Cre-mediated autoexision of selectable marker genes in soybean, cotton, canola and maize transgenic plants

Xudong Ye (xudong.ye@bayer.com)
Bayer Crop Science United States: Bayer CropScience LP https://orcid.org/0000-0001-5104-6947
Zarir Vaghchhipawala
Monsanto Chesterfield Village Research Center: Monsanto Co
Edwards J. Williams
University of Wisconsin-Madison
Changlin Fu
Bayer Crop Science United States: Bayer CropScience LP
Jinyuan Liu
Bayer Crop Science United States: Bayer CropScience LP
Fengming Lu
Bayer Crop Science United States: Bayer CropScience LP
Erin L. Hall
Bayer Crop Science United States: Bayer CropScience LP
Shirley X. Guo
Bayer Crop Science United States: Bayer CropScience LP
LaRee Frank
Bayer Crop Science United States: Bayer CropScience LP
Larry A. Gilbertson
Bayer Crop Science United States: Bayer CropScience LP

Research Article

Keywords: Marker gene removal, Cre/loxP autoexcision, Agrobacterium-mediated transformation, marker-free transgenic plant, plant transformation, transgenic soybean, transgenic cotton, transgenic canola, transgenic maize

Posted Date: June 16th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1731438/v1

Abstract

Selectable marker genes are often required for efficient generation of transgenic plants in tissue culture transformation systems but are not desired once the transgenic events are obtained. We have developed Cre/*loxP* autoexcision systems to remove selectable marker genes in soybean, cotton, canola and maize. We tested a set of vectors with diverse promoters and identified promising promoters to drive *cre* expression for each of the four crops. We evaluated both the efficiency of generating primary transgenic events with low transgene copy numbers, and the frequency of marker-free progeny in the next generation. The best performing vectors gave no obvious decrease in the transformation frequency in each crop and generated homozygous marker-free progeny in the next generation. We found that effective expression of Cre recombinase for marker gene autoexcision can be species dependent. Among the vectors tested, the best autoexcision frequency (41%) in soybean transformation came from using the soybean *RSP1* promoter for *cre* expression. The *cre* gene expressed by soybean *RSP1* promoter with an *Arabidopsis AtpE* intron delivered the best autoexcision frequency (69%) in cotton transformation. The *cre* gene expressed by the embryo specific *eUSP88* promoter from *Vicia faba* conferred the best marker excision frequency (32%) in canola transformation. Finally, the *cre* gene expressed by the rice *CDC45-1* promoter resulted in 44% autoexcision in maize transformation. The Cre/*loxP* recombinase system enables the generation of selectable marker-free transgenic plants for commercial product development in four agriculturally important crops and provides further improvement opportunities for more specific and better marker excision efficiency.

Key Message

Efficient selectable marker gene autoexcision in transgenic plants of soybean, cotton, canola, and maize is achieved by effective Cre recombinase expression.

Introduction

Plant transformation has provided a powerful tool for functional genomic research and biotechnology applications by integration of transgenes into chromosomes of a wide range of agriculturally relevant species. Most transgenic plants have been produced through tissue culture processes, in which a selectable marker gene (SMG) is co-transformed with genes of interest (GOI). The SMG confers a growth advantage of transformed cells over non-transformed cells under the corresponding selection reagent in media, which then leads to transgenic shoot regeneration. Approximately 50 SMGs including many that confer antibiotic resistance or herbicide tolerance have been applied for transgenic plant production (Miki and McHugh, 2004). The antibiotic selection marker genes are not desirable in commercial seed products due to public perception or regulatory concerns over food safety and some herbicide marker gene may be redundant for trait stacking (Hare &Chua, 2002; Rosellini, 2012). The SMGs can be removed or segregated away in progeny seeds by using vector designs with two T-DNAs in *Agrobacterium* (Huang et al., 2004), two DNA segment bombardment co-transformation (Shiva Prakash et al., 2009), or site-specific recombinase-mediated excision of marker genes (Hare & Chua 2002; Gilbertson 2003; Miki and McHugh 2004; Yau & Stewart 2013).

Among various recombinase systems, the Cre/*loxP* recombination system has been widely reported to remove SMGs in many species (Gilbertson 2003; Yau & Stewart 2013). In the Cre/*loxP* recombinase system, the Cre recombinase catalyzes a crossover between directly repeated *lox* (locus of crossover) recognition sites. To excise a SMG, the *lox* sites are designed to flank the SMG and any other accessory genes, and after successful recombination-mediated excision, the GOI(s) and one *lox* site would remain (Gilbertson 2003, Gidoni et al., 2008). The Cre/*loxP*-mediated SMG excision can be obtained either by sexual crossing with a Cre-expressing line (Bayley et al., 1992; Pradhan et al., 2016), or by autoexcision where the *cre* gene is introduced and removed together with the SMGs. The Cre recombinase expression can be regulated using heat-, cold-, drought- or chemical inducible promoters (Zuo et al., 2001; Zhang et al. 2003; Wang et al., 2005; Ma et al., 2009, Petri et al., 2012; García-Almodóvar et al., 2014; Mookkan et al., 2017; Éva et al., 2018) or developmentally regulated promoters (Li et al., 2007; Verweire et al., 2007; Moravčíková et al., 2008; Kopertekh et al., 2009; Luo et al., 2007; Van Ex et al., 2007). Autoexcision in which the *cre* expression is

driven by developmentally-regulated promoters is the most efficient application as no additional treatment is required to activate *cre* expression during tissue culture and plant growth. In contrast, the use of heat-, cold- or chemical treatments for *cre* expression involves the induction of stress conditions during tissue culture in certain transformation systems, which may impede transgenic plant recovery during early stages of plant regeneration.

Most studies on marker gene autoexcision have been performed in model plants such as *Arabidopsis thaliana* and tobacco, as well as in crops such as maize, rice, soybean and canola (Table S1). Marker excision has been reported during the tissue culture stage from inducible promoters (Sreekala et al., 2005; Cuellar et al., 2006; Petri et al., 2012; Du et al., 2019; García-Almodóvar et al., 2014; Mookkan et al., 2017), embryo-specific promoters (Li et al., 2007; Chong-Pe´rez et al., 2013), or constitutive promoter (Zou et al., 2013); during meiosis (Bai et al., 2008; Kopertekh et al., 2010; Polóniová et al., 2015); or in progenies (Verweire et al., 2007; Van Ex et al., 2007).

Plant transformation methods and explant types differ from species to species. Meristem explants excised from mature seeds are used for soybean and cotton transformation using a non-lethal *aadA*/spectinomycin selection system (Ye et al., 2008; Martinell et al., 2011; Chen et al., 2014). The hypocotyls from germinating seedlings are used for canola transformation (Radke et al. 1992). The immature embryos are used for maize transformation (Sidorov & Duncan, 2009). Such diverse transformation systems and tissue types require different marker gene excision systems. We have tested several promoters with different expression profiles to develop efficient Cre/*loxP*-mediated marker gene autoexcision systems in soybean, cotton, canola and maize, with a goal of high frequency of homozygous marker-free (MF) R1 seed production because hemizygous MF seeds take additional generation to propagate. In this report we disclose results of commercial-level marker gene autoexcision frequencies in these four major crops.

Materials And Methods

Vector design and construction

To search for developmentally regulated expression elements to drive *cre* specific expression in reproductive tissues, a set of putative meiosis-related *Arabidopsis* genes were manually identified from GenBank, including *Arabidopsis CDC45*. The maize and rice *CDC45* homologue sequences were identified by performing BLAST searches in the GenBank genomic sequences against the *Arabidopsis* CDC45 protein sequence. There were two versions of maize and rice *CDC45* promoters identified from GenBank by searching *Arabidopsis* for CDC45 protein homologues. Two promoters from monocot *CDC45* genes, one from maize at chromosome 3 and one from rice at chromosome 11 (Table 1), appear to be restricted in reproductive tissue, and were used to drive the *cre* expression in pMON138232 and pMON243847, respectively. The relevant expression elements of these genes were cloned by PCR. The corresponding species genomic DNA was used as a template for PCR amplification using Q5® Hot Start High-Fidelity DNA polymerase (NEB Cat. No. M0493) according to the manufacturer instructions. The corresponding genomic regions of these expression elements in GenBank and the primers used for amplifying them are listed in Table 1.

Table 1
Primers and DNA sources used for PCR amplification of *cre* autoexcision promoters and 3' UTRs

Expression element	Forward primer	Reverse primer	Size (bp)	GenBank No. / region
P-At.CDC45	5' ctaatacaaaggtgcatgagtagtagtaactg 3'	5' ttccgtgaaattgaatcacccagaagg 3'	1030	**CP002686.1** (9143262..9144291)
T-At.CDC45	5' catagtctcattgttcttcgattcagtg 3'	5' cacgagcttcaggtcataactctgg 3'	734	**CP002686.1** (9146083..9146816)
P-Gm.RSP1	5' aaataatatataaaaatattacaaaaatc 3'	5' tgaagcaaagtggttagagatgagaatg 3'	720	**NC_016091** (44628717..44629436)
P-Zm.CDC45-1	5' agccacatgcagtgaattctatactcg 3'	5' tgcctcatcaatcagctaggtcggatc 3'	2000	**CM007649.1** (234291301..234293300)
P-Os.CDC45-1	5' acatacatctgtctagattcattaatat 3'	5' tggcgcatcaatcgaagtggtgaattgg 3'	1957	**AP014967.1** (1304410..1306330)

Insert Table 1 **here**

The soybean *RSP1* promoter (*P-Gm.RSP1*) was initially nominated as a disease responsive gene promoter (Resistance Sensitive Protein 1) based on RNAseq data generated in-house. Upon testing this promoter to express a *gusA* transgene in soybean, we found that it was active at background levels in multiple tissues, except roots and top of hypocotyls (data not shown). Even though it does not seem to fit the developmental regulation pattern of other promoters that we tested, it was included to avoid a decrease in TF because of its low-level expression in most tissues. The same promoter comes from a gene that belongs to the BURP domain-containing protein family and was reported to be expressed in roots and hypocotyls, and is inducible by ABA, salt, and drought treatments (Gm04.3 gene, Xu et al. 2010). pMON263552 with the *RSP1* promoter (Fig. 1) was constructed to test SMG autoexcision in soybean.

