# Gene targeting in maize by somatic ectopic recombination

Ayhan Ayar, Sophie Wehrkamp-Richter, Jean-Baptiste Laffaire, Samuel Le Goff, Julien Lévy, Sandrine Chaignon, Hajer Salmi, Alexandra Lepicard, Christophe Sallaud, Maria Gallego, et al.

Plant Biotechnology Journal

*Plant Biotechnology Journal* (2013) **11**, pp. 305–314 doi: 10.1111/pbi.12014

# Gene targeting in maize by somatic ectopic recombination

Ayhan Ayar[1,2], Sophie Wehrkamp-Richter[1,2], Jean-Baptiste Laffaire[1], Samuel Le Goff[1], Julien Levy[1], Sandrine Chaignon[1], Hajer Salmi[1], Alexandra Lepicard[1], Christophe Sallaud[1], Maria E. Gallego[2], Charles I. White[2] and Wyatt Paul[1,*]

[1]*Biogemma, Centre de Recherche de Chappes, Chappes, France*
[2]*CNRS UMR 6293,Clermont Université INSERM U1103, Aubière, France*

*Received 13 June 2012;*
*revised 28 August 2012;*
*accepted 25 September 2012.*
**Correspondence* (Tel +(0)4 73 67 88 40; fax +(0)4 73 67 88 99; email wyatt.paul@biogemma.com)

**Keywords:** double-strand break, meganuclease, ectopic recombination, gene targeting, *nptII*, *Zea mays*.

## Summary

Low transformation efficiency and high background of non-targeted events are major constraints to gene targeting in plants. We demonstrate here applicability in maize of a system that reduces the constraint from transformation efficiency. The system requires regenerable transformants in which all of the following elements are stably integrated in the genome: (i) donor DNA with the gene of interest adjacent to sequence for repair of a defective selectable marker, (ii) sequence encoding a rare-cutting endonuclease such as I-*Sce*I, (iii) a target locus (*TL*) comprising the defective selectable marker and I-*Sce*I cleavage site. Typically, this requires additional markers for the integration of the donor and target sequences, which may be assembled through cross-pollination of separate transformants. Inducible expression of I-*Sce*I then cleaves the *TL* and facilitates homologous recombination, which is assayed by selection for the repaired marker. We used *bar* and *gfp* markers to identify assembled transformants, a dexamethasone-inducible I-*Sce*I::GR protein, and selection for recombination events that restored an intact *nptII*. Applying this strategy to callus permitted the selection of recombination into the *TL* at a frequency of 0.085% per extracted immature embryo (29% of recombinants). Our results also indicate that excision of the donor locus (*DL*) through the use of flanking I-*Sce*I cleavage sites may be unnecessary, and a source of unwanted repair events at the *DL*. The system allows production, from each assembled transformant, of many cells that subsequently can be treated to induce gene targeting. This may facilitate gene targeting in plant species for which transformation efficiencies are otherwise limiting.

## Introduction

Transgenesis offers the possibility to insert a known DNA sequence into the genome of an organism to introduce new heritable characters. It is commonly used in research to investigate gene function and in biotechnology to improve agronomic traits. However, random insertion of the transgene into the genome can result in mutations caused by the insertion into an endogenous gene (Krysan *et al.*, 1999), potential production of unintended peptides or variable expression due to the genomic environment of the transgene (Matzke and Matzke, 1998). Thus, there are currently considerable efforts worldwide to develop efficient technologies for gene targeting (GT) to produce genetically modified (GM) crops with transgenes located at predetermined positions in the plant genome. Exploiting the cellular homologous recombination (HR) machinery, GT allows the exchange of genetic information between homologous DNA sequences and can be used to precisely modify the genome. The integration of transgenes flanked by sequences homologous to the desired genomic insertion site permits efficient and routine gene targeting in prokaryotes and fungi, but GT is very inefficient in higher plants with frequencies of the order of $10^{-4}$ per transformant (Cotsaftis and Guiderdoni, 2005; Hanin *et al.*, 2001; Paszkowski *et al.*, 1988).

HR and non-homologous end-joining (NHEJ) are triggered to repair double-strand breaks (DSBs) of DNA. These lesions are formed accidentally by genotoxic stresses (Hanin and Paszkowski, 2003; Khanna *et al.*, 2001; Tuteja *et al.*, 2009) or in a programmed manner, for example during meiosis by the Spo11 complex (Grelon *et al.*, 2001). Repair through NHEJ links the two ends of the DSB and is frequently accompanied by the creation of mutations at the site of the repair. HR copies an endogenous (different allele or stably inserted transgene) or exogenous (non inserted transgene) sequence template with homology on either side of the break and allows a precise modification of the genome (Puchta *et al.*, 1996). Transgene integration is believed to generally involve insertion via NHEJ into a pre-existing DSB (Tzfira *et al.*, 2004) occurring randomly in the plant genome. A DSB at a precise genomic location presenting homologous sequence to the transgene significantly increases the recombination rate at this site (Puchta *et al.*, 1993; Szostak *et al.*, 1983). This has led to the development of tools such as meganucleases, zinc-finger nucleases and transcription activator-like effector nucleases for gene targeting (Christian *et al.*, 2010; Shukla *et al.*, 2009; Tzfira *et al.*, 2012). These endonucleases create a DSB at the target locus (*TL*) and have been used to modify the *TL* by mutation using NHEJ (De Pater *et al.*, 2009; Yang *et al.*, 2009) or by precise sequence modification using HR (Tzfira and White, 2005). For example, the mitochondrial I-*Sce*I meganuclease from *Saccharomyces cerevisiae* (Jacquier and Dujon, 1985) has been successfully used in plants to perform GT (D'Halluin *et al.*, 2008; Puchta *et al.*, 1996). In tobacco, cleavage of the *TL* containing an I-*Sce*I restriction site by I-*Sce*I increases recombination between the *TL* and the transforming T-DNA around 100-fold (Puchta

*et al.*, 1996). The enzyme required to produce the DSB can be introduced into the organism or cell via stable or transient transformation. For example, I-*Sce*I has been introduced into plants via *Agrobacterium*-mediated retransformation of a *TL* line or by crossing lines stably expressing I-*Sce*I to a *TL* line. In the latter case, the use of an inducible I-*Sce*I can allow the creation of the DSB at a predetermined moment. For example, application of the glucocorticoid, dexamethasone, induced the activity of an I-*Sce*I protein fused to the rat glucocorticoid receptor (GR) domain in *Arabidopsis* (Wehrkamp-Richter *et al.*, 2009). The GR domain sequesters the I-*Sce*I::GR complex in the cytosol. The addition of dexamethasone allows the dissolution of the complex (Aoyama and Chua, 1997), liberating the I-*Sce*I::GR protein which can move to the nucleus and produce a DSB at the *TL*. In plants, an inducible I-*Sce*I was used to enhance intrachromosomal recombination (Wehrkamp-Richter *et al.*, 2009) and to perform targeted mutagenesis (Yang *et al.*, 2009).

