

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 December 2008 (04.12.2008)

PCT

(10) International Publication Number
WO 2008/145757 A1

(51) International Patent Classification:
C12N 15/82 (2006.01) *A01H 5/00* (2006.01)

(21) International Application Number:
PCT/EP2008/056770

(22) International Filing Date: 2 June 2008 (02.06.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
07109346.2 31 May 2007 (31.05.2007) EP

(71) Applicant (for all designated States except US): **VRIJE UNIVERSITEIT BRUSSEL** [BE/BE]; Pleinlaan 2, B-1050 Brussel (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ANGENON, Geert** [BE/BE]; Sint-Denijslaan 257, B-9000 Gent (BE). **VERWEIRE, Dimitri** [BE/BE]; Merlostraat 8/D bus 18, B-1180 Ukkel (BE). **VAN EX, Frédéric** [BE/BE]; Vlisseingenstraat 33, B-1080 Sint-jans-molenbeek (BE).

(74) Agents: **BRANTS, Johan, Philippe, Emile** et al.; E. Gevaertdreef 10a, B-9830 Sint-martens-latem (BE).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv))

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

(54) Title: TARGETED GENOME MODIFICATIONS IN PLANTS

(57) Abstract: The invention provides materials and methods for performing targeted genome modification in plants and plant cells. More particularly, the invention allows to present a targeting cassette preferentially in specific cell types and/or in cells at particular developmental or cell division stages, which allows for more efficient gene targeting.

WO 2008/145757 A1

TARGETED GENOME MODIFICATIONS IN PLANTS

FIELD OF THE INVENTION

The invention relates to the field of molecular biology. More particularly, it concerns materials and methods for performing targeted genome modification in plants and plant cells.

5 BACKGROUND OF THE INVENTION

Targeted genome modification (or gene targeting) broadly refers to directed (non-random) alteration of specific DNA sequences in their genomic loci in cells or organisms. Gene targeting involves the transfer of genetic information from an exogenously provided nucleic acid molecule, usually referred to as a targeting cassette, to a specific target locus in the
10 host's nuclear or organellar (e.g., mitochondrial or plastidal) genome.

The targeting cassette typically comprises sequences homologous to the target locus. However, said sequences may be modified such as to alter the genetic information vis-à-vis the target locus, for example, through the insertion, deletion or substitution of one or more base pairs. Due to the introduction of such sequence changes, the functionality of sequences
15 at the target locus may be altered, such as, e.g., partly or wholly impaired, enhanced or functionally modified and/or new functionalities may be introduced, etc.

Given the ability to alter endogenous sequences, gene targeting may allow, *inter alia*, to study the function of genes and regulatory sequences; to disrupt or to modulate the level or specificity of gene expression; to eliminate, improve or otherwise alter the functional
20 properties of endogenous proteins; or to introduce transgenes of interest at predetermined loci, thereby ensuring their expression and/or avoiding possible detrimental effects of random transgene integration. Hence, improved gene targeting methods are of considerable value for research as well as industry, and there exists continual need for such.

An efficient and reproducible gene targeting system is lacking in higher plants, especially with
25 respect to modifying the nuclear genome. For example, previous studies have established that targeted nuclear genome modifications in plants occur at very low frequency, usually ranging between 10^{-4} to 10^{-5} targeted integrations per one random integration event (e.g., Halfter *et al.* 1992. Mol Gen Genet 231: 186-93).

Although recent strategies could increase the frequency of targeted integrations in plants, such as, *e.g.*, introduction of double strand breaks into target loci (Wright *et al.* 2005. Plant J 44: 693-705) or over-expression of the yeast RAD54 gene (Shaked *et al.* 2005. PNAS 102: 12265-9), these approaches entail caveats. For example, the routine use of the technique of Wright *et al.* 2005 may be hampered since it requires the inefficient, tedious and not generally applicable protoplast transformation and regeneration process. Moreover, the prolonged exposure of cells to zinc-finger nucleases used by Wright *et al.* 2005 may cause cell toxicity. Similarly, the method of Shaked *et al.* 2005 relies on over-expressing a protein that plays an important role in recombination and repair processes. This may alter the balance of DNA maintenance mechanisms and eventually introduce non-intended DNA modifications disrupting or altering various molecular or physiological processes in the cells.

SUMMARY OF THE INVENTION

The present invention provides materials and methods that address one or more of the above discussed problems of the art.

In particular, the inventors realised that while the overall frequency of targeted DNA modifications in plants and plant cells is low, more efficient gene targeting, especially nuclear gene targeting, can be achieved in specific plant cells, such as, *e.g.*, in select plant cell types, or in plant cells at particular developmental or cell division stages. Without wishing to be bound by theory, the inventors contemplate that the incidence of homologous recombination required for gene targeting may be elevated in said desired cell types and/or at said particular developmental or cell division stages. To take advantage hereof, the invention provides a strategy to present a targeting cassette preferentially in select cell types or at said particular developmental or cell division stages or, in a more general notion, preferentially in plant cells having increased frequency of homologous recombination, and thereby increase the frequency of targeted genome modifications in plants.

Accordingly, in an aspect, the invention provides a method for targeted genome modification in plants comprising:

(a) providing a plant comprising a locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:

(aa) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-

specific recombinase R thereon results in release of the targeting cassette TC from the locus L1, and

(ab) comprising at least one region of homology (RH) to a sequence (S) at a locus (L2) of said plant different from the locus L1; and

- 5 (b) expressing the site-specific recombinase R preferentially in cells of the plant having increased frequency of homologous recombination, thereby releasing the targeting cassette TC from the locus L1 in said cells,

whereby homologous recombination occurs in at least some of said cells between the at least one region of homology RH of the targeting cassette TC released in step (b) and the
10 sequence S at the locus L2.

Preferably, said method is aimed for targeted nuclear genome modification in plants. Then, the loci L1 comprising the targeting cassette TC and L2 are in the nuclear genome of said plant, *i.e.*, the loci L1 and L2 are nuclear genome loci.

In a further aspect, the invention provides a method for targeted genome modification in plant
15 cells *ex vivo* comprising:

(a) providing isolated plant cells comprising a locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:

(aa) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC from
20 the locus L1, and

(ab) comprising at least one region of homology (RH) to a sequence (S) at a locus (L2) of said plant cells different from the locus L1; and

- (b) expressing the site-specific recombinase R preferentially in those cells having increased
25 frequency of homologous recombination, thereby releasing the targeting cassette TC from the locus L1 in said cells,

whereby homologous recombination occurs in at least some of said cells between the at least one region of homology RH of the targeting cassette TC released in step (b) and the sequence S at the locus L2.

Preferably, said method is aimed for targeted nuclear genome modification in plant cells *ex vivo*. Then, the loci L1 comprising the targeting cassette TC and L2 are in the nuclear genome of said plant cell, *i.e.*, the loci L1 and L2 are nuclear genome loci.

Preferably, the cells in which the site-specific recombinase R is preferentially expressed in step (b) above, hence the desired cells contemplated to have increased incidence of homologous recombination:

- may be chosen from cells of meristem, preferably cells of apical meristem or stem cells; cells of microsporangia; or cells of megasporangia; and/or
- may be chosen from cells at one or more stages of mitotic cell cycle chosen from interphase, G₁ phase, G₁/S transition, S phase, S/G₂ transition, G₂ phase, G₂/M transition or M (mitosis), more preferably cells at G₂/M transition; and/or
- may be chosen from meiocytes, preferably microsporocytes and/or megasporocytes, more preferably wherein in said meiocytes the site-specific recombinase R is preferentially expressed at one or more stages of meiotic cell division chosen from interphase, G₁ phase, G₁/S transition, S phase, S/G₂ transition, G₂ phase, G₂/Meiosis I transition, Meiosis I or Meiosis II.

The inventors realised that presentation of the targeting cassette TC preferentially in any of the above recited cell types or at any of the above recited developmental or cell division stages can increase the efficiency of targeted genomic modification in plants.

In the above aspects, the targeting cassette TC resides in the genome, preferably nuclear genome, of the plant or plant cells and may be released, and thereby made available for homologous recombination with the target locus, by means of preferential expression of the recombinase R in one or more of the desired cell types and/or in cells at one or more of the desired developmental or cell division stages. Hence, the above aspects advantageously avoid the need to introduce the targeting cassette TC selectively in said cell types or at said developmental or cell division stages, which may be laborious and/or less efficient. Indeed, direct introduction of a targeting cassette to select cells and regenerating plants there from is quite unfeasible. For example, while meiocytes are very permissive to recombination, the introduction of nucleic acids into meiocytes and regeneration of plants there from has not been reported.

In a preferred embodiment, the plant or isolated plant cells may comprise a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements (e.g., comprising a promoter and possibly an enhancer), capable of effecting the preferential expression of said recombinase R in the desired cell types and/or in cells at one or more of the desired developmental or cell division stages as disclosed above. Preferably, said sequence may reside in the nucleus, e.g., in the nuclear genome, of said plant or plant cells. Advantageously, this avoids the need to introduce the recombinase R protein or an expression cassette encoding the recombinase R selectively in said cell types or at said developmental or cell division stages, which may be laborious and/or less efficient.

In a preferred embodiment, said regulatory elements, e.g., said promoter and optionally enhancer, do not drive or effect constitutive expression of the recombinase R, i.e., do not effect expression which would be substantially uniform throughout most or all cells and tissues of plants and/or throughout most or all developmental or cell cycle stages of said cells and tissues. Hence, in embodiment said regulatory elements, e.g., said promoter and optionally enhancer, is (are) not constitutive. Constitutive expression of the recombinase R would lead to undesirable excision and degradation or dilution (e.g., during subsequent cell division) of the targeting cassette before the cells reached a developmental or cell cycle stage, or before the cells gave rise to a cell type, which are permissive to homologous recombination. This would dramatically reduce the efficiency of the gene targeting.

In a preferred embodiment, said regulatory elements, e.g., said promoter and optionally enhancer, do not require external stimuli to drive or effect expression of the recombinase R. By external are meant stimuli which are not intrinsic or endogenous, i.e., which do not originate from within the plant or plant cells, but which originate or are applied from without or outside of the plant or plant cells. External stimuli may encompass without limitation chemical, biological and/or physical (e.g., heat) agents to which the plant or plant cells are exposed. Preferably, external stimuli as used here do not include factors required for normal plant development (such as, e.g., certain photoperiodic conditions). The present embodiment avoids the need to apply such external stimuli, which is laborious and can lead to suboptimal or indiscriminate induction of expression. For example, with external methods, such as physical induction methods including heat shock, it would not be possible to induce expression of the recombinase R in a particular cell type such as meiocytes, since it is not apparent from outside observation when cells in the desired stage are present, and such

target cells are not present at one particular moment – for example, floral meristems and meiocytes develop gradually in a plant. Hence, such external induction would only release the cassette in a small proportion of desired cells, thereby dramatically reducing the effectiveness of the gene targeting.

- 5 In a preferred embodiment, said regulatory elements driving or effecting expression of the recombinase R, *e.g.*, said promoter and optionally enhancer, may thus be denoted as cell-specific, tissue-specific and/or developmental stage-specific with respect to plant cells having increased frequency of homologous recombination as contemplated herein.

In a preferred embodiment, said regulatory elements, *e.g.*, said promoter and optionally
10 enhancer, can drive or effect expression of the recombinase R at levels which cause the excision of the targeting cassette TC from locus L1, in the desired cell types and/or in cells at one or more of the desired developmental or cell division stages, but substantially not in progenitors or predecessors of said desired cells. This ensures that the targeting cassette would still be present in the genome at the start of the desired expression of the recombinase
15 R.

Hence, the invention also provides the following particular aspects and embodiments:

(I) a method for targeted genome modification (preferably nuclear genome modification) in plants comprising:

- 20 (a) providing a plant comprising a genome (preferably nuclear genome) locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:

(aa) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC from the locus L1, and

- 25 (ab) comprising at least one region of homology (RH) to a sequence (S) at a genome (preferably nuclear genome) locus (L2) of said plant different from the locus L1; and

- (b) expressing the site-specific recombinase R at a level which causes the excision of the targeting cassette TC from locus L1 preferentially in cells of the plant having
30 increased frequency of homologous recombination but substantially not in progenitors

or predecessors of said cells, wherein the plant comprises a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements capable of effecting said expression of said recombinase R, thereby releasing the targeting cassette TC from the locus L1 in said cells,

- 5 whereby homologous recombination occurs in at least some of said cells between the at least one region of homology RH of the targeting cassette TC released in step (b) and the sequence S at the locus L2;

(II) a method for targeted genome modification (preferably nuclear genome modification) in plant cells *ex vivo* comprising:

- 10 (a) providing isolated plant cells comprising a genome (preferably nuclear genome) locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:

(aa) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC from the locus L1, and

15

(ab) comprising at least one region of homology (RH) to a sequence (S) at a genome (preferably nuclear genome) locus (L2) of said plant cells different from the locus L1; and

- (b) expressing the site-specific recombinase R at a level which causes the excision of the targeting cassette TC from locus L1 preferentially in those cells having increased frequency of homologous recombination but substantially not in progenitors or predecessors of said cells, wherein the cells comprise a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements capable of effecting said expression of said recombinase R, thereby releasing the targeting cassette TC from the locus L1 in said cells,
- 20
- 25

whereby homologous recombination occurs in at least some of said cells between the at least one region of homology RH of the targeting cassette TC released in step (b) and the sequence S at the locus L2;

- (III) the method as set forth in any one of (I) or (II), wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter which is not constitutive;
- 30

(IV) the method as set forth in any one of (I) or (II), wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter which does not require external stimuli to drive or effect expression of said recombinase R;

5 (V) the method as set forth in any one of (I) or (II), wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter which is cell-specific, tissue-specific and/or developmental stage-specific with respect to cells having increased frequency of homologous recombination.

In related aspects, the invention provides plants and plant cells useful in the methods of the invention and in particular a plant or plant cell comprising:

- 10 (i) a locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:
- (ia) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC from the locus L1, and
 - 15 (ib) comprising at least one region of homology (RH) to a sequence (S) at a locus (L2) of said plant or plant cell different from the locus L1; and
- (ii) the sequence S at the locus L2 different from the locus L1.

Preferably, such plants and plant cells are suited for targeted nuclear genome modification. Then, the loci L1 comprising the targeting cassette TC and L2 are in the nuclear genome of
20 said plant or plant cells, *i.e.*, the loci L1 and L2 are nuclear genome loci.

Preferably, the plant or plant cell may further comprise a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements capable of effecting the preferential expression of said recombinase R in the desired cell types and/or in cells at one or more of the desired developmental or cell division stages as disclosed above.

25 In further related aspects, the invention provides genetic constructs useful in the methods of the invention and in particular a genetic construct suitable for transformation, esp. of plants, comprising a targeting cassette (TC), said targeting cassette:

(I) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in excision of the targeting cassette TC from the genetic construct, and

- 5 (II) comprising at least one region of homology (RH) to an endogenous sequence of a plant or plant cell.

Preferably, said genetic construct may further comprise a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements which can effect the preferential expression of said recombinase R in the desired cell types and/or in cells at one or more of the desired developmental or cell division stages as disclosed above.

In a further aspect, the invention provides use of the genetic constructs as disclosed herein to transform a plant or plant cell.

Among others, the invention provides the use of

- a genetic construct comprising a targeting cassette (TC), said targeting cassette:

15 (I) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in excision of the targeting cassette TC from the genetic construct, and

(II) comprising at least one region of homology (RH) to an endogenous sequence, preferably endogenous nuclear genome sequence, of a plant or plant cell, and

- a genetic construct comprising a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements which can effect the preferential expression of said recombinase R in the desired cell types and/or in cells at one or more of the desired developmental or cell division stages as disclosed above to transform a plant or plant cell.

25 It shall be further appreciated that while the invention may preferably use a site-specific recombinase R and corresponding recombinase sites RS, in alternative aspects these elements may be substituted, respectively, by a site specific nuclease (e.g., an endonuclease, meganuclease, such as, e.g., I-Sce I, or a zinc finger nuclease – see below) and a pair of recognition sites specifically recognised by said nuclease, which recognition sites thus flank the targeting cassette TC. Hereby, the nuclease can facilitate the release of the targeting

cassette TC from the genome preferentially in specific cell types and/or in cells at particular developmental or cell division stages as explained above.

These and further aspects and preferred embodiments of the invention are described in the following sections and in the appended claims.

5 BRIEF DESCRIPTION OF FIGURES

Figure 1 illustrates schematic presentation of the TARGETING vectors TNG-1a, TNG-2l, TNG-3 and TNG-6. LOX: target sequence of the CRE recombinase; Pnos: *nopaline synthase* promoter; hpt: *hygromycin phosphotransferase*; Tnos: 3' polyA *nopaline synthase*; SDS: promoter sequence *solodancers* gene from *A. thaliana*; AP1: promoter sequence *apetala1* gene from *A. thaliana*; P35S: *CaMV* 35S promoter sequence; CRE-intron: *cre* recombinase with intron; T35S: 3' polyA 35S; codA: *cytosine deaminase*; pSDS is flanked upstream by the *attB1* and downstream by the *attB2* sequence (Gateway™, Invitrogen); lambda: a 5.4 kb fragment of phage lambda DNA; 5'delGUS-nptII: coding sequence of the *gus:nptII* fusion gene with a deletion at the 5'end of the *gus* gene; tg7: poly A signal of the tg7 gene of *Agrobacterium*; csr1-1: allele of the *A.thaliana als* gene containing an amino acid change Pro-197-Ser compared to the WT allele rendering resistance to the sulfonylurea herbicide chlorsulfuron; Pubi: Promoter of the *Helianthus annuus* Gubi gene coding for the hexaubiquitin protein. TNG-3 contains the SDS promoter, TNG-6 the AP1 promoter. TNG-1a contains one I-Sce I site. TNG-2l contains two I-Sce I sites.

Figure 2 illustrates schematic presentation of the K2L610 T-DNA, present in the FK24 plant line. A: original allele, B: recombined allele.

Primer 1: 5'-CCACACATTATACGAGCCGGAAGCAT-3',

Primer 2: 5'-GAGCGTCGCAGAACATTACA-3',

Primer 3: 5'-TGATCCATCTTGAGACCACAGGCCAC-3'.

Figure 3 illustrates results of PCR on T2 generation SDS-CRE1::FK24 and AP1-CRE1::FK24 plants. no exc = no recombined allele could be detected by PCR; part exc = both a recombined and a non-recombined allele could be detected; compl exc = only a recombined allele could be detected.

Figure 4 illustrates the results of PCR on T1 (A) and T2 (B) generation CRE3::FK24 plants, wherein CRE3 contains the respective promoters CLV3, LFY, AP1, AG or SDS. no exc = no

recombined allele could be detected by PCR; part exc = both a recombined and a non-recombined could be detected; compl exc = only a recombined allele could be detected.

Figure 5 illustrates the results after PCR on T1 (A) and T2 (B) generation CRE1::FK24 and CRE3::FK24 plants, wherein CRE1 and CRE3 contain the promoters AP1 or SDS. no exc =
 5 no recombined allele could be detected by PCR; part exc = both a recombined and a non-recombined could be detected; compl exc = only a recombined allele could be detected.

