Efficient Targeted Gene Insertion in Maize using *Agrobacterium*-Mediated Delivery

Sergei Svitashev (sergei.svitashev@corteva.com)
Corteva Agriscience https://orcid.org/0000-0003-1667-7430

Dave Peterson
Corteva Agriscience

Pierjuigi Barone
Corteva Agriscience https://orcid.org/0000-0002-2579-0913

Brian Lenderts
Corteva Agriscience

Chris Schwartz
Corteva Agriscience

Lanie Feigenbutz
Corteva Agriscience

Grace St. Clair
Corteva Agriscience

Spencer Jones
Department of Molecular Engineering, Corteva Agriscience™ Agriculture Division of DowDuPont™

Article

Keywords: Genome Editing, Gene Knock-outs, Nucleotide Replacement, Vector Design, Trait Product Development

Posted Date: December 22nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-99643/v1

Version of Record: A version of this preprint was published at Plant Biotechnology Journal on May 2nd, 2021. See the published version at https://doi.org/10.1111/pbi.13613.

Abstract

CRISPR-Cas is a powerful DNA double strand break technology with wide-ranging applications in plant genome modification. However, the efficiency of genome editing depends on various factors including plant genetic transformation processes and types of modifications desired. *Agrobacterium* infection is the preferred method of transformation and delivery of editing components into the plant cell. While this method has been successfully used to generate gene knock-outs in multiple crops, precise nucleotide replacement and especially gene insertion into a pre-defined genomic location remain highly challenging. Here we report an efficient, heritable, selectable marker-free site-specific gene insertion in maize using *Agrobacterium*-mediated delivery. Advancements in maize transformation and new vector design enabled targeted insertion with frequencies as high as 8–10%. Importantly, these advancements allowed not only an improvement of the frequency but also of the quality of generated events. These results further enable the application of genome editing for trait product development in a wide variety of crop species amenable to *Agrobacterium*-mediated transformation.

Introduction

Demonstration that the presence of a DNA double strand break (DSB) increases frequency of homologous recombination at the target site by more than 1,000-fold[1,2] has led to development of several classes of site-directed nucleases (SDNs) – homing endonucleases, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR (clustered randomly interspaced short polydromic repeats)-Cas (CRISPR associated) nucleases – capable of cutting genomic DNA and generating DSBs at predetermined locations[3,4]. Although all SDNs have been demonstrated to be useful DSB reagents, CRISPR-Cas quickly became the technology of choice for most laboratories due to its high activity, versatility, and low cost[5,6,7].

In eukaryotic cells, DSBs can be repaired via two highly-conserved pathways – non-homologous end joining (NHEJ) and homology directed repair (HDR)[8,9]. NHEJ is prone to imperfect repair and may lead to variable size insertions and/or deletions (INDELs) resulting in transcription frame-shifts and consequent gene knockouts. If two guide RNAs (gRNAs), with homology to sites flanking a gene or a DNA fragment of interest are used, it can result in a deletion of the entire sequence between the two sites[10,11]. HDR enables a precise insertion of a donor sequence to a specific target site and relies on the presence of homology regions to sequences upstream and downstream of the DSB. Thus, by using a repair template that contains nucleotide alterations, specific changes to the coding sequence of an endogenous gene can be introduced. Alternatively, if the repair template includes a new sequence (e.g. gene of interest), the DSB repair can result in the sequence insertion into a specific chromosomal site avoiding endogenous gene disruption often observed during random integration. Moreover, this approach allows for consecutive insertions of several genes into the same chromosomal region (gene stacking), which can significantly simplify breeding programs[12,13,14,15].

Recently, two new CRISPR-Cas-based technologies, base editing and prime editing, have been reported[16]. These technologies use either dead Cas9 or Cas9 nickase fusions with DNA deaminases (base editing), or Cas9 nickase fusion with reverse transcriptase (prime editing) and enable targeted nucleotide modifications without generating DSBs. Although these approaches open new exciting opportunities in gene editing, they are not allowing site-specific insertion of large DNA sequences.

HDR pathway is the preferred mechanism for targeted gene insertion, but its application remains limited due to low efficiency and a high attrition rate of recovered events. Moreover, in most successful gene insertion experiments reported in plants, repair template contains a selectable marker as the gene of interest[15,17,18,19,20,21,22]. With some rare exceptions[15], the presence of a selectable marker in the repair template, in addition to the gene of interest, is highly undesirable and needs to be removed in the consecutive generation(s). Additionally, such design increases the size of the repair template, which may further reduce the frequency of insertion events[23]. A targeted insertion with a selectable marker outside the repair template is more challenging as it requires two independent integration events in a single cell: HDR-mediated targeted gene insertion and NHEJ-mediated random integration of a selectable marker gene. Moreover, due to the DSB repair mechanisms, the selectable marker gene frequently co-integrates with repair template into the same locus, further lowering frequency of usable events[19]. As a result, very few examples of HDR-mediated selectable marker-free DNA insertions have been reported and usually relate to rather short DNA fragments[24,25,26,27].

During the past decade, several approaches to improve frequency of targeted insertions in plants have been exploited. One strategy is associated with modulating HDR and NHEJ DSB repair pathways. For example, mutations in key NHEJ pathway genes, *ku70* and *lig4*, in Arabidopsis and rice have been reported to increase HDR frequency[28,29]. Alternatively, overexpression of RAD54 gene in egg cells increased gene insertion frequency in Arabidopsis[30]. However, permanent knock-out of genes involved in the NHEJ repair pathway may lead to serious abnormalities, genome instability, chromosomal rearrangements, and even plant lethality. Although transient suppression of the NHEJ pathway is theoretically possible, this approach could significantly complicate the experimental design and may still lead to genome instability.

Another important factor influencing the efficiency of targeted insertions is the presence and amount of repair template in the nucleus. For this reason, most successful gene insertion experiments have used particle bombardment, which allows delivery of a higher copy number of DNA molecules in comparison to *Agrobacterium*-mediated transformation. Nevertheless, *Agrobacterium* infection remains the preferred method of plant transformation due to its least invasive nature, simplicity, and reproducibility. To increase repair template copy number using *Agrobacterium*-mediated delivery, several groups have used geminiviral replication system[17,31,32,33,34]. Although showing positive results, this approach requires more complicated vector designs, may lead to uncontrolled replication of the repair template, and have problems with plant regeneration[35].