Plant GT strategies are generally based on the positive selection for GT events which repair a defective selectable marker. A DSB is induced at the *TL* inducing HR between a defective *TL* selectable marker gene and the repair DNA. For example in *Zea mays* (maize), D'Halluin *et al.* (2008) re-transformed *TL* lines with a repair DNA and a construct encoding I-*Sce*I, either delivering the DNA via particle bombardment or *Agrobacterium*. The frequency of GT versus random insertion, measured by the acquisition of resistance to the herbicide BASTA, was up to 30% via particle bombardment and 3.7% using *Agrobacterium*. Shukla *et al.* (2009) have also reported efficient GT in maize using zinc-finger nucleases. Although these studies show that GT is now possible in a major crop plant, there is still the need to optimize GT to minimize the effort required to produce and identify GT events before GT becomes a routine tool for GM production. A major limiting factor is the need to deliver the repair DNA and nuclease-encoding sequence efficiently into a large number of cells, which in the case of maize transformation can involve the transformation of many thousands of immature embryos or calli to obtain a few GT events. An attractive alternative is to create a few transformation events where the repair DNA and I-*Sce*I-encoding sequence are stably integrated into the genome. The repair DNA is then controllably excised from the genome and acts as a template for GT at the *TL*. This system has the advantage that every cell contains the repair template, and thus, a single transformed individual can yield a potentially unlimited population of cells for GT. Such a GT strategy has been successfully implemented in *Drosophila*, with the repair DNA being excised from the genome using the FLP recombinase and then linearized using I-*Sce*I (Huang *et al.*, 2008; Rong, 2002) and has recently been reported also in *Arabidopsis* (Fauser *et al.*, 2012). The goal of the work presented here was to test a similar GT system in maize, using a dexamethasone-inducible I-*Sce*I both to excise the repair DNA from the genome and to induce a DSB at the *TL*.

## Results

### The GT test system

Two plant transformation constructs, the *TL* construct and the donor locus (*DL*) construct, were developed to test the GT strategy. The T-DNA of the *TL* construct contains the plant transformation selectable marker phosphinothricin acyl transferase (*bar)* gene followed by an I-*Sce*I restriction site and the 3′ part of the neomycin phosphotransferase II (*nptII*) gene (Figure 1b). The T-DNA of the *DL* construct contains a dexamethasone-activatable, maize codon-optimized I-*Sce*I (*I-SceI::GR*) gene and an *nptII* repair region bordered by two I-*Sce*I restriction sites (Figure 1a). The *nptII* repair region contains the *bar* gene, the green fluorescent protein (*gfp*) gene and a 5′ part of the *nptII* gene. The *gfp* gene here serves as a mock gene of interest to be inserted at the *TL* and additionally allows easy identification of *DL*-containing plants. The *nptII* repair region and the *TL* share common sequences of 2992 bp in the *bar* region and 1200 bp in the *nptII* region, provided largely by the insertion of a rice tubulin gene intron (*intTubI*) into the defective *nptII* genes. This homology should allow homologous recombination between these two sequences and the consequent repair of the *nptII* gene, resulting in kanamycin resistance (Figure 1c). The *TL* and the *DL* constructs were independently transformed into maize to generate *TL* and *DL* lines, respectively. Two intact *TL* (*TL1* and *TL2*) lines and one *DL* line, each containing a single copy of the transgene, were selected by Southern blotting analysis (not shown) and their genomic flanking sequences isolated (Figure S1). The *DL* line expressed both *gfp* and the *I-SceI::GR* transcript. The two *TL* lines were then selfed in order to isolate homozygous descendants which were then crossed with the *DL* line (Figure 2a). The F1 progenies and their descendents were selfed (Figure 2b). The segregation of the *TL* and the *DL* indicates that the two constructions were not genetically linked.

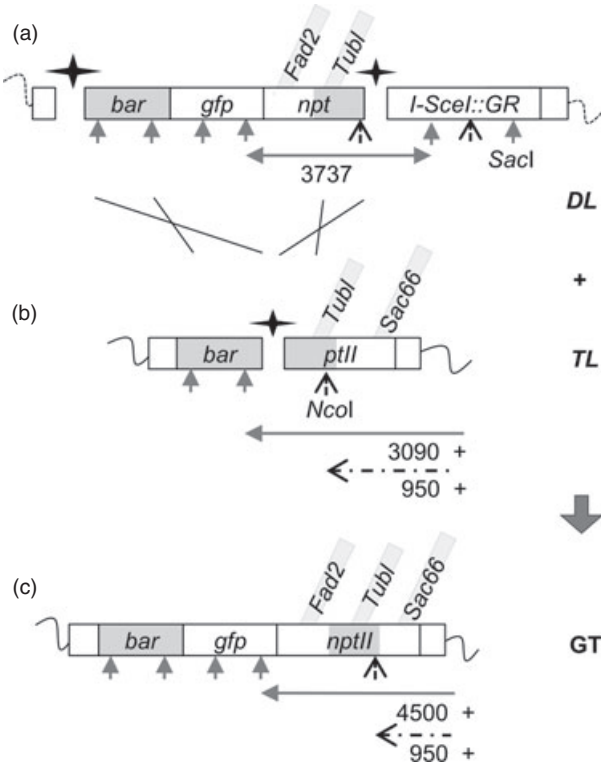


**Figure 1** Maps of the transgenic loci (not to scale). Structures of the donor locus (a), the target locus (b) and the expected gene targeting event (c) with the position of the I-*Sce*I cutting sites (black stars). For Southern blot analyses, the sizes of DNA fragments hybridized with *intTubI* (*TubI*), *intFad2* (*Fad2*) and *terSac66* (*Sac66*) are shown for digestion with *Sac*I (grey arrows) or *Nco*I (black dashed arrows). '+' indicates that the size of the fragment will be greater than indicated as it will extend to the first relevant restriction site in the flanking genomic DNA. The left border of the T-DNA is adjacent to the *bar* gene. Fragment lengths are given as bp.

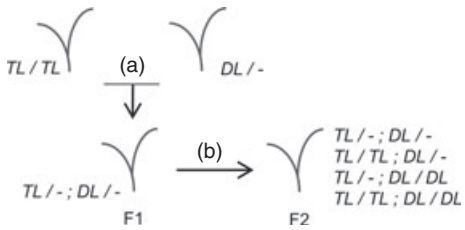


**Figure 2** Genealogy of maize Lines. The two homozygous target locus (*TL*) lines were crossed with a heterozygous donor locus (*DL*) line to obtain the F1 (a). The F1 progeny containing the *DL* and the *TL* were self-fertilized to obtain F2 descendants (b). '—' represents the original wild-type loci. Only the genotypes of interest (containing *TL* and *DL*) are shown.

## Detection of somatic repair of *nptII* in *TL*/*DL* leaves

For each *TL* line crossed with the *DL* line, the F1 progeny containing the *TL* and the *DL* were identified by PCR analysis and separated in two groups, seed of one group was treated with dexamethasone to induce I-*Sce*I activity and the other not (see Experimental procedures). To detect excision of the repair DNA from the *DL*, PCR was performed using primers positioned on either side of the *DL* I-*Sce*I restriction sites. A total of 12 untreated and seven dexamethasone-treated plants were analysed and three excision events were detected for each condition, indicating a basal activity of I-*Sce*I::GR and that dexamethasone treatment does not significantly induce I-*Sce*I::GR in these conditions.

The analysed F1 plants were then selfed to identify kanamycin-resistant plants among the F2 descendants. To detect the presence of the *TL* and the *DL*, 176 F2 (42 for the *TL1/DL* line and 134 for the *TL2/DL* line) plants were analysed by PCR; 21 *TL1/DL* and 55 *TL2/DL* descendants contained both *TL* and *DL*. Kanamycin was applied to the apical meristematic region of wild-type (WT) and the F2 plants. On WT plants and F2 descendants containing only either the *TL* or the *DL*, this resulted in bleaching of the developing leaf (Figure 3b). However, leaves with green sectors within the kanamycin-bleached zones (Figure 3c) were observed in 60% of plants carrying both the *TL* and *DL*, corresponding to 38% of *TL1/DL* and 70% of *TL2/DL* plants (Figure 3d). PCR analysis performed on the DNA extracted from the green sectors permitted the amplification and sequencing of the repaired *nptII* gene, and this was not so for DNA extracted from bleached or untreated (Figure 3a) leaf sectors. Other batches of F2 seeds were sown, and none of the additional 504 descendants were fully kanamycin-resistant; however, green kanamycin-resistant sectors were again observed in *TL*/*DL* plants.