Figure 6 illustrates schematic presentation of the TARGET T-DNA. P35S: *CaMV* 35S promoter sequence; T35S: 3' polyA 35S; phage lambda: 5.4kb fragment of lambda DNA; GUS-3'delNPTII: coding sequence for the *gus:nptII* fusion gene with a deletion at the 3' end
 10 of the *nptII* gene; BAR: coding sequence of the *bar* gene rendering resistance to phosphinotrycin. Primers are represented by arrows. Primer 1: TDN-F; Primer 2: TDN-R; Primer 3: TDN-NesF and Primer 4: 35S-NesR (see Example 1). E: GUS1 probe; D: GUS2 probe; C: lambda1 probe; B: lambda2 probe; A: BAR probe (see Example 1). Lambda2 probe and GUS2 probe overlap with the *BspHI* restriction site.

Figure 7 illustrates PCR confirmation of the insertion of TARGET T-DNA at chromosome 2 (17993801) in line A and at chromosome 1 (1728401) in line D. Lanes A = PCR on DNA from line A; Lanes D = PCR on DNA from line A. Left panel = PCR using primers genA3F2 and pptProbe-R; expected product of about 1834 bp for line A, no product for line D. Right panel =
 15 PCR using primers genD5R2 and pptProbe-R; no product for line A, expected product of about 1087 bp for line D.

Figure 8 illustrates an overview of the gene targeting strategy in which the CRE is delivered in *trans*. The figure refers to TARGET and TARGETING vectors as described, respectively, in Fig. 6 and Example 2 and in and Fig. 1 and Example 3.

Figure 9 illustrates overview of the target locus and the recombined allele. B: GUS1 probe ;
 25 A: BAR probe. Primers:

Primer 1: NPTII 3'del: 5'-CCAAGCTCTTCAGCAATATCACGGG-3'

Primer 2: GUS 5'del: 5'-CAGTCTGGATCGCGAAACTGTGG-3'

Primer 3: NPTII 3'del-R: 5'-CCCGTGATATTGCTGAAGAGCTTGG-3'

Primer 4: Target-F1: 5'-CCATGTTGGCAAGCTGCTCTAGC-3'

30 Primer 5: TDN-R: 5'-GCAAGGCGATTAAGTTGGGTAACGCCAGGG-3'

Primer 6: 35S-NesR: 5'-CCCACTATCCTTCGCAAGACCC-3'

Primer 7: genA5R2: 5'-ACGTTCTTAATTCGTATCGACAAAGTGACG-3'

Figure 10 illustrates an overview of the gene targeting strategy in which the CRE is delivered in *cis*. The figure refers to TARGET and TARGETING vectors as described, respectively, in Fig. 6 and Example 2 and in Fig. 1 and Example 3.

5 **Figure 11** illustrates Southern blot screening of recombination events.

DETAILED DESCRIPTION OF THE INVENTION

As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

10 The terms "comprising", "comprises" and "comprised of" as used herein are synonymous with "including", "includes" or "containing", "contains", and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps.

The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

15 The term "about" as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, is meant to encompass variations of +/-20% or less, preferably +/-10% or less, more preferably +/-5% or less, even more preferably +/-1% or less, and still more preferably +/-0.1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. It is to be understood that the value to which the modifier "about" refers is itself also specifically, and preferably, disclosed.

20 Unless otherwise defined, all terms used in disclosing the invention, including technical and scientific terms, have the meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. By means of further guidance, term definitions are included to better appreciate the teaching of the present invention. When specific terms are defined in connection with a particular aspect or embodiment, such denotation or connotation is meant
25 to apply throughout this specification, *i.e.*, also in the context of other aspects or embodiments, unless otherwise stated.

All documents cited in the present specification are hereby incorporated by reference in their entirety. In particular, the teachings of all documents herein specifically referred to are incorporated by reference.

Definitions

The following terms and definitions shall be used to better appreciate the specification and claims.

As used herein, the term "plant" includes plant organisms, as well as parts of plants such as pollen, flowers, seeds, leaves, stems, and the like. The term may also encompass plant cells, plant protoplasts, plant cells or tissue culture from which plant organisms can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or plant parts. It also should be appreciated that plant cells, tissue or organism can be from various plant species, such as by way of example and not limitation: maize, wheat, rice, barley, sorghum, tobacco, tomato, potato, Brassica spp., soybean, pea, sunflower, cotton, peanut, Arabidopsis thaliana, Nicotiana, and Medicago.

The terms "a part" or "a portion" denote any smaller part of that referred to.

The term "ex vivo" denotes outside, or external to, the body of a plant organism or a plant part, and specifically refers to plant tissues or cells removed from the body of a plant organism or a plant part and maintained and/or propagated (*i.e.*, cultured) outside of said body, *e.g.*, in a culture vessel.

The term "isolated cell" generally denotes a cell that is not associated with one or more cells or one or more cellular components with which said cell is associated *in vivo* or *in planta*. Typically, an isolated cell may have been removed from its natural environment, *e.g.*, from a plant organism, plant part or a plant tissue, or may result from propagation, *e.g.*, *ex vivo* propagation, of a cell that has been removed from its natural environment.

As used herein, the term "locus" (plural "loci") refers to a specific location or site on a chromosome. In relation to genetically transformed organisms, such as plants, the term may denote one or more sites on one or more chromosomes, where the transformed nucleic acid(s) were inserted. The term "chromosome" may encompass nuclear chromosomes, nuclear episomes (*e.g.*, plasmids, artificial chromosomes, *etc.*), as well as organellar chromosomes, such as, *e.g.*, mitochondrial and plastidal chromosomes. In the context of the present disclosure, this term preferably refers to nuclear chromosomes or episomes, more preferably to nuclear chromosomes.

The term "sequence" as used herein refers to a nucleic acid sequence. Where the context implies that a sequence forms part of the genomic material, the term refers to a DNA

sequence, *i.e.*, a sequence consisting essentially of the deoxyribonucleotides A, G, T and C, and derivatives thereof (*e.g.*, methylated derivatives) as can be expected to occur in the genomic material of a cell, tissue or an organism.

The term "encoding" refers to the inherent property of a given sequence of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as a template for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (*esp.*, rRNA, tRNA, micro-RNA and mRNA) or a defined sequence of amino acids. Thus, a sequence encodes a protein if transcription and translation of mRNA corresponding to that sequence produces the protein in a cell or other biological system. Both the coding strand and the non-coding strand can be referred to as encoding the protein or other product of that sequence.

Placing an expressed sequence "operably" under the control of regulatory elements, such as, *e.g.*, a promoter (and optionally an enhancer), means positioning the said sequence such that its expression is controlled by the said elements. A person skilled in molecular expression is aware of the importance of these considerations.

Reference herein to a "promoter" or "enhancer" is to be taken in its broadest context and includes transcriptional regulatory sequences required for accurate transcription initiation and for the spatial and/or temporal control of gene expression or its response to, *e.g.*, external stimuli. More particularly, "promoter" may depict a region on a nucleic acid molecule, preferably DNA molecule, to which an RNA polymerase binds and initiates transcription.

A promoter is usually, but not necessarily, positioned upstream or 5', of the sequence the transcription of which it regulates. Regulatory elements may however be present in the coding sequence, introns, 5' or 3' untranslated sequence(s).

The term "terminator" refers to a sequence element at the end of a transcriptional unit which signals termination of transcription. Terminators active in plant cells are known and well-documented in the literature, such as, *e.g.* the nopaline synthase (NOS) gene terminator of *Agrobacterium tumefaciens*, the terminator of the Cauliflower mosaic virus CaMV35S gene, the zein gene terminator from *Zea mays*, the ribulose-1,5-biphosphate carboxylase small subunit gene (*rbcS la*) terminator, and the isopentenyladenine transferase (*ipt*) terminator, amongst others.

"Gene" refers to a nucleic acid segment that encodes and is capable of expressing a specific product. This product can be an mRNA, protein or polypeptide or a structural or functional nucleic acid. Functional or structural nucleic acids, for example, rRNA, tRNA, ribozymes, antisense RNA, co-suppression molecule or interfering RNA (*e.g.*, siRNA) may also be considered gene products. A "gene" may also contain sequences containing regulatory elements, such as without limitation, promoters, enhancers and terminators, operably linked to the expressed sequence to facilitate the transcription of the expressed sequence. The expressed sequence may also contain introns.

As used herein, the term "transgene" refers to a gene that has been introduced into a cell, tissue or organism by artifice, *e.g.*, by transformation. A transgene may be inserted into the genome and stably maintained. The transgene may be partly or entirely exogenous (non-endogenous, heterologous), *i.e.*, foreign, to the cell, tissue or organism transformed therewith. The transgene may in principle also comprise a sequence substantially identical or similar to an endogenous or native gene.

The term "transgene of interest" as used herein generally refers to a transgene the presence (and expression) of which is desired in the resulting transgenic organism, *e.g.*, a transgenic plant or transgenic plant cultivar.

Without limitation, a transgene of interest may be, *e.g.*, a functional gene that provides a new plant trait, enhances an existing plant trait, or otherwise modifies expression of plant phenotypes exhibited by the plant. Such traits may include, *e.g.*, herbicide resistance, pest resistance, disease resistance, environmental tolerance (*e.g.*, heat, cold, drought, salinity), morphology, growth characteristics, nutritional content, taste yield, horticultural characteristics, consumer (quality) traits, and the like.

"Transformation" refers to the transfer of a foreign sequence or gene into the genome of a host organism, where the foreign gene is preferably stably maintained.

The terms "transformed", "transformant" and "transgenic" refer to cells, tissues or organisms, *e.g.*, plants or calli, that have undergone the transformation process, or that descend from cells, tissues or organisms that have undergone the transformation process, and comprise a transgene.

The term "targeting cassette" (TC) herein broadly refers to a non-endogenous (*i.e.*, non-native, exogenously introduced) nucleic acid molecule serving as a substrate for targeted

modification of a genome of a plant or plant cell. Advantageously, a targeting cassette may comprise sequences the targeted introduction of which into the genome of a plant or plant cell is desired. While a targeting cassette as a whole is normally foreign or non-native to the plant or plant cells into which it is introduced, it may comprise sequences identical or substantially identical to endogenous or native sequences of said plant or plant cell.

The term "homozygous" refers to a genetic condition of a cell, tissue or organism where identical alleles of a genetic trait reside at corresponding loci on homologous chromosomes. For instance, a plant which is said to be homozygous at a given locus for a targeting cassette TC contains the said targeting cassette at the said locus in each chromosome of the respective pair of homologous chromosomes.

As used herein, the term "hemizygous" refers to a genetic condition of a cell, tissue or organism where a genetic trait is present on one of the homologous chromosomes, but is not found at the corresponding locus of the other homologous chromosome. For example, a plant which is said to be hemizygous at a given locus for a targeting cassette TC only contains said targeting cassette TC at that locus of one but not the other one of the respective pair of homologous chromosomes.

When a particular sequence element is said to be at the locus L1 or within the genetic construct of the invention, but "outside" of the targeting cassette TC flanked by the recombinase sites RS, this denotes that said sequence element is not located in between the pair of recombination sites RS, but is instead external or adjacent thereto. Hence, release of the targeting cassette will not eliminate said sequence element from the locus L1.

The term "recombinase site" refers to a specific nucleic acid sequence, esp. DNA sequence, that a recombinase, esp. a site-specific recombinase, will recognise and bind to. This may be a wild-type or altered (e.g., mutant) recombinase site, as long as functionality is maintained and the recombinase still recognises and binds to the site. Typically, the recombinase will catalyse recombination between two nearby recombination sites.

The term a recombination site "responsive to" a site-specific recombinase reflects the fact that the action of the recombinase in a cell can lead to recombination at those sites. Hence, the term may reflect the specific relationship between a given site-specific recombinase and the respective recombination sites it acts upon.

A "site-specific recombinase" is understood to mean an enzyme or polypeptide capable of binding to a specific nucleotide sequence ("recombination site") in a nucleic acid molecule, preferably a DNA molecule, and induce a cross-over event in the nucleic acid molecule in the vicinity of the said recombination site. Preferably, a site-specific recombinase will induce excision of a nucleic acid sequence located between two such, suitably oriented, recombination sites.

Non-limiting examples of site-specific recombinases and their corresponding recombination sites include: the CRE recombinase and the *lox* (e.g., *loxP*, *loxB*, *loxL*, *loxR*) sites from bacteriophage P1 (see, e.g., Russell et al. 1992. Mol Gen Genet 234: 49-59 and WO 93/01283 for use thereof in plants); the FLP recombinase and the *frt* sites from *Saccharomyces cerevisiae* (see, e.g., Sonti et al. 1995. Plant Mol Biol 28: 1127-1132 for use thereof in plants); the pSR1-R recombinase and the pSR1-RS sites from *Zygosaccharomyces rouxii* (see, e.g., Sugita et al. 2000. Plant J 22: 461-469 for use thereof in plants) (the designations "R" and "RS" as used throughout this specification and claims are intended to refer generally to any site-specific recombinase and its corresponding recombination sites, and do not refer specifically to the R/RS system of *Z. rouxii*); the GIN recombinase and the *gix* sites from bacteriophage Mu (see Maeser and Kahmann. 1991. Mol Gen Genet 230: 170-176); the XcerC-encoded recombinase and the *cer* sites from *E. coli* (see Colloms et al. 1990. J Bacteriol 172: 6973-80); and the integrase and the *attP* and *attB* sites from the *Streptomyces* phage phiC31 (Groth et al. 2000. PNAS 97: 5995-6000). It shall be appreciated that functional (i.e., retaining an adequate level of activity) fragments or variants (e.g., by amino acid deletions, additions, substitutions or derivations, etc.) of these and other native site-specific recombinases are also useful in the invention. The *cre/lox* system is commonly employed in the art and may also be preferred in the present invention.

It may be desirable that the recombinase is localised in the nucleus of a the plant cell. Accordingly, the recombinases as described herein may be further modified to include a heterologous nuclear localisation signal, as known in the art.

The relative orientation of two recombination sites in a nucleic acid molecule influences whether the sequences interposed there between are released or excised or, alternatively, inverted when a site-specific recombinase acts on the recombination sites. Hence, the recombination sites may be so "oriented" or "configured" relative to each other such as to promote the release or excision of the interposed sequences by the action of the site-specific

recombinase upon, or in the vicinity of, said recombination sites. These considerations will be available to a skilled molecular biologist. By means of example, he will understand that orientation of some recombination sites (such as, *e.g.*, lox sites) in a direct repeat configuration may facilitate excision of the interposed sequences.

- 5 The term "expressing the recombinase" generally refers to expressing a functional recombinase in cells, tissues or organism, *e.g.*, in plants, plant parts or plant cells. Typically, said expression may be facilitated by providing in said cells, tissues or organism a sequence encoding the recombinase operably linked to suitable regulatory elements, such as without limitation a promoter, optionally an enhancer and a terminator.
- 10 As used herein, the phrase "expressing the recombinase R preferentially in" refers to a situation where the recombinase R is expressed – at levels which cause the excision of the targeting cassette TC from locus L1 – substantially only, only or exclusively in the recited cell type(s) and/or in cells at the recited developmental or cell division stage(s). Preferably, less than about 30%, more preferably less than about 20%, *e.g.*, less than about 15%, even more
- 15 preferably less than about 10%, *e.g.*, less than about 7%, still more preferably less than about 5%, *e.g.*, less than 4%, less than 3% or less than 2%, and very preferably less than 1%, *e.g.*, less than 0.5%, less than 0.1%, or less than 0.01% or even less of cells belonging to cell types different from the recited cell type(s) and/or of cells at developmental or cell division stages other than the recited ones would express the recombinase R at levels causing
- 20 excision of the targeting cassette from locus L1. The presence or absence of expression of recombinase R in cells at a level adequate to cause the excision of a sequence bordered by sites responsive to said recombinase R, such as for example excision of the targeting cassette TC, can be determined using tests and molecular analysis assays generally available in the art, such as for example PCR, Southern blotting, etc.
- 25 The phrase "expressing the recombinase R preferentially in" may preferably also encompass situations where the recombinase R is expressed – at levels which cause the excision of the targeting cassette TC from locus L1 – in the recited cell type(s) and/or in cells at the recited developmental or cell division stage(s), but substantially not expressed or not expressed at said levels in progenitors or predecessors of said desired cells. For example, the level of
- 30 expression of recombinase R may be, individually and/or on average, at least 2-fold, or at least 3-fold, preferably at least 5-fold, more preferably at least 10-fold, or at least 20, even more preferably at least 50-fold, still more preferably at least 100-fold, or even 200-fold, 300-

fold, 500-fold or 1000-fold lower in said progenitors or predecessors than in said desired cells derived there from. Otherwise, the frequency of expression of recombinase R at levels which cause the excision of the targeting cassette TC from locus L1 may be at least 2-fold, or at least 3-fold, preferably at least 5-fold, more preferably at least 10-fold, or at least 20, even more preferably at least 50-fold, still more preferably at least 100-fold, or even 200-fold, 300-fold, 500-fold or 1000-fold lower in said progenitors or predecessors than in said desired cells derived there from. This can ensure that the targeting cassette would still be present in the genome at the start of the desired expression of the recombinase R. Level and/or frequency of expression may be determined by conventional assays, such as, *e.g.*, PAGE-immunoblotting, RIA, ELISA, FACS-immunostaining, or enzymatic activity determination.

Moreover, the recombinase R may be preferably expressed – at levels which cause the excision of the targeting cassette TC from locus L1 – in at least about 5%, *e.g.*, in at least about 10%, preferably in at least about 20%, *e.g.*, in at least about 30%, more preferably in at least about 40%, *e.g.*, in at least about 50%, even more preferably in at least about 60%, *e.g.*, in at least about 70%, and still more preferably in substantially all cells belonging to the recited cell type(s) and/or cells at the recited developmental or cell division stage(s) (“substantially all” as used throughout the specification refers to about 70% or more, *e.g.*, about 75% or more, preferably about 80% or more, *e.g.*, about 85% or more, more preferably about 90% or more, even more preferably about 95% or more, and most preferably at least about 96%, at least about 97%, at least about 98%, at least about 99% or even 100%).

The phrase “expressing the recombinase R preferentially in” may preferably also encompass situations where the level of expression of recombinase R is at least 2-fold, or at least 3-fold, preferably at least 5-fold, more preferably at least 10-fold, or at least 20, even more preferably at least 50-fold, still more preferably at least 100-fold, or even 200-fold, 300-fold, 500-fold or 1000-fold higher in the recited cell type(s) and/or in cells at the recited developmental or cell division stage(s) as compared to other cells, as determined by conventional assays, such as, *e.g.*, PAGE-immunoblotting, RIA, ELISA, or enzymatic activity determination.

Furthermore, the phrase “expressing the recombinase R preferentially in” may preferably also encompass such preferential expression which is not constitutive, or which is not dependent on external stimuli, or which can be denoted as cell-specific, tissue-specific and/or developmental stage-specific with respect to plant cells having increased frequency of homologous recombination.

In an embodiment, such preferential expression of the recombinase R is achieved by placing the sequence encoding the recombinase R under the control (operably linked to) suitable regulatory elements, encompassing *inter alia* promoter and optionally enhancer.