The 1.2 kp λ phage segment corresponding to GenBank accession No. J02459.1, region 21042 to 22237, was synthesized in Bio Basic Inc. (Markham, ON, Canada) and used in pMON243107 as a spacer sequence. The 754 bp of *Arabidopsis AtpE* intron corresponding to GenBank accession No. LR699765.1, region 17923873 to 19152863, was amplified by PCR and cloned into pMON291996 to test enhance *P-Gm.RSP1* expression. The 804 bp of maize *DnaK* intron, which was previously annotated as *ZmHSP70* intron and disclosed in GenBank Accession No. KX640115.1, was cloned after *P-Zm.CDC45-1* in pMON138232 to enhance expression.

The *cre* coding sequence used is as previously described and is interrupted by 189 bp IV2 intron from the potato *ST-LS1* gene (Vancanneyt et al. 1990; Zhang et al. 2003). The promoters described above were used to drive expression of the *cre* coding sequence. For 3'UTR, *T-At.CDC45* (Table 1) was used for soybean, cotton, and canola, and *Agrobacterium nos* transcription terminator (Depicker et al. 1982) was used for corn.

All dicotyledonous transformation vectors were built on *ori* pRi vector backbone with kanamycin resistance gene, and the maize transformation vectors were on RK2 *oriV* replicon with spectinomycin selection (Ye et al., 2011). The right and left borders sequences were described previously (Ye et al., 2008). The dicotyledonous *gusA* cassette driven by the CaMV 35S promoter and terminated by the *Agrobacterium nos* transcription terminator was described (Vancanneyt et al. 1990, Ye et al. 2008). The *gusA* cassette in maize vector pMON138232 was driven by the rice actin1 (*Os.Act1*) promoter with an additional 333 bp CaMV enhancer sequence in front of the *Os.Act1* first intron (McElroy et al. 1990). The *gusA* cassette in maize vector pMON243847 was driven by a 2181 bp rice tublin-3 (*TubA-3*) promoter (GenBank accession No. MH931401). In dicotyledonous transformation, the *aadA* gene with the chloroplast target sequence *ctp2* (Chen et al. 2014) driven by *Arabidopsis* actin 7 (*At.Act7*) promoter (GenBank accession No. JN400384) and terminated by the *Agrobacterium nos* transcription terminator was

used with spectinomycin for plant selection. In maize transformation, the *cp4 epsps* coding sequence with the *Os.Act1* promoter and the *Agrobacterium nos* terminator was used for glyphosate selection as described previously (Ye et al. 2011). In some vector designs, the *splA* (sucrose phosphorylase-like gene, GenBank accession No. AE007871, region 153761..155218) gene derived from *Agrobacterium* Ti plasmid driven by enhanced USP88 (eUSP88) promoter (Bäumlein et al., 1991; Wang et al. 2006) and terminated by the *nos* transcription terminator in three dicotyledonous constructs was included to reduce R1 seed screening due to seed abortion phenotype, similar to the approach taken for 2 T-DNA transformation (Fig S1). Some vector designs were simplified to omit this negative selection. The SMG and *cre* genes were flanked by *lox* sites for autoexcision. The genetic elements and the T-DNA structure of all binary vectors are depicted in Fig. 1.

Standard cloning procedures were applied for all binary vector construction (Sambrook et al. 1989). For seamless fusion between a promoter and *cre* elements, the hot fusion cloning protocol was used with PCR products bearing 20–25 bp element junction overlaps (Fu et al. 2014).

Insert Fig. 1**here**

Agrobacterium **preparation and plant transformation**

A single binary vector was transfected into a nopaline type of *Agrobacterium tumefaciens* strain by electroporation as described previously (Ye et al. 2008). The ABI strain containing gentamicin and kanamycin resistance was used for maize vector transfection using spectinomycin for *Agrobacterium* selection. The AB30 strain (Ye et al. 2016), which is derived from ABI with deletion of kanamycin resistance gene, was used for soybean and canola binary vector transfection which contains kanamycin-resistant gene in the vector backbone for *Agrobacterium* selection. The AB33 strain, derived from AB30 with $VirG^{I77V}$ mutation (Ye et al. 2016) was used for cotton binary vector transfection with kanamycin for *Agrobacterium* selection (Chen et al. 2014).

For soybean (*Glycine max*) transformation, the dry meristem explants from the cultivar A3555 were mechanically excised (Calabotta et al. 2013). The explants were imbibed for 30 min in inoculation buffer, inoculated with *Agrobacterium* AB30 suspension containing corresponding binary vectors at OD_{600} = 0.3 and sonicated for 20 seconds (Ye et al. 2008). The explant co-culture, plant regeneration and growth in green house were described previously (Martinell et al. 2002; Ye et al. 2008), except that 150 mg/L spectinomycin instead of glyphosate was used for selection during shoot elongation.

For cotton transformation, the dry meristem explants from cotton cultivar DP393 seeds were excised mechanically (Dersch et al. 2015). The explants were imbibed in inoculation buffer for 30 min, inoculated with AB33 strain containing binary vectors, and co-cultured for 3–5 days. Plant regeneration was obtained with 150 mg/L spectinomycin selection, which was described in detail previously (Chen et al. 2014).

Canola hypocotyls explants from canola (*Brassica napus* L.) cultivar Ebony were used for canola transformation (Radke et al. 1992; Ye et al. 2011). Spectinomycin at 100 mg/L instead of glyphosate in the regeneration media was used to recover transgenic canola shoots.

The immature embryos of maize elite cultivar LH244 were used for generating maize transgenic plants with glyphosate selection as previous described (Sidorov & Duncan, 2009).

Molecular analyses

R0 regenerants were analyzed for transgene copy number and vector backbone presence or absence by TaqMan® technology (Applied Biosystems). Leaf samples were collected for DNA extraction (Dellaporta et al. 1983). For dicotyledonous transgenic plant analysis, the *gusA* gene as a GOI, the *aadA*, as well as *cre* were analyzed for copy number. For maize transgenic plants, the *gusA*, *cp4 epsps* and *cre* were analyzed for copy number. The T-DNA left border (LB) was also detected in all constructs for T-DNA intactness. The TaqMan® detection probes of the *gusA*, *cp4 epsps*, LB, and the backbone oriRi in dicotyledonous vectors or RK2 oriV in maize vectors were described previously (Ye et al. 2011).

The primers 5'- AGCTAAGCGCGAACTGCAAT-3' (forward) and 5'- GGCTCGAAGATACCTGCAAGA-3' (reverse) amplifying the *aadA* gene in the dicotyledonous binary vectors, and further detected by minor grove binding (MGB) TaqMan® probe 6FAM-TGGAGAATGGCAGCGCAATGACA, were used for the dicotyledonous selectable marker gene copy number assay. The primers 5'-CAAGTGACAGCAATGCTGTTTCA-3' (forward) and 5'-GTCGAAATCAGTGCGTTCGAA-3' (reverse) amplifying a *cre* fragment, and the TaqMan® probe 6FAM-CGGTGAACGTGCAAAA were used for *cre* cassette presence.

R1 plants are defined as the progeny produced from self-pollinating the R0 plant, i.e. the primary transformant derived from tissue culture. For R1 progeny screening, leaf samples from the green house grown plantlets were collected for DNA extraction. The GOI (*gusA*), marker gene (*aadA* for dicotyledonous, *cp4 epsps* for maize) and *cre* gene were assayed for copy number with TaqMan® analysis. The GOI TaqMan® detection positive, but marker and *cre* TaqMan® detection negative plants were counted as MF lines, and a subset of these marker free lines were partially verified by Southern blot with DIG-labeled probes (Ye et al. 2011; Chen et al. 2014). In general, for a population of 100 R1 plants, we project a total 75 R1 plants that are positive for the GOI, either as hemizygotes or homozygotes, and 25 null plants, assuming Mendelian segregation of a single locus (1:3 transmission; 1 null : 2 hemizygous : 1 homozygous transgene segregation). The R1 MF frequency is calculated as percentage of the projected transgenic R1 plants. If all 75 of these hemizygous and homozygous R1 plants are negative for the SMG, the calculated marker gene excision frequency would be 100%.

Results

1. Soybean marker gene autoexcision

P **-** At.CDC45 **-** cre **showed severe leaky expression in leaves during plant production in soybean transformation**

The *Arabidopsis CDC45* promoter was the first promoter we tested for marker gene autoexcision in soybean transformation. However, plants with the *P-At.CDC45-cre* expression in pMON131703 showed a severe mottled leaf phenotype during spectinomycin selection, which suggested that there was premature SMG excision from leaky expression of *cre* in vegetative tissues (Fig. S2, B). The mottled leaf phenotype could be recovered in soil after removing spectinomycin selection (Fig. S2, C). A decrease in the transformation frequency (TF) was observed in pMON131703 compared to the control plasmid pMON131702 without *cre* cassette (average 2.3% vs. 5% in 3 parallel comparison experiments, Table S2).

Eighty R0 events from pMON131703 with single copy insertion of the GOI were advanced for R1 seed setting. Only one MF R0 event was found to produce *aadA* marker negative and *gusA* positive plants by molecular analysis of R1 seeds.