In planta [36,37,38,39] and intra-genomic[14,40] homologous recombination approaches have been used to address plant transformation inefficiency and quality of targeted insertion events. These approaches rely on stable random integration of the repair template, flanked with SDN target sites, into the genome. Introduction of SDN (usually by crossing) releases repair template and simultaneously generates DSB at the intended target site. Long generation time and low efficiency have been two major limitations of this method in plants. However, a promising new approach with high efficiency intra-genomic targeted gene insertion has recently been reported[41].

It has been suggested that linear DNA can be a better substrate for HDR-mediated DSB repair and increase frequency of targeted gene insertions[42,43]. Flanking repair template with target sites and release of linear repair template has been shown to provide a 2 to 5-fold improvement of targeted gene insertion in human culture cells in comparison to circular plasmid DNA[44]. However, in several experiments conducted in plant species using particle bombardment delivery of editing components, no such effect was demonstrated[25,45,46,47].

Here we report on HDR-mediated targeted gene insertion in maize with frequencies as high as 3% using advanced Corteva Agriscience *Agrobacterium*-mediated transformation. We also demonstrate that flanking repair template with Cas9-gRNA target sites to release the repair template from the T-DNA further increases frequency of targeted insertions by approximately 3-fold. Advances in plant transformation combined with improvement in vector design allowed us to significantly increase not only frequencies (up to 8–10%) but also improve the quality of HDR-mediated gene insertion events. Moreover, our results demonstrate that targeted insertion of a repair template with and without selectable marker gene occur with comparable frequencies. These results further enable the application of genome editing for trait product development in a wide variety of crop species amenable to *Agrobacterium*-mediated transformation.

Results

Vector design

Three different T-DNA vectors were used in this study (Fig. 1). Two nearly identical vectors (Fig. 1A and 1B) included morphogenic genes, *Wus* and *Bbm*, regulated by the Axig1 and PLTP promoters, respectively[48], maize Ubiquitin promoter regulated *Streptococcus pyogenes Cas9*[19], Polymerase III U6 promoter regulated gRNA for genomic target site TS45[15], the first selectable marker gene – acetolactate synthase (*Als*) promoter regulating highly herbicide resistant *Als* (*Hra*) gene[49], and the repair template comprised of the *NptII* (*neomycin phosphotransferase II*) gene as the second selectable marker gene under the maize Ubiquitin promoter[50] flanked with regions of homology (HR1 and HR2). The only difference between the two vectors is the presence of TS45 sequences with protospacer adjacent motif (PAM) flanking HR1 and HR2 fragments in the second vector leading to release of the repair template from the T-DNA upon target sites cleavage by Cas9 nuclease (Fig. 1B). This vector design allowed for

evaluation of the repair template release effect on the frequency of the HDR-mediated targeted gene insertion. In addition, the presence of two selectable marker genes, one inside the repair template (*NptII*) and one outside (*Hra*), allowed to compare the effect of a selectable marker position on the frequency and quality of insertion events.

The third vector (Fig. 1C) contained two morphogenic genes, *Wus2* and *Bbm*, regulated by Axig1 and PLTP promoters, maize Ubiquitin promoter regulated *Streptococcus pyogenes Cas9*, Polymerase III U6 promoter regulated gRNA for genomic target site TS45, the repair template comprised of the trait gene, Arabidopsis carboxylesterase 20 (CXE-20) [51] under rice Actin promoter[52] flanked with homology regions (HR1 and HR2) and TS45 target sequences with PAM, and selectable marker gene *NptII* under maize Ubiquitin promoter. In this vector, a selectable marker gene is placed downstream from the repair template. With this configuration, cleavage of the target sites flanking repair template at transient stage results not only in the release of repair template, but also in separation of selectable marker gene (*NptII*) from the rest of T-DNA (Fig. 1C). This design allowed us to compare selectable marker gene position effect on both transformation and targeted gene insertion frequencies.

Described vectors were transformed into *Agrobacterium* LBA4404 cells containing pVIR9 accessory plasmid, which provides increased virulence and T-DNA transmission to maize embryo cells[53].

Targeted gene insertion

Results of the four experiments conducted using constructs described above are summarized in Table 1. In Experiment 1, 889 and 865 maize immature embryos (PH184C genotype) were transformed using *Agrobacteria* carrying T-DNA constructs shown in Figs. 1A and 1B, respectively. Embryos were then subjected to tissue culture and regeneration procedures described in Methods section. A total of 343 and 425 T0 plants, respectively, were regenerated on G418 (Geneticin) containing media using selectable marker gene (*NptII*) within the repair template and analyzed as illustrated in Fig. 2. HDR-mediated insertion events were first detected by diagnostic HR1 and HR2 junction qPCR (Fig. 2, A and B). Putative insertion events were further validated by long PCR spanning the entire insertion using genomic primers outside of HR1 and HR2 homology regions (Fig. 2, A and C). This approach replaced Southern blot hybridization analysis commonly used to identify T0 plants with putative perfect insertion among HR1/HR2 junction PCR positive events[19]. Long PCR products were analyzed by agarose gel electrophoresis (Fig. 2D shows a group of 15 T0 plants with representative patterns as an example). The presence of approximately 4.6 kb band indicates events with putative perfect HDR-mediated gene insertions. Lower, approximately 1.0 kb band, represents PCR product of the allele without insertion. Size variation of this band indicates various INDELs as the result of the DSB repair via NHEJ pathway. Three bands detected in plant 4, approximately 4.6 kb, 1.0 kb and 0.8 kb, likely indicate the chimeric nature of this insertion event. Plant 15 with only 4.6 kb PCR band may be bi-allelic for insertion or have one allele with an INDEL of a substantial size, which prevented amplification of the lower fragment. For T-DNA vector without target sites flanking repair template (Fig. 1A), 12 out of 343 T0 plants (3.5%) were HR1/HR2 junction qPCR positive, while for the construct with flanking target sites (Fig. 1B), 39 out of 425

T0 plants (9.2%) were HR1/HR2 positive when analyzed by junction qPCR (Table 1, Experiment 1). Long PCR analysis demonstrated that for the two treatments, 6 (1.7%) and 26 (6.1%) T0 plants, respectively, had putative perfect gene insertions. In addition, different proportions of putative perfect events (50% and 67%) were observed among HR1/HR2 junction positive events for constructs without and with flanking target sites, respectively (Table 1, Experiment 1).

