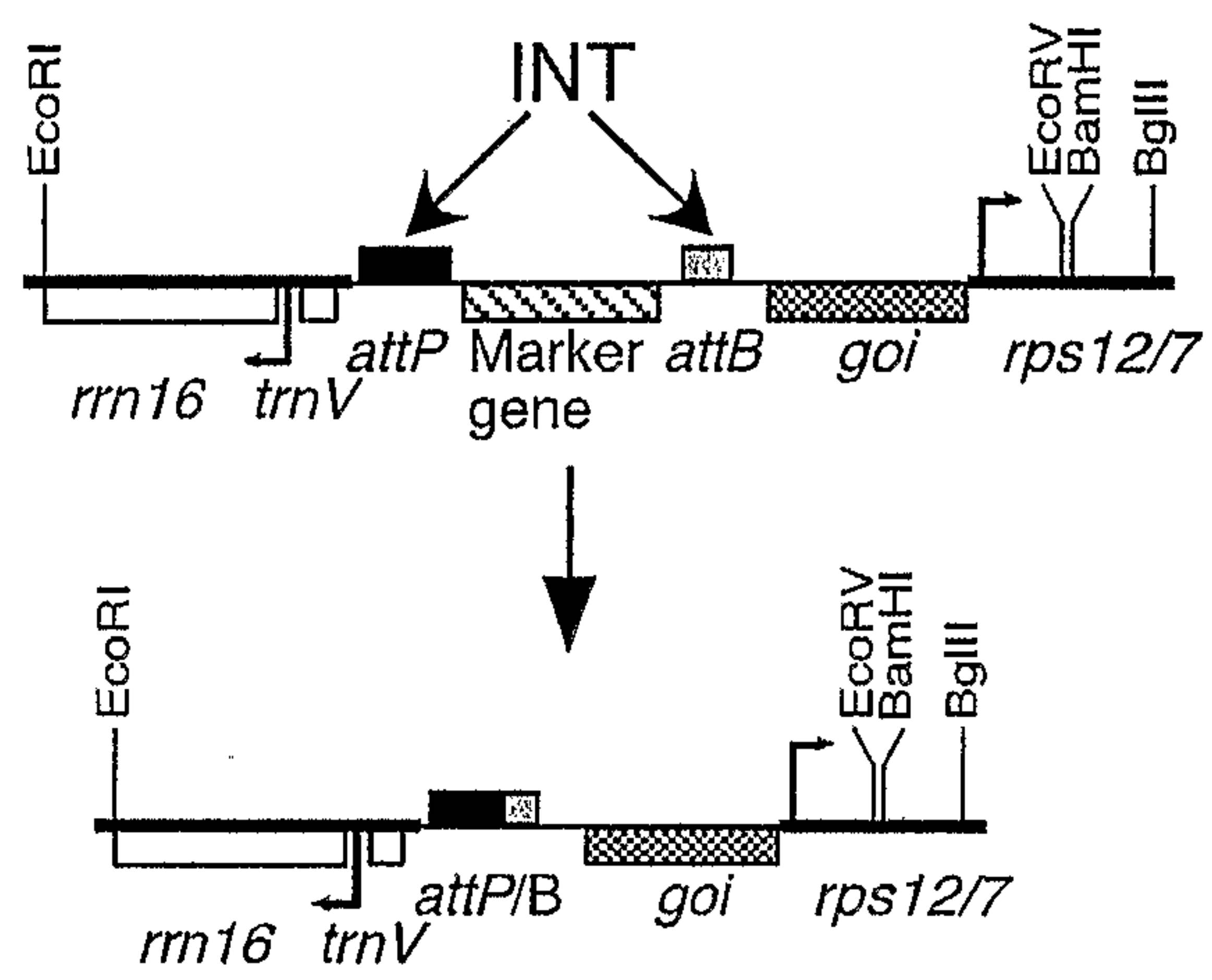


Figure 6

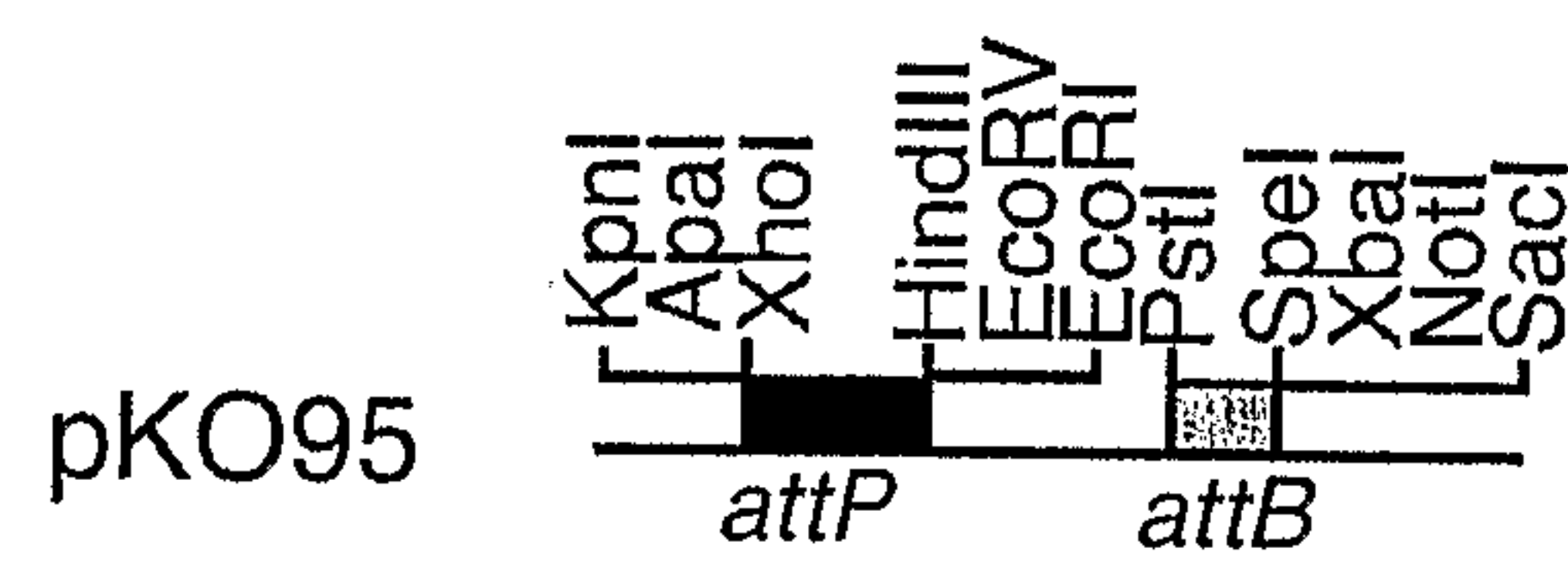