To reduce potential *cis* element impact from the adjacent *P-CaMV-gusA* cassette, we inserted a 1.2 kb λ-phage fragment between the *P-CaMV 35S-gusA* and the *aadA* SMG cassettes in pMON131703 (Fig. 1), which resulted in pMON243107. In total 159 phenotypically normal shoots and 80 mottled shoots were produced with combined TF of 2.36%, which is comparable to pMON131703. Twenty-two single copy, backbone free events were selected for R1 marker analysis. Only 6 out of the 22 analyzed R0 soybean events produced MF progeny (Table 2).

Table 2
R1 MF progeny from soybean plants transformed with pMON243107 (*P-At.CDC45-cre*)

Soybean R0 line	R1 plants	Total *gusA*+, MF R1 plants	Homozygous *gusA*+ MF R1 plants
GM_At.CDC45-1	44	23	3
GM_At.CDC45-2	35	26	5
GM_At.CDC45-3	45	6	2
GM_At.CDC45-4	46	10	5
GM_At.CDC45-5	45	7	3
GM_At.CDC45-6	39	17	6

Note: MF: marker free. GM_At.CDC45: *Glycine max* plant with *cre* driven by *Arabidopsis CDC45* promoter. If 100% marker gene excision is present in the R1 population, we expect 34.5 *gusA*+ and MF plants from a total of 46 R1 plants (i.e. single locus transmission, 1 null: 2 hemizygous : 1 homozygous), and 11.5 homozygous *gusA*+ MF plants out of the total 46 R1 plants.

Insert Table 2**here**

P-Gm.RSP1 **-** cre **showed efficient marker gene excision in soybean transformation**

The *cre* expression driven by *P-Gm.RSP1* in pMON263552 (Fig. 1) showed no apparent reduction in TF and no abnormal phenotypes in soybean transformation, suggesting a lack of significant premature excision with this promoter. R1 seeds from 49 R0 lines that had germline transmission of a single copy transgene were analyzed for marker segregation by TaqMan® assay. As shown in Table 3, all 49 events produced *aadA* negative, *gusA* positive R1 plants, indicating that efficient SMG excision occurred in this construct. In addition, 30 of the 49 events generated three or more MF, homozygous seeds out of 46 planted seeds. Overall, the frequency of MF R1 transgenic plants is 41% [686/(2226x75%)] out of the total projected *gusA* positive plants (2226x75%), with the same homozygous MF frequency 41% [228/(2226x25%)] out of the total projected homozygous *gusA* positive plants (2226x25%).

Table 3
R1 MF progeny from soybean plants transformed with pMON263552 (*P-Gm.RSP1-cre*)

Soybean R0 line	Total R1 analyzed	Total *gusA* + MF R1 plants	Homozygous *gusA* + MF R1 plants
GM_RSP-1	46	6	1
GM_RSP-2	41	7	0
GM_RSP-3	46	7	1
GM_RSP-4	39	14	3
GM_RSP-5	46	12	2
GM_RSP-6	43	12	7
GM_RSP-7	46	21	4
GM_RSP-8	42	14	4
GM_RSP-9	46	4	1
GM_RSP-10	43	12	4
GM_RSP-11	46	18	6
GM_RSP-12	45	18	10
GM_RSP-13	46	6	1
GM_RSP-14	40	9	3
GM_RSP-15	46	26	11
GM_RSP-16	41	24	8
GM_RSP-17	46	5	4
GM_RSP-18	44	21	11
GM_RSP-19	46	17	2
GM_RSP-20	45	15	11
GM_RSP-21	45	22	5
GM_RSP-22	43	12	2
GM_RSP-23	46	7	3
GM_RSP-24	45	13	3
GM_RSP-25	46	5	0
GM_RSP-26	44	Epidermal	Epidermal
GM_RSP-27	46	15	5
GM_RSP-28	44	9	2
GM_RSP-29	44	10	1

Note: GM_RSP: *Glycine max* plant with *cre* driven by soybean *RSP1* promoter. Epidermal-all transgene detection negative in all R1 plants, suggesting that only epidermal cells were transformed in R0. If 100% marker gene excision is present in the R1 population, out of a total 46 R1 plants we expect 34.5 MF *gusA* + plants and 11.5 homozygous *gusA* + MF plants.

Soybean R0 line	Total R1 analyzed	Total *gusA* + MF R1 plants	Homozygous *gusA* + MF R1 plants
GM_RSP-30	44	16	2
GM_RSP-31	46	12	2
GM_RSP-32	44	22	4
GM_RSP-33	46	22	4
GM_RSP-34	44	10	5
GM_RSP-35	45	12	2
GM_RSP-36	42	9	2
GM_RSP-37	46	17	10
GM_RSP-38	45	12	1
GM_RSP-39	46	19	6
GM_RSP-40	43	10	3
GM_RSP-41	46	15	9
GM_RSP-42	41	17	7
GM_RSP-43	46	11	4
GM_RSP-44	43	29	15
GM_RSP-45	46	8	0
GM_RSP-46	45	26	15
GM_RSP-47	45	13	6
GM_RSP-48	45	19	13
GM_RSP-49	46	14	1
GM_RSP-50	46	12	2
Total	2226	686	228

Note: GM_RSP: *Glycine max* plant with *cre* driven by soybean *RSP1* promoter. Epidermal-all transgene detection negative in all R1 plants, suggesting that only epidermal cells were transformed in R0. If 100% marker gene excision is present in the R1 population, out of a total 46 R1 plants we expect 34.5 MF *gusA* + plants and 11.5 homozygous *gusA* + MF plants.

Insert Table 3**here**

The MF progenies in the R1 generation were also further confirmed by Southern blot analysis. Twenty-nine R1 plants from 24 R0 single copy lines were randomly selected for DNA extraction and tested with both *aadA* and *gusA probes*. As shown in Fig. 2, the selected R1 events were further confirmed to be MF by Southern blot. Lanes 5a, 5b, and 21 showed 2 bands, which indicated an error call in R0 TaqMan® copy number assay. We determined that the R0 copy number assay is approximately 90% accuracy as revealed by R1 MF copy number assay, which is more accurate to distinguish hemi- or homozygous transgenic plants in R1 plants. The faint or no signal bands of *gusA* probe were largely due to uneven DNA loads and were confirmed by extended film exposure. These results further confirm that TaqMan® analyzed R1 plants were *aadA* marker negative and *gusA* positive.

Insert Fig. 2**here**

A construct with Arabidopsis AtpE **intron at** P-Gm.RSP1 **3' showed reduced marker gene excision efficiency in soybean transformation**

Considering the efficient marker gene autoexcision with *P-Gm.RSP1-cre* in pMON263552 (Table 3), we hypothesized that adding an intron at 3' end of the promoter may further increase marker gene excision frequency. The *Arabidopsis* intron, *I-At.AtpE*, was selected to enhance the *P-Gm.RSP1* expression, and an *P-eUSP88-splA* expression cassette, which is a seed lethal cassette, was placed between the *gusA* and *cre* cassettes to reduce R1 analysis (Fig. S1, D), which resulted in pMON291996.

No obvious TF decrease or any abnormal phenotype was observed in the transgenic shoots from pMON291996 transformation compared to regular constructs. Seeds from 15 R0 single copy events were planted for MF excision to be confirmed by TaqMan® analysis for *gusA* and *aadA* transgene probes. In pMON291996, we observed a drastically reduced marker gene excision efficiency compared to pMON263552 (Table 3). Only 10 out of 15 R0 events produced marker free progeny, with few hemizygous MF *gusA* positive R1 plants, and only two of those events produced one homozygous R1 plant (Table 4). Overall, the frequency of MF *gusA* positive R1 plants is 7.6% [37/(651x75%)] out of total 488 projected *gusA* + plants (homozygous and hemizygous) (651x75%), with a poor homozygous MF frequency of 1.2% [2/(651x25%)] out of the total 162 projected homozygous *gusA* + plants (651x25%).

Table 4
R1 MF progeny from soybean plants transformed with pMON291996 (P-Gm.RSP1 + I- At.AtpE-cre**)**

Soybean R0 line	Analyzed R1 plants	*gusA* + MF R1 plants	Homozygous *gusA* + MF R1 plants
GM_RSP-Int-1	42	4	0
GM_RSP-Int-2	42	0	0
GM_RSP-Int-3	45	11	1
GM_RSP-Int-4	44	1	0
GM_RSP-Int-5	44	1	0
GM_RSP-Int-6	44	1	0
GM_RSP-Int-7	45	0	0
GM_RSP-Int-8	40	0	0
GM_RSP-Int-9	45	6	1
GM_RSP-Int-10	44	4	0
GM_RSP-Int-11	40	1	0
GM_RSP-Int-12	43	0	0
GM_RSP-Int-13	45	0	0
GM_RSP-Int-14	40	4	0
GM_RSP-Int-15	39	4	0
Total	651	37	2

Note: GM_RSP-Int: *Glycine max* plant with *cre* driven by soybean *RSP1* promoter and *Arabidopsis AtpE* intron.

If 100% marker gene excision presents in R1 population, out of a total 45 R1 plants we expect 33.75 MF *gusA* + plants and 11.25 homozygous *gusA* + MF plants.

Insert Table 4**here**

2. Cotton marker gene autoexcision

The Arabidopsis CDC45 **promoter enabled efficient marker gene excision in cotton but decreased TF**

Because of promising marker gene autoexcision result in the initial soybean transformation with *P-At.CDC45-cre* (Table 2), we tested pMON131703 (*P-At.CDC45-cre*) with the control binary vector pMON131702 (without the *cre* cassette) side-by-side in cotton transformation for autoexcision efficiency and TF impact. The autoexcision construct pMON131703 also showed lower transformation frequencies and had mottled leaf phenotypes in some cotton events (Fig. S2, E). On average a TF of 1.25% was observed compared to 4.2% with the control construct in four separate experiments (Fig. S3). Approximately 23% of R0 events are chimeric or epidermal transformation, and 45% (41/91) single copy R0 lines were following the Mendelian segregations (Table S3).