As used herein, the term "homology" denotes structural similarity between two macromolecules, particularly between two polynucleotides, irrespective of whether said similarity is or is not due to shared ancestry.

The phrase "region of homology" (RH) as used herein denotes a portion, preferably a contiguous portion, of the targeting cassette TC which is homologous to, *i.e.*, the nucleotide sequence of which is similar to, a sequence S or a part thereof at the locus L2.

Preferably, the length of the region of homology RH may be at least 10 bp, *e.g.*, at least 15 bp or at least 20 bp, more preferably at least 30 bp, *e.g.*, at least 40 bp, even more preferably at least 50 bp, *e.g.*, at least 60 bp, at least 70 bp, at least 80 bp or at least 90 bp, still more preferably at least 100 bp, *e.g.*, at least 200 bp, at least 300 bp or at least 400 bp, and very preferably may be even at least 500 bp, at least 1 kb, at least 2 kb, at least 5 kb, at least 10 kb, or more.

Also preferably lengths of the region of homology RH may be: 10 bp $\leq$ LRH $\leq$ 10 kb, or 10 bp $\leq$ LRH $\leq$ 5 kb, or 10 bp $\leq$ LRH $\leq$ 2 kb, or 10 bp $\leq$ LRH $\leq$ 1 kb, or 15 bp $\leq$ LRH $\leq$ 10 kb, or 15 bp $\leq$ LRH $\leq$ 5 kb, or 15 bp $\leq$ LRH $\leq$ 2 kb, or 15 bp $\leq$ LRH $\leq$ 1 kb, or 20 bp $\leq$ LRH $\leq$ 10 kb, or 20 bp $\leq$ LRH $\leq$ 5 kb, or 20 bp $\leq$ LRH $\leq$ 2 kb, or 20 bp $\leq$ LRH $\leq$ 1 kb, more preferably 30 bp $\leq$ LRH $\leq$ 10 kb, or 30 bp $\leq$ LRH $\leq$ 5 kb, or 30 bp $\leq$ LRH $\leq$ 2 kb, or 30 bp $\leq$ LRH $\leq$ 1 kb, or 40 bp $\leq$ LRH $\leq$ 10 kb, or 40 bp $\leq$ LRH $\leq$ 5 kb, or 40 bp $\leq$ LRH $\leq$ 2 kb, or 40 bp $\leq$ LRH $\leq$ 1 kb, even more preferably 50 bp $\leq$ LRH $\leq$ 10 kb, or 50 bp $\leq$ LRH $\leq$ 5 kb, or 50 bp $\leq$ LRH $\leq$ 2 kb, or 50 bp $\leq$ LRH $\leq$ 1 kb, or 60 bp $\leq$ LRH $\leq$ 10 kb, or 60 bp $\leq$ LRH $\leq$ 5 kb, or 60 bp $\leq$ LRH $\leq$ 2 kb, or 60 bp $\leq$ LRH $\leq$ 1 kb, or 70 bp $\leq$ LRH $\leq$ 10 kb, or 70 bp $\leq$ LRH $\leq$ 5 kb, or 70 bp $\leq$ LRH $\leq$ 2 kb, or 70 bp $\leq$ LRH $\leq$ 1 kb, or 80 bp $\leq$ LRH $\leq$ 10 kb, or 80 bp $\leq$ LRH $\leq$ 5 kb, or 80 bp $\leq$ LRH $\leq$ 2 kb, or 80 bp $\leq$ LRH $\leq$ 1 kb, or 90 bp $\leq$ LRH $\leq$ 10 kb, or 90 bp $\leq$ LRH $\leq$ 5 kb, or 90 bp $\leq$ LRH $\leq$ 2 kb, or 90 bp $\leq$ LRH $\leq$ 1 kb, or still more preferably 100 bp $\leq$ LRH $\leq$ 10 kb, or 100 bp $\leq$ LRH $\leq$ 5 kb, or 100 bp $\leq$ LRH $\leq$ 2 kb, or 100 bp $\leq$ LRH $\leq$ 1 kb, or 200 bp $\leq$ LRH $\leq$ 10 kb, or 200 bp $\leq$ LRH $\leq$ 5 kb, or 200 bp $\leq$ LRH $\leq$ 2 kb, or 200 bp $\leq$ LRH $\leq$ 1 kb, or 300 bp $\leq$ LRH $\leq$ 10 kb, or 300 bp $\leq$ LRH $\leq$ 5 kb, or 300 bp $\leq$ LRH $\leq$ 2 kb, or 300 bp $\leq$ LRH $\leq$ 1 kb, or 400 bp $\leq$ LRH $\leq$ 10 kb, or 400 bp $\leq$ LRH $\leq$ 5 kb, or 400 bp $\leq$ LRH $\leq$ 2 kb, or 400 bp $\leq$ LRH $\leq$ 1 kb, such as very

preferably $500 \text{ bp} \leq \text{LRH} \leq 10 \text{ kb}$, or $500 \text{ bp} \leq \text{LRH} \leq 5 \text{ kb}$, or $500 \text{ bp} \leq \text{LRH} \leq 2 \text{ kb}$, or $500 \text{ bp} \leq \text{LRH} \leq 1 \text{ kb}$; where "LRH" stands for length of the region of homology RH.

Also preferably, sequence identity between the region of homology RH of the targeting cassette TC and its corresponding sequence S or a part thereof at the locus L2 may be at least 50%, e.g., at least 60%, more preferably at least 70%, e.g., at least 75%, even more preferably at least 80%, e.g., at least 85%, still more preferably at least 90%, yet more preferably at least 95%, e.g., at least 96%, at least 97%, at least 98%, at least 99%, or even 100%.

Sequence identity between two nucleic acid sequences may be determined by optimally aligning the nucleic acid sequences and scoring, on one hand, the number of positions in the alignment at which the two nucleic acids contain the same nucleotide and, on the other hand, the number of positions in the alignment at which the two nucleic acids differ in their sequence. The two nucleic acids differ in their sequence at a given position in the alignment when the nucleic acids contain different nucleotide at that position (substitution), or when one of the nucleic acids contains a nucleotide at that position while the other one does not or *vice versa* (insertion or deletion). Sequence identity is calculated as the proportion (percentage) of positions in the alignment at which the nucleic acids contain the same nucleotide *versus* the total number of positions in the alignment.

Optimal alignment of sequences for determination of sequence identity may be conducted using a variety of algorithms, such as the local homology algorithm of Smith and Waterman 1981 (Adv Appl Math 2: 482-9), the search for similarity method of Pearson and Lipman 1988 (PNAS 85: 2444), and the computerised implementations of these algorithms (such as GAP, BESTFIT, FASTA and TFASTA in the GCGTM v. 11.1.2 package from Accelrys).

Further suitable algorithms for performing sequence alignments and determination of sequence identity include those based on the Basic Local Alignment Search Tool (BLAST) originally described by Altschul et al. 1990 (J Mol Biol 215: 403-10), such as the "Blast 2 sequences" algorithm described by Tatusova and Madden 1999 (FEMS Microbiol Lett 174: 247-250), for example using the published default settings or other suitable settings (such as, e.g., for the BLASTN algorithm: cost to open a gap = 5, cost to extend a gap = 2, penalty for a mismatch = -2, reward for a match = 1, gap x_dropoff = 50, expectation value = 10.0, word size = 11).

Alternatively or in addition, sequence identity between the region of homology RH of the targeting cassette TC and its corresponding sequence S or a part thereof at the locus L2 may be such that said nucleic acids can hybridise under stringent conditions. "Hybridisation" refer to a process by which a nucleic acid strand anneals with complementary or substantially complementary sequence(s) comprised in the same or another polynucleotide strand through base pairing, preferably Watson-Crick base pairing. "High stringency" conditions include conditions equivalent to the following exemplary conditions: binding at 65 °C in a solution consisting of 5×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), 0.1% SDS, 5×Denhardt's reagent (50×Denhardt's contains 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 5×SSPE, 0.1% SDS at 65 °C when a probe of about 500 nucleotides in length is employed. Other exemplary conditions for hybridisation at "high stringency" for nucleic acid sequences over approximately 20-100 nucleotides in length, preferably about 30-100 nucleotides in length, e.g., between 30-50 nucleotides in length, include conditions equivalent to hybridisation in 6×SSC at 45°C, followed by one or more washes in 0.2×SSC, 0.1% SDS or even 0.1×SSC, 0.1% SDS at 65°C. Numerous equivalent conditions may be employed to vary stringency conditions; factors such as the length and nature (DNA, RNA, base composition) of the probe and nature of the target (DNA, RNA, base composition, present in solution or immobilised, *etc.*) and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulphate, polyethylene glycol) are considered and the hybridisation solution may be varied to generate conditions of low or high stringency hybridisation different from, but equivalent to, the above listed conditions. In addition, the art knows conditions that promote hybridisation under conditions of high stringency (e.g., increasing the temperature of the hybridisation and/or wash steps, the use of formamide in the hybridisation solution, *etc.*). Guidance for performing hybridisation reactions can be found, for example, in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 6.3.1-6.3.6, 1989, and more recent updated editions, all of which are incorporated by reference.

The term "homologous recombination" as used herein has its art-established meaning and generally refers to complementary base-pairing and crossing-over between nucleic acid regions containing identical or similar (homologous) sequences.

The term "cells having increased frequency of homologous recombination" refers to any plant cells, such as, e.g., cells of specific cell types and/or cells at particular developmental or cell division stages, in which the incidence of homologous recombination is increased relative to the remaining cells of a plant or to the remaining isolated plant cells. By means of preferred examples and not limitation, the frequency of homologous recombination in such cells may be about 1×10^{-4} or more, preferably about 5×10^{-4} or more, more preferably about 1×10^{-3} or more, even more preferably about 5×10^{-3} or more, still more preferably about 1×10^{-3} or more, such as, e.g., about 2×10^{-2} or more, about 4×10^{-2} or more, about 6×10^{-2} or more, about 8×10^{-2} or more, and very preferably about 1×10^{-1} or more targeted integration events per one random integration event, as measured by techniques derivable from the art, e.g., from Halfter et al. 1992 (*supra*).

The term "selectable marker gene" shall be taken to refer to any gene which confers a phenotype to a cell in which it is expressed to permit the identification and/or selection of cells which are transformed with a genetic construct encoding the said marker.

Preferred selectable marker genes include those which can confer resistance to a substance, e.g., an antibiotic, herbicide or a metabolic inhibitor, which is detrimental to the growth or survival of wild-type cells. These markers are herein termed "positive" selectable markers. A particularly effective and thus preferred positive selectable marker gene is the neomycin phosphotransferase gene *nptII*, which confers resistance to neomycin, kanamycin and related compounds, or hygromycin phosphotransferase gene (*hpt*).

Another kind of selectable marker genes is termed herein "reporter genes", i.e., genes whose expression, while usually not conferring resistance, can yet be assayed, e.g., by visual inspection. Preferred but non-limiting examples of these markers include the bacterial chloramphenicol acetyl transferase (*cat*) gene, bacterial glucuronidase (*gus*) gene, firefly luciferase (*luc*) gene, green fluorescent protein (*gfp*) gene, etc.

A further type of selectable marker genes that may be useful in the present invention are "negative" selectable markers. Such markers typically confer a toxic property on a cell in which they are expressed and thereby permit counter-selection against cells containing the marker gene. A negative selectable marker may encode, e.g., an enzyme which metabolises a specific substrate to a cytotoxic product. An illustrative, non-limiting negative selectable marker is the bacterial cytosine deaminase (*codA*) gene, which confers sensitivity to 5-fluorocytosin.

Those skilled in the art will be aware of other selectable marker genes useful in the present invention and will understand that the subject invention is not limited by the nature of the selectable marker gene.

Preferred embodiments

- 5 As detailed in the Summary section, aspects of the invention relate to methods for targeted genome modification in plants or isolated plant cells employing a targeting cassette TC integrated in a locus L1 of the genome of said plant or plant cells and suitably released by the action of site-specific recombinase preferentially in a specific subset of cells of said plant or isolated plant cells, such as to increase the probability of obtaining targeted genome
- 10 modification therein. Related aspects relate to the plants or isolated plant cells configured for use in the methods of the invention, as well as to genetic constructs suitable for obtaining said plants or isolated plant cells. Unless otherwise stated or apparent from the context, a skilled person can appreciate that the herein disclosed preferred embodiments can be directed to any of the above related aspects of the invention.
- 15 Hence, the invention provides, and the methods of the invention may depart from, a plant or isolated plant cells comprising
- a locus L1 and a targeting cassette TC at said locus L1, said targeting cassette TC:
 - o being flanked by a pair of recombinase sites RS responsive to a site-specific recombinase R, the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC
 - 20 from the locus L1, and
 - o comprising at least one region of homology RH to a sequence S at a locus L2 of said plant or plant cell different from the locus L1; and
 - the sequence S at the locus L2 different from the locus L1.
- 25 A skilled person can appreciate that plants and plant cells as above can be readily obtained by transformation methods known in the art, wherein plants or plant cells are transformed with the targeting cassette TC or with a genetic construct comprising a targeting cassette TC as defined herein. In transformants, the targeting cassette TC or genetic construct comprising such may be integrated into and stably maintained in the genome. Transformants can be
- 30 identified and selected, e.g., using suitable molecular genetic screens or by means of a

selectable marker gene, preferably a positive selectable marker gene or a reporter gene, more preferably a positive selectable marker. Preferably, a selectable marker gene may be comprised within the targeting cassette TC, or in said genetic construct within or outside of the targeting cassette TC (whereby following integration of the genetic construct at locus L1, the selectable marker gene will be outside of the targeting cassette TC). The targeting cassette TC or the genetic construct comprising the targeting cassette TC may include one or more selectable marker genes. However, in other embodiments, introduction of a selectable marker gene on a distinct genetic construct can be envisaged by, e.g., co-transformation.

A skilled person is well-aware of techniques suitable for plant transformation. Preferred methods include, for instance, *Agrobacterium*-mediated transformation (Deblaere *et al.* 1987. Meth Enzymol 143: 277); particle-accelerated or "gene gun" transformation (Klein *et al.* 1987. Nature 327: 70-73; US 4,945,050); direct microinjection into plant cells using micropipette (Crossway. 1985. Mol Gen Genetics 202: 179-185); gene transfer induced by polyethylene glycol (Krens *et al.* 1982. Nature 296: 72-74); fusion of plant protoplasts with other fusible, lipid-surfaced bodies (mini-cells, cells, liposomes, etc.), where the latter bodies contain the nucleic acids whose introduction is desired (Fraley *et al.* 1982. PNAS 79: 1859-1863); or electroporation of plant protoplasts with the to-be-introduced nucleic acid (Fromm *et al.* 1985. PNAS 82: 5824).

For *Agrobacterium*-mediated transformation, plant cells or tissues are infected with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* previously transformed with the nucleic acid to be introduced. To facilitate *Agrobacterium*-mediated transformation, a nucleic acid to be introduced (e.g., the targeting cassette or a genetic construct comprising such) is usually placed between the left and/or right T-DNA border sequences, as known in the art. *Agrobacterium* can be employed to introduce the desired nucleic acid, for instance, into individual plant cells, plant protoplasts, or plant tissue culture, or to parts of a growing plant, i.e., *in planta*, such as, e.g., in the '*Agrobacterium* vacuum infiltration' method of Bechtold *et al.* 1993 (C R Acad Sci Paris, Life Sciences 316: 1194-1199) or the 'floral-dip' method described by Clough and Bent. 1998 (Plant J 16: 735-743).

Following transformation and the selection of successful transformants by means of a selectable marker gene, some above techniques may require that the plant is regenerated from the obtained transformed cells or cell lines. By means of example, regeneration of plants from, e.g., cultured protoplasts, is described in Evans *et al.* 1983 (Handbook of Plant Cell

Cultures, Vol. 1, MacMillan Publishing Co. New York) and Vasil I.R. (ed.) Cell Culture and Somatic Cell Genetics of Plants, Acad. Press, Orlando, Vol. I, 1984, and Vol. II, 1986). It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to, all major species of sugarcane, sugar beet, cotton, fruit trees, and legumes, and the conditions for the said regeneration are generally known to a skilled person.

In an embodiment, the plant or plant cells may comprise the targeting cassette TC at only one locus L1 ("single-site" integration). This can reduce the chances of unwanted genetic alterations being introduced to said plant or plant cells by insertion of the targeting cassette TC or of the genetic construct comprising the targeting cassette TC at multiple sites within the genome. Moreover, this may also reduce the chances of chromosomal aberrations resulting from the reaction between the recombinase sites RS at different loci.

In another embodiment, the plant or plant cells may comprise the targeting cassette TC at more than one different loci L1 ("multiple-site" integration). This can increase the number of copies of the targeting cassette TC released by the recombinase R in the method of the invention and thereby improve the probability of homologous recombination occurring between the so released targeting cassette TC and the target sequences at the locus L2.

In an embodiment, the one or more loci L1 may each comprise only one targeting cassette TC ("single-copy" integration). This reduces the chances that the recombinase R would catalyse reaction between recombinase sites RS from nearby but different targeting cassettes, releasing substrates that could not undergo the desired homologous recombination with the locus L2.

In an alternative embodiment, the one or (at least some of) more loci L1 may comprise more than one (e.g., 2, 3, 4 or more) targeting cassettes TC ("multi-copy" integration). For example, more than one copy of the targeting cassette TC or of the genetic construct comprising the targeting cassette TC may integrate at said one or (at least some of) more loci L1, and/or the integrated genetic construct may comprise more than one copy of the targeting cassette TC. This can increase the number of copies of the targeting cassette TC released by the recombinase R in the method of the invention and thereby improve the probability of homologous recombination occurring between the so released targeting cassette TC and the target sequences at the locus L2.

A person skilled in the art knows to identify single-site, multiple-site, single-copy and/or multi-copy transformants by suitable methods, *e.g.*, (Real-Time) PCR, RFLP, out-crossing, Southern blot hybridization, *etc.*

In a further embodiment, the plant or plant cells may be hemizygous for the targeting cassette TC at the one or more loci L1. In another embodiment, the plant or plant cells may be homozygous for the targeting cassette TC at the one or more loci L1. The latter situation may advantageously yield more copies of targeting cassette TC available for homologous recombination with the locus L2. A skilled person is generally capable of generating plants and plant cells homozygous for a desirable trait, *e.g.*, by crossing experiments.

As explained, the plant or plant cells of the invention further comprise a sequence S at a locus L2 different from the locus (or loci) L1, and the targeting cassette TC comprises at least one region of homology RH as defined herein to said sequence S at the locus L2.

As can be understood by a skilled person, the invention endows no principal restrictions on the nature of the sequence S at the locus L2. By means of example and not limitation, said sequence S may be endogenous, *i.e.*, native to the plant or plant cells or may be exogenous. By means of further examples and not limitation, said sequence S may encompass one or more genes, regulatory elements or parts thereof or the sequence S may be from an intergenic region.

In a preferred embodiment, the sequence S may be from an intergenic region. This may advantageously allow to introduce sequences, *e.g.*, expressed sequences, from the targeting cassette TC to said intergenic region substantially reducing the risk of disturbing the function of the plant's native genes. Preferably, the sequence S may be from a transcriptionally accessible (*e.g.*, euchromatin) intergenic region, which may facilitate expression of the sequences introduced thereto from the targeting cassette.