The exact same constructs and experimental design were used in Experiment 2 to validate high frequency and quality of *Agrobacterium*-mediated gene insertion observed in Experiment 1. A total of 1155 and 1185 embryos were transformed and produced 591 and 342 T0 plants using constructs without and with target sites flanking the repair template, respectively. For T-DNA vector without target sites, 14 T0 plants (2.4%) were HR1/HR2 junction qPCR positive, while for the construct with flanking target sites, 33 T0 plants (9.6%) were HR1/HR2 positive when analyzed by junction qPCR (Table 1, Experiment 2). Notably, long PCR analysis performed on HR1/HR2 junction positive T0 plants again demonstrated higher proportion of putative perfect HDR-mediated gene insertion events (64%) for the construct with repair template flanked with target sites. For the construct without flanking sites, this number was 50% (Table 1, Experiment 2). Thus, Experiments 1 and 2 showed very close results confirming efficient targeted gene insertion using *Agrobacterium*-mediated delivery. In addition, flanking repair template with Cas9-gRNA target sites not only improved HDR frequency by approximately 3-fold, but also the quality of the insertion events as indicated by the long PCR data.

To compare insertion frequencies of repair templates with and without selectable marker gene, we used T-DNA construct shown in Fig. 1B. In this experiment (Experiment 3), after *Agrobacterium* infection, a total of 2884 transformed embryos were divided in two groups. For the group of 909 embryos, tissue culture steps and T0 plant regeneration were performed on selective media containing G418 using *NptII* as selectable marker gene (inside the repair template), while for 1975 embryos, these steps were conducted using imazapyr as selective agent using *Hra* as selectable marker gene (outside of the repair template). The number of embryos used in the treatment with imazapyr selection was intentionally doubled to compensate for potentially lower frequency of the HDR-mediated insertion events. A total of 320 and 766 plants were regenerated using G418 and imazapyr selection, respectively. With selectable marker gene (*NptII*) inside the repair template, 33 (10.3%) regenerants were HR1/HR2 positive by junction qPCR, while 64 (8.3%) plants were HR1/HR2 positive when selectable marker gene (*Hra*) outside the repair template was used (Table 1, Experiment 3). Long PCR analysis demonstrated lower proportion of putative perfect insertion events in the group of T0 plants with selectable marker gene outside the repair template (53% vs. 60%).

Experiment 4 was designed to test frequency of targeted insertion using a trait gene, CXE-20, and effect of selectable marker gene position (downstream of repair template, Fig. 1C) on transformation and trait gene insertion frequencies. In addition, to test potential differences in gene insertion frequency between genotypes, this experiment was conducted using a different Corteva Agriscience inbred line, PH1V5T, which has higher regenerability. One thousand immature embryos were transformed and 1403 T0 plants regenerated in this experiment. A total of 116 (8.3%) T0 plants were positive for insertion based on

HR1/HR2 junction qPCR analysis, and 55 (3.9%) plants were identified to have putative perfect integration events by long genomic PCR (Table 1, Experiment 4).

It is important to emphasize, that in the four Experiments conducted, the proportion of long PCR positive events varied and probably reflects differences in vector design and position of selectable marker genes (Table 1). For the construct without flanking target sites, this number was 50% in both Experiments 1 and 2. In Experiments 1, 2, and 3 when the repair template was flanked with TS45 target sites and selectable marker was inside the repair template, 60 to 67% of all HR1/HR2 positive events were also long PCR positive. When selectable marker gene (*Hra*) upstream of the repair template was used, this number was 53% (Experiment 3). In Experiment 4 when selectable marker gene (*NptII*) was placed downstream from the repair template, the proportion of long PCR positive events dropped to 47%.

Transmission to next generation and segregation analysis

To evaluate transmission of targeted gene insertions to next generation, segregation pattern, and attrition rates, a group of 67, long PCR-positive T0 plants from Experiments 1 and 3, were sent to the greenhouse. A total of 49 T0 plants grew to maturity, were fertile and used as pollen donors to cross with wild-type PH184C plants to generate T1 progeny (Table 2).

Fifteen to forty T1 plants (based on seed availability) for each of the 49 T0 events were then analyzed for the presence of targeted gene insertion and T-DNA components (*Cas9*, gRNA, *Bbm*, *Wus2*, *Hra*, and *NptII*). Out of 49 T0 events, 33 showed 1:1 segregation for the targeted gene insertion, consistent with stable Mendelian inheritance of a mono-allelic locus. From the remaining 16 events, 8 transmitted gene insertion allele to 20–25% of T1 progeny likely due to chimeric nature of the T0 plants. Other 5 T0 plants had insertions in both alleles – one allele containing a putative perfect insertion, while the second one was either HR1 or HR2 junction-qPCR positive. Progeny of one T0 plant was 100% positive for gene insertion likely indicating a perfect bi-allelic insertion event. The remaining 2 events did not show transmission of the insertion to T1 generation also likely due to chimeric nature of the original T0 plants. Positive for gene insertion and T-DNA-free progenies were obtained from 39 T0 plants. In the remaining 8 events, random T-DNA integration likely occurred either into the same chromosome as targeted gene insertion or into more than one genomic site. In these cases, additional round(s) of backcross are need to complete segregation of T-DNA components.

Gene insertion verification by sequencing

T1 progeny plants of 13 insertion-positive and properly segregated T0 events from Experiment 1 were further analyzed by sequencing to verify quality and integrity of the insertions. Long PCR products from 13 selected T1 plants (one plant representing each of the 13 T0 events) were sequenced by NGS and assembled into contigs. Comparison of all 13 contigs showed no sequence variation confirming that every plant had precise, HDR-mediated targeted gene insertion.