## Recovery of fully kanamycin-resistant plants via *in vitro* tissue culture

Notwithstanding the presence of kanamycin-resistant leaf sectors (and thus GT), no fully kanamycin-resistant progeny were identified in 680 F2 plantlets. We thus used a tissue culture approach. Plant regeneration from maize leaves has not been reported, but calli derived from immature maize embryos are routinely used to regenerate plants (Lu *et al.*, 1983). Embryos isolated from immature kernels of selfed F2 plants containing the *TL* and the *DL* were placed on callus induction medium with dexamethasone at 0 μM (control), 30 or 50 μM (Figure 4). From 2356 extracted embryos (619 from the *TL1/DL* and 1737 from the *TL2/DL* plants), seven independent kanamycin-resistant GT events

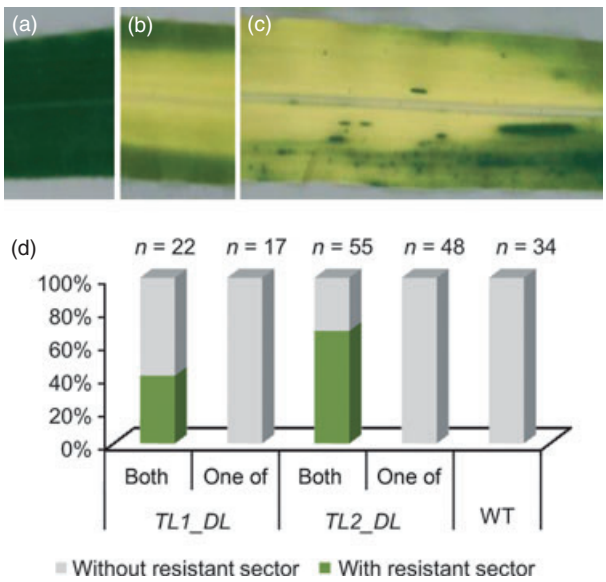


**Figure 3** Frequency of green sectors in kanamycin-treated plants. Leaves of F2 maize plants untreated with kanamycin (a), sensitive to kanamycin (b) and sensitive to kanamycin with kanamycin-resistant green sectors (c). The histogram (d) shows the percentage of sensitive plants with green kanamycin-resistant sectors for the different genotypes (n is the number of plants analysed). 'Both' indicates plants possessing a target locus (*TL*) and a donor locus (*DL*); 'one of' indicates plants possessing either only a *TL* or a *DL*.

(Table 1) were recovered and shown to carry a repaired *nptII* gene which was amplified by PCR and sequenced. Two were obtained from the *TL1/DL* embryos (GT1 and GT2) and five from the *TL2/DL* embryos (GT3, GT4, GT5, GT6 and GT7). GT efficiencies calculated as the number of GT events per immature embryo range from 0.13% to 0.55% (Table 1).

Only one of the seven GT events was obtained from dexamethasone non-treated control embryos; thus, dexamethasone treatment appears to increase the number of GT events. The number of events are, however, low and as we observed somatic recombination during the development of F2 plants in the absence of dexamethasone treatment, some GT events probably come from a basal, leaky I-*Sce*I::GR activity.

To clarify the question of the inducibility of I-*Sce*I::GR activity, we tested the effect of dexamethasone treatment on *TL* DSB

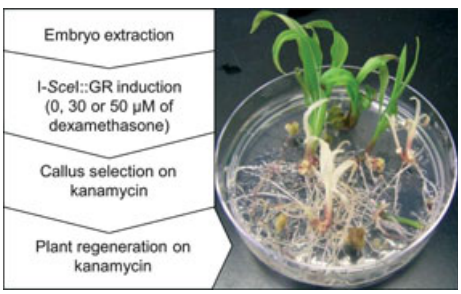


**Figure 4** Strategy for *in vitro* culture and regeneration. Immature embryos were collected and separated into three groups, non-treated, treated with 30 μM and treated with 50 μM of dexamethasone. Kanamycin selection allowed the regeneration of kanamycin-resistant plants containing a repaired *nptII* gene (green plantlets).

**Table 1** Summary of gene targeting experiments

| Lines | Extracted embryos | Dexamethasone concentration (μM) | Number of kanamycin-resistant events | Name of kanamycin-resistant events | GT frequency (kanamycin-resistant/ total embryos) |
|---|---|---|---|---|---|
| *TL1* × *DL* | 183 | 0 | 0 | / | 0% |
| | 255 | 30 | 1 | GT1 | 0,39% |
| | 181 | 50 | 1 | GT2 | 0,55% |
| *TL2* × *DL* | 800 | 0 | 1 | GT4 | 0,13% |
| | 563 | 30 | 3 | GT3, GT5, GT6 | 0,53% |
| | 374 | 50 | 1 | G7 (+) | 0,27% |

+ indicates a sterile plant.

GT, gene targeting; *TL*, target locus; *DL*, donor locus.

induction through the measurement of mutations in the *TL* I-*Sce*I site. Embryos of F2 plants were extracted (14 from *TL1/DL* and 69 from *TL2/DL* lines) for somatic embryogenesis. A sample of the callus formed from each embryo was analysed by PCR to determine the presence of the *DL* and the *TL*. Each callus was divided into three parts, one part placed for 1 week on medium without dexamethasone, one on medium with 30 μM and one with 50 μM dexamethasone. Samples of *TL/DL* calli from each treatment and of the *TL* calli were pooled separately for DNA extraction (Table 2). A 400-bp region around the I-*Sce*I *TL* restriction site was amplified from each pool and sequenced by 454 sequencing (Genome Sequencer FLX by Roche). Approximately 15 000 sequences were obtained and analysed to estimate the mutation rate at the I-*Sce*I restriction site due to NHEJ repair. No mutations were detected in the absence of the *DL* (*TL* controls that do not carry the *I-SceI* gene). In the *TL/DL* lines, mutations of the *TL* I-*Sce*I site were detected, with the number of independent mutations increasing 3.5 to 5-fold with 30 μM and 5 to 6-fold with 50 μM of dexamethasone. These data thus confirm a basal activity of I-*Sce*I::GR in inducing mutations in the I-*Sce*I target site and that dexamethasone treatment increases I-*Sce*I::GR activity in calli (Table 2). The presence of mutations in the target in the absence of dexamethasone, however, confirms the leakiness in this system.

## Analysis of GT obtained from *TL1/DL* plants

Two GT events were identified from calli from *TL1/DL* plants. These were regenerated to give plants GT1 and GT2. Southern blot analysis was carried out on *Sac*I-digested genomic DNA of these plants, the parent line (*TL1/DL*) and control lines carrying only the target locus (*TL1*) or the *DL*. Three different probes were used: *Arabidopsis AtFad2* gene intron (*intFad2*, present in the *DL* and predicted to be present in a GT locus), *intTubI* (common to the *DL*, *TL* and predicted to be in the GT locus) and *Arabidopsis AtSac66* terminator (*terSac66*, present in the *TL* and predicted to be in the GT locus). The Southern blot results with the *intTubI* probe are shown in Figure 5b. A band of 3.7 kbp was detected in the *DL* lane and a 4.1-kbp band in the *TL1* lane; both bands were observed in the *DL/TL1* control lane. As expected, in GT1 and GT2 lanes, the *TL1* band disappeared and a new 5.5-kbp band, also observed with *intFad2* and *terSac66* probes (Figure S2), was detected, confirming *nptII* repair at the *TL*. The non-excised *DL* band was observed at 3.7 kbp with an intensity consistent with a homozygous state.