In another preferred embodiment, the sequence S may encompass a gene or any part thereof. Without limitation, the sequence S may be from and/or may encompass one or more exons, introns, exon-intron boundaries and/or regulatory sequences (*e.g.*, promoters, enhancers, terminators, *etc.*) of said gene. This may advantageously allow to introduce sequences from the targeting cassette TC into said gene or part thereof, thereby advantageously affecting the expression of the gene and/or the function of the gene products in any of various ways, such as, without limitation, disrupting the expression or modulating the

strength or temporal or tissue specificity of the expression, disrupting or modulating (e.g., increasing, decreasing or qualitatively altering) the function of the expressed gene products, *etc.*

It will also be appreciate that the sequence S may encompass two or more genes or parts thereof.

In an embodiment, the targeting cassette TC, as used in the methods, plants, plants cells and constructs of the invention, may comprise only one region of homology RH to the sequence S at the locus L2. This may preferably lead to a single homologous recombination (single crossing-over) between the region of homology RH and the sequence S at the locus L2, but may otherwise lead to multiple, e.g., two, three, four or more, preferably two, crossing-overs between the region of homology RH and the sequence S at the locus L2. Single crossing-over may be particularly preferred in situations when the targeting cassette TC is circular and when insertion of substantially the entire targeting cassette TC into the locus L2 is desired. On the other hand, multiple crossing-overs, and most practically double crossing-over may be particularly preferred (irrespective of whether the targeting cassette TC is circular or linear) when the introduction of only some sequences of the targeting cassette TC (*i.e.*, sequences residing between said crossing-overs) into the locus L2 is desired.

In another embodiment, the targeting cassette TC may comprise two or more, e.g., two, three, four or more (RH¹, RH², RH³, RH⁴,...) more preferably two (RH¹, RH²), regions of homology RH to the sequence S at locus L2. Typically, said regions of homology RH (e.g., RH¹, RH²) in the targeting cassette TC may be separated by sequences for which no homologous counterparts are found in the sequence S at the locus L2. Alternatively, said regions of homology RH (e.g., RH¹, RH²) in the targeting cassette TC may be directly adjacent to each other in the targeting cassette TC but their homologous counterparts may be non-adjacent in the sequence S at the locus L2. The stretches of the sequence S at the locus L2 which are homologous to the respective regions of homology RH (e.g., to RH¹ and RH²) may be directly adjacent to each other or may be separated by sequences having no homologous counterparts in the targeting cassette TC. Desirably, this configuration may lead to a single crossing-over between each of the multiple regions of homology RH and its respective homologous sequence within the sequence S at the locus L2, and thus overall to multiple, e.g., two, three, four or more, preferably two (e.g., at RH¹ and at RH²) crossing-overs between the targeting cassette TC and the sequence S at the locus L2. Such multiple

crossing-overs, and most practically double crossing-over may be particularly preferred (irrespective of whether the targeting cassette TC is circular or linear) when the introduction of only some sequences of the targeting cassette TC (*i.e.*, sequences residing between said crossing-overs and thus between said regions of homology RH) into the locus L2 is desired.

5 Hence, in the present method sequence information from the targeting cassette TC may be directedly introduced into the locus L2. As can be understood by a skilled person, the invention endows no principal restrictions on the nature of sequences intended to be so-introduced to the locus L2. Nevertheless, it can be appreciated that introduction of said sequences advantageously induces an alteration of some kind in the sequence S at the locus
10 L2.

By means of preferred example and not limitation, the to-be-introduced sequences may comprise one or more transgenes of interest or parts thereof, one or more regulatory sequences (*e.g.*, promoters, enhancers, terminators, *etc.*) or parts thereof, one or more coding sequences or parts thereof, and/or one or more selection marker genes, *etc.* In a
15 preferred embodiment, when the sequence S is from an intergenic region, the to-be-introduced sequences may comprise one or more transgenes of interest and/or one or more selection marker genes. In a further preferred embodiment, when the sequence S encompasses a gene or part thereof, the targeting cassette TC may be configured to directedly introduce one or more insertions, deletions and/or substitutions which may affect,
20 *e.g.*, eliminate, diminish, enhance or qualitatively alter the function of the gene or part thereof, or gene product thereof. The above examples are non-limiting and analogous considerations are available to a skilled person.

In further preferred embodiments, the methods, plants, plant cells and genetic constructs of the invention may benefit from situations where the targeting cassette TC further comprises a
25 positive selection marker gene and/or reporter gene. When the targeting cassette TC is configured such that the positive selection marker gene and/or reporter gene is introduced into the locus L2 upon homologous recombination, this may allow to identify and select plants or plant cells in which said recombination occurred.

In further preferred embodiments, the methods, plants, plant cells and genetic constructs of
30 the invention may benefit from situations where the targeting cassette TC further comprises a negative selection marker gene. When the targeting cassette TC is configured such that the negative selection marker gene is not introduced into the locus L2 upon homologous

recombination (but would expectedly be introduced if the targeting cassette TC integrated randomly or ectopically elsewhere in the genome), this may allow to further identify and select plants or plant cells in which homologous recombination occurred (as opposed to random or ectopic integration, or even as opposed to plants or plant cells where the targeting cassette TC was not released from the locus L1). In a preferred, yet merely exemplary configuration, the targeting cassette TC may comprise two regions of homology (RH^1 , RH^2) to the sequence S at the locus L2, and further comprise a negative selection marker outside of the sequence bordered by said regions of homology RH^1 , RH^2 . The expected homologous recombination would then introduce into the locus L2 the sequence of the targeting cassette TC in between the regions of homology RH^1 , RH^2 , but would eliminate the negative selection marker gene. Other configurations will be evident to a skilled person.

In yet other preferred embodiments, the methods, plants, plant cells and genetic constructs of the invention may benefit from situations where the targeting cassette TC further comprises a means that allows for inclusion of a step of linearising the targeting cassette TC after having been released from the locus L1. As can be appreciated, the released targeting cassette TC may often be initially circular. To promote homologous recombination (in instances where said recombination does not require a circular targeting cassette TC), the targeting cassette TC may be advantageously linearised.

By means of example and not limitation, the targeting cassette TC may be engineered to comprise a recognition sequence for a site specific nuclease (preferably endonuclease or meganuclease), such as I-Sce I. The I-Sce I endonuclease may be expressed in the plants or plant cells, preferably transiently after the targeting cassette TC has been released, to effect the linearisation thereof. In a preferred example, the I-Sce I endonuclease may be expressed in the same cells as the site-specific recombinase R to ensure the co-occurrence of the release and linearisation of the targeting cassette TC. This can be realised, e.g., by placing a transgene encoding the I-Sce I endonuclease under the control of the same regulatory elements as used for the site-specific recombinase R. Also see WO 03/004659 and WO 2006/032426 for guidance on the use of I-Sce I in plants.

In another possibility, the targeting cassette TC may be engineered to comprise a recognition site for a specific zinc-finger nuclease (ZFN). The corresponding ZFN may be expressed in the plants or plant cells, preferably transiently after the targeting cassette TC has been released, to effect the linearisation thereof. In a preferred example, the ZFN may be

expressed in the same cells as the site-specific recombinase R to ensure the co-occurrence of the release and linearisation of the targeting cassette TC. This can be realised, e.g., by placing a transgene encoding the ZFN endonuclease under the control of the same regulatory elements as used for the site-specific recombinase R. Also see Wright *et al.* 2005 (*supra*) for
5 guidance on the use of zinc-finger nucleases to introduce double strand breaks into DNA molecules in plants.

In yet other preferred embodiments, the methods, plants, plant cells and genetic constructs of the invention may benefit from situations where the locus L2 further comprises a means that allows for inclusion of a step of introducing a double-strand break within or adjacent to the
10 sequence S at the locus L2. It has been reported that double stranded breaks (DSB) in target sequences may augment the frequency of homologous recombination in plants (Puchta. 1999. Genetics 152: 1173-1181).

For example, without limitation, the locus L2 may be engineered to comprise a recognition sequence for I-Sce I site-specific endonuclease. The I-Sce I endonuclease may be expressed
15 in the plants or plant cells, preferably transiently after the targeting cassette TC has been released (e.g., as taught above) to introduce a DSB within or adjacent to the sequence S at the locus L2, and thereby promote the homologous recombination thereof with the targeting cassette TC.

In another possibility, a zinc finger nuclease (ZFN) configured to specifically recognise a site
20 which is either engineered into or endogenously present at the locus L2, may be expressed in the plants or plant cells, preferably transiently after the targeting cassette TC has been released, (e.g., as taught above) to introduce a DSB within or adjacent to the sequence S at the locus L2, and thereby promote the homologous recombination thereof with the targeting cassette TC.

As noted, the methods of the invention may further comprise expressing the site-specific recombinase R that can act upon the recombination sites RS preferentially in cells of the plant or of the isolated plant cells having increased frequency of homologous recombination, thereby releasing the targeting cassette TC from the locus L1 in said cells.

Preferably, the plant or plant cells of the invention may comprise a sequence (SR) encoding
30 the site-specific recombinase R operably linked to adequate 5' and 3' regulatory elements ("a

recombinase R gene") capable of effecting the preferential expression of said recombinase R according to the methods of the invention.

In a preferred embodiment, the sequence SR and regulatory elements thereof (the recombinase R gene) may be located at the locus L1 of the plant or plant cells of the invention, for example within or outside of the targeting cassette TC (*i.e.*, in *cis*). This arrangement may be obtained, *e.g.*, by transforming plants or plant cells using a genetic construct of the invention comprising said recombinase gene within or outside of the targeting cassette TC. This may advantageously reduce handling required to introduce the recombinase R gene.

Advantageously, the recombinase R gene may also be placed within the targeting cassette TC, in which case it will be eliminated together with the targeting cassette TC from the locus L1 upon excision of the latter ("auto-excision", see, *e.g.*, WO 97/37012). In some preferred embodiments, the recombinase R gene may be so positioned (*e.g.*, positioned similar to the negative selection marker as discussed above) within the targeting cassette TC such as not to be re-integrated by homologous recombination between the cassette TC and the locus L2.

In another embodiment, the recombinase R gene may be introduced on a genetic construct different from the one comprising the targeting cassette TC. In this situation, the recombinase R gene may co-integrate into the locus L1 (in *cis*), or may integrate into a locus (L3) different from the locus L1. (in *trans*). For example, the recombinase R may be introduced by methods known in the art, such as, *e.g.*, super-transformation, sexual crosses or by agro-infiltration (see, *e.g.*, WO 93/01283 and Russell. 1992. Mol Gen Genet 234: 49-59).

Preferential expression of site-specific recombinase R in specific cell types or in cells at specific developmental and/or cell division stages can be obtained by placing the sequence encoding the recombinase R under regulatory elements (*e.g.*, promoters and/or enhancers, at least promoters) capable of effecting said preferential expression in methods, plants, plant cells and genetic constructs of the invention.

In a preferred embodiment, the recombinase R may be preferentially expressed in cells of meristem. Preferably, the recombinase R may be preferentially expressed in cells of root and/or shoot meristem, more preferably at least shoot meristem. Further, the recombinase R may be preferentially expressed in apical meristem and/or in primary meristems including protoderm, procambium and ground meristem. More preferably, the recombinase R may be

preferentially expressed in apical meristem, even more preferably in the cells of the central zone of said apical meristem. Also very preferably, the recombinase R may be preferentially expressed in floral meristem.

5 In a further preferred embodiment, the recombinase R may be preferentially expressed in stem cells. The term "stem cells" generally refers to cells having substantially unlimited self-renewal and differentiation potential. Usually, plant stem cells are present within the apical meristem.

10 In a further preferred embodiment, the recombinase R may be preferentially expressed in cells of microsporangia. The term "microsporangium" generally refers to a plant structure producing and containing male spores. By means of example and not limitation, in flowering plants, microsporangia include anthers.

15 In a further preferred embodiment, the recombinase R may be preferentially expressed in cells of megasporangia. The term "megasporangium" generally refers to a plant structure producing and containing female spores. By means of example and not limitation, in flowering plants, megasporangia include nucelli of the ovules.

In another preferred embodiment, the recombinase R may be preferentially expressed in cells at one or more stages of mitotic cell cycle chosen from interphase (including any one or more of G₁ phase, G₁/S transition, S phase, S/G₂ transition, G₂ phase, G₂/M transition) or Mitosis (M).

20 More preferably, the recombinase R may be preferentially expressed in cells at G₂/M transition. The inventors contemplate that frequency of targeted genome modification is particularly increased at this stage.

25 In a further preferred embodiment, the recombinase R may be preferentially expressed in meiocytes. The term "meiocyte" generally refers to a cell destined to undergo or undergoing meiosis. In an embodiment, the meiocyte may be a microsporocyte. In another embodiment, the meiocyte may be a megasporocyte.

In a preferred embodiment, the recombinase R may be preferentially expressed in meiocytes at one or more stages of meiotic cell division chosen from interphase (including any one or

more of G₁ phase, G₁/S transition, S phase, S/G₂ transition, G₂ phase, G₂/Meiosis I transition) or Meiosis I or Meiosis II. Preferably, the recombinase R may be preferentially expressed in cells at G₂/Meiosis I transition.

5 The inventors realised that preferential expression of the recombinase R in the above cell types and/or cell stages facilitates improved gene targeting.

In a preferred embodiment, the sequence encoding the recombinase R may be placed under the control of (operably linked to) regulatory elements, such as encompassing promoter and optionally an enhancer, to drive or effect the preferential expression of said recombinase R as intended herein.

10 In a preferred embodiment, the sequence encoding the recombinase R may be placed under the control of (operably linked to) a promoter chosen from promoters of genes: *SOLO DANCERS*, *APETALA1*, *CLAVATA3*, *LEAFY*, *AGAMOUS*, *SPO11-1*, *DMC1* or *MSH4* in methods, plants, plant cells and genetic constructs of the invention. This can achieve desired preferential expression of said recombinase R.

15 In an embodiment, native promoters of said genes may be used. The term "native promoter" refers to a promoter the nucleotide sequence of which is identical to that of a promoter present in nature. The modifier "native promoter" thus relates to the sequence of the promoter and is not to be construed as requiring that the promoter be obtained or produced in any particular manner. By means of example and not limitation, the term would thus encompass
20 promoters in their natural hosts, isolated there from, cloned and propagated using recombinant DNA technology, produced by an amplification method or generated by synthetic means, *etc.*, insofar the sequence of such promoters would be the same as of their counterparts occurring in nature.

A skilled person understands that the native sequence of the promoter of a given gene may
25 differ between different plant species or and/or between different subspecies within a single plant species, due to natural genetic divergence between the said species or subspecies. Thus, such divergent but found-in-nature promoter sequences would be considered native.

A skilled person is in general capable of predicting and identifying natural promoters of genes, using suitable sequence analysis software as well as experimental procedures, such as, *e.g.*,
30 Northern blotting, RT-PCR, 5'-rapid amplification of cDNA ends method (5'-RACE), assaying

sequences overlapping with or upstream to the transcription initiation site in reporter gene assays, *etc.*

The use of functional variants of the native promoters is also contemplated. The term "variant" refers to a sequence which is substantially identical (*i.e.*, largely but not wholly identical) to a corresponding native sequence, in particular, at least 70% identical, preferably at least 80% identical, more preferably at least 90% identical, *e.g.*, at least 91% identical, 92% identical, yet more preferably at least 93% identical, *e.g.*, 94% identical, still more preferably at least 95% identical, *e.g.*, at least 96% identical, even more preferably at least 97% identical, *e.g.*, at least 98% identical, and most preferably at least 99% identical.

The use of functional fragments of native promoters is also contemplated. As used herein, the term "fragment" refers to a sequence that has a 5' and/or 3' deletion of one or more nucleotides as compared to a native promoter sequence or a variant thereof, but where the remaining nucleic acid sequence of the fragment is identical to the corresponding positions in the sequence of the native promoter or a variant thereof. The remaining sequence of a fragment can represents preferably at least 50%, *e.g.*, at least 60%, even more preferably at least 70%, *e.g.*, at least 80% or at least 85%, and still more preferably at least 90%, *e.g.*, at least 95% or more of the nucleic acid sequence of the respective native promoter or variant thereof.

The term "functional" with reference to the variants and fragments of promoters as above refers to the fact that the particular variants and fragments will have substantially retained or may even have improved the activity of the corresponding native promoter, *i.e.*, the capability to bind RNA polymerase and initiate transcription in the desired temporal and/or spatial pattern.

In an embodiment, the promoters used by the invention may be derived from a plant species which is to be modified in the methods of the invention.

In another embodiment, said promoters may be derived from *Arabidopsis thaliana*. Said promoters may thus be used in methods of the invention to achieve gene targeting in *Arabidopsis thaliana* or in other plant species where the promoters have the desired activity.

For example, the Gene ID numbers uniquely identifying the said genes from *Arabidopsis thaliana* as listed in the "Entrez Gene" database of NCBI (described in Maglott et al. 2005

Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res. 33 (Database Issue): D54–D58) are as follows: *APETALA1* (AP1): 843244; *LEAFY* (LFY): 836307; *CLAVATA3* (CLV3): 817267; *SOLO DANCERS* (SDS): 838040; *AGAMOUS* (AG): 827631, *SPO11-1*: 820506, *DMC1*: 821860 and *MSH4*: 827450. On the basis hereof, a skilled person can
5 identify and isolate corresponding native promoters of said genes from *Arabidopsis*, as well as homologues and orthologues from further plant species of interest, for use in methods, plants, plant cells and constructs of the invention.

Hence, the present invention teaches materials and methods to achieve more efficient gene targeting in plants or plant cells. As will be appreciated, suitable strategies may be applied to
10 identify and isolate plants or plant cells in which the desired gene targeting event has occurred.

In an embodiment, said identification and isolation may involve molecular genetic screening for the desired genetic modification, such as, e.g., using Southern blotting, PCR, or microarray hybridisation.

15 In another embodiment, the selection may involve screening for the presence or absence of selection marker genes, e.g., for the presence of a positive selection marker or reporter gene and/or for the absence of a negative selection marker, preferably as taught elsewhere in this specification. In a non-limiting example, the homologous recombination event may re-constitute a positive selection marker gene or may remove a negative selection marker gene
20 otherwise present in the targeting cassette TC.

Further, where the gene targeting is expected to introduce a detectable phenotypic change, such change may be used as a selection marker. Other modalities will be available to a skilled person.

When the present method induces gene targeting in select cells of plants, the above
25 identification and selection processes may be suitably applied on said plants or parts thereof, on plant parts, tissues or cells isolated from said plants, or on the vegetative or generative progeny of said plants. Advantageously, when the gene targeting occurs in cells which, or the descendants of which, give rise to gametes (such as, without limitation, in meiocytes or in floral meristem, etc.), the occurrence of gene targeting event may be assayed in the seeds or
30 in the generative progeny of said plants. As can be understood by a skilled person, whole plants may be suitably regenerated from the so-selected cells, tissues or parts.

When the method is applied to isolated cells, the identification and selection process may be applied to said cells *ex vivo*. As can be understood by a skilled person, whole plants may be suitably regenerated from the so-selected cells, tissues or parts.