Discussion

Targeted gene insertion is the most challenging type of genome editing in plants. Its success depends on multiple factors, including transformation and regeneration processes, efficiency of genome editing components delivery, copy number of the repair template and its availability at the DSB site, size of the DNA sequence to be inserted, presence or absence of a selectable marker gene in the repair template, and HDR efficiency at a given target site. In addition, gene insertion experiments usually have high attrition rate of generated events due to the complexity of the insertion, plant chimerism, transmission to the next generation, and ability to segregate intended edits from helper genes (e.g. Cas9, gRNA, and selectable marker). Different approaches have been suggested and tested to overcome some of these issues and increase the overall success rate of targeted gene insertions: particle bombardment and *Agrobacterium* transformation, delivery of editing components as DNA, RNA, and ribonucleoprotein (RNP), suppression of NHEJ and activation of HDR pathways, viral replication to increase repair template copy number, linear and circular repair template molecules, tethering repair template to Cas9-gRNA RNP complexes, inducible, cell cycle-specific and developmental promoters, intragenomic and *in planta* approaches, and exploiting different insertion mechanisms – HDR, NHEJ, homology mediated end joining (HMEJ), and microhomology mediated end joining (MMEJ)[54,55,56]. However, despite new genome editing approaches and better understanding of underlying mechanisms of DSB repair, the progress in developing a simple, robust, and reliable process for targeted gene insertion in plants at a higher efficiency has remained limited.

Agrobacterium-mediated transformation is the preferred method for delivery of editing components into the plant cell[57]. This system is reproducible, amenable for a broad variety of crop species (both monocots and dicots) and cost effective since it doesn't require special equipment and associated consumables. In addition, *Agrobacterium*-mediated delivery is less invasive, has lower attrition rate during regeneration stage[58,59], and results in lower frequency of potential chromosomal rearrangements observed in particle bombardment experiments[60,61]. The major limitation of the *Agrobacterium*-mediated approach is a low number of T-DNA molecules transmitted into the plant cell during infection resulting in low copy number of repair template and, therefore, low frequency of targeted gene insertion. This disadvantage can be overcome by repair template copy number increase using geminivirus replication system[56]. However, this approach requires a more complicated vector design, results in uncontrolled DNA replication, may lead to unintended random integration of DNA into the genome, and can result in lower regeneration frequency[35].

Due to low frequencies of genome editing via HDR, targeted insertion experiments require hundreds of regenerated events to recover plants with the desired outcome. Many important crop species, especially their elite genotypes, have low regenerability and are extremely challenging for complex HDR-based edits. A rapid maize transformation system[48] that relies on expression of two morphogenic genes, *Bbm* and *Wus2*[62], allows recovery of transgenic plants from transformed immature embryos at high frequencies. The use of morphogenic genes has also significantly improved transformation efficiency of elite maize genotypes[62], allowing a faster trait product development. Furthermore, it has been suggested that DSB repair via homologous recombination occurs predominantly during the late S, G2 and early M phases of the cell cycle, while being actively suppressed at G1 stage[63,64,65,66]. Consequently, efficient HDR-based

genome editing is largely restricted to actively proliferating cells. Therefore, an additional benefit of using morphogenic genes in genome editing experiments is associated with an effect of BBM and WUS2 proteins on stimulation of cell division, providing an HDR-friendly cellular environment.

Another improvement of maize *Agrobacterium*-mediated transformation has been accomplished by developing a new ternary vector system that utilizes an optimized accessory plasmid pVIR9[53]. This system increased transient T-DNA delivery and recovery of stable callus events, resulting in six-to-seven-fold improvement over conventional random transformation using the plasmid pSB1 in an elite maize inbreds. The combination of morphogenic genes and new ternary vector design resulted in high transformation efficiency, increased number of T-DNA molecules transmitted, and high frequency of plant regeneration. Implementation of this enhanced *Agrobacterium* transformation protocol allowed us to generate HDR-based insertion events at target site TS45 with frequencies up to approximately 3% (Table 1). However, further optimization of *Agrobacterium*-mediated gene insertion was advantageous.

Target sites flanking repair template was described in intragenomic gene targeting experiments[36,37,38], where they were used to release the repair template from random integration loci. This design has an additional advantage as linear repair template potentially increases frequency of HDR-based gene insertion[42,43]. It might be even more advantageous considering recent results indicating that damaged DNA might be transported to specific loci at the nuclear for further repair[67]. Target sites flanking repair template have been used in recent particle bombardment experiments in Arabidopsis and rice[25,45,46,47]. However, a positive effect of this vector design on frequency of gene insertion could not be verified as no controls were provided in these experiments. Contrary, introduction of flanking target sites has resulted in 2-5-fold increase of integration frequencies relative to circular plasmids in human culture cells[44]. In our experiments, direct comparison of T-DNA vectors with and without target sites flanking repair templates, consistently demonstrated an approximately 3-fold increase of targeted gene insertion frequency, resulting in about 8–10% frequency of HDR-mediated gene insertion events based on the number of T0 plants analyzed (Table 1).

Target site TS45 used in this study was identified as one of the best sites with high HDR frequency in previous particle bombardment experiments[15]. Selection of this site was relevant for this study as it allowed to generate high number of HDR-based gene insertion events, compare results of different experiments and the corresponding controls, and draw reliable conclusions.

Another important factor that can significantly affect the outcome of gene insertion experiments, but rarely discussed, is the event attrition rate. For example, in this report, the frequency of HR1/HR2 junction qPCR positive events (approximately 8–10%) drops by approximately 50% after long PCR analysis. Most likely, this is the outcome of the concurrent action of HDR and NHEJ pathways often resulting in co-integration of various DNA molecules (vector DNA and/or genomic DNA sequences) into the target site[19,68]. In addition, our results indicate that attrition rate may depend on different parameters, including T-DNA vector configuration, such as position of selectable marker gene, regenerants survival rate, plant

fertility, and chimerism. In this report, the combined attrition rate ranges from 64 to 80%, with the highest lost observed in the experiment with selectable marker gene outside the repair template (Table 2).