Ninety-one R0 events were harvested with seeds. In total, 58 out of 91 R0 events of one or two copy *gusA* transgene inserts produced MF R1 progeny, among which 41 R0 events produced 55–95% MF *gusA* positive R1 plants of the projected total *gusA* positive R1 plants (assuming 75% of total seeds for single locus).

The MF autoexcision from pMON131703 in cotton was further verified by Southern blot. Eight R1 events negative for marker and four R1 events positive for marker identified by the TaqMan® assay were selected. The total genomic DNA was digested with *HindIII*, and hybridized with DIG-labeled *aadA* (marker) or *gusA* (GOI) probe.

As shown in Fig. 3, the Southern blot confirmed that all eight events detected to be marker negative by TaqMan® were negative by Southern analysis, and the four control events detected to be marker positive by Taqman® marker positive were also *aadA* positive by Southern blot, suggesting that the Taqman® assay is accurate and consistent for transgene presence. Overall, the results from pMON131703 in cotton transformation indicate that *Arabidopsis CDC45* promoter driving *cre* expression in cotton was suitable for efficient marker gene removal.

Insert Fig. 3**here**

We tested more constructs using this promoter in cotton with different expression cassette configurations to mitigate the TF reduction, including pMON243107 (with 1.2 kb l phage spacer sequence and marker gene between *gusA* and *cre* cassettes) and pMON244545 (no spacer sequence) with the same *Arabidopsis CDC45* promoter (Fig. 1). Both pMON243107 and pMON244545 still showed reduced transformation frequencies (1.45% and 0.89%, respectively) compared to the control construct pMON131702 or other constructs without the *cre* cassette (3–5%) (data not shown). Abnormal leaf phenotype was observed in half of the regenerating shoots in the two constructs (Fig. S2, E), suggesting that leaky *cre* expression can be causing the premature marker excision. All shoots were recovered to normal growth phenotype in soil and set seeds as we used *aadA*/spectinomycin no-lethal selection system.

Fifteen R0 single copy event progeny seeds for either construct were planted to test for marker autoexcision. Leaf samples from 40 R1 plantlets per R0 event were assayed for the *gusA* and *aadA* copy number. Efficient marker gene autoexcision was observed in both constructs with the best performance of 50–100% projected *gusA* + transgenic plants (i.e. 75% of total seeds) showing marker gene absence and high numbers of homozygous MF plants in the R1 generation among the majority of germline transmission events (Table 5). On average, 68% total projected *gusA* + transgenic R1 cotton seedlings [578/(1122 x75%)] are GOI positive, MF, and 67.7% [190/(1122 x25%)] of the projected R1 *gusA* + homozygous seeds were confirmed as homozygous MF in R1 population.

Table 5
R1 MF progeny from cotton plants transformed with pMON243107 (with spacer) and pMON244545 (no spacer) containing *P-At.CDC45-cre* cassette

Construct	Cotton R0 event	R1 total	*gusA* + MF R1 plants	Homozygous *gusA* + MF R1 plants
pMON243107	GH_At.CDC45-1	40	Epidermal	Epidermal
pMON243107	GH_At.CDC45-2	40	0	0
pMON243107	GH_At.CDC45-3	40	4	1
pMON243107	GH_At.CDC45-4	36	Epidermal	Epidermal
pMON243107	GH_At.CDC45-5	40	24	6
pMON243107	GH_At.CDC45-6	40	30	8
pMON243107	GH_At.CDC45-7	40	24	7
pMON243107	GH_At.CDC45-8	40	28	9
pMON243107	GH_At.CDC45-9	40	23	5
pMON243107	GH_At.CDC45-10	40	19	6
pMON243107	GH_At.CDC45-11	40	29	8
pMON243107	GH_At.CDC45-12	40	26	7
pMON243107	GH_At.CDC45-13	40	25	7
pMON243107	GH_At.CDC45-14	40	29	8
pMON243107	GH_At.CDC45-15	40	19	4
pMON244545	GH_At.CDC45-16	40	22	7
pMON244545	GH_At.CDC45-17	40	35	14
pMON244545	GH_At.CDC45-18	40	22	8
pMON244545	GH_At.CDC45-19	38	25	9
pMON244545	GH_At.CDC45-20	40	30	8
pMON244545	GH_At.CDC45-21	40	24	10
pMON244545	GH_At.CDC45-22	35	33	17
pMON244545	GH_At.CDC45-23	34	23	5
pMON244545	GH_At.CDC45-24	35	15	3
pMON244545	GH_At.CDC45-25	37	28	8
pMON244545	GH_At.CDC45-26	27	22	18
pMON244545	GH_At.CDC45-27	40	14	5
pMON244545	GH_At.CDC45-28	40	5	2
pMON244545	GH_At.CDC45-29	40	0	0
Total		1122	578	190

Note: GH_At.CDC45: *Gossypium herbaceum* plant with *cre* driven by *Arabidopsis CDC45 promoter*. Epidermal-no R1 germline transgene transmission. If 100% marker gene excision presents in R1 population, we expect total 30 MF R1 plants and 10 homozygous MF R1 plants from total 40 R1 plants.

Insert Table 5**here**

pMON291996 with P-Gm.RSP1 + I-At.AtpE-cre **confers highly efficient marker gene excision in cotton transformation**

We tested pMON291996 which included *P-Gm.RSP1 + I-At.AtpE-cre* autoexcision cassette and a *P-eUSP88-splA* seed abortion expression cassette (Fig. S1) to reduce R1 analysis. While we observed that the reduced soybean marker gene autoexcision in this construct compared to pMON263552, this construct showed the highest efficiency of marker gene excision for cotton that we have tested in this study without obvious negative TF compromise [TF = 5.65%, which was comparable to non-autoexcision control constructs (data not shown)]. On average, 69% of the total projected R1 *gusA* + transgenic lines [392/(750 x75%)] are GOI positive, MF, and 53% of the projected R1 *gusA* + homozygous transgenic plants [100/(750 x25%)] were confirmed to be homozygous MF (Table 6).

Table 6
R1 MF progeny from cotton plants transformed with pMON291996 (*P-Gm.RSP1 + I-At.AtpE-cre*)

Cotton R0 event	**R1 plant sample #**	***gusA* + MF R1 plants**	**Homozygous *gusA* + MF R1 plants**
GH_RSP-Int-1	50	29	2
GH_RSP-Int-2	50	20	7
GH_RSP-Int-3	50	0	0
GH_RSP-Int-4	50	30	8
GH_RSP-Int-5	50	20	11
GH_RSP-Int-6	50	Epidermal	Epidermal
GH_RSP-Int-7	50	33	6
GH_RSP-Int-8	50	34	6
GH_RSP-Int-9	50	28	9
GH_RSP-Int-10	50	29	6
GH_RSP-Int-11	50	33	5
GH_RSP-Int-12	50	32	11
GH_RSP-Int-13	50	37	8
GH_RSP-Int-14	50	31	12
GH_RSP-Int-15	50	36	9
Total	750	392	100

Note: GH_RSP-Int: *Gossypium herbaceum* plant with *cre* driven by soybean *RSP1* promoter and *Arabidopsis AtpE* intron. If 100% marker gene excision presents in R1 population, we expect total 37.5 MF R1 plants (75%), and 12.5 homozygous MF R1 plants (25%) from total 50 R1 plants.

Insert Table 6**here**

3. Canola marker gene autoexcision

Four *cre* autoexcision cassettes with *P-At.CDC45*, *P-Gm.RSP1*, *P-Br.nap* (canola napin gene), or *P-Vf.eUSP88* promoter (Fig. 1) were tested in canola transformation for SMG autoexcision based on our promising results from soybean for the *CDC45* and *RSP1* promoters as well as the embryo specific autoexcision in canola (Kopertekh et. al., 2009). The construct with *P-At.CDC45* promoter expressing the *cre* cassette showed a decreased TF and speckling phenotype, while the transformation frequencies for

the other three constructs were comparable to our regular construct transformation (data not shown). Eight to 12 single copy R0 events and 88 R1 per event were evaluated for MF progeny for each of the four constructs.

Seven out of nine R0 events from the construct containing *At.CDC45* promoter produced MF R1 plants. However, on average only 4.8% total projected *gusA*+ seeds (29/(792x75%) were confirmed to be MF by TaqMan® assay (Table 7).

Table 7
R1 MF progeny from canola plants transformed with pMON243107 (*P-At.CDC45-Cre*)

R0 event	R1 total	Total *gusA*+ MF R1 plants	Homozygous *gusA*+ MF R1 plants
BN_At.CDC45-1	88	7	2
BN_At.CDC45-2	88	3	0
BN_At.CDC45-3	88	11	6
BN_At.CDC45-4	88	0	0
BN_At.CDC45-5	88	2	3
BN_At.CDC45-6	88	4	0
BN_At.CDC45-7	88	1	0
BN_At.CDC45-8	88	0	0
BN_At.CDC45-9	88	1	0
Total	792	29	11

Note: BN_At.CDC45: *Brassica napus* plant with *Arabidopsis CDC45* promoter driven *cre*. If 100% marker gene excision presents in R1 population, we expect total 66 MF R1 plants (75%) and 22 homozygous MF plants (25%) from total 88 plantlets

Insert Table 7**here**

Only one out of 12 R0 events from the construct with soybean *RSP1* promoter [pMON417179 (*P-Gm.RSP-1-cre*)] generated MF R1 progeny after analyzing 88 plantlets from each R0 event (Table S4). No MF R1 plant was obtained from the construct containing napin promoter [pMON263567 (P-*Br.nap-cre*-T-*Br.nap*)] after analyzing R1 progenies from nine single copy R0 events (Table S5).