The GT1 and GT2 plants were backcrossed twice to wild-type plants, and kanamycin resistance was inherited as a single Mendelian locus. In the first backcross, 53% and 55% of GT1 and GT2 descendants were kanamycin resistant and all presented the non-excised *DL*, confirming that GT1 and GT2 are heterozygous for the reconstructed (by GT) *nptII* gene at the *TL*. For the second backcross, 35% and 36% of GT1 and GT2 descendants were kanamycin resistant. All resistant plants expressed *gfp,* and PCR amplification confirmed both the presence of the *intFad2* and *terSac66* regions and the absence of the *TL1*-specific fragment containing the I-*Sce*I *TL1* site (Figure S3), which together confirm the expected reconstitution of *nptII*. Among the kanamycin-resistant descendants, 61% of GT1 and 55% of GT2 also contained all sequences specific to the *DL* [left border (LB) I-*Sce*I site, right border (RB) I-*Sce*I site and *I-SceI::GR*].

Finally, to confirm that the modified *TL* in plants GT1 and GT2 are the result of homologous recombination on both sides of the break in the *TL* with the donor, we also sequenced the junction fragments amplified by PCR with primers to the *terSac66* and the genomic flanking sequence of the *TL1* LB. PCR fragments were amplified from kanamycin-resistant GT1 and GT2 plants containing only the GT locus and sequenced. The sequence of the

**Table 2** Quantification of mutations at the I-*Sce*I site of target locus (*TL*)

| Lines | Number of Embryos | Genotype | Number of reads | Dexamethasone concentration (μM) | Mutations in *TL* I-*Sce*I site |
|---|---|---|---|---|---|
| *TL1* × *DL* | 14 | *TL1* | 15370 | / | 0 |
| | | *TL1/DL* (*I-SceI::GR*) | 11575 | 0 | 15 |
| | | | 17685 | 30 | 76 |
| | | | 23279 | 50 | 76 |
| *TL2* × *DL* | 69 | *TL2* | 15597 | / | 0 |
| | | *TL2/DL* (*I-SceI::GR*) | 16293 | 0 | 6 |
| | | | 19583 | 30 | 15 |
| | | | 17931 | 50 | 36 |

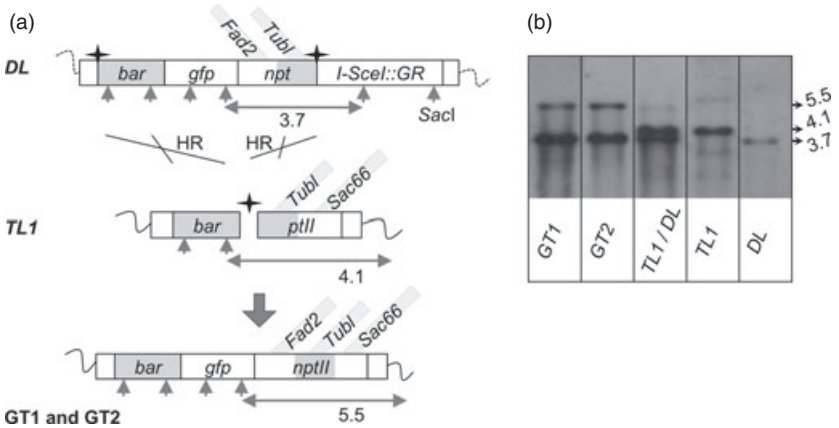


**Figure 5** Southern blot analysis of gene targeting (GT)1 and GT2 events. (a) Schema of the GT event occurring in the GT1 and GT2 events. The black stars show I-*Sce*I restriction sites, and cleavage of the site in target locus (*TL*)*1* is indicated by the separation of the two halves of the locus. Grey arrows indicate *Sac*I restriction sites and the size of DNA fragments detected by the *intTubI* probe. (b) Southern blot analysis of control parental plants (*TL1*, donor locus (*DL*), *TL1/DL* plants) and of recombinant GT plants from *TL1/DL* tissue culture (GT1 and GT2 lines) with *Sac*I-digested DNA and the *intTubI* probe. A fragment around 3.7 kbp, indicative of the original *DL*, was detected in the GT1 and GT2 events. A 4.1-kbp fragment, indicative of *TL1,* was not observed in the GT1 and GT2 events. The expected fragment of 5.5 kbp, indicative of *nptII* reconstitution due to HR between the *DL* and *TL1*, was detected in the GT1 and GT2 events. The *DL* band is more intense than the *TL1* band for the lanes GT1 and GT2, indicating the homozygous state of the *DL* and the non-excision of the *DL*. Fragment lengths are given as kbp.

amplified fragment obtained (Figure S4) was identical to that predicted for HR between the *DL* and *TL1* resulting in the repair of *nptII* at the *TL1*.

Analyses of the GT1 and GT2 events thus showed that they are true GT events at the *TL* and that the GT was not associated with excision of the donor sequence from the *DL* in either case, suggesting that they arose through ectopic recombination (Puchta, 1999) between the *TL* and *DL* (Figure 5a).

## Analysis of GT from *TL2/DL* plants

Five GT events were identified from calli from *TL2/DL* plants. These were regenerated to give plants GT3, GT4, GT5, GT6 and GT7. Analysis of these plants revealed a second class of GT events involving reconstitution of *nptII* at the *DL*, rather than at the *TL* (Figure 6a).

Southern blot analysis was carried out on *Sac*I-digested genomic DNA of these plants, the parent line (*TL2/DL*) and control lines carrying only the target locus (*TL2*) or the *DL*, hybridized with *intFad2*, *intTubI* and *terSac66* probes. The results with the *intTubI* probe are presented in Figure 6b. The original *DL* and *TL2* bands of 3.7 and 5.7 kbp respectively were detected in the GT samples, except for GT3 that lacked the original *DL*. However, the predicted GT-specific band of 7.1 kbp for GT at the *TL* (4.5-kbp plus 2.6-kbp *TL2* flanking sequence) was not detected in the GT lanes. Instead, a band was observed at 5.5 kbp for GT3 and around 4.6 kbp for GT4, GT5, GT6 and GT7, indicating a different mechanism of *nptII* repair. This band was also observed with the *intFad2* and *terSac66* probes (Figure S2). The *Sac*I digestion results were confirmed by Southern blotting of *Nco*I-digested DNA hybridized with the *terSac66* probe (Figure 6c). The *TL* band of 2.0 kbp was found unchanged in all GT lanes, and an additional band was observed in GT3, GT5 and GT6 lanes. The presence of the *terSac66* on two different DNA fragments indicates that either one copy of a potentially homozygous *TL2* was modified, but not via the expected double crossover, or that the *TL2* was used as template by HR to repair an I-*Sce*I::GR-induced DSB in the *DL* (Figure 6a). The fact that PCR of the GT3 line could not detect a *DL* lacking the *nptII* repair fragment (data not shown) and that GT3 lacks a band specific to the original *DL* supports the idea that in GT3 at least, the *DL* has been modified.