In summary, the advancements in *Agrobacterium*-mediated maize transformation[48,53,62] combined with optimized vector design, enabled efficient targeted gene insertions in maize. Our experiments showed reliable and reproducible results for two genotypes with different transformability, two different genes, and for constructs with a selectable marker both inside and outside repair templates. These results open new opportunities for accelerated precision breeding in a wide range of crop species amenable to *Agrobacterium*-mediated transformation.

Methods

Plant material

Maize (*Zea mays* L.) inbred lines PH184C and PH1V5T were obtained from internal Corteva Agriscience sources.

Plasmids and reagents used for plant transformation

Cas9 expression cassette (Svitashev et al., 2015), target site TS45 and guide RNA cassette (Gao et al., 2020b), morphogenic transcription factors, PLTP promoter regulated *Babyboom* (*Bbm*), and Axig1 promoter regulated *Wuschel2* (*Wus2*)[62], acetolactate synthase (*Als*) promoter regulated highly herbicide resistant *Als* (*Hra*) gene[49], and new ternary vector and pVIR9 accessary plasmid[53] have been previously described. Repair template consisted of either Ubiquitin promoter driven *NptII* (*neomycin phosphotransferase II*) gene[50], or Arabidopsis carboxylesterase 20 (CXE-20) gene under rice Actin promoter[51] flanked with HR1 and HR2 homology regions (411 bp and 417 bp, respectively). For vectors depicted on Fig. 1 (B and C), HR1 and HR2 fragments were flanked with target site TS45 sequences with PAM.

Maize transformation

Two Corteva Agriscience inbred lines used in this study, PH184C and PH1V5T, are proprietary. All plants used as a source of immature embryos were grown in greenhouse conditions. Ear harvest, immature embryo isolation, *Agrobacterium*-mediated transformation, and plant regeneration were performed as previously described[69].

T0 and T1 plant analysis

Genomic DNA was extracted from leaf tissue as previously described[70]. Quantitative PCR (qPCR) was used to detect T-DNA components (*Cas9*, gRNA, *Bbm*, *Wus2*, *Hra*, and *NptII*) and mutation frequency at endogenous target site TS45 using Qiagen QuantiTect Multiplex PCR Master Mix (Qiagen, Germany) with the primers and probes listed in supplementary Table S1. Junction PCR assays were used to detect gene insertion at the target site TS45. In this assay, PCR amplification of the target region was coupled with a

nested qPCR using primers and probes listed in supplementary Table S2. PCR was performed using 2x Extract-N-amp PCR Ready Mix (Sigma, USA) or 2x Phusion Flash High-fidelity PCR Master Mix (Thermo Fisher Scientific, USA) according to the manufacturer's recommendations. For nested PCR used in screening of the insertion events, the first PCR was carried out in 5 µL of reaction mixtures for 20 cycles. Fifteen µL of the reaction mixture containing 2xTaqMan Master Mix (LGC, USA) and primers then were added, and the second PCR was performed using LightCycler 480 (Roche Life Science, Germany) for 30 cycles. Data were analyzed using the Endpoint Genotyping Software (Roche Life Science, Germany). Long PCR was performed using Extensor Master Mix (Thermo Fisher Scientific, USA) or LongAmp® Hot Start *Taq* 2X Master Mix (NEB, USA) with primers and probes listed in supplementary Table S2. All qPCR copy number assays were calibrated using Corteva Agriscience proprietary housekeeping gene assay (VIC).

Integrity of the gene insertion was confirmed in T1 plants by NGS sequencing of long genomic PCR amplicons. Two sets of sequencing primers (22 forward and 22 reverse) were used in the analysis (Supplementary Table S3). Approximately 250 bp-long reads were then assembled into contigs and compared using Sequencher v. 4.8.

Data Availability

The authors declare that the data supporting the findings are available within the paper or are available from the corresponding author upon reasonable request. Corteva Agriscience™ will provide plasmids to academic investigators for non-commercial research under an applicable material transfer agreement subject to proof of permission from any third-party owners of all or parts of the material and to governmental regulation considerations. Completion of a stewardship plan is also required. The Corteva Agriscience inbred lines PH184C and PH1V5T described in this research are proprietary.

Declarations

COMPETING INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

S.S. conceived the project, D.P., P.B. and S.S. designed experiments; D.P., P.B., L.F., S.J. and G.S-C. conducted transformation and plant regeneration; B.L. and C.S. analyzed plant materials; D.P., P.B. and S.S. analyzed data; P.B. and S.S. wrote the manuscript.

ACKNOWLEDGEMENTS

We thank Lijuan Wang for vector construction, Anne Ulm for assistance with transformation experiments, Tyler Engelhart for cultivation of the plants in the greenhouse, Clara Alarcon, Todd Jones, Maria Fedorova, Sendil Devadas and Nancy Vosnidou for critical reading of the manuscript. We also thank Clara Alarcon and Doane Chilcoat for leadership and support for the project.

References

1. Puchta, H., Dujon, B. & Hohn, B. 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 (1993).
2. Smih, F., Rouet, P., Romanienko, P.J. & Jasin, M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res.* **23**, 5012–5019 (1995).
3. Voytas, D.F. Plant genome engineering with sequencespecific nucleases. *Annu. Rev. Plant Biol.* **64**, 327–350 (2013).
4. Chandrasegaran, S. & Carroll, D. Origins of programmable nucleases for genome engineering. *J. Mol. Biol.* **428**, 964–989 (2016).
5. Jinek, M. et al. A programmable dual-RNA–guided DNA Endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
6. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **109**, 2579–2586 (2012).
7. Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* **34**, 933–941 (2016).
8. Puchta, H., Dujon, B. & Hohn, B. 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 (1996).
9. Puchta, H. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J. Exp. Bot.* **56**, 1–14 (2005).
10. Srivastava, V., Underwood, J.L. & Zhao, S. Dual-targeting by CRISPR/Cas9 for precise excision of transgenes from rice genome. *Plant Cell Tiss. Organ Cult.* **129**, 153–160 (2017).
11. Gao, H. et al. Superior field performance of waxy corn engineered using CRISPR–Cas9. *Nat. Biotechnol.* **38**, 579–581 (2020a).
12. Que, Q. et al. Trait stacking in transgenic crops: challenges and opportunities. *GM Crops* **1**, 220–229 (2010).
13. Ainley, W.M. et al. Trait stacking via targeted genome editing. *Plant Biotech. J.* **11**, 1126–1134 (2013).
14. Kumar, S., Barone, P. & Smith, M. Gene targeting and transgene stacking using intra genomic homologous recombination in plants. *Plant Methods* **12**, 11 (2016a).