The most effective marker gene excision was recovered from pMON420845 with *eUSP88* promoter. Seeds from eight single copy R0 events were planted and sampled for R1 MF progeny screening with GOI and *aadA* TaqMan® probes. All eight R0 events produced MF R1 progeny with seven R0 events producing homozygous MF plants. On average 32.7% of the projected *gusA*+ transgenic R1 [173/(704x75%)] are verified to be truly MF by TaqMan® assay, among which 37.5% of the projected homozygous *gusA*+ transgenic R1 progenies [66/(704x25%)] were confirmed to be MF by TaqMan® assay (Table 8).

Table 8
R1 MF progeny from canola plants transformed with pMON420845 (*P-Vf.eUSP88-cre*)

R0 event	R1 total	Total *gusA*+ MF R1 plants	Homozygous *gusA*+ MF R1 plants
BN_eUSP88-1	88	12	4
BN_eUSP88-2	88	24	12
BN_eUSP88-3	88	1	0
BN_eUSP88-4	88	33	13
BN_eUSP88-5	88	42	15
BN_eUSP88-6	88	5	4
BN_eUSP88-7	88	32	6
BN_eUSP88-8	88	24	12
Total	704	173	66

Note: BN_eUSP88: *Brassica napus* plant with *Vicia faba* enhanced *USP88* promoter driven *cre*. If 100% marker gene excision presents in R1 population, we expect total 66 MF R1 plants (75%) and 22 homozygous MF plants (25%) from total 88 plantlets

Insert Table 8**here**

4. Marker-gene autoexcision in maize transformation

The initial promising SMG autoexcision in soybean transformation with the *Arabidopsis CDC45* promoter directly inspired us to search for monocot homologous promoters for maize SMG autoexcision. These vectors showed no obvious impact on maize TF when they were used for marker gene autoexcision in maize. However, other *CDC45* promoters tested, including one at maize chromosome 10 (GenBank accession NC_050105.1, region 2024417..2026424) and one at rice chromosome 12 (GenBank accession NC_029267.1, region 1187144..1189143), showed a TF drop when used for marker gene autoexcision in maize transformation, suggesting leaky expression in vegetative tissue (data not shown).

The rice and maize *CDC45-1* promoters showed similar marker gene autoexcision efficiency in maize LH244 immature embryo transformation. All 9 R0 events from pMON243847 with rice *CDC45-1* promoter showed on average 44% marker removal from total projected *gusA*+ transgenic R1 [113/(339x75%)], and 40% [34/(339x25%)] marker removal in the projected homozygous *gusA*+ transgenic R1 plants (Table 9).

Table 9
R1 MF progeny from maize plants transformed with pMON243847 (*P-Os.CDC45-Cre*)

Maize R0 event	R1 total	Total *gusA*+ MF R1 plant	Homozygous *gusA*+ MF R1 plants
ZM_Os.CDC45-1	41	14	6
ZM_Os.CDC45-2	18	9	1
ZM_Os.CDC45-3	41	13	2
ZM_Os.CDC45-4	34	19	7
ZM_Os.CDC45-5	41	12	5
ZM_Os.CDC45-6	41	14	4
ZM_Os.CDC45-7	41	10	1
ZM_Os.CDC45-8	41	6	0
ZM_Os.CDC45-9	41	16	8
Total	339	113	34

Note: ZM_Os.CDC45: *Zea mays* plant with rice *CDC45* promoter driven *cre*. If 100% marker gene excision presents in R1 population, we expect total 30.75 MF R1 plants (75%) and 10.25 homozygous MF plants (25%) from total 41 plantlets (i.e. single locus transmission, 1 null: 2 homozygous: 1 hemizygous).

Insert Table 9**here**

Similarly, R1 seeds from 10 single copy R0 events from pMON138232 with maize *CDC45-1* promoter were analyzed for MF segregation (Table 10). All 10 R0 events produced MF progeny at 41% [101/(328x75%)] frequency of the total projected *gusA*+ transgenic plants, and 19.5% [16/(328x25%)] of the projected homozygous plants.

Table 10
R1 MF progeny from maize plants transformed with pMON138232 (*P-Zm.CDC45-Cre*)

Maize R0 event	R1 total	Total *gusA*+ MF R1 plant	Homozygous *gusA*+ MF R1 plants
ZM_Zm.CDC45-1	41	9	2
ZM_Zm.CDC45-2	40	8	0
ZM_Zm.CDC45-3	39	10	2
*ZM_Zm.CDC45-4	40	14	3
ZM_Zm.CDC45-5	17	6	0
ZM_Zm.CDC45-6	17	7	1
ZM_Zm.CDC45-7	22	8	3
ZM_Zm.CDC45-8	41	13	3
ZM_Zm.CDC45-9	41	12	0
ZM_Zm.CDC45-10	30	14	2
Total	328	101	16

Note: ZM_Zm.CDC45: *Zea mays* plant with maize *CDC45* promoter driven *cre*. *ZM_ Zm.CDC45-4 has 45 green plants and 13 albinos dying later.

Insert Table 10**here**

The MF nature in pMON138232 R1 plants was further confirmed by Southern blot. Twenty TaqMan® screened MF R1 plants from Table 10 were sampled for DNA extraction and analyzed by Southern blot. The same Southern blot membrane was hybridized with DIG labeled *gusA* probe, then stripped and re-hybridized with DIG labeled *cre* probe. As shown in Fig. 4, all 20 selected lines did not hybridize with the *cre* probe, whereas all showed a clear signal with the *gusA* probe, which is consistent with the TaqMan® screening results.

Insert Fig. 4**here**

Discussion

Many efficient transformation systems require tissue culture and plant regeneration under antibiotic or herbicide selection. The Cre*/loxP* recombination system, which allows for transgenic plant selection under tissue culture conditions followed by selectable marker removal, has been widely reported in many species. The SMG autoexcision to generate MF plants can be obtained by using heat-, cold-, drought- or chemically inducible promoters. An ideal promoter to drive the *cre* recombinase for marker gene autoexcision should have no leaky expression in vegetative tissues during the tissue culture steps to allow transgenic shoot development under selection. We envisioned that reproductive tissue specific (floral-, meiosis-, microspore-, pollen- or egg-specific) promoters may be the most suitable for this purpose. Since transgenic plants with hemizygous alleles must take one more generation for homozygous line production, floral or meiosis promoters working in both male and female germline cells are highly desirable for marker gene autoexcision. Other promoters, such as soybean *RSP1* with background level expression in vegetative tissues but significant expression in germline cells, may be a good option as well if the TF is not impacted due to premature excision.

Premature marker gene excision and transformation system

The *Arabidopsis CDC45* promoter was found to be promising for SMG autoexcision in our soybean transformation system. The *CDC45* gene is reported to be required for initiation of DNA replication and mainly upregulated at the G1/S transition and in young meiotic flower buds (Stevens et al., 2004). The promoter has enabled us to obtain MF transgenic plants in soybean, cotton, and canola, albeit with decreased TF and mottle leaf phenotype. It is also surprising to see the dicotyledonous *CDC45* promoter is active in maize (Table S6) which produced some MF lines without a decrease in the TF, suggesting that a common expression motif may be present in the promoter for both dicotyledonous and monocot expression.

We observed a decrease in the TF and mottled phenotype in leaf tissues with the *At.CDC45* promoter driving *cre* expression in our dicotyledonous transformation system. One explanation is that the expression of *cre* with the *At.CDC45* promoter in vegetative tissues could have resulted in premature marker gene removal during plant regeneration under selection.

However, we were still able to recover some of the events with the mottled leaf phenotype. We use an *aadA/spectinomycin* non-lethal selection system for dicotyledonous transformation (Martinell et al., 2011; Chen et al., 2014). Spectinomycin binds to the 16S rRNA, which blocks translation on the prokaryotic type 70S plastid ribosomes and usually induces albino leaves in dicotyledonous transformation (Svab et al., 1990). Premature marker removal during tissue culture in this non-lethal selection system still allows a portion of the chimeric transgenic/non-transgenic tissue to survive for shoot regeneration. The MF areas on leaves are bleached by spectinomycin and produce a mottled phenotype. The transgenic plants with such phenotypes can be fully recovered in soil as chlorophyll biosynthesis resumes after the spectinomycin selection is removed. Therefore, the success of MF seed production with the leaky *CDC45* promoter is unique to our non-lethal selection system, which may not be replicated in other lethal selection transformation systems.

Several developmentally regulated promoters have been used in floral dip transformation system (Van Ex et al., 2007; Verweire et al., 2007). In such a non-tissue culture system, vegetative leaky expression is not a concern as the T1 transformants directly come from the egg cells during flowering. Leaky expression can be an advantage to remove the SMG in T1 during seed germination. The *Arabidopsis CLV3* has been reported to generate 100% marker removal in T2 plants (Van Ex et al., 2007), which is not feasible for transformation system based on tissue culture, because the *CLV3* gene is expressed mainly in shoot

apical meristems, which will lead to marker excision during shoot regeneration under selection. Indeed, when we used the *Arabidopsis* meristem-specific *Erecta* promoter (Yokoyama et al., 1998) for autoexcision in soybean, the TF dropped greatly (data not shown).

The system using BBM/WUS2 for plant transformation is unique, in which plant regeneration depends on BBM/WUS2 expression in initial tissue culture materials after embryo induction followed by removal of both embryogenic genes. Both developmentally regulated and inducible promoters have been used to excise the BBM/WUS2 expression to enable plant regeneration and the excision frequency was counted as R0 MF events (Wang et al. 2020). Such a system may not be applied on other tissue culture-based transformation systems as a selectable marker gene is required for transgenic tissue proliferation.