To resolve this question, the GT events were backcrossed twice to the wild type and analysed by PCR (Figure S3). Kanamycin resistance was inherited as a single locus and was not correlated with the presence of the LB *TL* amplicon, which is specific to the *TL* and predicted to be present in a true GT event at the *TL*. Amplification of the LB *TL* and *TL I-SceI* amplicons in 46% of the GT4, 44% of the GT5 plants and about 17% of the GT6 events can thus be attributed to the presence of a segregating unmodified *TL* in these plants. Kanamycin resistance was strictly correlated with the presence of *DL* sequences on either side of the I-*Sce*I restriction site next to the defective *nptII* in the *DL*. However, a PCR fragment of the expected size across this I-*Sce*I site could not be amplified. This suggests either deletion around this I-*Sce*I site or the insertion of a sequence including the *TL terSac66* into this I-*Sce*I site. This latter hypothesis was confirmed by amplification and sequencing of the GT loci using primers located in the *intFad2* and in the *I-SceI::GR* gene. Analysis of the amplified sequence (Figure S4) showed an HR event restoring the *nptII* gene on the one side and a NHEJ or microhomology-mediated end-joining (MMEJ) event copying and linking a part of the *TL2* flanking sequence to the *I-SceI::GR* promoter on the other side. For the GT3 event, after the region of homology in the *nptII* gene, 909 bp of the *TL2* corresponding to the missing part of the defective *nptII* (including the *terSac66*) and 502 bp of the genomic flanking sequence of the *TL2* RB were linked by non-homologous recombination to the other side of the break, which had lost 52 bp of *DL* sequence. For GT5, the event is similar to the GT3 event, but only 82 bp of the *TL2* flanking sequence was copied and 9 bp of the break was deleted to repair this side by NHEJ including 115 bp of mitochondrial DNA in the junction. In the GT4 event, a microhomology of 4 bp is present at the junction, and 882 bp of the *TL2* comprising the missing part of *nptII* (including the *terSac66*) was copied into the repair sequence with 55 bp deleted from the *DL*. The lengths of these sequences

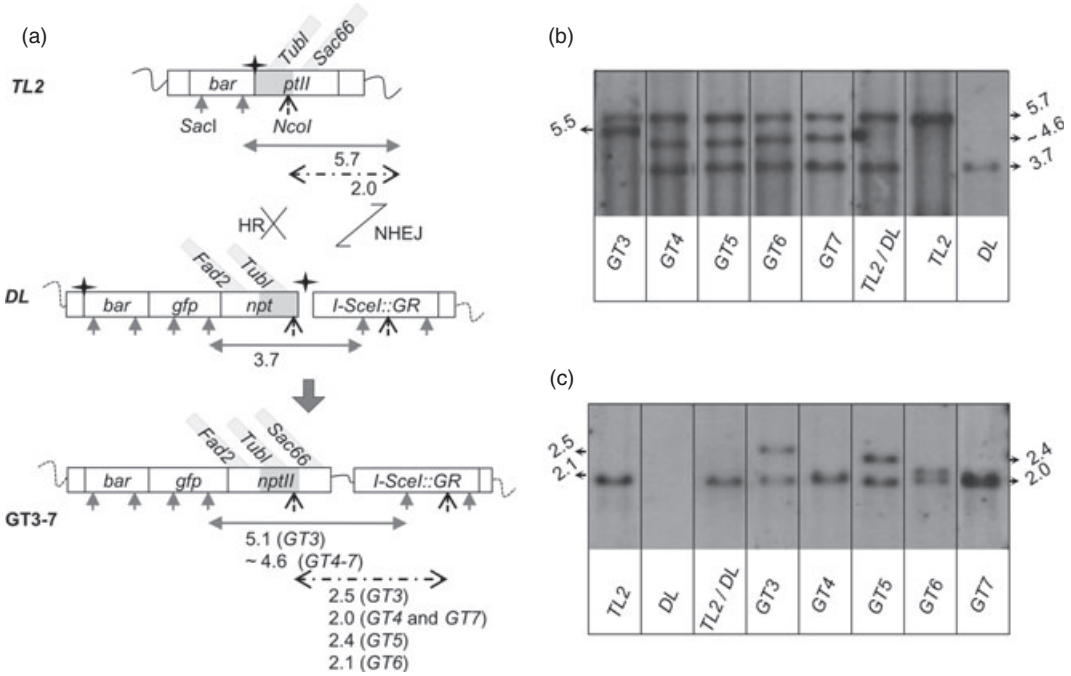


**Figure 6** Southern blot results of gene targeting (GT)3-7 events. (a) Schema of the GT events occurring in GT3-7. The black stars show I-*Sce*I restriction sites, and cleavage of the site in donor locus (*DL*) is indicated by the separation of the two halves of the locus. The grey arrows and black dashed arrows indicate respectively the *Sac*I and *Nco*I restriction site positions and the size of DNA fragments detected by the probes used. (b) Southern blot analyses of control parental plants [target locus (*TL*)*2*, *DL*, *TL2*/*DL*] and of kanamycin-resistant events derived from the *TL2*/*DL* (GT3-7) with *Sac*I-digested genomic DNA and the *intTubI* probe. A 3.7-kbp band, indicative of the non-excised *DL*, was detected in the GT4-7 events, but not in the GT3 event. The 5.7-kbp fragment indicative of native *TL2* was observed in the GT3–7 events. All putative GT events show an additional band around 5.5 kbp for GT3 and around 4.6 kbp for GT4–7 events (but not at the expected size of 7.1 kbp). (c) Southern blot analysis of the same events with *Nco*I-digested DNA blotted with a *terSac66*-specific probe. The signal at 2.0 kbp was observed for the *TL2* and *DL*/*TL2* controls. For the GT3, GT5 and GT6 events, unexpected bands appeared at respectively 2.5, 2.4 and 2.1 kbp. Fragment lengths are given as kbp.

of these GT events correspond to the observed sizes of bands seen on the Southern blots.

Analyses of the GT3-7 events thus showed that they are GT events, but involve the modification of the *DL* using the *TL* as template. Cleavage of the I-*Sce*I site of the *DL* adjacent to the *nptII* sequence, followed by recombination of the *nptII* side of the break with the homologous *TL2* as donor, creates a functional *nptII* at the *DL*. The other side of the break in the *DL* does not carry homology to the *TL2*, and thus must be repaired by NHEJ or MMEJ, resulting in variable lengths of *TL2* sequence integrated into the *DL* (Figure 6a).

## Discussion

The goal of this study was to develop a tool for precise remobilization of a transgene randomly inserted into the maize genome by its excision and insertion into a defined genomic site using homologous recombination. This strategy was tested by crossing of stably transformed *TL* and *DL* maize lines containing 3′ and 5′ overlapping regions of an *nptII* gene, respectively. Induction of I-*Sce*I activity in these lines with dexamethasone was used both to create a DSB at the *TL* and also to release the *nptII* repair DNA from the *DL*. HR of the liberated *nptII* repair DNA with the *TL* would then reconstitute the *nptII* gene and also mobilize a *gfp* gene into the *TL*. Kanamycin selection allows the selection of putative GT events.

Testing of 680 F2 progeny carrying the *TL* and *DL* did not permit the identification of any kanamycin-resistant plants, suggesting that germinal or early meristematic GT events are very rare under the conditions tested. However, in the course of testing these plants for kanamycin resistance, we noted the presence of green kanamycin-resistant sectors on the kanamycin-bleached leaves, suggesting the presence of somatic HR events between the *TL* and *DL*. DNA extracted from these green sectors, but not bleached leaf regions, could be used to amplify a restored functional *nptII* gene. Such green kanamycin-resistant sectors on bleached plants have previously been described in tobacco plants carrying an intrachromosomal HR reporter based on *nptII* reconstitution (Peterhans *et al.*, 1990) and also in *Arabidopsis* (Assaad and Signer, 1992). Other studies of GT based on *nptII* restoration and selection of resistant plants through the addition of kanamycin to the culture medium in tobacco (Puchta, 1999) and *Arabidopsis* (Vergunst *et al.*, 1998) did not, however, report green kanamycin-resistant sectors. In maize, we show here that application of kanamycin to the apex permits the detection and quantification of somatic GT events in leaves without affecting the survival of the sensitive plants. Multiple kanamycin treatments are possible and progeny can be obtained from treated plants. This assay, which should be applicable to other plant species, is currently being used to test and optimize GT frequencies.