15. Gao, H. et al. Complex trait loci in maize enabled by CRISPR-Cas9 mediated gene insertion. *Front. Plant Sci.* **11**, 535 (2020b).
16. Anzalone, A.V., Koblan, L.W. & Liu, D.R. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020).
17. Čermák, T., Baltes, N.J., Čegan, R., Zhang, Y. & Voytas, D.F. High-frequency, precise modification of the tomato genome. *Genome Biol.* **16**, 232 (2015).
18. Li, Z. et al. Cas9-guide RNA directed genome editing in soybean. *Plant Physiol.* **169**, 960–970 (2015).
19. Svitashev, S. et al. Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* **169**, 931–945 (2015).
20. Begemann, M.B. et al. Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci. Rep.* **7**, 11606 (2017).
21. Wang, M. et al. Gene targeting by homology-directed repair in rice using a geminivirus-based CRISPR/Cas9 system. *Mol. Plant* **10**, 1007–1010 (2017).
22. Lee, L. et al. CRISPR/Cas9-mediated targeted T-DNA integration in rice. *Plant Mol. Biol.* **99**, 317–328 (2019).
23. Van Vu, T. et al. Challenges and perspectives in homology-directed gene targeting in monocot plants. *Rice* **12**, 95 (2019).
24. Li, J. et al. Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat. Plants* **2**, 16139 (2016).
25. Zhao, Y. et al. An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Sci. Reports* **6**, 23890 (2016).
26. Shi, J. et al. ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol. J.* **15**, 207–216 (2017).
27. Miki, D., Zhang, W., Zeng, W., Feng, Z. & Zhu, J. K. CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat. Commun.* **9**, 1967 (2018).
28. Qi, Y. et al. Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res.* **23**, 547–554 (2013).
29. Endo, M., Mikami, M. & Toki, S. Biallelic gene targeting in rice. *Plant Physiol.* **170**, 667–677 (2016).
30. Even-Faitelson, L., Samach, A., Melamed-Bessudo, C., Avivi-Ragolsky, N. & Levy, A.A. Localized egg-cell expression of effector proteins for targeted modification of the Arabidopsis genome. *Plant J.* **68**, 929–937 (2011).
31. Baltes, N.J., Gil-Humanes, J., Čermák, T., Atkins, P.A. & Voytas, D.F. DNA replicons for plant genome engineering. *Plant Cell* **26**, 151–163 (2014).
32. Gil-Humanes, J. et al. High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J.* **89**, 1251–1262 (2017).
33. Dahan-Meir, T. et al. Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *Plant J.* **95**, 5 – 16 (2018).

34. Van Vu, T. et al. Highly efficient homology-directed repair using CRISPR/ Cpf1-geminiviral replicon in tomato. *Plant Biotechnol. J.* **18**, 1985–2143 (2020).
35. Afzal, S., Sirohi, P. & Singh, N.K. A review of CRISPR associated genome engineering: application, advances and future prospects of genome targeting tool for crop improvement. *Biotechnol. Lett.* **42**, 1611–1632 (2020).
36. Fauser, F. et al. In planta gene targeting. *Proc. Natl. Acad. Sci. USA* **109**, 7535–7540 (2012).
37. Puchta, H. & Hohn, B. In planta somatic homologous recombination assay revisited: a successful and versatile, but delicate tool. *Plant Cell* **24**, 4324–4331 (2012).
38. Ayar, A. et al. Gene targeting in maize by somatic ectopic recombination. *Plant Biotechnol. J.* **11**, 305–314 (2013).
39. Wolter, F., Klemm, J. & Puchta, H. Efficient in planta gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *Plant J.* **94**, 735–746 (2018).
40. Kumar, S., Worden, A., Novak, S., Lee, R. & Petolino, J. F. A trait stacking system via intra-genomic homologous recombination. *Planta* **244**, 1157–1166 (2016b).
41. Barone, P. et al. Efficient gene targeting in maize using inducible CRISPR-Cas9 and marker-free repair template. *Mol. Plant* **13**, 1219–1227 (2020).
42. Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K. & Carroll, D. Efficient gene targeting in Drosophila with zinc-finger nucleases. *Genetics* **172**, 2391–2403 (2006).
43. Song, F. & Stieger, K. Optimizing the DNA donor template for homology-directed repair of double-strand breaks. *Mol. Ther. Nucleic Acids* **7**, 53–60 (2017).
44. Zhang, J.-P. et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **18**, 35 (2017).
45. Sun, Y. et al. Engineering herbicide-resistant rice plants through CRISPR/ Cas9-mediated homologous recombination of acetolactate synthase. *Mol. Plant* **9**, 628–631 (2016).
46. Yao, X. et al. Homology-mediated end joining-based targeted integration using CRISPR/Cas9. *Cell Res.* **27**, 801–814 (2017).
47. Dong, O.X. et al. Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat. Commun.* **11**, 1178 (2020).
48. Lowe, K. et al. Rapid genotype “independent” Zea mays L. (maize) transformation via direct somatic embryogenesis. *In Vitro Cell. Dev. Biol. Plant* **54**, 240–252 (2018).
49. Green, J.M., Hale, T., Pagano, M. A., Andreassi, J.L. & Gutteridge, S.A. Response of 98140 corn with gat4621 and hra transgenes to glyphosate and ALS-inhibiting herbicides. *Weed Sci.* **57**, 142–148 (2009).
50. Anand, A. et al. High efficiency *Agrobacterium*-mediated site-specific gene integration in maize utilizing the FLP-FRT recombination system. *Plant Biotechnol. J.* **17**, 1636–1645 (2019).
51. Allen, S.M. et al. Drought tolerant plants and related constructs and methods involving genes encoding DTP4 polypeptides. *International Patent WO 2015/102999 a9* (2015).