Figure 7

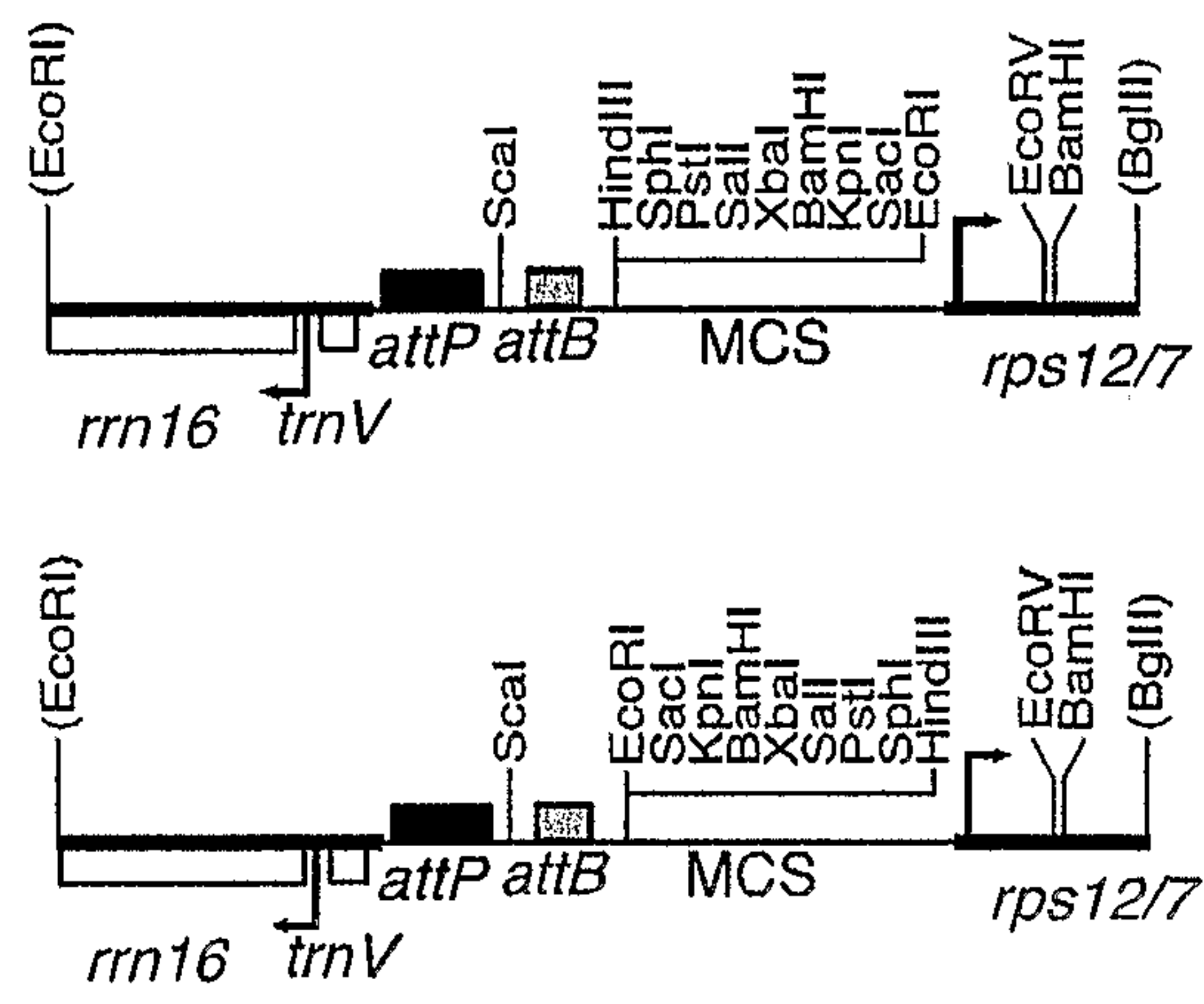


Figure 8