In tobacco lines containing the equivalent of our *TL*, retransformed with a repair sequence and constitutive I-*Sce*I, the observed GT frequency increased proportionally with the expression level of I-*Sce*I (Puchta *et al.*, 1996). Similarly, in our maize plants, the frequency of green kanamycin-resistant sectors gives direct information about I-*Sce*I::GR activity. Given that we

observed *nptII* repair sequence excision from *DL* in equivalent proportions from maize plants grown in the absence or the presence of dexamethasone treatment, there is clearly basal activity of I-*Sce*I::GR in the maize leaves and dexamethasone does not further induce I-*Sce*I::GR in the tested conditions. In our previous study with I-*Sce*I::GR in *Arabidopsis*, basal activity was found, but the expression could be induced around 25- to 200-fold when dexamethasone was supplied in the growth medium (Wehrkamp-Richter *et al.*, 2009). We speculate that the dexamethasone applied to maize germinating seed does not penetrate into the seed in sufficient quantities to further induce I-*Sce*I::GR activity. Dexamethasone treatment does, however, induce I-*Sce*I::GR activity when added to the callus growth medium, where a 3.5 to 6.0-fold increase in the number of mutations at the *TL* was observed with dexamethasone (Table 2).

Notwithstanding the GT observed in somatic tissues, no kanamycin-resistant plants were found in the 680 tested F2 progeny of the *TL/DL* lines. We thus tested a strategy based on tissue culture selection and regeneration of kanamycin-resistant plants from *TL/DL* calli. This approach permitted the selection of seven independent GT events in two separate experiments involving a total of 2356 embryos (Table 1). Two of these, GT1 and GT2, were generated from embryos from *TL1/DL* plants; molecular and genetic analyses confirmed that they are true GT in which the *TL* has been modified by ectopic recombination using the *DL* as template on both sides. The overall frequency of obtaining true GT events at the *TL* from the two experiments was therefore 0.085% (29% of recombinants). The remaining five events (GT3-7) were generated from *TL2/DL* line embryos, and Southern blot and sequence analyses showed that they result from the modification of the *DL*, using the *TL* as template. The mechanism appears to be the creation of a DSB by I-*Sce*I::GR in the *DL* I-*Sce*I site next to the 5′ *nptII* region. Recombination of the *nptII* side of the break with the homologous *TL2* region as donor creates a functional *nptII* at the *DL*. However, the other side of the break does not carry homology to the *TL2*, and thus must be repaired by NHEJ or MMEJ (GT4), resulting in variable lengths of *TL2* sequence integrated into the *DL*. Such HR + NHEJ gene conversion events have been previously reported in plants (Puchta, 1999).

This surprising difference in the nature of the GT events identified in the calli from the two parent lines led us to resequence the *TL* and *DL* of these lines. This analysis identified a mutation which eliminates the right side I-*Sce*I cut site of the *DL* in the F1 *TL1/DL* plant (between *nptII* and *I-SceI::GR* – see Figure 7a). In the *TL1/DL* calli therefore, and in contrast to the *TL2/DL* calli, I-*Sce*I::GR can only cleave the *DL* once (to the left of the *bar* marker). Although the numbers of GT events analysed are low, this difference very probably explains the different types of GT events identified in calli from the two lines. In the *TL1/DL* calli, recombination initiated by I-*Sce*I cleavage of the *DL* would not generate a functional *nptII* gene and so only events initiated by cleavage in the *TL* would be selected. In the *TL2/DL* calli however, recombination initiation through cleavage adjacent to the *nptII* sequences in either the *TL* or the *DL* would result in the reconstruction of *nptII* (Figure 7b). In the *TL2/DL* calli, identification of recombination events in which only the *DL* was recipient clearly shows that single, incomplete I-*Sce*I cleavage of the *DL* is frequent in these cells.

These data thus show that only cleavage of the *TL* is needed for successful GT in these plants, as well as providing a clear illustration of the risk of including multiple I-*Sce*I restriction sites in plants in which I-*Sce*I expression or activity is limiting. The basal level of I-*Sce*I cleavage in the absence of dexamethasone induction further compounds this risk, through increasing levels of mutation in the I-*Sce*I sites of the *DL*. The dependence of this problem on limited I-*Sce*I activity would thus explain the difference with the recent study in *Arabidopsis* using a comparable strategy, in which only clean GT events were found (Fauser *et al.*, 2012). They observed efficient repair DNA excision, probably due to efficient activity of the I-*Sce*I, and GT was observed in up to 1% of the progeny. Limiting endonuclease activity is, however, a common problem in experiments of this type (Puchta *et al.*, 1996). In *Drosophila*, Gong *et al.*, (Gong and Golic, 2003) also reported low I-*Sce*I-mediated repair fragment excision and estimated that excision occurred in 7% of cells. They

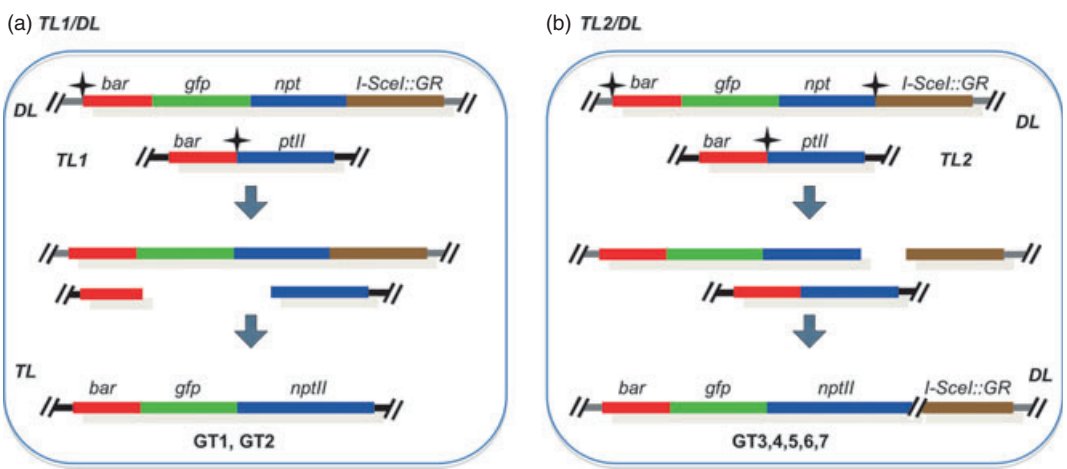


**Figure 7** Model. (a) For the target locus (*TL*)*1*/donor locus (*DL*)-derived embryos, I-*Sce*I cleavage in the *TL* initiated recombination with the *DL*, using homology in both *bar* and *nptII* regions. This resulted in gene targeting (GT) at the *TL*, using the *DL* as template. (b) For the *TL2/DL*-derived embryos, I-*Sce*I cleavage in the *DL* (right I-*Sce*I site in the *DL* between the *nptII* and *I-SceI::GR* cassette) initiated recombination with the *TL*, using homology in only the *npII* region. The other side of the break was repaired by non-homologous recombination. This resulted in GT at the *DL* using the *TL* as template. The I-*Sce*I sites are shown as black stars. The rightmost I-*Sce*I in the *DL* is mutated in the *TL1/DL* line, as indicated by the absence of the corresponding star.

thus used the FLP recombinase to excise a circular repair DNA from the genome that was subsequently linearized by I-*Sce*I. This system gave good GT rates in *Drosophila* but has not thus far been shown to work in maize (Yang *et al.*, 2009).

The goal of this work was to test in maize a GT system based on I-*Sce*I-mediated cleavage of the target, and excision of the *nptII* repair region from the genome. In previously described GT systems in maize, GT can occur only within a limited period after transformation of donor sequences and site-specific nucleases (D'Halluin *et al.*, 2008; Shukla *et al.*, 2009). In contrast, in our GT system, transformation of components required for GT is uncoupled from recombination and GT. This allows multiplication of the cells carrying the *TL*, the *DL* and I-*Sce*I-encoding sequences and permits the selection of rare ectopic recombination events. With this approach, a few stable transformation events can be used to generate a large population of cells from which to select GT events, of particular interest in cases where plant transformation frequencies are a limiting factor. In the work reported here, GT efficiencies range from 0.13% to 0.55% GT events per immature embryo. The actual number of cells screened is of course much higher than the number of calli, but the calculation with respect to calli expresses best the employed human effort. Our results furthermore show that cleavage of the *DL* is both unnecessary for targeted recombination, and a source of unwanted events when endonuclease cleavage is limiting.