52. McElroy, D., Zhang, W., Cao, J. & Wu, R. Isolation of an efficient Actin promoter for use in rice transformation. *Plant Cell* **2**, 163–171 (1990).
53. Anand, A. et al. An improved ternary vector system for *Agrobacterium*-mediated rapid maize transformation. *Plant Mol. Biol.* **97**, 187–200 (2018).
54. Huang, T.-K. & Puchta, H. CRISPR/Cas-mediated gene targeting in plants: finally a turn for the better for homologous recombination. *Plant Cell Rep.* **38**, 443–453 (2019).
55. Mao, Y., Botella, J.R., Liu, Y. & Zhu, J.-K. Gene editing in plants: progress and challenges. *Nat. Sci. Rev.* **6**, 421–437 (2019).
56. Atkins, P.A.P. & Voytas, D.F. Overcoming bottlenecks in plant gene editing. *Curr. Opp. Plant Biol.* **54**, 79–84 (2020).
57. Gordon-Kamm, W. et al. Strategies for CRISPR/Cas9-mediated genome editing: from delivery to production of modified plants. In Genome Editing for Precision Crop Breeding, M.R. Willmann, ed. (Burleigh Dodds Science Publishing Limited, UK), in press (2021).
58. Shou, H., Frame, B.R., Whitham, S.A. & Wang, K. Assessment of transgenic maize events produced by particle bombardment or *Agrobacterium*-mediated transformation. *Mol. Breeding* **13**, 201–208 (2004).
59. Jackson, M.A., Anderson, D.J. & Birch, R.G. Comparison of *Agrobacterium* and particle bombardment using whole plasmid or minimal cassette for production of high-expressing, low-copy transgenic plants. *Transgenic Res.* **22**, 143–151 (2013).
60. Svitashev, S., Ananiev, E., Pawlowski, W.P. & Somers, D.A. Association of transgene integration sites with chromosome rearrangements in hexaploid oat. *Theor. Appl. Genet.* **100**, 872–880 (2000).
61. Liu, J. et al. Genome-scale sequence disruption following biolistic transformation in rice and maize. *Plant Cell* **31**, 368–383 (2019).
62. Lowe, K. et al. Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *Plant Cell* **28**, 1998–2015 (2016).
63. Saleh-Gohari, N. & Helleday, T. Conservative homologous recombination preferentially repairs DNA DSBs in the S phase of the cell cycle in human cells. *Nucleic Acids Res.* **32**, 3683–3688 (2004).
64. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902–2906 (2008).
65. Heyer, W.D., Ehmsen, K.T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **44**, 113–139 (2010).
66. Yang, D. et al. Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci. Rep.* **6**, 21264 (2016).
67. Caridi, C.P. et al. Nuclear F-actin and myosins drive relocalization of heterochromatic breaks. *Nature* **559**, 54–60 (2018).
68. Svitashev, S.K., Pawlowskiy, W.P., Makarevitch, I., Plank, D.W. & Somers, D.A. Complex transgene locus structures implicate multiple mechanisms for plant transgene rearrangement. *Plant J.* **32**,

433–445 (2002).

69. Jones, T. et al. Maize transformation using the morphogenic genes Baby Boom and Wuschel2. *Methods Mol. Biol.* **1864**, 81–93 (2019).