Accordingly, the invention also provides methods as disclosed herein further comprising
5 identifying and/or selecting plants or plant cells in which the targeted genome modification occurred.

As will be immediately apparent to a skilled person, the nature of sequences of the targeting cassette TC, the materials and methods of the invention may be adapted to achieve a plethora of different goals.

10 By means of example and not limitation, the invention may allow to:

- introduce transgenes of interest in predetermined, preferably transcriptionally active chromosomal loci;
- introduce transgenes of interest in substantially free of other ancillary sequences, such as, *e.g.*, selection markers, *etc.*;
- 15 - inactivate genes, *e.g.*, endogenous genes (loss-of-function) of plants, *e.g.*, by deleting said genes or parts thereof or introducing other mutations (*e.g.*, insertions or substitutions) in their coding sequence or regulatory sequences (*e.g.*, promoters, enhancers);
- modulate, *e.g.*, eliminate, increase or decrease or change the temporal or spatial patter of
20 the expression of (endogenous) genes, *e.g.*, by modifying the sequence of their respective promoters or enhancers, or by "swapping" such control elements for other regulatory sequences;
- studying expression of (endogenous) genes by fusing them to reporter genes;
- altering splicing, *e.g.*, by deleting exons by or altering exon-intron boundaries or other
25 splicing regulating sequences;
- qualitatively or quantitatively modulate the function of (endogenous) gene products by altering the sequence encoding such, *e.g.*, using deletion, insertions or substitution;

and the like.

Accordingly, the invention also provides methods and materials as disclosed herein for the above specified as well as further uses.

In addition, the invention also relates to plants and progeny thereof, as well as plant parts, plant tissues and plant cells of such, obtainable or directly obtained by the gene targeting methods of the invention.

The invention is further explained with reference to the following examples and figures, which are in no way intended as limiting.

EXAMPLES

The invention is herein demonstrated using the following constructs in *Arabidopsis*: a “TARGET” vector; “targeting” vectors TNG-1a, TNG-3, TNG-6 or TNG-2I; CRE expression cassette encoding the CRE recombinase under the control of *SOLO DANCERS* (SDS), *APETALA1* (AP1), *CLAVATA3* (CLV3), *LEAFY* (LFY), *AGAMOUS* (AG), *SPO11-1*, *DMC1* or *MSH4* promoters.

The insert of the TARGET vector is integrated into the plant's genome and provides an experimental target locus for homologous recombination with a targeting cassette encompassed within the above targeting vectors. The inserts of the targeting vectors are also integrated into the plant's genome and comprise a portion flanked by a pair of *lox* sites, which – when released by the CRE recombinase – serves as the experimental targeting cassette. The target locus derived from the TARGET vector comprises a GUS:3' del NPTII fusion and a 5.4-kb lambda sequence. In turn, the targeting cassette derived from the targeting vectors comprises a 5'del GUS:NPTII fusion and the same 5.4-kb lambda sequence. These sequence elements are configured such that exemplary double homologous recombination involves the lambda sequence and the 3' GUS portion, reconstituting GUS:NPTII fusion at the target locus.

The CRE expression cassette is provided either in *trans* (for targeting vectors TNG-1a and TNG-2I) on constructs named CRE1 or CRE3, or is provided in *cis* in the targeting vectors TNG-3 and TNG-6.

The following sections give a more detailed account of the above nucleic acids and their experimental use.

Example 1: General materials and protocolsPlant Material

The plant material used is *Arabidopsis thaliana* Col/O. Plants were grown for 3-5 weeks under sterile conditions at 25°C, in a light/dark cycle of 16h/8h (3500lx). After 3-5 weeks plants were transferred to the greenhouse, where the growth continued in soil.

The FK24 plant line contains the K2L610 T-DNA, which contains a P35S driven *gus* gene flanked by *lox* sites, as a single copy and in a homozygous way (De Buck *et al.* 2004. Cell. Mol. Life Sci 61: 2632-2645).

Surface sterilization of seeds

Seeds were packed in Miracloth (Merck Biosciences Ltd. Nottingham, UK) and immersed in a 75% EtOH solution. After 2 minutes the EtOH solution was replaced by a 4% sodium hypochlorite (NaOCl) solution containing 0.0005% (v/v) Tween 20 (Polyoxyethylene sorbitan mono-laurate). After 12 minutes the Sodium hypochlorite solution was discarded and the seed packages were washed with sterilised bidistilled H₂O until pH7 was reached. The surface sterilized seeds were incubated for 1 hour in the last washing water.

Plant Media

K1 medium: 1x Murashige & Skoog (MS) micro, macro salts and vitamins (Duchefa, Haarlem, The Netherlands), 3% (m/v) Sucrose, 0.05% (m/v) 4-Morpholino-ethanesulfonic acid (MES) (Duchefa, Haarlem, The Netherlands), 0.8% plant agar (Duchefa, Haarlem, The Netherlands).

Selection media: K1 supplemented with 15mg/l hygromycin (Duchefa, The Netherlands), 75 mg/l kanamycin (Sigma, Germany), 10mg/l phosphinotrycin (Duchefa, The Netherlands) or 500mg/l 5-Fluorocytosin (Sigma, Germany).

Southern blot

DNA from plant material was purified by using the Nucleon™ Phytopure DNA extraction kit (RPN8511, GE Healthcare). Genomic DNA was cut with the respective restriction enzymes and separated on a 1% TAE agarose gel. The DNA was transferred to a Hybond Nylon filter (GE Healthcare, RPN203B) by upward capillarity by using a high salt buffer, 20xSSC (3M NaCl; 0.3M Na-citrate; pH7). The probe obtained after PCR and purification was labelled with

fluorescein by using the Gene images Random Prime Labeling kit (GE Healthcare, RPN3520). For obtaining the GUS1 probe the primers: GUS probe F (5'-GGTTAATCAGGAACTGTTGGCCC-3') and GUS probe R (5'-CAAGAAAGCCGGGCAATTGCTGTG-3') were used. For obtaining the GUS2 probe the primers: BspHI GUS probe F (5'-CGCAGAACATTACATTGACGCAGGTG-3') and BspHI GUS probe R (5'-GCTATACGCCATTTGAAGCCGATGTCACG-3') were used. For obtaining the BAR probe the primers: PPT probe F (5'-CAAATCTCGGTGACGGGCAGG-3') and PPT probe R (5'-GCACCATCGTCAACCACTACATC-3') were used. For obtaining the lambda1 probe the primers: lambda probe F (5'-GGAGACGGGCAATCAGTTCATC-3') and lambda probe R (5'-CGTGACAAGTCCACGTATGACC-3') were used. For obtaining the lambda2 probe the primers: BspHI lambda probe F (5'-GAACTGTGCAGGCAGCATACGCATG-3') and BspHI lambda probe R (5'-GAACATAACGGTGTGACCGTCACGC-3') were used.

After hybridization of the probe two stringency washes were performed both at 68°C for 10 minutes using a first washing solution containing 1x SSC and 0.1% SDS and a second washing solution containing 0.5x SSC and 0.1% SDS.

The detection was performed by using the CDP-star Detection kit (GE healthcare, RPN3550) and Hyperfilm ECL (RPN2103K) to detect the chemiluminescent signal.

Inverse-PCR (I-PCR)

100ng of genomic DNA of each line was digested with the restriction enzyme *NcoI*. After heat inactivation and re-circularization a first PCR reaction was performed on the ligation mixture using primers TDN-F and TDN-R followed by a nested-PCR using TDN-NesF and 35S-NesR as primers.

TDN-F: 5'-GAAGAGGCCCGCACCGATCGCCCTT-3'

TDN-R: 5'-GCAAGGCGATTAAGTTGGGTAACGCCAGGG-3'

TDN-NesF: 5'-GTCGTTTCCCGCCTTCAGTTTAAAC-3'

35S-NesR : 5'-CCCACTATCCTTCGCAAGACCC-3'

genA3F2: 5'-TCGAGAATTTTCGCCATGCCACCAAACAACC-3'

pptProbeR : 5'-GCACCATCGTCAACCACTACAT-3'

genD5R2: 5'-TTGGGTCTCCAGATCCTCTGTTATTAGTTGGTGACG-3'

Example 2: Cloning of the TARGET vector (Fig. 6)

The TARGET vector was obtained after the introduction of the *P35S-GUS-NPTII3'del* sequence and a 5.4kb lambda DNA fragment in the pCAMBIA 3300 (Cambia, Australia).

Subcloning the *35S-GUS-NPTII3'del* sequence: the *P35S-GUS-NPTII3'del* fragment, obtained after cutting the pBI426 (Datla *et al.* 1991. Gene 101(2): 239-46), a plasmid containing the *P35S-GUS:NPTII-Tnos* sequence, with *HindIII* and *RsrII*, was blunted (T4 DNA polymerase) and cloned into a pUC18 vector (*SmaI* site). Subsequently the *35S-GUS-NPTII3'del* sequence was cloned into the pCAMBIA 3300 as a *HindIII-EcoRI* fragment. The obtained vector was called "pCAMBIA 3del".

Introducing the 5.4kb lambda sequence: a 5.4kb lambda DNA fragment was picked-up from intact lambda DNA (Invitrogen) by PCR using the LamF3Paul (5'-CCGCGCGCGCATTATGGGGATCCTCAACTGTG-3') and LamR3 (5'-CGGAATTCGCCTGATTGCCGGAAGGACCAG-3') primers (GC content: 53.3%) and was cloned in the pCAMBIA 3del vector as an *EcoRI-Paul* fragment. The obtained plasmid was called "TARGET".

Example 3: Cloning of the targeting vectors (Fig. 1)

The *P35S-GUS-NPTII3'del* and the *P35S-bar-T35S* sequence were removed from the TARGET vector in two steps. After each removal the plasmid was circulated again by adding an adaptor (see further). This was followed by the introduction of the respective cassettes necessary to make TNG-1a, TNG-3, TNG-6 and TNG-2I (see figure 1).

Step 1: The *P35S-GUS-NPTII3'del* sequence was removed by cutting the TARGET with *Paul* and *HindIII*. The vector was circulated again by introducing an adaptor, compatible with *Paul* and *HindIII* overhangs (the *HindIII* site was not recovered). The adaptor contained from 5' to 3' the *I-SceI* site, the *MfeI* site, *lox* site, *BglII* site, *XbaI* site, *AfeI* site and the *SbfI* site. This resulted in a plasmid herein called "TNG lox1".

Step 2: The *P35S-bar-T35S* sequence was eliminated by cutting the TNG lox1 with *Asel* and *EcoRI*. The vector was circulated by introducing an adaptor that is compatible with *Asel* and *EcoRI* overhangs. This adaptor contained from 5' to 3' an *Acc65I* site, *SwaI* site, *PacI* site, *lox* site, *AvrII*, *SpeI*, *HindIII*, *Sall* and *EcoRI* site. This resulted in a plasmid herein called "TNG lox2DR" (lox2DR standing for the presence of 2 lox sites in Direct Repeat orientation).

Step 3: Step 3a: "TNG lox2DR" was cut by *MfeI*. Step 3b: pBI-426 (Datla *et al.* 1991, *supra*) was cut by *MfeI* and *EcoRI*. *MfeI* cuts the *gus* sequence 148bp downstream from the start codon. *EcoRI* cuts after the *Tnos* sequence which is placed downstream of the *nptII* sequence. In this way the "5'delGUS-NPTII-Tnos fragment" is formed and can be ligated in the TNGlox2DR vector. Cutting with *EcoRI* and *MfeI* resulted in compatible cohesive ends. As a consequence the "5'delGUS-NPTII-Tnos fragment" could be introduced in two possible orientations. The orientation in which 5'delGUS is placed closest to the RB was retained. This resulted in the restoration of the *MfeI* site but not of the *EcoRI* site. The vector obtained in this way is called "TNG lox2DR 5'd-GN" (GN stands for GUS-NPTII; 5'd: deletion at the 5' prime end of the *gus* gene).

Step 4: The *P35S-codA-Tnos* sequence was introduced in the TNG lox2DR 5'd-GN vector as a *HindIII-EcoRI* fragment from a subclone in pUC19. This resulted in the "TNG lox2DR 5'-GN-codA" vector.

At this stage, the vector that can be used as a starting product for both the TNG-1a and the TNG-3 vectors.

TNG-1a

Step 5: The *pUbi-csr1-1-tg7* sequence was introduced as a *PacI* fragment outside the lox cassette in the TNG lox2DR 5'-GN-codA vector. The orientation in which the cassette is directed towards the LB was retained. As a result the vector TNG-1a was formed. The *csr1-1* allele of the *A. thaliana acetolactate synthase (als)* gene contains an amino acid change Pro-197-Ser compared to the WT allele rendering resistance to the sulfonylurea herbicide chlorsulfuron (Haughn *et al.* 1988. *Mol Gen Genet* 211: 266-271). *Pubi*: Promoter of the *Helianthus annuus* GUBi gene coding for the hexaubiquitin protein. *tg7*: Poly(A) signal from the *tg7* gene.

TNG-3, TNG-6

Step 6: The *Pnos-hpt-Tnos* sequence was introduced in the TNG lox2DR 5'-GN-codA vector as a *HindIII* fragment. The orientation in which the cassette is oriented towards the LB was retained.

Step 7: An *attB1-SequenceX-attB2-CREi-T35S* sequence was introduced as a *PacI* fragment outside the lox sites. The orientation in which the cassette is oriented towards the LB was

retained. After the introduction of this fragment the SequenceX was exchanged *via* a BP reaction (Gateway™, Invitrogen) rendering the GMR destination vector.

Step 8: The SDS promoter and AP1 promoter were introduced as promoters upstream of the *cre* gene containing an intron (*cre-i*) (Joubès *et al.* 2004. Plant J 37: 889-896) gene *via* an LR reaction (Gateway™, Invitrogen) resulting in the TNG-3 and TNG-6 vectors respectively. This methodology was used to avoid cloning in a DB3.1 background, which appeared less efficient.

TNG-2I

Step 9: A second *I-SceI* site was introduced in the TNG lox2DR 5'-GN-codA vector via an *EcoRI* adaptor, which restored the *EcoRI* site at both ends.

Step 10: The *pUbi-csr1-1-tg7* sequence was introduced as a *PacI* fragment outside the lox cassette as described in step 5, resulting in the TNG-2I vector.

Example 4: Cloning of the CRE vectors

CRE1

CRE1 vector is equivalent to pHGWC (Marjanac, G. 2006. "Development of enabling technologies to modify gene expression in *Arabidopsis thaliana*", PhD thesis, University of Ghent, Ghent, Belgium; available via Central Library of UG under accession number BIB.GTH.031229) and contains the *cre* gene downstream of a Gateway™ cassette, allowing introduction of different promoters via an LR reaction (Gateway™, Invitrogen). This vector contains two upstream ATG codons in its 108-bp leader sequence, which is as follows:

5'attB2-

TCG**ATG**TGCCATTTAAATAAGCTTGAGCTCTCCCAT**ATG**GTCGACCTGCAGGCGGCCGC
ACTAGTGATATCCCGCGGGTTGACATG, wherein the underlined ATG marks the expected

start codon of the *cre* gene and the bold ATG mark two upstream ATG in the leader. attB2 is 5'-ACCCAGCTTTCTTGTACAAAGTGGT. Free energy of the secondary structure of the leader is -10.5kcal/mol.

CRE3

CRE3 vector is analogous to CRE1 except for its leader sequence, which contains no ATG upstream to the expected *cre* start codon: The 92-bp leader is as follows:

5'-attB2-

- 5 TCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCGGCCGCACTAGTGATATCCCC
CGGGTTGACATG, wherein the underlined ATG is the *cre* start codon and attB2 is as above.
Free energy of the secondary structure of the leader is -11.7 kcal/mol.

Promoter-Entry clones

- The desired promoters were obtained by amplification from genomic plant DNA using a
10 forward primer containing an attB1 sequence and a reverse primer containing an attB2
sequence and cloned into pDONR-ZEO™ via a BP reaction (Invitrogen). From these
promoter-entry clones, the promoters were easily introduced into CRE1 or CRE3 vectors. The
primers used were as follows:

SOLO DANCERS (SDS)

- 15 SDS F: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTAGGAAGCGTATTGCTCGACTC-3'
SDS R: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTTCTCCGTACGAAAGCTTG-3'

AP1 (APETALA1)

AP1-F: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCTTGGGATGTTGTCTTCAAGG-3'
AP1-R: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCCAAACAAAACAAAGACCCCC-3'

- 20 CLV3 (*CLAVATA3*)

Clv-GW-F1:

5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCGGATTATCCATAATAAAAACAAA-3'

Clv-GW-R1:

5'-GGGGACCACTTTGTACAAGAAAGCTGGGTAGAGAAATATAGAACTGTTCTTTACT-3'

- 25 ClvER-Term-F1: 5'-GAATTCATATCTCTTGTTGCTTTAAATTA-3'

ClvER-Term-R1: 5'-GAATTCATTTTTTTAAAAAATAGTTAATTATC-3'

LFY (LEAFY)

Lfy-GW-F1: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGATCCATTTTTTCGCAAAGG-3'

Lfy-GW-R1:

5'-GGGGACCACTTTGTACAAGAAAGCTGGGTAATCTATTTTTCTCTCTCTCTATCA-3'

Ag (AGAMOUS)

Ag-ENH-F1 : 5'-**CCCGGG**GTTTCTTCTTCTTCGTGCTC-3'

5 Ag-ENH-R1 : 5'-**CCCGGG**CTGCAATAATTTTTTTAAAGGAATT-3'

min35S-1:

5'-

GGTACCCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAG**GGTACC**-3'

min35S-2:

10 5'-

GGTACCCTCTCCAAATGAAATGAACTTCCTTATATAGAGGAAGGGTCTTGCG**GGTACC**-3'

AtDMC3.5 (cloned from genomic DNA of *A. thaliana* Landsberg Erecta)

DMC3.5-GW-F:

5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCCCTCTTATAAGCTTCAAGCTGCTGA-3'

15 DMC3.5-GW-R:

5'-GGGGACCACTTTGTACAAGAAAGCTGGGCTTTCTCGCTCTAAGACTCTCTAAGCT-3'

AtMSH4

MSH4-GW-F

5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGTGAGCTGTGTGACGTTATTGTT-3'

20 MSH4-GW-R

5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTTCGCTCCACAGATCAGAA-3'

AtSPO11-1

SPO11-GW-F

5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTTCGCACGCACCTCCGAT-3'

25 SPO11-GW-R

5'-GGGGACCACTTTGTACAAGAAAGCTGGGCTCTTTCGAGTTTCAAACCTGAAAAA-3'

The sequence of the minimal 35S promoter was placed upstream of the reverse oriented *AGAMOUS* enhancer sequence.

Sequence that is underlined is the promoter specific part of the primer.

Example 5: Functional characterisation of the SDS and AP1 promoters in CRE1 vector

SDS and AP1 promoters were introduced in the CRE1 vector as above.