Autoexcision activity of a cre **promoter may be different among species and germplasms**

A specific *cre* promoter often shows very different autoexcision pattern in different species, dependent on the transformation and selection methods. In addition to the *At.CDC4*5 promoter mentioned above, the best canola autoexcision promoter from *Vicia faba, eUSP88*, produces much less marker gene autoexcision in soybean. The *eUSP88* promoter appeared to be highly specific in embryos in soybean and canola, and no expression in callus (Bäumlein et al., 1991; Wang et al. 2006). In our canola experiments, hypocotyl materials were used for transformation and plant regeneration requires callus formation. When canola embryonic materials were used for transformation, we noticed a large reduction with the same autoexcision construct, suggesting that the *eUSP88* drives expression in embryo tissue which caused premature marker gene excision. Previously, an embryo-specific *app1* promoter from *Arabidopsis pei1* gene was reported to drive marker gene autoexcision in soybean embryogenic culture transformation by particle bombardment and resulted in 30% R0 events with MF progeny (Li et al., 2007). The *eUSP88* promoter expression is similar to the *app1* promoter in the early heart stage to the late cotyledon stage of embryo development (Bäumlein et al., 1991).

The *Gm.RSP1* promoter showed the best marker autoexcision frequency among the tested promoters in soybean transformation. However, it performed poorly in canola, suggesting that the expression is limited to the species or the transformation system. The construct with the *Gm.RSP1* promoter and *At.AtpE* intron has very different excision frequency in soybean and cotton. It gave poor autoexcision in soybean (Table 4) but showed the best autoexcision frequency in cotton transformation (Table 6). It is plausible that the *cre* cassette with the *Arabidopsis* intron is poorly expressed in soybean by not properly splicing.

A napin promoter has been reported to enable efficient marker gene excision in *Brassica napus* (Kopertekh et al., 2009). We tested a seed-specific napin promoter (Fig. S1, A, B) in canola for marker autoexcision and did not observe any marker gene excision in R1 progeny. Sequence BLAST analysis revealed different napin promoters in GenBank are highly similar in the last 300 bp (data not shown). The marker excision difference in the two napin promoters may be due to different germplasms being used in the experiments or different expression patterns between the two similar embryo-specific napin promoters.

Chimera marker gene excision may occur in R0 and but not in progeny

Non-specific *cre* expression in vegetative tissues and non-germline cell expression may be the major reason for chimeric marker excision in R0 plants, which is often associated with a decrease in TF and potential phenotypes such as mottle leaves in our soybean and cotton transformation systems with non-lethal *aadA* selectable marker.

In addition to reported strong meiotic cell expression, the *At.CDC45* promoter appears to be leaky in leaf tissue in all three dicotyledonous species we tested, which formed chimera leaves in R0 events (Fig. S2 B, E). However, the R1 seeds are derived from single reproductive cells, in which the marker is excised either in vegetative tissues or during flowering due to the strong activity at meiosis (Stevens et al., 2004), and for this reason, no chimera excision has been observed in R1 MF plants.

Cre driven by embryo-specific promoters are more likely to form chimeric R0 events. Li et al. (2007) reported that the 13% complete excision and 31% chimeric excision were observed in R0 soybean plants when the *app1* promoter was used for autoexcision. Moravcíková et al. (2008) reported that the *cre* driven by *Arabidopsis* cruciferin C promoter regenerated chimeric

T0 plant, only 10.2% T1 plant showed complete marker excision, and the excision rate was increased by repeated *cre* activation in T2 plants, which may indicate that the cruciferin promoter is expressed in late embryo stage and not a vegetative or germline preferred promoter. We did not observe chimeric marker excision in canola plants from *eUSP88-cre* autoexcision after two generation observation, indicating earlier embryo expression present in this promoter as confirmed in Fig S1, D, which caused embryo complete abortion when *splA* is expressed.

Conclusion Remark

The Cre/*loxP* system is an effective tool for removal of SMGs in transgenic plants. Efficient marker gene autoexcision by a developmentally regulated promoter is often species dependent. We tested a diverse set of promoters in multiple agriculturally important crop species, and identified promising promoters for soybean, cotton, canola and maize marker gene autoexcision. Testing of specific promoters for good marker excision efficiency remains an option for many plant species.

Declarations

Author contribution statement

XY nominated promoters, constructed plasmids, collected data, coordinated work, and drafted the manuscript. ZV nominated promoters, constructed plasmids, did all Southern blot analyses. EJW performed portion of soybean transformation. FC constructed plasmids. LJ conducted maize progeny molecular analyses. FL performed canola transformation. ELH and SXG organized canola experiments. LF designed TaqMan® assays. LG introduced Cre/*loxP* technology to us and critically revised the manuscript. All authors reviewed and approved the manuscript.

Acknowledgment

We sincerely thank numerous former Monsanto colleagues in Middleton, Wisconsin; Davis, California; Mystic, Connecticut and St. Louis, Missouri for transgenic plant production, greenhouse care and molecular analyses. Special thanks to Drs. David Somers and Doug Boyes for supporting this research and Drs. Jenn To, Miguel Vega-Sanchez, Bertho Lieselot and Julie Francois for critical review of the manuscript.

Conflict of interest

The authors are employees of Bayer Crop Science, a manufacturer of seeds produced by conventional and biotechnology methods. A relevant US patent application has been submitted and assigned to Bayer Crop Science.

References

1. Bai X, Wang Q, Chu C (2008) Excision of a selective marker in transgenic rice using a novel Cre/*loxP* system controlled by a floral specific promoter. Transgenic Res 17:1035–1043. https://doi:10.1007/s11248-008-9182-7
2. Bäumlein H, Boerjan W, Nagy I, Bassüner R, Van Montagu M, Inzé D, Wobus U (1991) A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and *Arabidopsis* plants. Mol Gen Genet 225:459–467. https://doi:10.1007/BF00261688
3. Bayley CC, Morgan M, Dale EC, Ow DW (1992) Exchange of gene activity in transgenic plants catalyzed by the Cre-*lox* site-specific recombination system. Plant Mol Biol 18(2):353–361. https:// doi:10.1007/BF00034962
4. Calabotta BJ, Deppermann KL, Dersch E, Heise JD, Koestel AR, Ludwig CL, Martinell BJ, Williams E (2013) Preparation and use of plant embryo explants for transformation. US Pat. US8362317B2
5. Chen Y, Rivlin A, Lange A, Ye X, Vaghchhipawala Z, Eisinger E, Dersch E, Paris M, Martinell B, Wan Y (2014) High throughput *Agrobacterium tumefaciens*-mediated germline transformation of mechanically isolated meristem explants of cotton (*Gossypium hirsutum L.*). Plant Cell Rep 33(1):153–164. https:// doi:10.1007/s00299-013-1519-x
6. Chong-Pérez B, Reyes M, Rojas L, Ocaña B, Ramos A, Kosky RG, Angenon G (2013) Excision of a selectable marker gene in transgenic banana using a Cre/*lox* system controlled by an embryo specific promoter. Plant Mol Biol 83(1–2):143–152.

https:// doi:10.1007/s11103-013-0058-8

7. Cuellar W, Gaudin A, Solórzano D, Casas A, Nopo L, Chudalayandi P, Medrano G, Kreuze J, Ghislain M (2006) Self-excision of the antibiotic resistance gene *nptII* using a heat inducible Cre-*loxP* system from transgenic potato. Plant Mol Biol 62(1–2):71–82. https://doi:10.1007/s11103-006-9004-3
8. Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1:19–21. https://doi.org/10.1007/BF02712670
9. Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM (1982) Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet 1:561–573
10. Dersch E, Heinzen RJ, Martinell BJ, Rivlin A, Wan Y, Ye X (2015) Method of meristem excision and transformation. US Pat. No. US8937216B2
11. Du D, Jin R, Guo J, Zhang F (2019) Construction of marker-free genetically modified maize using a heat-inducible autoexcision vector. Genes 10:374. https://doi:10.3390/genes10050374
12. Éva C, Téglás F, Zelenyánszki H, Tamás C, Juhász A, Mészáros K, Tamás L (2018) Cold inducible promoter driven Cre-*lox* system proved to be highly efficient for marker gene excision in transgenic barley. J Biotechnol 265:15–24. https://doi:10.1016/j.jbiotec.2017.10.016
13. Fu C, Donovan WP, Shikapwashya-Hasser O, Ye X, Cole RH (2014) Hot Fusion: an efficient method to clone multiple DNA fragments as well as inverted repeats without ligase. PLoS One 9(12):e115318. https://doi.org/10.1371/journal.pone.0115318
14. García-Almodóvar RC, Petri C, Padilla IMG, Burgos L (2014) Combination of site-specific recombination and a conditional selective marker gene allows for the production of marker-free tobacco plants. Plant Cell Tiss Organ Cult 116:205–215. https://doi.org/10.1007/s11240-013-0396-x
15. Gidoni D, Srivastava V, Carmi N (2008) Site-specific excisional recombination strategies for elimination of undesirable transgenes from crop plants. In Vitro Cell Dev Biol Plant 44:457–467. https://doi.org/10.1007/s11627-008-9140-3
16. Gilbertson L (2003) Cre-*lox* recombination: Cre-ative tools for plant biotechnology. Trends Biotechnol 21:550–555. https://. doi
17. Hare P, Chua NH (2002) Excision of selectable marker genes from transgenic plants. Nat Biotechnol 20(6):575–580. https:// doi:10.1038/nbt0602-575
18. Huang S, Gilbertson LA, Adams TH, Malloy KP, Reisenbigler EK, Birr DH, Snyder MW, Zhang Q, Luethy MH (2004) Generation of marker-free transgenic maize by regular two-border *Agrobacterium* transformation vectors. Transgenic Res 13:451–461. https://doi:10.1007/s11248-004-1453-3
19. Kopertekh L, Broer I, Schiemann J (2009) Developmentally regulated site-specific marker gene excision in transgenic *B. napus* plants. Plant Cell Rep 28:1075–1083. https://doi:10.1007/s00299-009-0711-5
20. Kopertekh L, Schulze K, Frolov A, Strack D, Broer I, Schiemann J (2010) Cre-mediated seed-specific transgene excision in tobacco. Plant Mol Biol 72:597–605. https://doi:10.1007/s11103-009-9595-6
21. Li Z, Xing A, Moon BP, Burgoyne SA, Guida AD, Liang H, Lee C, Caster CS, Barton JE, Klein TM, Falco SC (2007) A Cre/*loxP*-mediated self-activating gene excision system to produce marker gene free transgenic soybean plants. Plant Mol Biol 65:329–341. https://doi:10.1007/s11103-007-9223-2
22. Luo K, Duan H, Zhao D, Zheng X, Deng W, Chen Y, Stewart CN Jr, McAvoy R, Jiang X, Wu Y, He A, Pei Y, Li Y (2007) 'GM-gene-deletor': fused *loxP-FRT* recognition sequences dramatically improve the efficiency of FLP or CRE recombinase on transgene excision from pollen and seed of tobacco plants. Plant Biotechnol J 5:263–274. https://doi:10.1111/j.1467-7652.2006.00237.x
23. Ma BG, Duan XY, Niu JX, Ma C, Hao QN, Zhang LX, Zhang HP (2009) Expression of stilbene synthase gene in transgenic tomato using salicylic acid-inducible Cre/*loxP* recombination system with self-excision of selectable marker. Biotechnol Lett 31:163–169. https://doi:10.1007/s10529-008-9843-x