## Experimental procedures

### Production of GT constructs and maize transgenic lines

The binary vectors for the creation of the target locus, pBIOS-*TL,* and donor locus, pBIOS-*DL,* were constructed in the following manner. First in order to extend the region of homology between the truncated *nptII* genes in the *TL* and *DL* lines, an 886-bp rice tubulin intron (GenBank, *AJ488063*) was introduced into the coding sequence of the *nptII* gene at position 204 bp downstream of the ATG. A 5′ truncated *nptII-intTubI* fragment lacking the first 150 bp of the *nptII* coding region was cloned between an I-*Sce*I site and in front of the *Arabidopsis Sac66* polyadenylation sequence (GenBank, *AJ002532*). The *I-SceI-3′nptII-intTubI-terSac66* fragment was then cloned into an SB11-based plant binary vector (Komari *et al.*, 1996) containing a rice actin promoter (*pAct*) (McElroy *et al.*, 1991) linked to the *bar* selectable marker gene (White *et al.*, 1990) and a nopaline synthase terminator (*terNos*), forming pBIOS-*TL*. To produce pBIOS-*DL,* a 3′ truncated *nptII-intTubI* fragment lacking the last 227 bp of the *nptII* coding region was cloned behind the constitutive *SC4* promoter (*pSC4*) (Schünmann *et al.*, 2003). A pSB11-based binary vector was created that contained the *pAct-bar-terNos* gene cassette and a cassava vein mosaic virus (*CsVMV*) promoter (Verdaguer *et al.*, 1998) linked to *gfp*, with both gene cassettes flanked by I-*Sce*I restriction sites. The *pSC4-5′nptII-intTubI* fragment was then cloned between the terminator of the *gfp* gene and the 3′ I-*Sce*I site to complete the *nptII* repair region. Next, the *NLS::I-SceI::GR* gene (Wehrkamp-Richter *et al.*, 2009), codon optimized for maize expression, was cloned between a *CsVMV* promoter and 35S cauliflower mosaic virus terminator. This cassette was then cloned between the *nptII* repair region 3′ I-*Sce*I site and the *RB* to form pBIOS-*DL*. *Agrobacterium tumefaciens* strain LBA 4404 (pSB1) (Hoekema *et al.*, 1983) was transformed with pBIOS-*TL* and pBIOS-*DL*. For each construction, a clone containing the recombinant plasmid was selected. Embryos of maize inbred A188 were transformed with each construction and transformed plants were regenerated according to Ishida *et al.* (Ishida *et al.*, 1996) using glufosinate selection.

### Plant analysis

Genomic DNA was extracted from the leaves by using the DNeasy 96 plant kit (Qiagen, Valencia, CA). Genomic DNA (10 μg) was digested, separated on 1% agarose gel by electrophoresis, transferred to nylon membrane and hybridized to $^{32}$P-marked probes following standard procedures (Sambrook and Russell, 2006). The genomic sequences flanking the transgenes were amplified using an adapter-anchor PCR method according to the method of Balzergue *et al.* (Balzergue *et al.*, 2001), with previously described modifications (Sallaud *et al.*, 2003), using DNA digested with *Ssp*I or *Pvu*II. Plant genotyping was performed by PCR. To amplify the fragments longer than 2.0 kbp, a Takara La Taq kit (Takara Bio, Shiga, Japan) was used. GFP fluorescence of sampled plant leaves was visualized under a fluorescence stereomicroscope (Leica MZ16F) using a GFP2 (Leica, Bannockburn, IL) filter.

### Crossing, culture and treatment of transformed plants

Plants were grown in the glasshouse with a 16-h day at 26 °C, 400 μE/m$^2$/s and an 8-h night at 18 °C. Dexamethasone treatments on seed were performed by immersion of the seed during 2 days in an aqueous solution of 30 μM dexamethasone. Kanamycin treatments were performed by the application of 50 μL of a solution at 200 mg/L kanamycin and 1% (v/v) Tween-20 on the apical region of 2-week-old plants.

### Somatic embryogenesis

Embryos isolated from selfed plants containing the *TL* and *DL* were placed onto medium according to Ishida *et al.* (Ishida *et al.*, 1996), lacking the kanamycin selective agent. For the first experiment, LS-AS medium was complemented by 0, 30 or 50 μM dexamethasone, and after 1 week, plantlets were transferred sequentially to LSD1.5, LSZ and 1/2LSF media lacking dexamethasone and containing 50 mg/L of kanamycin. For the second experiment, callus was developed for 3 days on LS-AS, 1 week on LSD1.5 and 3 weeks on LSZ medium. Then, callus was cultivated 1 week on LSZ medium containing 0, 30 or 50 μM of dexamethasone. GFP fluorescence of calli was visualized under the fluorescence stereomicroscope (Leica MZ16F) with a GFP2 filter.

### Callus analysis

PCR was performed directly on callus tissues using Terra direct PCR polymerase (Clontech Inc., Palo Alto, CA) in 20 μL with specific *TL* (forward: GTGGCGGACCGCTATCAG and reverse: ACATGTATTAAGAAGCAATGCATGTAGTAC) and *DL* (forward: TGGCAATCCCTTTCACAACC and reverse: CCCAGTCATAGCCGAATAGCC) primers. Genomic DNA was extracted from pooled calli, of the same genotype and dexamethasone treatment, with the DNeasy 96 plant kit (Qiagen). Primers were designed according to GS FLX Titanium emPCR LIBL kit (Roche Applied Science, Mannheim, Germany) with a specific TAG for each condition (forward: CCATCTCATCCCTGCGTGTCTCCGACTCAG-X-TCATCCCTACCCGTTCGTT and reverse: CCTATCCCCTGTGTGCCTTGGCAGTCTCAG-X-ATCACCCAGATCCACCCA, X represents the specific TAG of 10 bp for each condition). PCR was performed using Platinum Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA). Emulsion PCR was realized on the obtained PCR products with the emPCR Emulsion kit (Roche). PCR products were

sequenced with the emPCR sequencing kit (Roche) with Genome Sequencer FLX (Roche). Sequences from each condition were independently assembled and aligned to the reference sequence containing the non-mutated I-*Sce*I site. Sequence differences of 1 or 2 bp with the reference sequence found in the *TL* genotype, lacking *I-SceI::GR*, were discarded as these are likely to be sequencing errors. Sequence differences of three or more base pairs encompassing the I-*Sce*I site were identified and manually regrouped per condition to identify the number of independent mutations per condition.

## Acknowledgements

This work was supported by the EU FP6 project TAGIP; by the EU FP7 project RECBREED; and by an ANRT CIFRE grant to AA. We thank Friedrich Fauser and Holger Puchta for sending us their manuscript prior to publication. The authors are grateful for the support of the Biogemma Cloning, Transformation, Analysis, Greenhouse and Upstream Genomic teams and to Dr. O. Da Ines and Dr. S. Amiard for reading of the manuscript.

## References

Aoyama, T. and Chua, N.-H. (1997) A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.*, **11**, 605–612.

Assaad, F.F. and Signer, E.R. (1992) Somatic and germinal recombination of a direct repeat in *Arabidopsis*. *Genetics*, **132**, 553–566.

Balzergue, S., Dubreucq, B., Chauvin, S., Le-Clainche, I., Le Boulaire, F., de Rose, R., Samson, F., Biaudet, V., Lecharny, A., Cruaud, C., Weissenbach, J., Caboche, M. and Lepiniec, L. (2001) Improved PCR-walking for large-scale isolation of plant T-DNA borders. *Biotechniques*, **30**, 496–498.502,504.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. and Voytas, D.F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, **186**, 757–761.