An *Arabidopsis thaliana* Col/O plant line FK24 (De Buck *et al.* 2004, *supra*) containing the K2L610 T-DNA (Fig. 2), which contains a P35S driven *gus* gene flanked by *lox* sites, as a single copy and in a homozygous way, was supertransformed by using the floral dip method (Clough and Bent 1998, *supra*) with respectively the SDS-CRE1 and AP1-CRE1 vector.

Fourteen FK24::SDS-CRE1 and fifteen FK24::AP1-CRE1 primary transformants (T1), selected on hygromycin containing medium, were analyzed by PCR, using primers 1, 2 and 3 (see Fig. 2 and legend) capable of detecting the CRE/*lox* recombination event. Three T2 plants for each primary transformant were analyzed by PCR as well.

8 of 14 (57%) FK24::SDS-CRE1 T1 plants and 11 of 15 (73%) of FK24::AP1-CRE1 T1 plants were heterozygous for the recombined and non-recombined alleles. The results for T2 generation are shown in Fig. 3.

These results demonstrate that no a-zygous plants for the GUS allele were found in the T2 generation of the FK24::SDS-CRE1 plants which suggest that the CRE recombinase was functional in only one germline, which is in contrast to what was expected from literature (Azumi *et al.* 2002. EMBO J 21: 3081-95; WO0174144). Of the analyzed FK24::SDS-CRE1 T2 plants 80% showed to be hemizygous. Furthermore, 52% of the T2 generation of FK24::AP1-CRE1 plants were a-zygous for the GUS allele and 32% showed to be hemizygous. "Germline" refers to cells of which descendants are destined to be gametes and the gametes themselves. The genetic material of germline cells is heritable.

To assess to which germline the functionality of the SDS promoter was confined, the following experiment was conducted. T2 generation FK24::pSDS-CRE1 plants have been crossed with a line that is homozygous for the *bar* gene and doesn't contain the *cre* gene. FK24::pSDS-CRE1 plants were used both as pollen donor and as pollen acceptor. The seeds, which resulted from the respective crosses, were sown on selective medium (selective agent corresponding to the selective marker present in the pollen donor in order to distinguish real F1s from selfed progeny) and analyzed by histochemical GUS staining. From each T2 generation FK24::pSDS-CRE1 plant used in the crossing, seed from selfing was also

collected in order to perform GUS stainings on T3 plants and in this way deduce the status of the T2 plants, homozygous or heterozygous. We used T2 plants originating from two independent T1 transformants referred to as SDS-CRE1 4a and SDS-CRE1 7b. The results are in Table 1.

- 5 Table 1. Results from the GUS assays on progeny from reciprocal crosses of T2 FK24::SDS-CRE1 (*gus* +/+ or *gus* +/-) lines crossed with *bar* +/+ lines. GUS++: homozygous for the lox-flanked *gus* gene, GUS+-: hemizygous for the lox-flanked *gus* gene, GUS--: a-zygous for the lox-flanked *gus* gene.

T2 (pollen acceptor x pollen donor)	T3		
	# plants	# blue (<i>gus</i> + -)	# white (<i>gus</i> - -)
SDS-CRE1-4a line			
SDS-CRE1 GUS++ x <i>bar</i> ++	357	357	0
<i>bar</i> ++ x SDS-CRE1 GUS++	193	61	132
SDS-CRE1 GUS+- x <i>bar</i> ++	97	49	48
<i>bar</i> ++ x SDS-CRE1 GUS+-	10	0	10
SDS-CRE1-7b line			
SDS-CRE1 GUS++ x <i>bar</i> ++	52	52	0
<i>bar</i> ++ x SDS-CRE1 GUS++	29	1	28
SDS-CRE1 GUS+- x <i>bar</i> ++	78	39	39
<i>bar</i> ++ x SDS-CRE1 GUS+-	76	3	73

- This data shows that CRE under control of the SDS promoter, when introduced in the CRE1 vector, is functional in only one germline; the results of the reciprocal crosses showed that this functionality is limited to the male germline. Moreover, the effectiveness is rather high in both plant lines. For the 4a line 68.4% (132/193) and for the 7b line 96.6% (28/29) of the analyzed progeny plants starting from T2 plants that are homozygous for the *gus* allele, became GUS-negative. Starting from T2 plants that are heterozygous for the *gus* allele, 96.1% (73/76) of the analyzed progeny plants of the 7b line were GUS-negative.

Example 6: Functional characterisation of the CLV3, LFY, AP1, AG and SDS promoters in CRE3 vector

- Modified CRE vector was created to assess the functionality of the meristematic (CLV3, LFY, AP1 and AG) and meiotic (SDS) promoters and to compare this to CRE1 vector (CRE1-AP1 and SDS as controls).

The CLV3, LFY, AP1, AG and SDS promoters were independently introduced in the CRE3 vector which was in turn introduced in the FK24/1 plant genome (see Example 5). For each construct primary transformants were obtained after selection on hygromycin and subsequently analyzed for excision events by PCR using primers 1, 2 and 3 on T1 and T2 plants (as shown in Fig. 2).

The results for T1 and T2 generations are shown in Fig. 4A and B, respectively. These results clearly demonstrate an efficient CRE/Lox recombination. All tested promoters can be considered apt to drive CRE expression in order to obtain sufficient CRE/Lox recombination.

Influence of the leader sequence on CRE functionality

Comparison of the AP1-CRE1 vs. AP1-CRE3 and SDS-CRE1 vs. SDS-CRE3 results in function of the total number excision events (in T1 and T2 FK24 supertransformants), shows that these efficiencies are enhanced in the CRE3 vector for both promoters (Fig. 5), presumably due to the absence of leader ATG codons and/or to the presence of a shorter leader sequence in the CRE3 vector. Also the secondary structure of the leader may have an impact on the expression level.

As was demonstrated in Example 5, the SDS promoter in the CRE1 vector is exclusively functional in the male germline. Results herein concerning the SDS-CRE3 vector indicate the functionality of the SDS promoter (in the CRE3 vector) to a certain extent in both germ lines. On the other hand the amount of background excision in the T1 was lower when the SDS-CRE1 and AP1-CRE1 were used instead of the SDS-CRE3 and AP1-CRE3.

It may be concluded that modulation of the CRE leader – by adding/removing one or more upstream ATG codons to/from the leader, and/or by lengthening/shortening the leader sequence, and/or by changing the secondary structure of the leader sequence – can lead either to a shift in CRE/Lox recombination efficiency and/or a shift in germline specificity. For example, if one wants to lower the background excision, lowering the expression levels of the CRE by introducing upstream ATG(s) may be an option.

Example 7: Developing the TARGET lines

Arabidopsis thaliana ColO WT plants were transformed with the TARGET construct (Fig. 6). After selection of primary transformants on phosphinotrycin (10mg/l), we collected seeds from

the first sexual progeny. After segregation analysis we retained only single locus plants, which were characterized both molecularly and phenotypically.

Molecular characterization

Southern blot analysis was performed in order to obtain information on the integration pattern of the TARGET T-DNA and to verify the integrity of the T-DNA.

DNA was cut with BspHI and probed with the GUS1 probe to collect information around the RB. The probe was washed away and the blot was hybridized with the BAR probe which rendered information around the LB. Based hereon, three integration lines were retained: 7A ("line A"), 6D ("line D") and 1G ("line G"). The integrity of the T-DNA in these lines was verified using two additional probes: the GUS2 probe and the lambda2 probe. All three lines proved to contain an intact T-DNA. Lines A and D contain a single copy complete T-DNA. Line G contains more than one copy.

The genomic position of the TARGET T-DNA in lines A and D was determined via I-PCR as described in Example 1, using TDN-F and TDN-R primers and nested TDN-NesF and 35S-NesR primers. The PCR products obtained after the nested PCR were sequenced, which allowed to pinpoint the genomic location of the T-DNA. The RB is inserted at chromosome 2 (17993801) and chromosome 1 (1728401) for target line A and D respectively. These results coincide with the data obtained from the Southern blot and were confirmed by performing PCR with one primer corresponding with the T-DNA (pptProbe-R) and another primer (genA3F2/genD5R2) corresponding with the adjacent genomic region of both target lines, producing the expected products of about 1843 bp and about 1087 bp, respectively (see Fig. 7).

Quantitative GUS analysis

The three target lines contain the same *gus:nptII 3'del* reporter gene under control of a CAMV35S promoter. To assess the GUS expression levels, the GUS activity was quantitated in protein extracts of five individual plants per line:

Line A: average 37,42 U Gus/mg protein (individual plants: 32,045; 55,74; 40,91; 28,9885; 29,405 U Gus/mg protein);

Line D: average 17,35 U Gus/mg protein (individual plants: 17,86; 17,365; 16,82 U Gus/mg protein);

FK24/1: average 361,727 U Gus/mg protein (individual plants: 510,32; 178,13; 330,73; 320,73; 468,725 U Gus/mg protein).

A control line (FK24) was included and showed elevated activity levels of GUS relative to the target lines. The target line G and two plants of target line D did not give GUS activity levels above the detection limit. The absence of GUS activity in target line G might be caused by lower GUS expression levels due to the presence of an inverted repeat of the T-DNA carrying the reporter gene in target line G.

Example 8: In *trans* acting CRE enzyme-SDS promoter

3 Target lines (G, D and A) were supertransformed with SDS-CRE1 (see Example 4) leading to TARGET/SDS-CRE1 plants. After selection on hygromycin of the T1 transformants we collected T2 seeds and did a segregation analysis by sowing 100 to 150 seeds on hygromycin containing medium.

Six single locus supertransformants per TARGET line were used as starting lines for the second supertransformation, with TNG-1a (Fig. 1, Example 3) resulting in TARGET/SDS-CRE 1/TNG-1a primary transformants, referred to as T1.

T1 seeds of this second supertransformation were collected and were sown on chlorsulfuron (Glean[®], Dupont, France) containing medium. Besides the presence of the TNG-1a we had to check the presence of the SDS-CRE1 as we used TARGET/SDS-CRE1 plants -of which we didn't know the zygosity of the SDS-CRE1- for the floral dip transformation with TNG-1a. This was achieved by PCR analysis using the following primers:

lambda-F: 5'-GTCATGGACTCCTCCACAGAGAAAC-3')

GUS-R: (5'-CCAGACAGAGTGTGATATCTACCCG-3')

An overview of the strategy is given in Fig. 8.

For the G, D and A line T2 seeds were collected from respectively 41, 47 and 37 TARGET/SDS-CRE 1/TNG1a T1 lines.

Selection of T2 seeds on medium containing kanamycin (50mg/l) revealed that in total 157 Km^R seeds were obtained out of 585317 seeds or 1 out of every 3728 seeds corresponding with a frequency of 0.27×10^{-3} . The results for the three TARGET lines individually are:

line G: 33 Km^R seeds/182000 seeds= 1/5515 → 0.18 x 10⁻³

line D: 40 Km^R seeds/222964 seeds= 1/5574 → 0.18 x 10⁻³

line A: 84 Km^R seeds/180353 seeds= 1/2147 → 0.47 x 10⁻³

Molecular characterization of the Km^R plants

- 5 Kanamycin resistant plants were analyzed molecularly by PCR. An overview of the different PCR reactions, performed on DNA purified from leaf material from the T2 kanamycin resistant plants, is given in Figure 9. An overview of the results of the PCR are listed in Table 2.

Table 2. PCR1 = Primer 1+2, PCR3 = Primer 1+5, PCR4 = Primer 1+6 (wherein the primers are numbered as in Fig. 9). +/Km^R: number of positive PCR results out of total of kanamycin resistant (Km^R) plants; +/PCR1+: number of positive PCR results for the respective PCR performed only on DNA from plants having a positive result in PCR1. *In line A 55 of the 71 Km^R plants originated from one T1 plant. If we analyze the results without this line 16 plants were kanamycin resistant of which 5 scored positive on PCR1.

	PCR1 (+/Km^R)	PCR3 (+/PCR1+)	PCR4 (+/PCR1+)
line G	2/13	2/2	2/2
line D	3/30	2/3	1/3
line A*	59/71	54/59	59/59
	5/16	5/5	5/5

- 15 Three control lines (Col0, TARGET A, TARGET D) and 9 Kanamycin resistant lines (putative GT events) were analysed by Southern blot (Fig. 11). Genomic DNA was prepared from plants pooled per line and digested with SphI(PaeI) and a gus-35S probe (GT1a-probe) was used.

GT1a-probe-F: 5'-CAATTGCCCGGCTTTCTTGTAACG-3'

GT1a-probe-R: 5'-CCTCGGATTCCATTGCCAGCT-3'

- 20 Figure 11 shows Southern blot results of three control lines (Col0, Target line A, Target line D) and nine kanamycine resistant lines. Genomic DNA was digested with SphI and a gus-35S probe was used (GT1a-probe). The expected fragment for the original target and recombined target allele (respectively 9916 bp and 3075 bp) are both visible in the kanamycin resistant lines in contrast to the control lines, Target A and D, where only the original target allele is visible. The results corresponding to Figure 11 are also represented in Table 3 below:
- 25

No	LINE	Original Allele	Recombined
1	ColO	No	No
2	A	Yes	No
3	D	Yes	No
4	149.1	Yes	Yes
5	149.27	Yes	Yes
6	150.1	Yes	Yes
7	149.20	Yes	Yes
8	149.12	Yes	Yes
9	149.17	Yes	Yes
10	149.21	Yes	Yes
11	149.23	Yes	Yes
12	149.26	Yes	Yes

These results clearly show that one copy of the TARGET allele was restored by homologous recombination.

Example 9: In *cis* acting CRE enzyme-SDS promoter

- 5 The 3 TARGET lines (G, D and A) were supertransformed with TNG-3. Primary transformants were selected on hygromycin. T2 seeds were collected and sown on medium containing kanamycin (50mg/l).

An overview of the strategy is given in Fig. 10.

- 10 Selection of T2 seeds on medium containing kanamycin (50mg/l) revealed that in total 33 Km^R seeds were obtained out of 213854 seeds or 1 out of every 6480 seeds corresponding with a frequency of 0.15×10^{-3} . Results for the three TARGET lines individually:

line G: 11 Km^R seeds / 79124 seeds $\rightarrow 0.14 \times 10^{-3}$

line D: 12 Km^R seeds / 80318 seeds $\rightarrow 0.15 \times 10^{-3}$

line A: 10 Km^R seeds / 54412 seeds $\rightarrow 0.18 \times 10^{-3}$

- 15 Preliminary PCR analysis on a limited number of T2 transformants indicated that 9 out 22 investigated kanamycin resistant plants scored positive on PCR1(Primer 1+2, as in Fig. 9).

Example 10: *In cis* acting CRE enzyme- AP1 promoter

The 3 TARGET lines (G,D and A) were supertransformed with TNG-6. Primary transformants were selected on hygromycin. T2 seeds were collected and sown on medium containing kanamycin.

- 5 Selection on medium containing kanamycin (50mg/l) of T2 seeds revealed that in total 20 Km^R seeds were obtained out of 104276 seeds or 1 out of every 52132 seeds corresponding with a frequency of 0.19×10^{-3} . Results for the three TARGET lines individually:

line G: 9 Km^R seeds / 31968 seeds $\rightarrow 0.28 \times 10^{-3}$

line D: 9 Km^R seeds / 48464 seeds $\rightarrow 0.19 \times 10^{-3}$

- 10 line A: 9 Km^R seeds / 23844 seeds $\rightarrow 0.38 \times 10^{-3}$

Preliminary PCR analysis on a limited number of T2 transformants indicated that 6 out of 16 investigated kanamycin resistant plants scored positive on PCR1(Primer 1+2, as in Fig. 9).

Example 11: *In trans* acting CRE enzyme- DMC3.5 promoter

- 2 Target lines (D and A) were supertransformed with TNG-2I (Fig. 1, Example 3) resulting in
15 TARGET/TARGETING primary transformants after selection on chlorsulfuron (Glean[®], Dupont, France) containing medium. We collected T2 seeds and did a segregation analysis by sowing 100 to 150 seeds on chlorsulfuron containing medium. Eight single locus lines and two lines with multiple loci were used as starting lines for the second supertransformation with DMC3.5-CRE3 resulting in TARGET/TNG2I/DMC3.5-CRE3 primary transformants, referred to
20 as T1.

- T1 seeds of this second supertransformation were collected and were sown on hygromycin containing medium. Besides the presence of the DMC3.5-CRE3 we had to check the presence of the TNG-2I as we used TARGET/TNG2I plants – of which we didn't know the zygosity of the TNG2I – for the floral dip transformation with DMC3.5-CRE3. This was
25 achieved by PCR analysis.

An overview of the strategy is given in Fig. 8.

For the D and A line T2 seeds were collected from respectively 24 and 18 TARGET/TNG2I/DMC3.5-CRE3 T1 lines.

Selection of T2 seeds on medium containing kanamycin (50mg/l) revealed that in total 26 Km^R seeds were obtained out of 77172 seeds or 1 out of every 3215 seeds corresponding with a frequency of 0.33×10^{-3} . The results for the two TARGET lines individually are:

line A: 11 Km^R seeds/31168 seeds= 1/2833 $\rightarrow 0.35 \times 10^{-3}$

5 line D: 15 Km^R seeds/46004 seeds= 1/3066 $\rightarrow 0.32 \times 10^{-3}$

Example 12: In *trans* acting CRE enzyme- MSH4 promoter

2 Target lines (D and A) were supertransformed with TNG-2I (Fig. 1, Example 3) resulting in TARGET/TARGETING primary transformants after selection on chlorsulfuron (Glean[®], Dupont, France) containing medium. We collected T2 seeds and did a segregation analysis
10 by sowing 100 to 150 seeds on chlorsulfuron containing medium. Eight single locus lines and two lines with multiple loci were used as starting lines for the second supertransformation with MSH4-CRE3 resulting in TARGET/TNG2I/ MSH4-CRE3 primary transformants, referred to as T1.

T1 seeds of this second supertransformation were collected and were sown on hygromycin
15 containing medium. Besides the presence of the MSH4-CRE3 we had to check the presence of the TNG-2I as we used TARGET/TNG2I plants – of which we didn't know the zygosity of the TNG2I – for the floral dip transformation with MSH4-CRE3. This was achieved by PCR analysis.

An overview of the strategy is given in Fig. 8.

20 For the D and A line T2 seeds were collected from respectively 10 and 11 TARGET/TNG2I/MSH4-CRE3 T1 lines.

Selection of T2 seeds on medium containing kanamycin (50mg/l) revealed that in total 555 Km^R seeds were obtained out of 51344 seeds or 1 out of every 92 seeds corresponding with a frequency of 10.8×10^{-3} . The results for the two TARGET lines individually are:

25 line A: 73 Km^R seeds/25816 seeds= 1/353 $\rightarrow 2.83 \times 10^{-3}$

line D: 482 Km^R seeds/25528 seeds= 1/53 $\rightarrow 18.9 \times 10^{-3}$

Example 13: In *trans* acting CRE enzyme- Spo11 promoter

2 Target lines (D and A) were supertransformed with TNG-2I (Fig. 1, Example 3) resulting in TARGET/TARGETING primary transformants after selection on chlorsulfuron (Glean[®], Dupont, France) containing medium. We collected T2 seeds and did a segregation analysis
 5 by sowing 100 to 150 seeds on chlorsulfuron containing medium. Eight lines with a single locus and two with multiple loci were used as starting lines for the second supertransformation with Spo11-CRE3 resulting in TARGET/TNG2I/Spo11-CRE3 primary transformants, referred to as T1.