70. Gao, H. et al. Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.* **61**, 176–187 (2010).

Tables

Tables 1 and 2 are available in the Supplementary Files

Figures

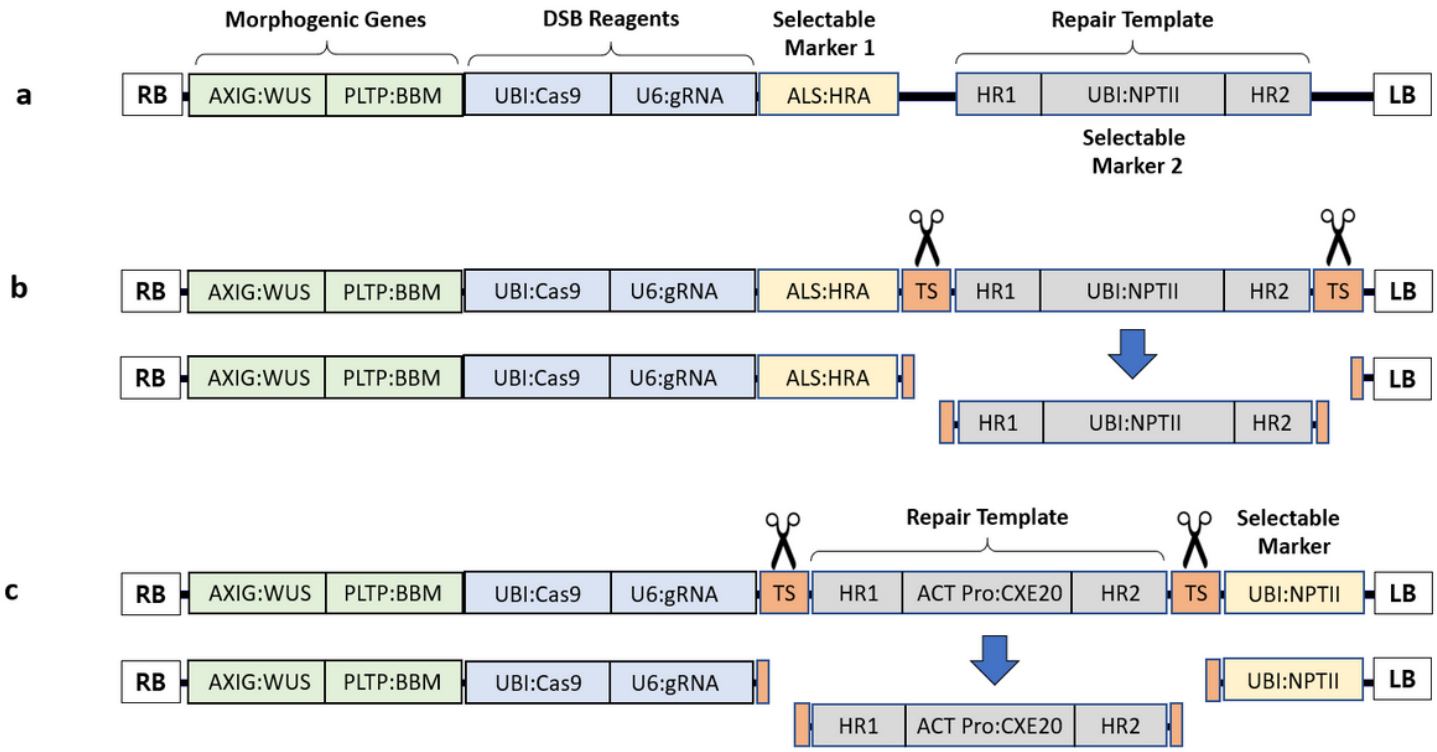

Figure 1

Schematic description of T-DNA constructs. a T-DNA vector containing morphogenic genes Wus2 and Bbm, Cas9, gRNA, selectable marker gene 1 (Hra) and repair template, comprising selectable marker gene 2 (NptII) flanked with homology arms (HR1 and HR2). b T-DNA vector as on figure 1A supplemented with Cas9/gRNA TS45 target sites (TS) flanking homology arms. c T-DNA vector containing morphogenic genes Wus2 and Bbm, Cas9, gRNA, repair template (CXE-20, homology arms, HR1 and HR2) flanked with TS45 target sequences, and selectable marker gene (NptII).

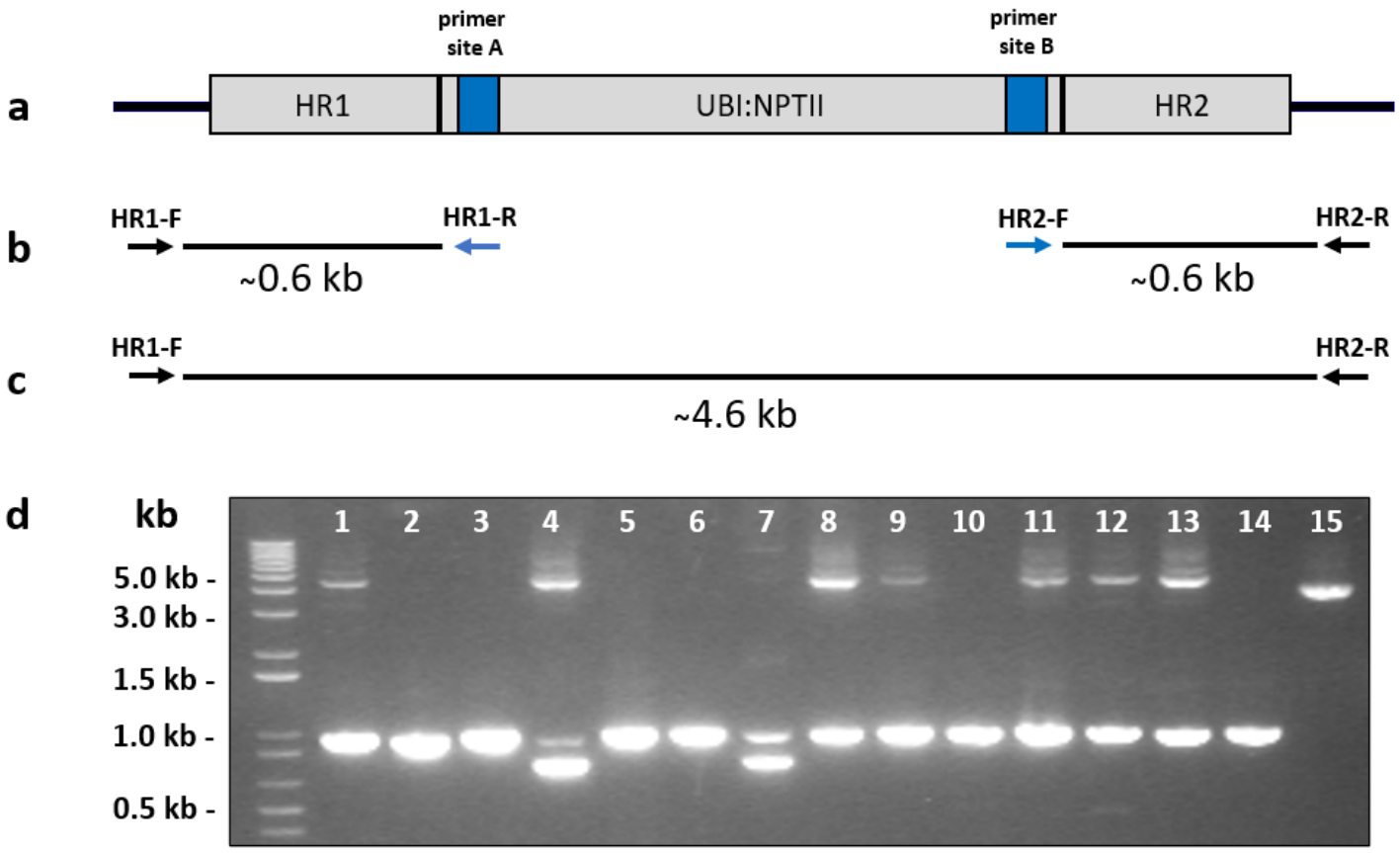

Figure 2

Plant analysis for gene insertion events. a Schematic depiction of HDR-mediated targeted gene insertion locus. b Diagnostic qPCR primers and expected products from 5' and 3' PCR indicating gene insertion. c PCR primers and expected product of long PCR used for further characterization of HR1/HR2 junction qPCR positive gene insertion events. d Example of long PCR diagnostic agarose gel. Upper, approximately 4.6 kb band, corresponds to a putative perfect single copy gene insertion. Lower, approximately 1 kb band, corresponds to amplification of the allele without insertion. Size variation of the shorter fragment reflects the presence of various size deletions and insertions. Plant 4 has 3 bands indicating that the plant is chimeric. Plant 15 has only one band indicating either bi-allelic gene insertion or a large insertion or deletion in the second allele resulting in no product amplification.

Supplementary Files

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

- PetersonetalTable1NATCOMM.pptx
- PetersonetalTable2NATCOMM.pptx
- PetersonetalSupplementaryTablesNATCOMM.pptx