Cotsaftis, O. and Guiderdoni, E. (2005) Enhancing gene targeting efficiency in higher plants: rice is on the move. *Transgenic Res.* **184**, 1–14.

D'Halluin, K., Vanderstraeten, C., Stals, E., Cornelissen, M. and Ruiter, R. (2008) Homologous recombination: a basis for targeted genome optimization in crop species such as maize. *Plant Biotechnol. J.*, **6**, 93–102.

Fauser, F., Roth, N., Pacher, M., Ilg, G., Sànchez-Fernàndez, R., Biesgen, C. and Puchta, H. (2012) In planta gene targeting. *Proc. Natl Acad. Sci. USA*, **109**, 7535–7540.

Gong, W.J. and Golic, K.G. (2003) Ends-out, or replacement, gene targeting in *Drosophila*. *Proc. Natl Acad. Sci. USA*, **100**, 2556–2561.

Grelon, M., Vezon, D., Gendrot, G. and Pelletier, G. (2001) AtSPO11-1 is necessary for efficient meiotic recombination in plants. *EMBO J.*, **20**, 589–600.

Hanin, M. and Paszkowski, J. (2003) Plant genome modification by homologous recombination. *Curr. Opin. Plant Biol.*, **6**, 157–162.

Hanin, M., Volrath, S., Bogucki, A., Briker, M., Ward, E. and Paszkowski, J. (2001) Gene targeting in *Arabidopsis*. *Plant J.*, **28**, 671–677.

Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J. and Schilperoort, R.A. (1983) A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, **303**, 179–180.

Huang, J., Zhou, W., Watson, A.M., Jan, Y.N. and Hong, Y. (2008) Efficient ends-out gene targeting in *Drosophila*. *Genetics*, **180**, 703–707.

Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T. and Kumashiro, T. (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. *Nat. Biotechnol.*, **14**, 745–750.

Jacquier, A. and Dujon, B. (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, **41**, 383–394.

Khanna, K.K., Lavin, M.F., Jackson, S.P. and Mulhern, T.D. (2001) ATM, a central controller of cellular responses to DNA damage. *Cell Death Differ.*, **8**, 1052–1065.

Komari, T., Hiei, Y., Saito, Y., Murai, N. and Kumashiro, T. (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. *Plant J.*, **10**, 165–174.

Krysan, P.J., Young, J.C. and Sussman, M.R. (1999) T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell*, **11**, 2283–2290.

Lu, C., Vasil, V. and Vasil, I.K. (1983) Improved efficiency of somatic embryogenesis and plant regeneration in tissue cultures of maize (*Zea mays* L.). *Theor. Appl. Genet.*, **66**, 285–289.

Matzke, A.J.M. and Matzke, M.A. (1998) Position effects and epigenetic silencing of plant transgenes. *Curr. Opin. Plant Biol.*, **1**, 142–148.

McElroy, D., Blowers, A.D., Jenes, B. and Wu, R. (1991) Construction of expression vectors based on the rice actin 1 (Act1) 5′ region for use in monocot transformation. *Mol. Gen. Genet.*, **231**, 150–160.

Paszkowski, J., Baur, M., Bogucki, A. and Potrykus, I. (1988) Gene targeting in plants. *EMBO J.*, **7**, 4021–4026.

De.Pater, S., Neuteboom, L.W., Pinas, J.E., Hooykaas, P.J.J. and Van Der Zaal, B.J. (2009) ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. *Plant Biotechnol. J.*, **7**, 821–835.

Peterhans, A., Schlupmann, H., Basse, C. and Paszkowski, J. (1990) Intrachromosomal recombination in plants. *EMBO J.*, **9**, 3437–3445.

Puchta, H. (1999) Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics*, **152**, 1173–1181.

Puchta, H., Dujon, B. and Hohn, B. (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res.*, **21**, 5034–5040.

Puchta, H., Dujon, B. and Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc. Natl Acad. Sci. USA*, **93**, 5055–5060.

Rong, Y.S. (2002) Gene targeting by homologous recombination: a powerful addition to the genetic arsenal for *Drosophila* geneticists. *Biochem. Biophys. Res. Commun.*, **297**, 1–5.

Sallaud, C., Meynard, D., van Boxtel, J., Gay, C., Bes, M., Brizard, J.P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., Ouwerkerk, P.B., Rueb, S., Delseny, M. and Guiderdoni, E. (2003) Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor. Appl. Genet.*, **106**, 1396–1408.

Sambrook, J. and Russell, D.W. (2006) *The condensed protocols from Molecular cloning: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schünmann, P.H.D., Llewellyn, D.J., Surin, B., Boevink, P., Feyter, R.C.D. and Waterhouse, P.M. (2003) A suite of novel promoters and terminators for plant biotechnology. *Funct. Plant Biol.*, **30**, 443–452.

Shukla, V.K., Doyon, Y., Miller, J.C., DeKelver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X., Choi, V.M., Rock, J.M., Wu, Y.Y., Katibah, G.E., Zhifang, G., McCaskill, D., Simpson, M.A., Blakeslee, B., Greenwalt, S.A., Butler, H.J., Hinkley, S.J., Zhang, L., Rebar, E.J., Gregory, P.D. and Urnov, F.D. (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*, **459**, 437–441.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and Stahl, F.W. (1983) The double-strand-break repair model for recombination. *Cell*, **33**, 25–35.

Tuteja, N., Ahmad, P., Panda, B.B. and Tuteja, R. (2009) Genotoxic stress in plants: shedding light on DNA damage, repair and DNA repair helicases. *Mutation Res.*, **681**, 134–149.

Tzfira, T. and White, C. (2005) Towards targeted mutagenesis and gene replacement in plants. *Trends Biotechnol.*, **23**, 567–569.

Tzfira, T., Li, J., Lacroix, B.Ì.t. and Citovsky, V. (2004) *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet.*, **20**, 375–383.

Tzfira, T., Weinthal, D., Marton, I., Zeevi, V., Zuker, A. and Vainstein, A. (2012) Genome modifications in plant cells by custom-made restriction enzymes. *Plant Biotechnol. J.*, **10**, 373–389.

Verdaguer, B., de Kochko, A., Fux, C.I., Beachy, R.N. and Fauquet, C. (1998) Functional organization of the cassava vein mosaic virus (CsVMV) promoter. *Plant Mol. Biol.*, **37**, 1055–1067.

Vergunst, A.C., Jansen, L.E. and Hooykaas, P.J. (1998) Site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. *Nucleic Acids Res.*, **26**, 2729–2734.

Wehrkamp-Richter, S., Degroote, F., Laffaire, J.B., Paul, W., Perez, P. and Picard, G. (2009) Characterisation of a new reporter system allowing high throughput in planta screening for recombination events before and after controlled DNA double strand break induction. *Plant Physiol. Biochem.*, **47**, 248–255.

White, J., Chang, S.Y. and Bibb, M.J. (1990) A cassette containing the bar gene of *Streptomyces hygroscopicus*: a selectable marker for plant transformation. *Nucleic Acids Res.*, **18**, 1062.

Yang, M., Djukanovic, V., Stagg, J., Lenderts, B., Bidney, D., Carl Falco, S. and Alexander Lyznik, L. (2009) Targeted mutagenesis in the progeny of maize transgenic plants. *Plant Mol. Biol.*, **70**, 669–679.

## Supporting information

Additional Supporting information may be found in the online version of this article:

**Figure S1** Genomic DNA sequences of flanking regions of *TL1*, *TL2* and *DL*.
**Figure S2** Additional Southern analysis of gene targeting events.
**Figure S3** Association between gene targeting events and amplified PCR fragments.
**Figure S4** Sequence of gene targeting events.
**Table S1** Sequences of PCR primers.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.