T1 seeds of this second supertransformation were collected and were sown on hygromycin
 10 containing medium. Besides the presence of the Spo11-CRE3 we had to check the presence of the TNG-2I as we used TARGET/TNG2I plants – of which we didn't know the zygosity of the TNG2I – for the floral dip transformation with Spo11-CRE3. This was achieved by PCR analysis.

An overview of the strategy is given in Fig. 8.

15 For the D and A line T2 seeds were collected from respectively 8 and 10 TARGET/TNG2I/Spo11-CRE3 T1 lines.

Selection of T2 seeds on medium containing kanamycin (50mg/l) revealed that in total 1 Km^R seed was obtained out of 63316 seeds corresponding with a frequency of 0.016×10^{-3} . The results for the two TARGET lines individually are:

20 line A: $1 \text{ Km}^{\text{R}} \text{ seeds} / 23696 \text{ seeds} = 1/23696 \rightarrow 0.042 \times 10^{-3}$

line D: $0 \text{ Km}^{\text{R}} \text{ seeds} / 39620 \text{ seeds} = \rightarrow 0$

CLAIMS

1. A method for targeted nuclear genome modification in plants comprising:

(a) providing a plant comprising a nuclear genome locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:

5 (aa) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC from the locus L1, and

10 (ab) comprising at least one region of homology (RH) to a sequence (S) at a nuclear genome locus (L2) of said plant different from the locus L1; and

(b) expressing the site-specific recombinase R at a level which causes the excision of the targeting cassette TC from locus L1 preferentially in cells of the plant having increased frequency of homologous recombination but substantially not in progenitors or predecessors of said cells, wherein the plant comprises a sequence (SR) encoding the site-specific
15 recombinase R operably linked to regulatory elements capable of effecting said expression of said recombinase R, thereby releasing the targeting cassette TC from the locus L1 in said cells,

whereby homologous recombination occurs in at least some of said cells between the at least one region of homology RH of the targeting cassette TC released in step (b) and the
20 sequence S at the locus L2.

2. The method according to claim 1, wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter which is not constitutive.

3. The method according to claim 1, wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter which does not require external stimuli to
25 drive or effect expression of said recombinase R.

4. The method according to claim 1, wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter which is cell-specific, tissue-specific and/or

developmental stage-specific with respect to cells having increased frequency of homologous recombination.

5. The method according to any of claims 1 to 4, wherein cells in which the site-specific recombinase R is preferentially expressed in step (b) are chosen from cells of meristem, preferably cells of apical meristem or stem cells; cells of microsporangia; or cells of megasporangia.

6. The method according to any of claims 1 to 4, wherein cells in which the site-specific recombinase R is preferentially expressed in step (b) are chosen from cells at one or more stages of mitotic cell cycle chosen from interphase, G₁ phase, G₁/S transition, S phase, S/G₂ transition, G₂ phase, G₂/M transition or M (mitosis).

7. The method according to claim 6, wherein cells in which the site-specific recombinase R is preferentially expressed in step (b) are at G₂/M transition.

8. The method according to any of claims 1 to 4, wherein cells in which the site-specific recombinase R is preferentially expressed in step (b) are meiocytes, preferably microsporocytes and/or megasporocytes.

9. The method according to claim 8, wherein in said meiocytes the site-specific recombinase R is preferentially expressed at one or more stages of meiotic cell division chosen from interphase, G₁ phase, G₁/S transition, S phase, S/G₂ transition, G₂ phase, G₂/Meiosis I transition, Meiosis I or Meiosis II.

10. The method according to claim 1, wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter chosen from promoters of genes: *SOLO*, *DANCERS*, *APETALA1*, *CLAVATA3*, *LEAFY*, *AGAMOUS*, *SPO11-1*, *DMC1* or *MSH4*.

11. The method according to any of claims 1 to 10 further comprising linearising the targeting cassette TC released in step (b) and/or introducing a double strand break within or adjacent to the sequence S at the locus L2, in order to augment homologous recombination there between.

12. The method according to any of claims 1 to 11, wherein the targeting cassette TC further comprises a negative selection marker.

13. A plant comprising:

(i) a nuclear genome locus (L1) and a targeting cassette (TC) at said locus L1, said targeting cassette TC:

(ia) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in release of the targeting cassette TC from the locus L1, and

(ib) comprising at least one region of homology (RH) to a sequence (S) at a nuclear genome locus (L2) of said plant or plant cell different from the locus L1; and

(ii) the sequence S at the locus L2 different from the locus L1.

14. The plant according to claim 13, further comprising a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements capable of effecting the preferential expression of said recombinase R in cells as defined in step (b) of any of claims 1 to 10.

15. The plant according to claim 14, wherein the sequence SR and regulatory elements thereof are at the locus L1, optionally within the targeting cassette TC.

16. The plant according to any of claims 14 or 15, wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter chosen from promoters of genes: *SOLO DANCERS*, *APETALA1*, *CLAVATA3*, *LEAFY*, *AGAMOUS*, *SPO11-1*, *DMC1* or *MSH4*.

17. The plant according to any of claims 13 to 16, wherein the targeting cassette TC further comprises a means allowing for linearisation of said targeting cassette TC after having been released from the locus L1, and/or wherein the locus L2 comprises a means allowing for introduction of a double strand break within or adjacent to the sequence S at the locus L2.

18. The plant according to any of claims 13 to 17, wherein the targeting cassette TC further comprises a negative selection marker.

19. A genetic construct suitable for transformation, esp. of plants, comprising a targeting cassette (TC), said targeting cassette:

(I) being flanked by a pair of recombinase sites (RS) responsive to a site-specific recombinase (R), the recombinase sites RS configured so that the action of the site-specific recombinase R thereon results in excision of the targeting cassette TC from the genetic construct, and

- 5 (II) comprising at least one region of homology (RH) to an endogenous nuclear sequence of a plant or plant cell.

20. The genetic construct according to claim 19 further comprising a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements which can effect the preferential expression of said recombinase R in cells as defined in step (b) of any of claims 1
10 to 9.

21. The genetic construct according to claim 20, wherein the sequence SR and regulatory elements thereof are within the targeting cassette TC.

22. The genetic construct according to any of claims 20 or 21, wherein the sequence SR encoding the site-specific recombinase R is operably linked to a promoter chosen from
15 promoters of genes: *SOLO DANCERS*, *APETALA1*, *CLAVATA3*, *LEAFY*, *AGAMOUS*, *SPO11-1*, *DMC1* or *MSH4*.

23. The genetic construct according to any of claims 19 to 22, wherein the targeting cassette TC further comprises a means allowing for linearisation of said targeting cassette TC after having been released from the genetic construct.

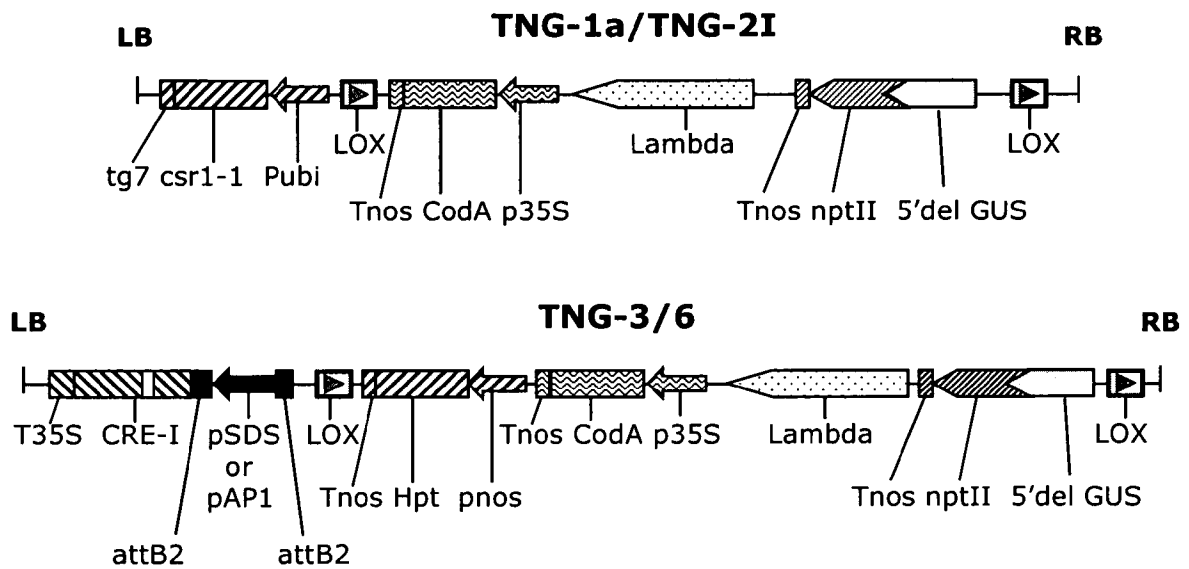
20 24. The genetic construct according to any of claims 19 to 23, wherein the targeting cassette TC further comprises a negative selection marker.

25. Use of the genetic construct according to any of claims 19 to 24 to transform a plant or plant cell.

25 26. Use of the genetic construct according to claim 19 and a genetic construct comprising a sequence (SR) encoding the site-specific recombinase R operably linked to regulatory elements which can effect the preferential expression of said recombinase R in cells as defined in step (b) of any of claims 1 to 9 to transform a plant or plant cell.