24. Martinell BJ, Julson LS, Emler CA, Huang Y, McCabe DE, Williams EJ (2022) Soybean agrobacterium transformation method. US Patent No. US6384301B1
25. Martinell BJ, Petersen MW, Somers DA, Wan Y, Williams E, Ye X (2011) Methods for plant transformation using spectinomycin selection. US Patent No. US8030544B2
26. McElroy D, Zhang W, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. Plant Cell 2:163–171. https://doi:10.1105/tpc.2.2.163
27. Miki B, McHugh S (2004) Selectable marker genes in transgenic plants: applications, alternatives and biosafety. J Biotech 107:193–232. https://doi:10.1016/j.jbiotec.2003.10.011
28. Mookkan M, Nelson-Vasilchik K, Hague J, Zhang ZJ, Kausch AP (2017) Selectable marker independent transformation of recalcitrant maize inbred B73 and sorghum P898012 mediated by morphogenic regulators BABY BOOM and WUSCHEL2. Plant Cell Rep 36:1477–1491. https://doi:10.1007/s00299-017-2169-1
29. Moravcíková J, Vaculková E, Bauer M, Libantová J (2008) Feasibility of the seed specific cruciferin C promoter in the self excision Cre/*loxP* strategy focused on generation of marker-free transgenic plants. Theor Appl Genet 117:1325–1334. https://doi:10.1007/s00122-008-0866-4
30. Petri C, López-Noguera S, Wang H, García-Almodóvar C, Alburquerque N, Burgos L (2012) A chemical-inducible Cre-*LoxP* system allows for elimination of selection marker genes in transgenic apricot. Plant Cell Tiss Organ Cult 110:337–346. https://doi.org/10.1007/s11240-012-0155-4
31. Polóniová Z, Jopčík M, Matušíková I, Libantová J, Moravčíková J (2015) The pollen- and embryo-specific *Arabidopsis DLL* promoter bears good potential for application in marker-free Cre/*loxP* self-excision strategy. Plant Cell Rep 34:469–481. https://doi:10.1007/s00299-014-1726-0
32. Pradhan S, Chakraborty A, Sikdar N, Chakraborty S, Bhattacharyya J, Mitra J, Manna A, Dutta Gupta S, Sen SK (2016) Marker-free transgenic rice expressing the vegetative insecticidal protein (Vip) of *Bacillus thuringiensis* shows broad insecticidal properties. Planta 244:789–804. https://doi:10.1007/s00425-016-2535-1
33. Radke SE, Turner JC, Facciotti D (1992) Transformation and regeneration of *Brassica rapa* using *Agrobacterium tumefaciens*. Plant Cell Rep 11:499–505. https://doi:10.1007/BF00236265
34. Rosellini D (2012) Selectable markers and reporter genes: a well furnished toolbox for plant science and genetic engineering. Crit Rev Plant Sci 31:401–453. https://doi.org/10.1080/07352689.2012.683373
35. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
36. Sidorov V, Duncan D (2009) *Agrobacterium*-mediated maize transformation: immature embryos versus callus. Methods Mol Biol 526:47–58. https://doi:10.1007/978-1-59745-494-0_4
37. Shiva Prakash N, Bhojaraja R, Shivbachan SK, Hari Priya GG, Nagraj TK, Prasad V, Srikanth Babu V, Jayaprakash TL, Dasgupta S, Spencer TM, Boddupalli RS (2009) Marker-free transgenic corn plant production through co-bombardment. Plant Cell Rep 28:1655–1668. https://doi:10.1007/s00299-009-0765-4
38. Sreekala C, Wu L, Gu K, Wang D, Tian D, Yin Z (2005) Excision of a selectable marker in transgenic rice (*Oryza sativa L.*) using a chemically regulated Cre/*loxP* system. Plant Cell Rep 24:86–94. https://doi:10.1007/s00299-004-0909-5
39. Stevens R, Grelon M, Vezon D, Oh J, Meyer P, Perennes C, Domenichini S, Bergounioux C (2004) A *CDC45* homolog in *Arabidopsis* is essential for meiosis, as shown by RNA interference-induced gene silencing. Plant Cell 16:99–113. https://doi:10.1105/tpc.016865
40. Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. Proc Natl Acad Sci USA 87:8526–8530. https://doi:10.1073/pnas.87.21.8526
41. Van Ex F, Verweire D, Claeys M, Depicker A, Angenon G (2009) Evaluation of seven promoters to achieve germline directed Cre-*lox* recombination in *Arabidopsis thaliana*. Plant Cell Rep 28:1509–1520. https://doi:10.1007/s00299-009-0750-y
42. Vancanneyt G, Schmidt R, O'Connor-Sanchez A, Willmitzer L, Rocha-Sosa M (1990) Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in *Agrobacterium*-mediated

plant transformation. Mol Genet Gen 220:245–250. https://doi:10.1007/BF00260489

43. Verweire D, Verleyen K, De Buck S, Claeys M, Angenon G (2007) Marker-free transgenic plants through genetically programmed autoexcision. Plant Physiol 145:1220–1231. https://doi:10.1104/pp.107.106526
44. Wang N, Arling M, Hoerster G, Ryan L, Wu E, Lowe K, Gordon-Kamm W, Jones TJ, Chilcoat ND, Anand A (2020) An Efficient Gene Excision System in Maize. Front Plant Sci 11:1298. https://doi:10.3389/fpls.2020.01298
45. Wang Y, Chen B, Hu Y, Li J, Lin Z (2005) Inducible excision of selectable marker gene from transgenic plants by the cre/*lox* site-specific recombination system. Transgenic Res 14:605–614. https:// doi:10.1007/s11248-005-0884-9
46. Wang Q, Fagaly T, Bassuner R, Liang J, Oulmassov TN, Dabrowski J (2006) Seed specific USP promoters for expressing genes in plants. US patent No. US7078588B2.
47. Xu H, Li Y, Yan Y, Wang K, Yai Y, Hu Y (2010) Genome-scale identification of Soybean BURP domain-containing genes and their expression under stress treatments. BMC Plant Biol 10:197. https://doi:10.1186/1471-2229-10-197
48. Yau Y, Stewart CN (2013) Less is more: strategies to remove marker genes from transgenic plants. BMC Biotechnol 13:36. https://doi:10.1186/1472-6750-13-36
49. Ye X, Williams EJ, Shen J, Esser JA, Nichols AM, Petersen MW, Gilbertson LA (2008) Plant development inhibitory genes in binary vector backbone improve quality event efficiency in soybean transformation. Transgenic Res 17(5):827–838. https:// doi:10.1007/s11248-008-9169-4
50. Ye X, Williams EJ, Shen J, Johnson S, Lowe B, Radke S, Strickland S, Esser JA, Petersen MW, Gilbertson LA (2011) Enhanced production of single copy backbone-free transgenic plants in multiple crop species using binary vectors with a pRi replication origin in *Agrobacterium tumefaciens*. Transgenic Res 20:773–786. https://doi:10.1007/s11248-010-9458-6
51. Ye X, Chen Y, Wan Y, Hong YJ, Ruebelt MC, Gilbertson LA (2016) Constitutive expression of the *tzs* gene from *Agrobacterium tumefaciens virG* mutant strains is responsible for improved transgenic plant regeneration in cotton meristem transformation. Plant Cell Rep 35:601–611. https://doi:10.1007/s00299-015-1906-6
52. Yokoyama R, Takahashi T, Kato A, Torii KU, Komeda Y (1998) The *Arabidopsis ERECTA* gene is expressed in the shoot apical meristem and organ primordia. Plant J 15(3):301–310. https://doi:10.1046/j.1365-313x
53. Zhang W, Subbarao S, Addae P, Shen A, Armstrong C, Peschke V, Gilbertson L (2003) Cre/*lox*-mediated marker gene excision in transgenic maize (*Zea mays* L.) plants. Theor Appl Genet 107:1157–1168. https://doi:10.1007/s00122-003-1368-z
54. Zuo J, Niu QW, Møller SG, Chua NH (2001) Chemical-regulated, site-specific DNA excision in transgenic plants. Nat Biotechnol 19:157–161. https:// doi:10.1038/84428
55. Zou X, Peng A, Xu L, Liu X, Lei T, Yao L, He Y, Chen S (2013) Efficient autoexcision of a selectable marker gene from transgenic citrus by combining the Cre/*loxP* system and *ipt* selection. Plant Cell Rep 32:1601–1613. https://doi:10.1007/s00299-013-1470-x