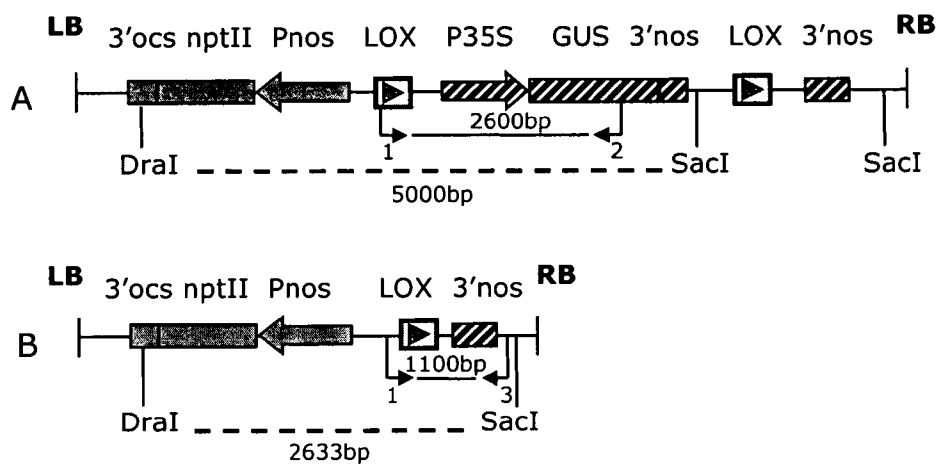
1/11

FIG 1



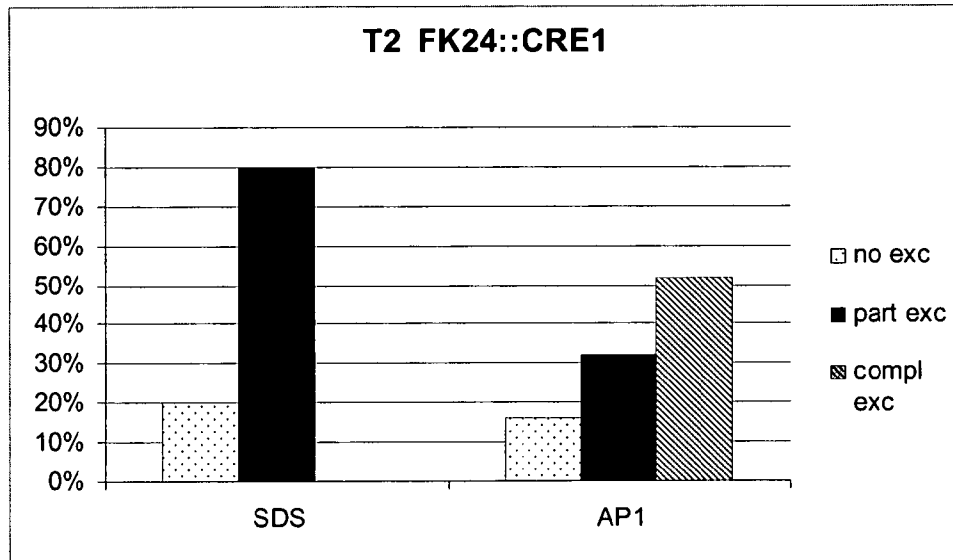
2/11

FIG 2



3/11

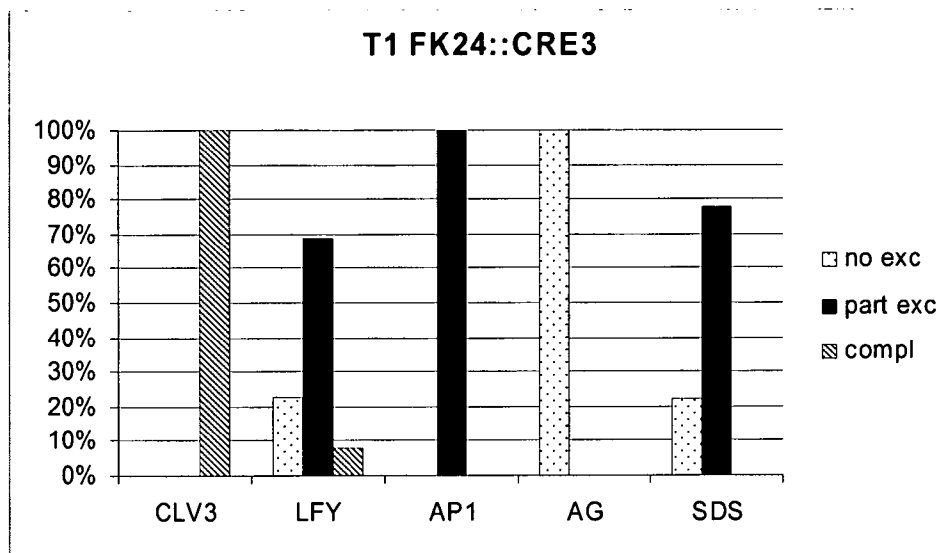
FIG 3



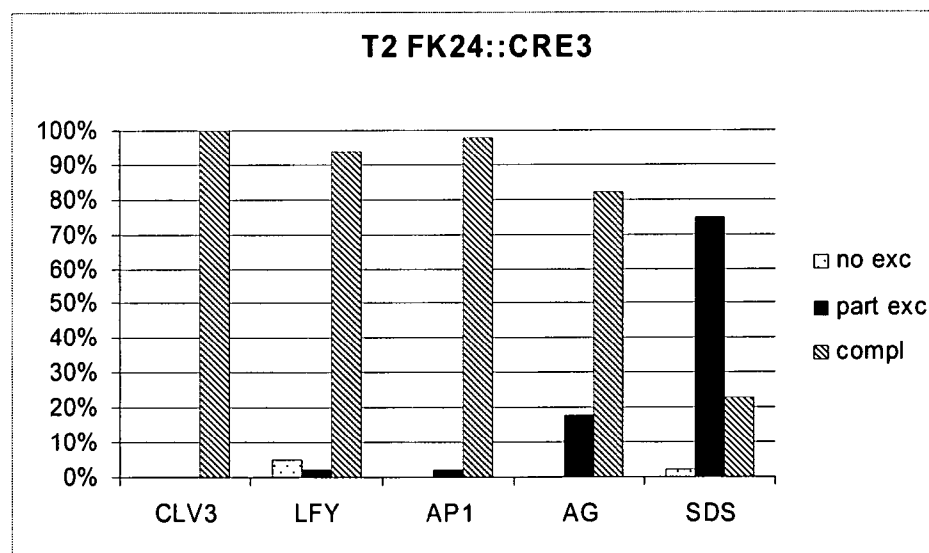
4/11

FIG 4

A



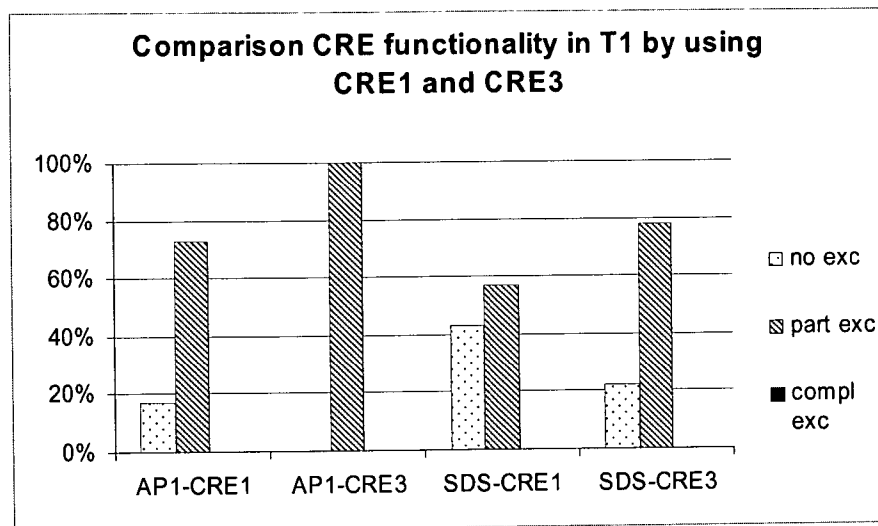
B



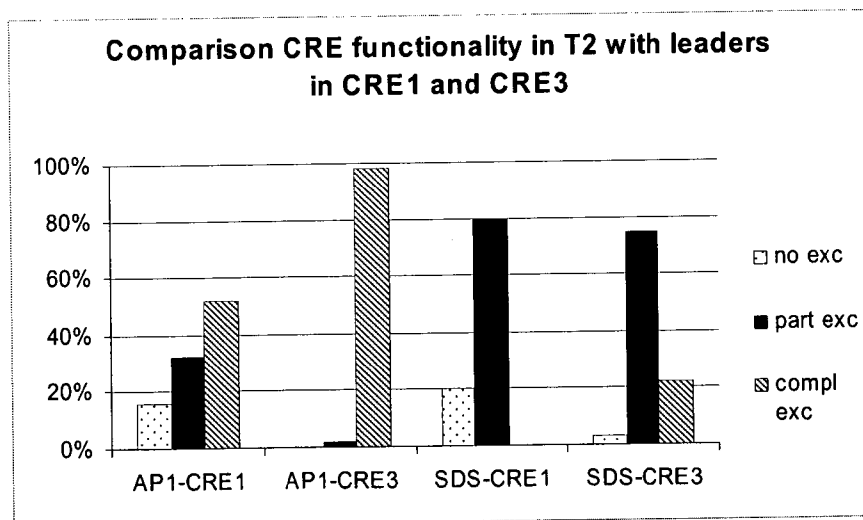
5/11

FIG 5

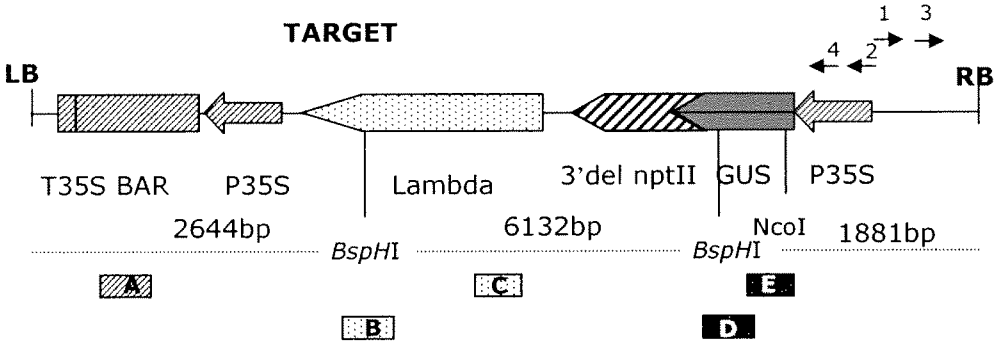
A



B

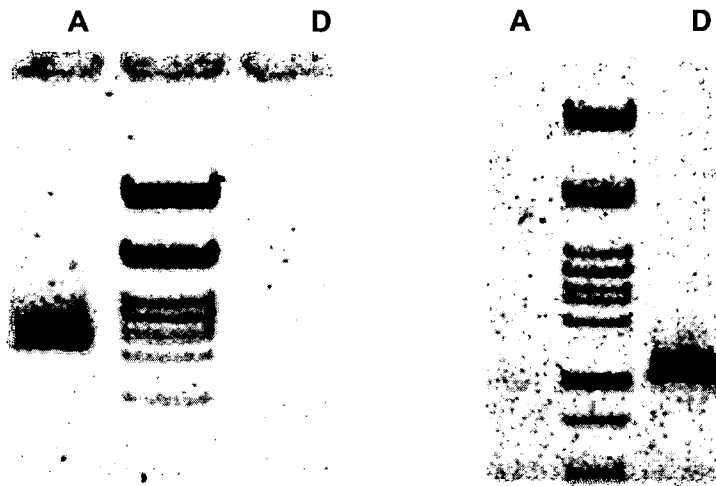


6/11
FIG 6



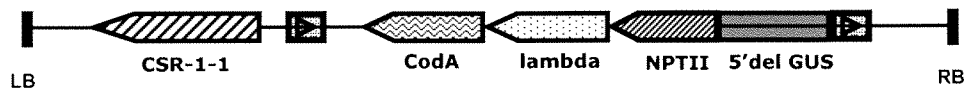
7/11

FIG 7

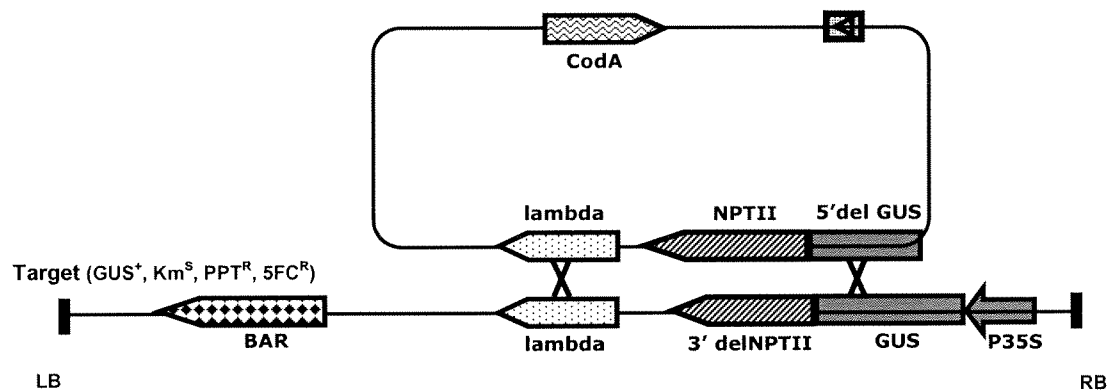


8/11
FIG 8

Targeting vector (GUS⁻, Km^S, Glean^R, 5FC^S)



↓ CRE *in trans* (SDS-CRE1)

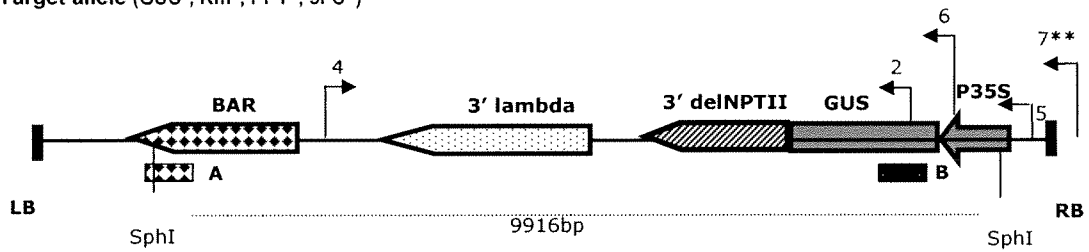


Recombined allele (GUS⁺, Km^R, PPT^R, 5FC^R)

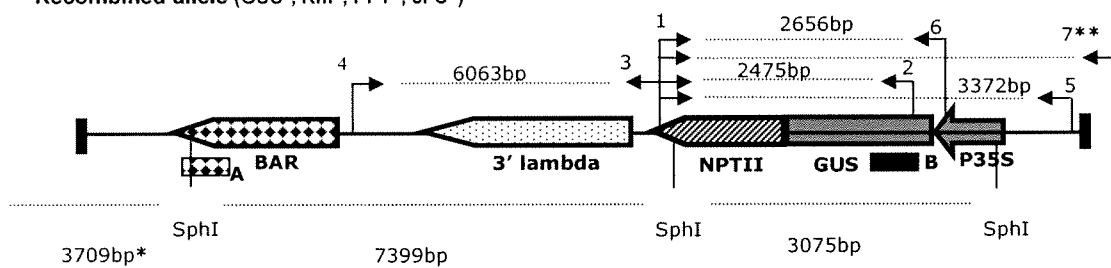


9/11
FIG 9

Target allele (GUS^+ , Km^S , PPT^R , $5FC^R$)

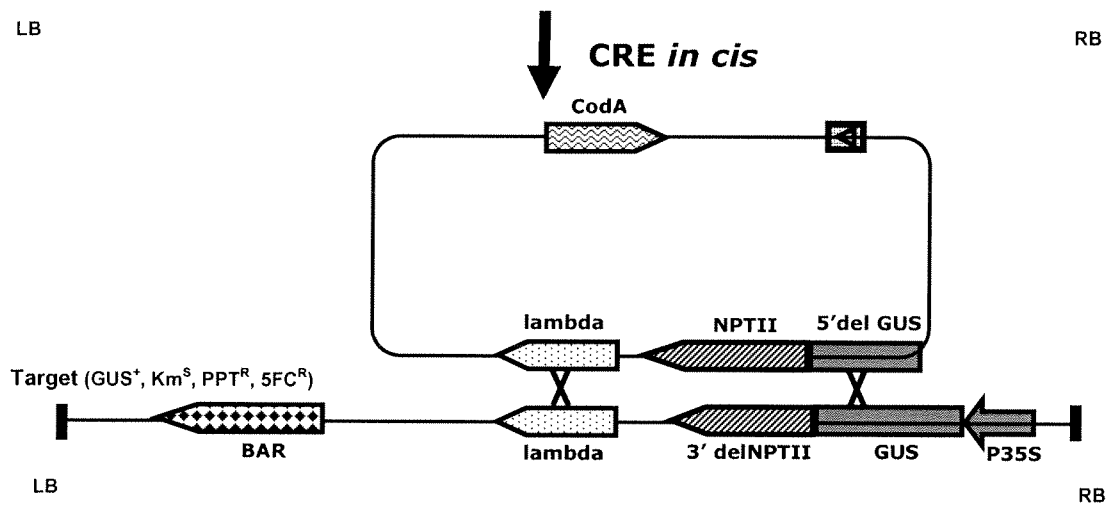
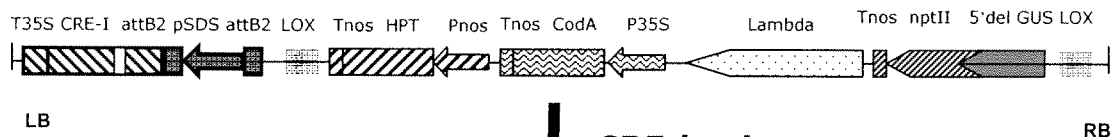


Recombined allele (GUS^+ , Km^R , PPT^R , $5FC^R$)



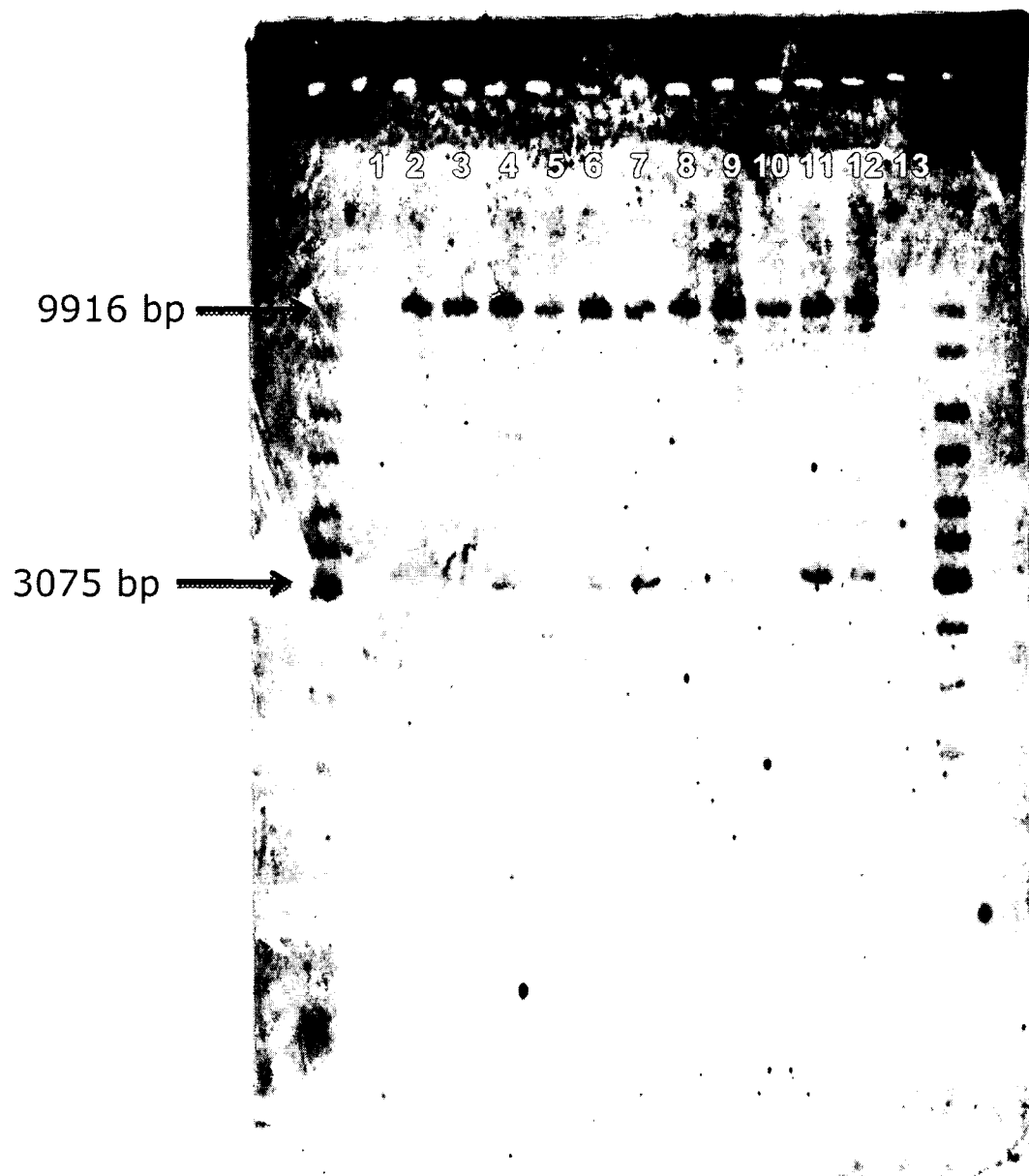
* 3709bp for A line

** xxx for A line

10/11
FIG 10Targeting vector (GUS^{-} , Km^S , Hyg^R , $5FC^S$)Recombined allele (GUS^{+} , Km^R , PPT^R , $5FC^R$)

11/11

FIG 11



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/056770

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/82 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KUMAR SANDEEP ET AL: "Controlling transgene integration in plants" TRENDS IN PLANT SCIENCE, vol. 6, no. 4, April 2001 (2001-04), pages 155-159, XP002457433 ISSN: 1360-1385 figure 1	13, 19, 25, 26
Y	page 157	1-12, 14-18, 20-24
X	WO 03/093428 A (PIONEER HI BRED INT [US]) 13 November 2003 (2003-11-13) claims 1, 18-33	13, 19, 25, 26
Y	page 23	1-12, 14-18, 20-24
	----- -/-	

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

* & * document member of the same patent family

Date of the actual completion of the international search

13 October 2008

Date of mailing of the international search report

28/10/2008

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Bucka, Alexander

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2008/056770

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MLYNÁROVÁ LUDMILA ET AL: "DIRECTED MICROSPORE-SPECIFIC RECOMBINATION OF TRANSGENIC ALLELES TO PREVENT POLLEN-MEDIATED TRANSMISSION OF TRANSGENES" PLANT BIOTECHNOLOGY JOURNAL, BLACKWELL, OXFORD, GB, vol. 4, no. 4, 1 July 2006 (2006-07-01), pages 445-452, XP008078342 ISSN: 1467-7644 the whole document	1-12, 14-18, 20-24
A	LUTZ KERRY A ET AL: "A novel approach to plastid transformation utilizes the phiC31 phage integrase." PLANT JOURNAL, vol. 37, no. 6, March 2004 (2004-03), pages 906-913, XP002457436 ISSN: 0960-7412 figure 6	1-26
A	WO 94/17176 A (PURDUE RESEARCH FOUNDATION [US]) 4 August 1994 (1994-08-04) figure 8	1-26
A	BIBIKOVA M ET AL: "Enhancing gene targeting with designed zinc finger nucleases" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 300, 2 May 2003 (2003-05-02), page 764, XP002974267 ISSN: 0036-8075 left-hand column	1-26
A	MOORE SARAH K ET AL: "Efficient deletion of transgenic DNA from complex integration locus of rice mediated by Cre/lox recombination system" CROP SCIENCE, vol. 46, no. 2, March 2006 (2006-03), pages 700-705, XP002457432 ISSN: 0011-183X figure 1	1-26
A	VERGUNST ANNETTE C ET AL: "Cre/lox-mediated recombination in Arabidopsis: Evidence for transmission of a translocation and a deletion event" CHROMOSOMA (BERLIN), vol. 109, no. 4, 2000, pages 287-297, XP002457435 ISSN: 0009-5915 figures 2,5	1-26
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/056770

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SHAKED HEZI ET AL: "High-frequency gene targeting in Arabidopsis plants expressing the yeast RAD54 gene" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC, US, vol. 102, no. 34, August 2005 (2005-08), pages 12265-12269, XP009087433 ISSN: 0027-8424 cited in the application the whole document	1-26
A	WO 2004/106496 A (TOSK INC [US]; FOGARTY PATRICK [US]) 9 December 2004 (2004-12-09) the whole document	1-26
A	MENGISTE T. ET AL: "PROSPECTS FOR THE PRECISE ENGINEERING OF PLANT GENOMES BY HOMOLOGOUS RECOMBINATION" BIOLOGICAL CHEMISTRY, WALTER DE GRUYTER GMBH & CO., BERLIN, DE, vol. 380, no. 7/8, 1999, pages 749-758, XP000885763 ISSN: 1431-6730 page 752 - page 753	1-26
A	VERGUNST A C ET AL: "RECOMBINATION IN THE PLANT GENOME AND ITS APPLICATION IN BIOTECHNOLOGY" CRITICAL REVIEWS IN PLANT SCIENCES, CRC PRESS, BOCA RATON, FL, US, vol. 18, no. 1, January 1999 (1999-01), pages 1-31, XP009038217 ISSN: 0735-2689 page 6 - page 7 page 9, right-hand column	1-26
A	VERGUNST ANNETTE C ET AL: "Cre/lox-mediated site-specific integration of Agrobacterium T-DNA in Arabidopsis thaliana by transient expression of cre" PLANT MOLECULAR BIOLOGY, SPRINGER, DORDRECHT, NL, vol. 38, no. 3, October 1998 (1998-10), pages 393-406, XP002159988 ISSN: 0167-4412 figure 2	1-26
	-/--	

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/056770

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WRIGHT D A ET AL: "High-frequency homologous recombination in plants mediated by zinc-finger nucleases" PLANT JOURNAL, BLACKWELL SCIENTIFIC PUBLICATIONS, OXFORD, GB, vol. 44, no. 4, November 2005 (2005-11), pages 693-705, XP002403198 ISSN: 0960-7412 cited in the application figure 2	1-26
A	PUCHTA H: "TOWARDS THE IDEAL GMP: HOMOLOGOUS RECOMBINATION AND MARKER GENE EXCISION" JOURNAL OF PLANT PHYSIOLOGY, FISCHER, STUTTGART, DE, vol. 160, no. 7, 2003, pages 743-754, XP009030464 ISSN: 0176-1617 page 745, left-hand column	1-26
A	RONG YIKANG S ET AL: "A targeted gene knockout in Drosophila" GENETICS, vol. 157, no. 3, March 2001 (2001-03), pages 1307-1312, XP002457434 ISSN: 0016-6731 figure 1	1-26
A	KOLB A F ET AL: "Insertion of a foreign gene into the beta-casein locus by Cre-mediated site-specific recombination" GENE, ELSEVIER, AMSTERDAM, NL, vol. 227, no. 1, 4 February 1999 (1999-02-04), pages 21-31, XP004155764 ISSN: 0378-1119 figure 1	1-26
A	US 7 126 041 B1 (HELMER GEORGIA L [US] ET AL) 24 October 2006 (2006-10-24) claim 1	1-26
P,X	VERWEIRE DIMITRI ET AL: "Marker-free Transgenic plants through genetically programmed auto-excision" PLANT PHYSIOLOGY (ROCKVILLE), vol. 145, no. 4, December 2007 (2007-12), pages 1220-1231, XP002498435 ISSN: 0032-0889 the whole document	1-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2008/056770

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 03093428	A	13-11-2003	AU 2003234330 A1	17-11-2003
			CA 2483891 A1	13-11-2003
			EP 1499720 A2	26-01-2005
			US 2003208794 A1	06-11-2003
			US 2007016980 A1	18-01-2007
WO 9417176	A	04-08-1994	AT 291092 T	15-04-2005
			DE 69434297 D1	21-04-2005
			DE 69434297 T2	29-12-2005
			EP 0686191 A1	13-12-1995
			US 5527695 A	18-06-1996
			US 5744336 A	28-04-1998
WO 2004106496	A	09-12-2004	NONE	
US 7126041	B1	24-10-2006	US 2007039074 A1	15-02-2007