Figures

Fig. 1

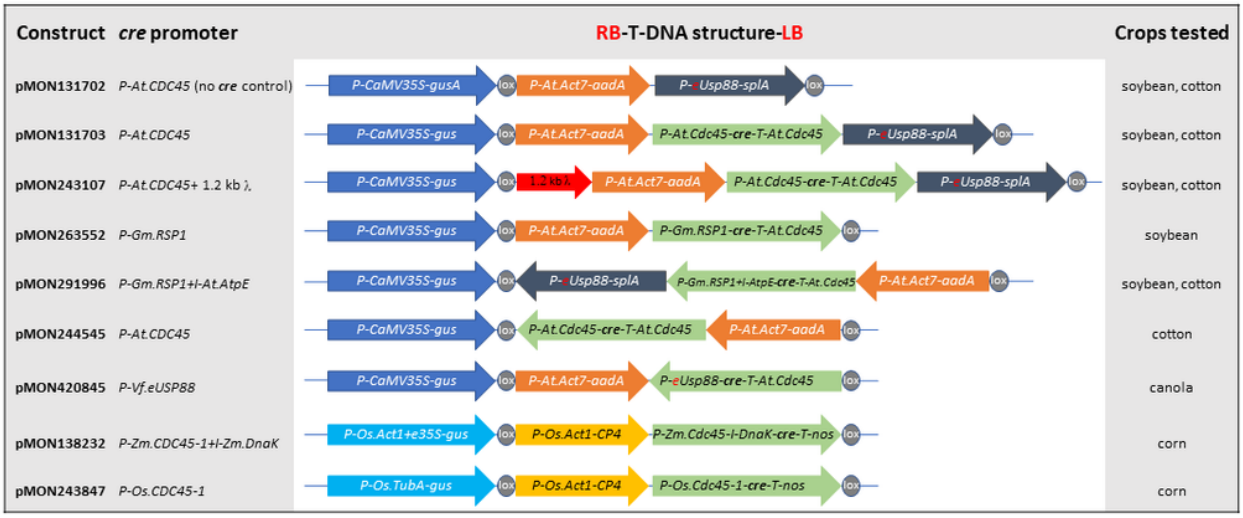

Figure 1

T-DNA structure between right and left borders of the autoexcision constructs. The arrows indicate cassette orientations from a promoter to a terminator. The genetic elements have been described in M&M, and Table 1.

Fig. 2

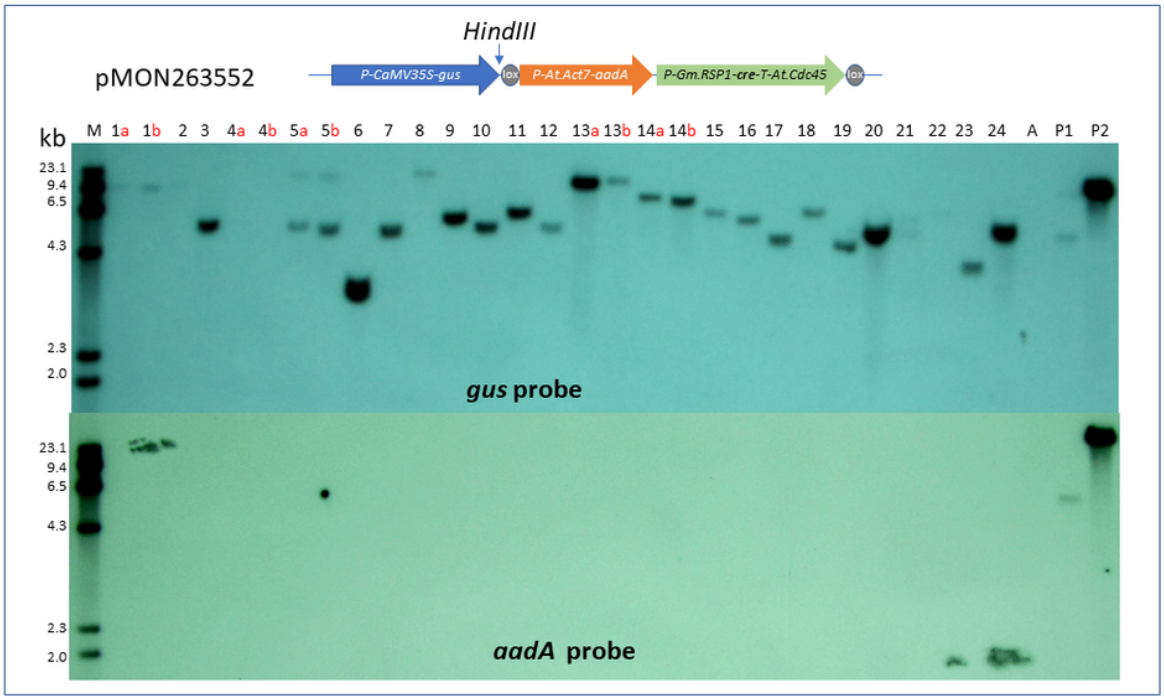

Figure 2

Marker free soybean R1 plants from pMON263552 (P-*GmRsp1-cre*) confirmed by Southern blot. The same membrane was detected sequentially by *gusA* and *aadA* probes according to DIG Southern blot instruction. M: l-*HindIII* DIG labeled marker, the band size is indicated on the left side as kb; Lane No. 1-29: 29 TaqMan® *aadA* negative and *gusA* positive R1 plants from 24 soybean R0 transgenic events, a and b indicate 2 R1 plants from the same R0 events; A: Soybean A3555 genomic DNA as a positive control; P1: 5 mg soybean genomic DNA from an *aadA* transgenic plant as a positive control; P2: 10 pg plasmid digested with *HindIII* as a positive control.

Fig. 3

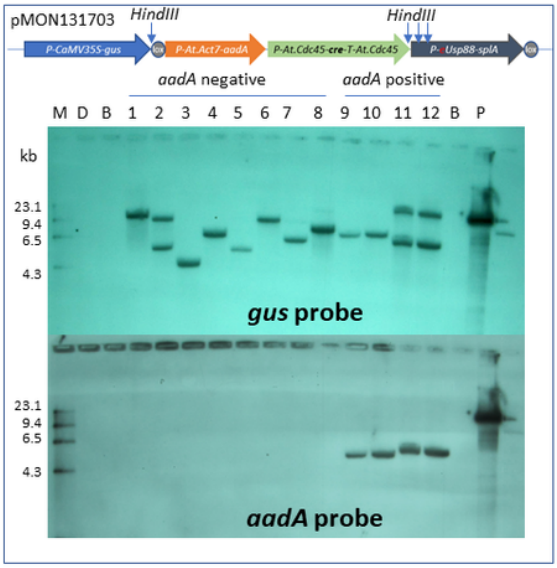

Figure 3

Southern analysis of the events from pMON131703 (P-*At.CDC45-cre*) with *gusA* and *aadA* probes. The *aadA* probe detected 5.56 kb fragment; the *gusA* probe detected minimal 3 kb fragment depending on inserts. The same membrane was detected sequentially by *gusA* and *aadA* probes, respectively, according to DIG Southern blot instruction for hybridization and probe stripping. M: l-*HindIII* DIG labeled marker, the band size is indicated on the left side as kb; D: 10 mg cotton DP393 genomic DNA as a negative control, B: blank lane; Lane No. 1-8: 8 TaqMan *aadA* negative and *gusA* positive R1 plants from 8 R0 transgenic cotton events; Lane No. 9-12: 4 TaqMan *aadA* positive R1 plants from 4 R0 transgenic cotton events as positive controls.

Fig. 4

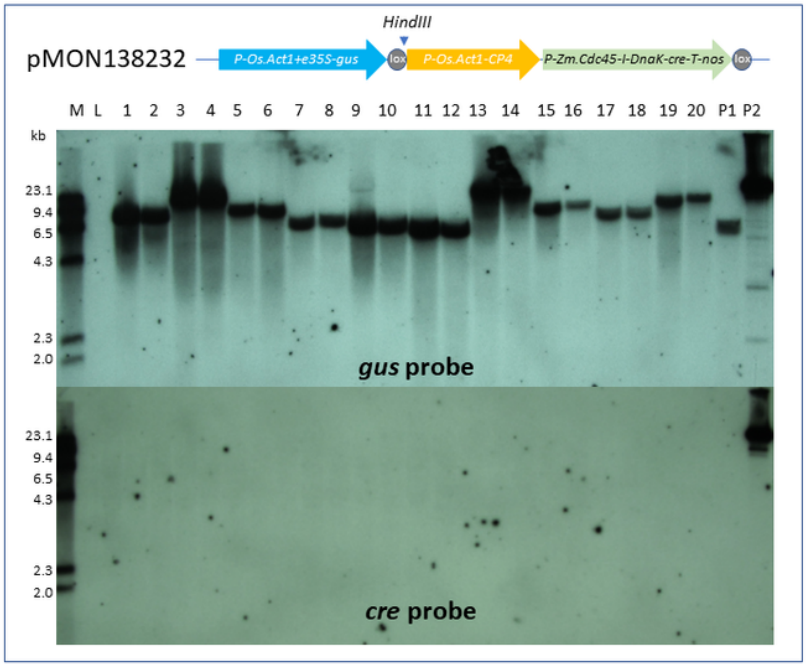

Figure 4

MF confirmation by Southern blot from maize R1 plants transformed with pMON138232 (maize *CDC45* promoter). The same membrane was detected sequentially by *gusA* and *cre* probes, respectively, according to DIG Southern blot instruction for hybridization and probe stripping. M: l-*HindIII* DIG labeled marker, the band size is indicated on the left side as kb; L: 10 mg maize genomic DNA as a negative control; Lane No. 1-20: 20 TaqMan® *CP4* negative, *gusA* positive R1 plants from 20 R0 transgenic maize events; P1: 10 mg maize genomic DNA from a plant transformed by *CP4* and *gusA* construct without *cre* as a positive control; P2: 10 pg pMON138232 plasmid digested with *HindIII* as a positive control.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx