

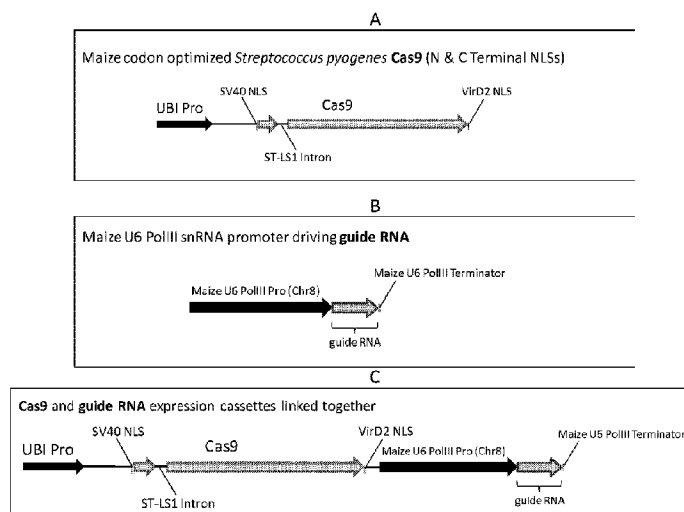


(86) Date de dépôt PCT/PCT Filing Date: 2014/08/20
(87) Date publication PCT/PCT Publication Date: 2015/02/26
(45) Date de délivrance/Issue Date: 2024/02/27
(85) Entrée phase nationale/National Entry: 2016/02/22
(86) N° demande PCT/PCT Application No.: US 2014/051781
(87) N° publication PCT/PCT Publication No.: 2015/026886
(30) Priorités/Priorities: 2013/08/22 (US61/868,706);
2013/09/25 (US61/882,532); 2014/02/07 (US61/937,045);
2014/03/14 (US61/953,090); 2014/07/11 (US62/023,239)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01)
(72) Inventeurs/Inventors:
CIGAN, ANDREW MARK, US;
FALCO, SAVERIO CARL, US;
LASSNER, MICHAEL, US;
LIU, ZHAN-BIN, US;
SVITASHEV, SERGEI, US
(73) Propriétaires/Owners:
E.I. DU PONT DE NEMOURS AND COMPANY, US;
PIONEER HI-BRED INTERNATIONAL, INC., US
(74) Agent: TORYS LLP

(54) Titre : PROCEDES VISANT A PRODUIRE DES MODIFICATIONS GENETIQUES DANS UN GENOME DE VEGETAL SANS INTRODUIRE DE MARQUEUR DE TRANSGENE SELECTIONNABLE, ET COMPOSITIONS CORRESPONDANTES

(54) Title: METHODS FOR PRODUCING GENETIC MODIFICATIONS IN A PLANT GENOME WITHOUT INCORPORATING A SELECTABLE TRANSGENE MARKER, AND COMPOSITIONS THEREOF



(57) **Abrégé/Abstract:**

Compositions and methods are provided for genome modification of a target sequence in the genome of a plant or plant cell, without incorporating a selectable transgene marker. The methods and compositions employ a guide polynucleotide / Cas endonuclease system to provide an effective system for modifying or altering target sites within the genome of a plant, plant cell or seed, without incorporating a selectable transgene marker. Once a genomic target site is identified, a variety of methods can be employed to further modify the target sites such that they contain a variety of polynucleotides of interest. Breeding methods and methods for selecting plants utilizing a guide polynucleotide / Cas endonuclease system are also disclosed. Compositions and methods are also provided for editing a nucleotide sequence in the genome of a cell, without incorporating a selectable transgene marker.

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

(19) World Intellectual Property
Organization
International Bureau



(10) International Publication Number
WO 2015/026886 A1

(43) International Publication Date
26 February 2015 (26.02.2015)

(51) International Patent Classification:
C12N 15/82 (2006.01)

(21) International Application Number:
PCT/US2014/051781

(22) International Filing Date:
20 August 2014 (20.08.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/868,706 22 August 2013 (22.08.2013) US
61/882,532 25 September 2013 (25.09.2013) US
61/937,045 7 February 2014 (07.02.2014) US
61/953,090 14 March 2014 (14.03.2014) US
62/023,239 11 July 2014 (11.07.2014) US

(71) Applicants: **E. I. DU PONT DE NEMOURS AND COMPANY** [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US). **PIONEER HI-BRED INTERNATIONAL, INC.** [US/US]; 7100 N.W. 62nd Avenue, P.O. Box 1014, Johnston, Iowa 50131-1014 (US).

(72) Inventors: **CIGAN, Andrew, Mark**; 5764 Chatham Circle, Johnston, Iowa 50131 (US). **FALCO, Saverio, Carl**; 1902 Millers Road, Wilmington, Delaware 19810 (US). **LASSNER, Michael**; 2935 133rd Way, Urbandale, Iowa 50323 (US). **LIU, Zhan-Bin**; 1357 Wooded Knoll, West Chester, Pennsylvania 19382 (US). **SVITASHEV, Sergei**; 9032 Timberwood Drive, Johnston, Iowa 50131 (US).

(74) Agents: **STOOP, Johan, M.** et al.; E. I. du Pont de Nemours and Company, Legal Patent Records Center, Chestnut Run Plaza 721/2340, 974 Centre Road, PO Box 2915 Wilmington, Delaware 19805 (US).

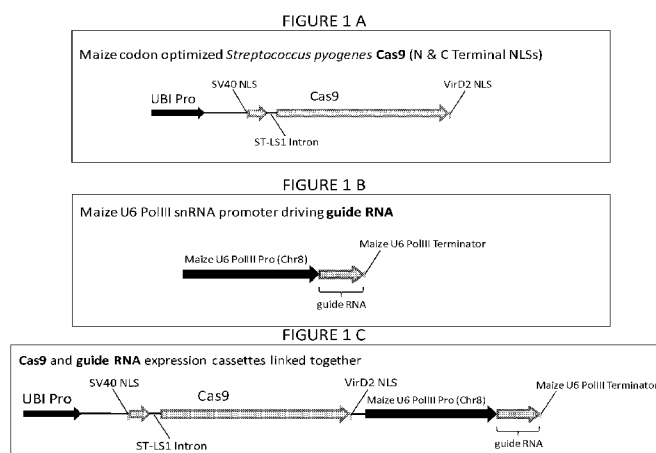
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP,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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: METHODS FOR PRODUCING GENETIC MODIFICATIONS IN A PLANT GENOME WITHOUT INCORPORATING A SELECTABLE TRANSGENE MARKER, AND COMPOSITIONS THEREOF



(57) Abstract: Compositions and methods are provided for genome modification of a target sequence in the genome of a plant or plant cell, without incorporating a selectable transgene marker. The methods and compositions employ a guide polynucleotide / Cas endonuclease system to provide an effective system for modifying or altering target sites within the genome of a plant, plant cell or seed, without incorporating a selectable transgene marker. Once a genomic target site is identified, a variety of methods can be employed to further modify the target sites such that they contain a variety of polynucleotides of interest. Breeding methods and methods for selecting plants utilizing a guide polynucleotide / Cas endonuclease system are also disclosed. Compositions and methods are also provided for editing a nucleotide sequence in the genome of a cell, without incorporating a selectable transgene marker.

METHODS FOR PRODUCING GENETIC MODIFICATIONS IN A PLANT GENOME WITHOUT INCORPORATING A SELECTABLE TRANSGENE MARKER, AND COMPOSITIONS THEREOF

This application claims the benefit of U.S. Provisional Application No. 61/868706, filed August, 22, 2013, U.S. Provisional Application No. 61/882532, filed September 25, 2013, U.S. Provisional Application No. 61/937045, filed February 07, 2014, U.S. Provisional Application No. 61/953090, filed March 14, 2014, and U.S. Provisional Application No. 62/023239, filed July 11, 2014.

FIELD

The disclosure relates to the field of plant molecular biology, in particular, to methods for altering the genome of a plant cell.

REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

The official copy of the sequence listing is submitted electronically via EFS-Web as an ASCII formatted sequence listing with a file named 20140814_BB2365PCT_ST25_SequenceListing created on August 14, 2014 and having a size of 584 kilobytes and is filed concurrently with the specification. The sequence listing contained in this ASCII formatted document is part of the specification.

BACKGROUND

Recombinant DNA technology has made it possible to insert foreign DNA sequences into the genome of an organism, thus, altering the organism's phenotype. The most commonly used plant transformation methods are *Agrobacterium* infection and biolistic particle bombardment in which transgenes integrate into a plant genome in a random fashion and in an unpredictable copy number. Thus, efforts are undertaken to control transgene integration in plants.

One method for inserting or modifying a DNA sequence involves homologous DNA recombination by introducing a transgenic DNA sequence flanked by sequences homologous to the genomic target. U.S. Patent No. 5,527,695 describes transforming eukaryotic cells with DNA sequences that are targeted to a predetermined sequence of the eukaryote's DNA. Specifically, the use of site-

specific recombination is discussed. Transformed cells are identified through use of a selectable marker included as a part of the introduced DNA sequences.

It was shown that artificially induced site-specific genomic double-stranded breaks in plant cells were repaired by homologous recombination with exogenously supplied DNA using two different pathways. (Puchta *et al.*, (1996) *Proc. Natl. Acad. Sci. USA* 93:5055-5060; U.S. Patent Application Publication No. 2005/0172365A1 published August 4, 2005; U.S. Patent Application Publication No. 2006/0282914 published December 14, 2006; WO 2005/028942 published June 2, 2005).

Since the isolation, cloning, transfer and recombination of DNA segments, including coding sequences and non-coding sequences, is most conveniently carried out using restriction endonuclease enzymes. Much research has focused on studying and designing endonucleases such as WO 2004/067736 published August 12, 2004; U.S. Patent No. 5,792,632 issued to Dujon *et al.*, August 11, 1998; U.S. Patent No. 6,610,545 B2 issued to Dujon *et al.*, August 26, 2003; Chevalier *et al.*, (2002) *Mol Cell* 10:895-905; Chevalier *et al.*, (2001) *Nucleic Acids Res* 29:3757-3774; Seligman *et al.*, (2002) *Nucleic Acids Res* 30:3870-3879.

Although several approaches have been developed to target a specific site for modification in the genome of a plant, there still remains a need for more efficient and effective methods for producing a fertile plant, having an altered genome comprising specific modifications in a defined region of the genome of the plant.

There remains a need to develop a genetic modification system that can modify a genomic location without incorporating a selectable transgene marker.

BRIEF SUMMARY

Compositions and methods are provided employing a guide polynucleotide /Cas endonuclease system in plants for genome modification of a target sequence in the genome of a plant or plant cell, for selecting plants, for gene editing, and for inserting a polynucleotide of interest into the genome of a plant without incorporating a selectable transgene marker.

The methods and compositions employ a guide polynucleotide/Cas endonuclease system to provide for an effective system for modifying or altering target sites and nucleotide of interest within the genome of a plant, plant cell or seed without incorporating a selectable transgene marker.

Once a genomic target site is identified, a variety of methods can be employed to further modify the target sites such that they contain a variety of polynucleotides of interest. Breeding methods and methods for selecting plants utilizing a two component RNA guide and Cas endonuclease system are also disclosed. Also provided are nucleic acid constructs, plants, plant cells, explants, seeds and grain having the guide polynucleotide/Cas endonuclease system. The methods and compositions employ a guide polynucleotide/Cas endonuclease system to provide for an effective system for modifying or altering target sites and editing nucleotide sequences of interest within the genome of a cell, wherein the guide polynucleotide is comprised of a RNA sequence, a DNA sequence, or a DNA-RNA combination sequence.

Thus in one embodiment of the disclosure, the method comprises a method for producing a genetic modification into a second gene of a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide polynucleotide, a polynucleotide modification template, a second guide polynucleotide, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be modified to confer herbicide resistance, wherein said first guide polynucleotide and Cas endonuclease are capable of forming a first complex that enables the Cas endonuclease to introduce a double strand break at a first target site, located in or near said first endogenous gene in the genome of said plant cell, wherein said second guide polynucleotide and Cas endonuclease are capable of forming a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant cell, wherein said polynucleotide modification template comprises at least one nucleotide alteration when compared to the first endogenous gene.

In one embodiment of the disclosure, the method comprises a method for introducing a polynucleotide of interest into a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be

modified to confer herbicide resistance, wherein said first guide RNA and Cas endonuclease are capable of forming a first complex that enables the Cas endonuclease to introduce a double strand break at a first target site located in or near said first endogenous gene in the genome of said plant cell, wherein said first
5 polynucleotide modification template comprises at least one nucleotide modification of said first endogenous gene to render said endogenous gene capable of conferring herbicide resistance to a plant cell, wherein said second guide RNA and Cas endonuclease are capable of forming a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the
10 genome of said plant cell, wherein said second polynucleotide modification template comprises at least one polynucleotide of interest to be introduced into said plant genome.

In one embodiment of the disclosure, the method comprises a method for editing a second gene of a plant genome without introducing an exogenous
15 selectable marker into said plant genome, the method comprising providing a first guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be modified to confer herbicide resistance, wherein said first guide RNA and Cas endonuclease are capable of
20 forming a complex that enables the Cas endonuclease to introduce a double strand break at a first target site (located in or near said first endogenous gene) in the genome of said plant cell, wherein said first polynucleotide modification template comprises at least one nucleotide modification of said first endogenous gene to render said endogenous gene capable of conferring herbicide resistance to a plant
25 cell, wherein said second guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at a second target site (located at a different locus than said first endogenous gene) in the genome of said plant cell, wherein said second polynucleotide modification template comprises at least one nucleotide alteration when compared to the second
30 gene to be edited.

Additional embodiments of the methods and compositions of the present invention are disclosed herein.

BRIEF DESCRIPTION OF THE DRAWINGS AND THE SEQUENCE LISTING

The disclosure can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing, which form a part of this application. The sequence descriptions and sequence listing attached hereto comply with the rules governing nucleotide and amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §§1.821-1.825. The sequence descriptions contain the three letter codes for amino acids as defined in 37 C.F.R. §§ 1.821-1.825.

Figures

Figure 1A shows a maize optimized Cas9 gene (encoding a Cas9 endonuclease) containing a potato ST-LS1 intron, a SV40 amino terminal nuclear localization sequence (NLS), and a VirD2 carboxyl terminal NLS, operably linked to a plant ubiquitin promoter (SEQ ID NO: 5). The maize optimized Cas9 gene (just Cas9 coding sequence, no NLSs) corresponds to nucleotide positions 2037-2411 and 2601-6329 of SEQ ID NO: 5 with the potato intron residing at positions 2412-2600 of SEQ ID NO: 5. SV40 NLS is at positions 2010-2036 of SEQ ID NO: 5. VirD2 NLS is at positions 6330-6386 of SEQ ID NO: 5. Figure 1 B shows a long guide RNA operably linked to a maize U6 polymerase III promoter terminating with a maize U6 terminator (SEQ ID NO: 12). The long guide RNA containing the variable targeting domain corresponding to the maize LIGCas-3 target site (SEQ ID NO: 8) is transcribed from/ corresponds to positions 1001-1094 of SEQ ID NO: 12. Figure 1 C shows the maize optimized Cas9 and long guide RNA expression cassettes combined on a single vector DNA (SEQ ID NO: 102).

Figure 2A illustrates the duplexed crRNA (SEQ ID NO:6)-tracrRNA (SEQ ID NO:7)/ Cas9 endonuclease system and target DNA complex relative to the appropriately oriented PAM sequence at the maize LIGCas-3 (SEQ ID NO: 18, Table 1) target site with triangles pointing towards the expected site of cleavage on both sense and anti-sense DNA strands. Figure 2 B illustrates the guide RNA / Cas9 endonuclease complex interacting with the genomic target site relative to the appropriately oriented PAM sequence (GGA) at the maize genomic LIGCas-3 target site (SEQ ID NO:18, Table 1). The guide RNA (shown as boxed-in in light gray,

SEQ ID NO:8) is a fusion between a crRNA and tracrRNA and comprises a variable targeting domain that is complementary to one DNA strand of the double strand DNA genomic target site. The Cas9 endonuclease is shown in dark gray. Triangles point towards the expected site of DNA cleavage on both sense and anti-sense DNA strands.

- 5 The sense genome sequence shown in Figures 2A and 2B is listed in SEQ ID NO: 555, while the complementary genome sequence shown in Figures 2A and 2B is listed in SEQ ID NO: 556.

Figure 3A-3B shows an alignment and count of the top 10 most frequent NHEJ mutations induced by the maize optimized guide RNA/Cas endonuclease system described herein compared to a LIG3-4 homing endonuclease control at the maize genomic Liguleless 1 locus. The mutations were identified by deep sequencing. The reference sequence represents the unmodified locus with each target site underlined. The PAM sequence and expected site of cleavage are also indicated. Deletions or insertions as a result of imperfect NHEJ are shown by a "-" or an italicized underlined nucleotide, respectively. The reference and mutations 1-10 of the LIGCas-1 target site correspond to SEQ ID NOs: 55-65, respectively. The reference and mutations 1-10 of the LIGCas-2 correspond to SEQ ID NOs: 65-75, respectively. The reference and mutations 1-10 of the LIGCas-3 correspond to SEQ ID NOs: 76-86, respectively. The reference and mutations 1-10 of the LIG3-4 homing endonuclease target site correspond to SEQ ID NOs: 76, 87-96, respectively.

Figure 4 illustrates how the homologous recombination (HR) repair DNA vector (SEQ ID NO: 97) was constructed. To promote site-specific transgene insertion by homologous recombination, the transgene (shown in light gray) was flanked on either side by approximately 1 kb of DNA with homology to the maize genomic regions immediately adjacent to the LIGCas3 and LIG3-4 homing endonuclease expected sites of cleavage.

Figure 5 illustrates how genomic DNA extracted from stable transformants was screened for site-specific transgene insertion by PCR. Genomic primers (corresponding to SEQ ID NOs: 98 and 101) within the Liguleless 1 locus were designed outside of the regions used in constructing the HR repair DNA vector (SEQ ID NO: 97) and were paired with primers inside the transgene (corresponding to SEQ ID NOs: 99 and 100) to facilitate PCR detection of unique genomic DNA junctions created by appropriately oriented site-specific transgene integration.

Figure 6 shows an alignment of the NHEJ mutations induced by the maize optimized guide RNA/Cas endonuclease system, described herein, when the short guide RNA was delivered directly as RNA. The mutations were identified by deep sequencing. The reference illustrates the unmodified locus with the genomic target site underlined. The PAM sequence and expected site of cleavage are also indicated. Deletions or insertions as a result of imperfect NHEJ are shown by a "-" or an italicized underlined nucleotide, respectively. The reference and mutations 1-6 for 55CasRNA-1 correspond to SEQ ID NOs: 104-110, respectively.

Figure 7 shows the QC782 vector comprising the Cas9 expression cassette.

Figure 8A shows the QC783 vector comprising the guide RNA expression cassette. Figure 8B show the DNA sequence (coding sequence) of the DD43CR1 (20 bp) variable targeting domain of the guide RNA, as well as the terminator sequence linked to the guide RNA. The 20 bp variable targeting domain DD43CR1 is in bold. The sequence shown in Figure 8B is listed in SEQ ID NO: 557.

Figure 9 shows the map of a linked soybean optimized Cas9 and guide RNA construct QC815.

Figure 10 A shows the DD20 soybean locus on chromosome 4 and the DD20CR1 and DD20CR2 genomic target sites (indicated by bold arrows). The sense sequence shown in Figure 10A is listed in SEQ ID NO: 558 and the complementary sequence shown in Figure 10A is listed in SEQ ID NO: 559. Figure 10 B shows the DD43 soybean locus on chromosome 4 and the DD43CR1 and DD43CR2 genomic target sites (indicated by bold arrows). The sense sequence shown in Figure 10B is listed in SEQ ID NO: 560 and the complementary sequence shown in Figure 10B is listed in SEQ ID NO: 561.

Figure 11A-11D. Alignments of expected target site sequences with mutant target sequences detected in four guide RNAs induced NHEJ experiments. Figure 11A shows the DD20CR1 PCR amplicon (reference sequence, SEQ ID NO:142, genomic target site is underlined) and the 10 mutations (SEQ ID NOs: 147-156) induced by the guideRNA/Cas endonuclease system at the DD20CR1 genomic target site. Figure 11B shows the DD20CR2 PCR amplicon (reference sequence, SEQ ID NO:143) and the 10 mutations (SEQ ID NOs 157-166) induced by the guide RNA/Cas endonuclease system at the DD20CR2 genomic target site. Figure 11C shows the DD43CR1 PCR amplicon (reference sequence, SEQ ID NO:144) and the 10 mutations (SEQ ID NOs:167-176) induced by the guide RNA/Cas endonuclease system at the DD43CR1 genomic target site. Figure 11D shows the DD43CR2 PCR amplicon (reference sequence, SEQ ID NO:145) and the 10 mutations (SEQ ID

NOs: 177-191) induced by the guide RNA/Cas endonuclease system at the DD43CR2 genomic target site. The target sequences corresponding different guide RNAs are underlined. Each nucleotide deletions is indicated by "-". Inserted and replaced sequences are in bold. The total number of each mutant sequence is
 5 listed in the last column.

Figure 12A-12B shows a schematic representation of the guide RNA/Cas endonuclease system used for editing a nucleotide sequence of interest. To enable specific nucleotide editing, a polynucleotide modification template that includes at least one nucleotide modification (when compared to the nucleotide sequence to be
 10 edited) is introduced into a cell together with the guide RNA and Cas endonuclease expression cassettes. For example, as shown herein, the nucleotide sequence to be edited is an endogenous wild type enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene in maize cells. The Cas endonuclease (shaded circle) is a maize optimized Cas9 endonuclease that cleaves a moCas9 target sequence within the
 15 *epsps* genomic locus using a guide RNA of SEQ ID NO:194. Figure 12-A shows a polynucleotide modification template that includes three nucleotide modifications (when compared to the wild type *epsps* locus depicted in Figure 12-B) flanked by two homology regions HR-1 and HR-2. Figure 12-B shows the guide RNA/maize optimized Cas9 endonuclease complex interacting with the *epsps* locus. The
 20 original nucleotide codons of the *EPSPS* gene that needed to be edited are show as aCT and Cca (Figure 12-B). The nucleotide codons with modified nucleotides (shown in capitals) are shown as aTC and Tca (Figure 12-B).

Figure 13 shows a diagram of a maize optimized Cas9 endonuclease expression cassette. The bacterial *cas9* coding sequence was codon optimized for
 25 expression in maize cells and supplemented with the ST-LS1 potato intron (moCas9 coding sequence, SEQ ID NO: 193). A DNA fragment encoding the SV40 nuclear localization signal (NLS) was fused to the 5'-end of the moCas9 coding sequence. A maize ubiquitin promoter (Ubi promoter) and its cognate intron (ubi intron) provided controlling elements for the expression of moCas9 in maize cells. The *pinII*
 30 transcription termination sequence (*pinII*) completed the maize *moCAS9* gene design.

Figure 14 shows some examples of the moCas9 target sequence (underlined), located on *EPSPS* DNA fragments, mutagenized by the introduction of double-strand breaks at the cleavage site of the moCas9 endonuclease (thick arrow) in maize cells.. In SEQ ID NO: 206, three nucleotides were deleted (dashes) next to the moCas9 cleavage site. SEQ ID NOs: 207- 208 indicate that the

5 nucleotide deletion can expand beyond the moCas9 cleavage site

Figure 15 depicts an *EPSPS* template vector used for delivery of the *EPSPS* polynucleotide modification template containing the three TIPS nucleotide modifications. The *EPSP* polynucleotide modification template includes a partial

10 fragment of the *EPSPS* gene. The vector was 6,475 bp in length and consisted of two homology regions to the *epsps* locus (*epsps*-HR1 and *epsps*-HR2). Two Gateway cloning sites (ATTL4 and ATTL3), an antibiotic resistance gene (KAN), and the pUC origin of replication (PUC ORI) completed synthesis of the *EPSPS* template vector1.

15 Figure 16 illustrates the PCR-based screening strategy for the identification of maize events with TIPS nucleotide modifications in maize cells. Two pairs of PCR primers were used to amplify the genomic fragments of the *epsps* locus (upper section). Both of them contained the TIPS specific primers (an arrow with a dot indicating the site of the three TIPS modifications). The shorter fragment (780 bp F-E2) was produced by amplification of the *EPSPS* polynucleotide modification

20 template fragment (template detection). The amplified *EPSPS* polynucleotide modification template fragment was found in all but 4 analyzed events (panel F-E2). The longer fragment (839 bp H-T) was produced by amplification of the genomic *EPSPS* sequence providing that the *epsps* locus contained the three nucleotide

25 modifications responsible for the TIPS modifications. Six events were identified as containing the three nucleotide modifications (panel H-T). The white arrows point to events that contain both the amplified *EPSPS* polynucleotide modification template and the nucleotide modifications responsible for the TIPS modification.

Figure 17 A shows a schematic diagram of the PCR protocol used to identify

30 edited *EPSPS* DNA fragments in selected events. A partial genomic fragment, comprising parts of Exon1, Intron 1 and Exon2 of the *epsps* locus, was amplified regardless of the editing product (panel A, 1050 bp F-E3). The amplification

products, representing only partial EPSPS gene sequences having one or more mutations, were cloned and sequenced. Figure 17-B shows 2 examples of sequenced amplification products. In some amplification products, the *epsps* nucleotides and the moCas9 target sequence (underlined) were unchanged
 5 indicating that one *EPSPS* allele was not edited (wild type allele; SEQ ID NO: 210). In other amplification products, three specific nucleotide substitutions (representing the TIPS modifications) were identified with no mutations at the moCas9 target sequence (underlined) (SEQ ID NO: 209).

Figure 18 shows the location of MHP14, TS8, TS9 and TS10 loci comprising
 10 target sites for the guide RNA/Cas endonuclease system near trait A (located at 53.14 cM) on chromosome 1 of maize.

Figure 19A shows the location of the MHP14Cas1 maize genomic target sequence (SEQ ID NO: 229) and the MSP14Cas-3 maize genomic target sequence (SEQ ID NO: 230) on the MHP14 maize genomic DNA locus on chromosome 1. The
 15 5'to 3' sequence. Figure 19B shows the location of the TS8Cas-1 (SEQ ID NO: 231) and TS8Cas-2 (SEQ ID NO: 232) maize genomic target sequences located on the TS8 locus. Figure 19-C shows the location of the TS9Cas-2 (SEQ ID NO: 233) and TS9Cas-3 (SEQ ID NO: 234) maize genomic target sequences located on the TS8 locus. Figure 19-D shows the location of the TS10Cas-1 (SEQ ID NO: 235), and
 20 TS10Cas-3 (SEQ ID NO: 236) maize genomic target sequences located on the TS10 locus. All these maize genomic target sites are recognized and cleaved by a guide RNA/Cas endonuclease system described herein. Each maize genomic target sequence (indicated by an arrow) is highlighted in bold and followed by the NGG PAM sequence shown boxed in.

Figure 20 shows a schematic of a donor DNA (also referred to as HR repair
 25 DNA) comprising a transgene cassette with a selectable marker (phosphomannose isomerase, depicted in grey), flanked by homologous recombination sequences (HR1 and HR2) of about 0.5 to 1 kb in length, used to introduce the transgene cassette into a genomic target site for the guide RNA/Cas endonuclease system.
 30 The arrows indicate the sections of the genomic DNA sequence on either side of the endonuclease cleavage site that corresponds to the homologous regions of the donor DNA. This schematic is representative for homologous recombination

occurring at any one of the 8 target sites (4 loci) located on chromosome 1 from 51.54cM to 54.56cM in maize genome.

Figure 21 shows the junction PCR screen for identification of insertion events. Primer 1 and 2 located on the transgene donor are common for all target sites. Primer TSHR1f is located on the genomic region outside of the homologous sequence HR1. Primer combination THR1f/primer1 amplify junction 1. Primer TSHR2r is located on the genomic region outside of the HR2 region. Primer combination primer2/TSHR2r amplify junction 2.

Figure 22 shows a junction PCR screen for identification of insertion events at the TS10Cas10 locus. A gel picture indicates the presence of insertion events at the TS10Cas10-1 target site (lane 02 A1). PCR reaction of HR1 and HR2 junction loaded next to each other (lane 02-white label and lane 02-gray label), with white label representing HR1 junction PCR, gray label representing HR2 junction PCR.

Figure 23 A-B. DNA expression cassettes used in gRNA/Cas9 mediated genome modification experiments. A) The Cas9 endonuclease cassette (EF1A2:CAS9) comprising a soybean EF1A2 promoter (GM-EF1A2 PRO) driving the soybean codon optimized Cas9 endonucleases (CAS9(SO)), a soybean optimized SV40 nuclear localization signal (SV40 NLS(SO)) and a PINII terminator (PINII TERM) was linked to a guide RNA expression cassette (U6-9.1:DD20CR1, comprising a soybean U6 promoter driving the DD20CR1 guide RNA) used in experiment U6-9.1DD20CR1 (Table 27). Other Guide RNA/Cas9 cassettes listed in Table 27 are identical except for the 20 bp variable targeting domains of the guide RNA targeting the genomic target sites DD20CR2, DD43CR1, or DD43CR2. B) The donor DNA cassette (DD20HR1-SAMS:HPT-DD20HR2) used in experiment U6-9.1DD20CR1 (Table 27). DD20HR1 and DD20HR2 homologous DNA regions between the donor DNA cassette and the genomic DNA sequences flanking the DD20 target site). Other Donor DNA cassettes listed in Table 27 are identical except for the DD43HR1 and DD43HR2 regions in two of them.

Figure 24 A-C. DD20 and DD43 soybean genomic target sites locations and qPCR amplicons. A) Diagram of Glycine max chromosome 04 indicating relative positions of DD20 and DD43 target sites. Genetic mapping positions of DD20 and DD43 sites are the positions of the most nearby genes Glyma04g39780.1 and

Glyma04g39550.1. B) DD20 qPCR 64 bp amplicon 45936307-45936370 from chromosome 04 (SEQ ID NO: 304). Relative positions of the target sites DD20-CR1 and DD20-CR2, qPCR primers and probe DD20-F, DD20-R, and DD20-T are marked. C) DD43 qPCR 115 bp amplicon 45731879-45731993 from chromosome
 5 04 (SEQ ID NO: 305). Relative positions of the target sites DD43-CR1 and DD43-CR2, qPCR primers and probe DD43-F2, DD43-F, DD43-R, and DD43-T are marked.

Figure 25 A-C. Schematic of guide RNA/Cas9 system mediated site-specific non-homologous end joining (NHEJ) and transgene insertion via homologous
 10 recombination (HR) at DD20CR1 site. A) Soybean plants are co-transformed with guide RNA/Cas9 and donor DNA cassettes as listed in Table 27. The DD20CR1 guide RNA/Cas9 complex transcribed from the linked guide RNA/Cas9 DNA cassettes will cleave specifically the DD20CR1 target site on chromosome 04 to make DNA double strand breaks. The breaks can be repaired spontaneously as
 15 NHEJs or repaired as a HR event by the donor DNA facilitated by the flanking homologous regions DD20-HR1 and DD20HR2. B) NHEJs are detected by DD20-specific qPCR and the mutated sequences are assessed by sequencing cloned HR1-HR2 PCR fragments. C) HR events are revealed by two border-specific PCR analyses HR1-SAMS and NOS-HR2, noting that the primers are only able to amplify
 20 DNA recombined between the DD20CR1 region of chromosome 04 and the donor DNA. Guide RNA/Cas9 mediated NHEJ and HR at DD20-CR2 site follow the same process except for using DD20-CR2 guide RNA. Guide RNA/Cas9 mediated site-specific NHEJ and HR at DD43CR1 and DD43CR2 sites follow the same process except for using guide RNA and homologous regions specific to the DD43 sites.

25 Figure 26 A-C. Sequences of gRNA/Cas9 system mediated NHEJs. Only 60 bp sequences surrounding the genomic target site shown in bold case are aligned to show the mutations. The PAM sequence is shown boxed in. Insertion sequences are indicated by symbol ^ marking the insertion position followed by the size of the insert. Actual insertion sequences are listed in the sequences listing. A) U6-
 30 9.1DD20CR1 sequences. Three colonies were sequenced for each of 54 events from experiment U6-9.1DD20CR1. A total of 150 sequences were returned, of which 26 were found to be short unique deletions while 2 of the events contained

small insertions. B) U6-9.1DD20CR2 sequences. Three colonies were sequenced for each of 28 events from experiment U6-9.1DD20CR2. A total of 84 sequences were returned, of which 20 were found to be short unique deletions while 1 of the events contained a single bp insertion. C) U6-9.1DD43CR1 sequences. Three colonies were sequenced for each of 46 events from experiment U6-9.1DD43CR1. A total of 132 sequences were returned, of which 18 were found to be short unique deletions while 10 of the events contained small insertions. D) U6-9.1DD43CR2 sequences.

Figure 27 A-C shows the ten most prevalent types of NHEJ mutations recovered based on the crRNA/tracrRNA/Cas endonuclease system. Figure 27A shows NHEJ mutations for LIGCas-1 target site, corresponding to SEQ ID NOs: 415-424), Figure 27B shows NHEJ mutations for LIGCas-2 target site corresponding to SEQ ID NOs: 425-434) and Figure 27V shows NHEJ mutations (for LIGCas-3 target site corresponding to SEQ ID NOs: 435-444).

Figure 28. Schematic representation of Zm-GOS2 PRO:GOS2 INTRON insertion in the 5'-UTR of maize ARGOS8 gene by targeting the guide RNA/Cas9 target sequence 1 (CTS1, SEQ ID NO: 1) with the gRNA1/Cas9 endonuclease system, described herein. HR1 and HR2 indicate homologous recombination regions.

Figure 29A-29C. Identification and analysis of Zm-GOS2 PRO:GOS2 INTRON insertion events in maize plants. (A) Schematic representation of Zm-GOS2 PRO:GOS2 INTRON insertion in the 5'-UTR of Zm-ARGOS8. CTS1 was targeted with the gRNA1/Cas9 endonuclease system, described herein. HR1 and HR2 indicate homologous recombination regions. P1 to P4 indicate PCR primers. (B) PCR screening of PMI-resistance calli to identify insertion events. PCR results are shown for 13 representative calli. The left and right junction PCRs were carried out with the primer pair P1+P2 and P3+P4, respectively. (C) PCR analysis of a T0 plant. A PCR product with the expected size (2.4 kb, Lane T0) was amplified with the primer P3 and P4.

Figure 30. Schematic representation of Zm-ARGOS8 promoter substitution with Zm-GOS2 PRO:GOS2 INTRON by targeting CTS3 (SEQ ID NO: 3) and CTS2 (SEQ ID NO:2). HR1 and HR2 indicate homologous recombination regions.

Figure 31A-31D. Substitution of the native promoter of the ARGOS8 gene with Zm-GOS2 PRO:GOS2 INTRON in maize plants. (A) Schematic representation of the Zm-GOS2 PRO:GOS2 INTRON:ARGOS8 allele generated by promoter swap. Two guide RNA/Cas9 target sites, CTS3 (SEQ ID NO:3) and CTS2 (SEQ ID NO:2), were targeted with a gRNA3/gRNA2/Cas9 system . HR1 and HR2 indicate homologous recombination regions. P1 to P5 indicate PCR primers. (B) PCR screening of PMI-resistance calli to identify swap events. PCR results are shown for 10 representative calli. One callus sample, 12A09, is positive for both left junction (L, primer P1+P2) and right junction (R, primer P5+P4) PCR, indicating that 12A09 is a swap event. (C) PCR analysis of the callus events identified in primary screening. PCR products with the expected size (2.4 kb) were amplified using the primer P3 and P4 from event #3, 4, 6, 8 and 9, indicating presence of the Zm-GOS2 PRO:GOS2 INTRON:ARGOS8 allele. (D) PCR analysis of a T0 plant. A PCR product with the expected size (2.4 kb, Lane T0) was amplified with the primer P3 and P4.

Figure 32A-32B. Deletion of the native promoter of the ARGOS8 gene in maize plants. (A) Schematic representation of promoter deletion. Two guide RNA's and a Cas9 endonuclease system, referred to as a gRNA3/gRNA2/Cas9 system, were used to target the CTS3 and CTS2 sites in Zm-ARGOS8. P1 and P4 indicate PCR primers for deletion event screening. (B) PCR screening of PMI-resistance calli to identify deletion events. PCR results are shown for 15 representative calli. A 1.1-kp PCR product indicates deletion of the CTS3/CTS2 fragment.

Figure 33. Schematic representation of enhancer element deletions using the guide RNA/Cas9 target sequence. The enhancer element to be deleted can be, but is not limited to, a 35S enhancer element.

Figure 34A-34C. Modification of a maize EPSPS polyubiquitination site. (A) The selected maize EPSPS polyubiquitination site is compared to the analogous sites of other plant species (SEQ ID NOs: 562-567). (B) The nucleotides to be edited in the maize EPSPS coding sequence (underlined, encoded amino acid shown in bold). The sequence shown in Figure 34B is listed in SEQ ID NO: 568. (C) The edited EPSPS coding sequence identified in the selected T0 plant. The sequence shown in Figure 34C is listed in SEQ ID NO: 569.

Figure 35A-35C. The intron mediated enhanced element (A). The 5' section of the first intron of the EPSPS gene (editing: substitutions underlined and deletions

represented by dots) (B) and its edited version conferring three IMEs elements (underlined). The edited nucleotides are shown in bold (C). The sequence shown in Figure 35B is listed in SEQ ID NO: 570. The sequence shown in Figure 35 C is listed in SEQ ID NO: 571.

5 Figure 36A-36B. Alternatively spliced EPSPS mRNA in maize cells. (A) left panel represents analysis of EPSPS cDNA. The lane I4 in Figure 36A shows amplification of the EPSPS pre-mRNA containing the 3rd intron unspliced (the 804bp diagnostic fragment as shown in Fig. 36 B indicates an alternate splicing event). Lanes E3 and F8 show the EPSPS PCR amplified fragments with spliced introns. 10 These diagnostic fragments are not amplified unless cDNA is synthesized (as is evident by the absence of bands in lanes E3, I4, and F8 comprising total RNA (shown in the total RNA panel on right of Fig 36A). The grey boxes in Fig. 36 B represent the eight EPSPS exons (their sizes are indicated above each of them).

 Figure 37. Splicing site at the junction between the second EPSPS intron 15 and the third exon (bolded). The nucleotide to be edited is underlined.

 Figure 38. Schematic representation of Southern hybridization analysis of T0 and T1 plants.

 Figure 39. Schematic representation of methods for producing a genetic modification (a gene insertion, a gene knockout via non-homologous-end-joining 20 (NHEJ), a gene editing, or any combination thereof) into a second gene of a plant genome without introducing an exogenous selectable marker into said plant genome. A first endogenous gene (located in a first locus-locus 1) can be modified to confer herbicide resistance using a first guide polynucleotide/Cas endonuclease system together with a first polynucleotide modification template. At the same time, 25 a second guide polynucleotide/Cas endonuclease system is provided to the same cell, with or without a second polynucleotide modification template to enable the second guide polynucleotide/Cas endonuclease system to introduce a double strand break at a second target site in the genome of said plant cell, resulting in either a gene insertion, gene knockout (NHEJ) or a gene editing. The second gene can be 30 located at different locus (say locus 2) than the first endogenous gene. A “ * “ represents at least one nucleotide alteration in the polynucleotide modification template (PMT). A second PMT can comprise a polynucleotide of interest (shown by a white box) for gene insertion.

Sequences

SEQ ID NO: 1 is the nucleotide sequence of the Cas9 gene from *Streptococcus pyogenes* M1 GAS (SF370).

SEQ ID NO: 2 is the nucleotide sequence of the potato ST-LS1 intron.

5 SEQ ID NO: 3 is the amino acid sequence of SV40 amino N-terminal.

SEQ ID NO: 4 is the amino acid sequence of *Agrobacterium tumefaciens* bipartite VirD2 T-DNA border endonuclease carboxyl terminal.

SEQ ID NO: 5 is the nucleotide sequence of an expression cassette expressing the maize optimized Cas9.

10 SEQ ID NO: 6 is the nucleotide sequence of crRNA containing the LIGCas-3 target sequence in the variable targeting domain.

SEQ ID NO: 7 is the nucleotide sequence of the tracrRNA.

SEQ ID NO: 8 is the nucleotide sequence of a long guide RNA containing the LIGCas-3 target sequence in the variable targeting domain.

15 SEQ ID NO: 9 is the nucleotide sequence of the Chromosome 8 maize U6 polymerase III promoter.

SEQ ID NO: 10 list two copies of the nucleotide sequence of the maize U6 polymerase III terminator.

20 SEQ ID NO: 11 is the nucleotide sequence of the maize optimized short guide RNA containing the LIGCas-3 variable targeting domain.

SEQ ID NO: 12 is the nucleotide sequence of the maize optimized long guide RNA expression cassette containing the LIGCas-3 variable targeting domain.

SEQ ID NO: 13 is the nucleotide sequence of the Maize genomic target site MS26Cas-1 plus PAM sequence.

25 SEQ ID NO: 14 is the nucleotide sequence of the Maize genomic target site MS26Cas-2 plus PAM sequence.

SEQ ID NO: 15 is the nucleotide sequence of the Maize genomic target site MS26Cas-3 plus PAM sequence.

30 SEQ ID NO: 16 is the nucleotide sequence of the Maize genomic target site LIGCas-2 plus PAM sequence.

SEQ ID NO: 17 is the nucleotide sequence of the Maize genomic target site LIGCas-3 plus PAM sequence.

SEQ ID NO: 18 is the nucleotide sequence of the Maize genomic target site LIGCas-4 plus PAM sequence.

SEQ ID NO: 19 is the nucleotide sequence of the Maize genomic target site MS45Cas-1 plus PAM sequence.

5 SEQ ID NO: 20 is the nucleotide sequence of the Maize genomic target site MS45Cas-2 plus PAM sequence.

SEQ ID NO: 21 is the nucleotide sequence of the Maize genomic target site MS45Cas-3 plus PAM sequence.

10 SEQ ID NO: 22 is the nucleotide sequence of the Maize genomic target site ALSCas-1 plus PAM sequence.

SEQ ID NO: 23 is the nucleotide sequence of the Maize genomic target site ALSCas-2 plus PAM sequence.

SEQ ID NO: 24 is the nucleotide sequence of the Maize genomic target site ALSCas-3 plus PAM sequence.

15 SEQ ID NO: 25 is the nucleotide sequence of the Maize genomic target site EPSPSCas-1 plus PAM sequence.

SEQ ID NO: 26 is the nucleotide sequence of the Maize genomic target site EPSPSCas-2 plus PAM sequence.

20 SEQ ID NO: 27 is the nucleotide sequence of the Maize genomic target site EPSPSCas-3 plus PAM sequence.

SEQ ID NOs: 28 -52 are the nucleotide sequence of target site specific forward primers for primary PCR as shown in Table 2.

SEQ ID NO: 53 is the nucleotide sequence of the forward primer for secondary PCR.

25 SEQ ID NO: 54 is the nucleotide sequence of Reverse primer for secondary PCR

SEQ ID NO: 55 is the nucleotide sequence of the unmodified reference sequence for LIGCas-1 and LIGCas-2 locus.

30 SEQ ID NOs: 56 -65 are the nucleotide sequences of mutations 1-10 for LIGCas-1.

SEQ ID NOs: 66 -75 are the nucleotide sequences of mutations 1-10 for LIGCas-2.

SEQ ID NO: 76 is the nucleotide sequence of the unmodified reference sequence for the LIGCas-3 and LIG3-4 homing endonuclease locus.

SEQ ID NOs: 77-86 are the nucleotide sequences of mutations 1-10 for LIGCas-3.

5 SEQ ID NOs: 88-96 are the nucleotide sequences of mutations 1-10 for LIG3-4 homing endonuclease locus.

SEQ ID NO: 97 is the nucleotide sequence of a donor vector referred to as an HR Repair DNA.

10 SEQ ID NO: 98 is the nucleotide sequence of forward PCR primer for site-specific transgene insertion at junction 1.

SEQ ID NO: 99 is the nucleotide sequence of reverse PCR primer for site-specific transgene insertion at junction 1.

SEQ ID NO: 100 is the nucleotide sequence of forward PCR primer for site-specific transgene insertion at junction 2.

15 SEQ ID NO: 101 is the nucleotide sequence of reverse PCR primer for site-specific transgene insertion at junction 2.

SEQ ID NO: 102 is the nucleotide sequence of the linked Cas9 endonuclease and LIGCas-3 long guide RNA expression cassettes

20 SEQ ID NO: 103 is the nucleotide sequence of Maize genomic target site 55CasRNA-1 plus PAM sequence.

SEQ ID NO: 104 is the nucleotide sequence of the unmodified reference sequence for 55CasRNA-1 locus.

SEQ ID NOs: 105-110 are the nucleotide sequences of mutations 1-6 for 55CasRNA-1.

25 SEQ ID NO: 111 is the nucleotide sequence of LIG3-4 homing endonuclease target site

SEQ ID NO: 112 is the nucleotide sequence of LIG3-4 homing endonuclease coding sequence.

30 SEQ ID NO: 113 is the nucleotide sequence of the MS26++ homing endonuclease target site.

SEQ ID NO: 114 is the nucleotide sequence of MS26++ homing endonuclease coding sequence

SEQ ID NO: 115 is the nucleotide sequence of the soybean codon optimized Cas9 gene.

SEQ ID NO: 116 is the nucleotide sequence of the soybean constitutive promoter GM-EF1A2.

5 SEQ ID NO: 117 is the nucleotide sequence of linker SV40 NLS.

SEQ ID NO: 118 is the amino acid sequence of soybean optimized Cas9 with a SV40 NLS.

SEQ ID NO: 119 is the nucleotide sequence of vector QC782.

10 SEQ ID NO: 120 is the nucleotide sequence of soybean U6 polymerase III promoter described herein, GM-U6-13.1 PRO.

SEQ ID NO: 121 is the nucleotide sequence of the guide RNA in Figure 8B.

SEQ ID NO: 122 is the nucleotide sequence of vector QC783.

SEQ ID NO: 123 is the nucleotide sequence of vector QC815.

15 SEQ ID NO: 124 is the nucleotide sequence of a Cas9 endonuclease (cas9-2) from *S. pyogenes*.

SEQ ID NO: 125 is the nucleotide sequence of the DD20CR1 soybean target site

SEQ ID NO: 126 is the nucleotide sequence of the DD20CR2 soybean target site

20 SEQ ID NO: 127 is the nucleotide sequence of the DD43CR1 soybean target site

SEQ ID NO: 128 is the nucleotide sequence of the DD43CR2 soybean target site

25 SEQ ID NO: 129 is the nucleotide sequence of the DD20 sequence in Figure 10A.

SEQ ID NO: 130 is the nucleotide sequence of the DD20 sequence complementary in Figure 10A.

SEQ ID NO: 131 is the nucleotide sequence of DD43 sequence.

30 SEQ ID NO: 132 is the nucleotide sequence of the DD43 complementary sequence.

SEQ ID NO: 133- 141 are primer sequences.

SEQ ID NO: 142 is the nucleotide sequence of the DD20CR1 PCR amplicon.

SEQ ID NO: 143 is the nucleotide sequence of the DD20CR2 PCR amplicon.

SEQ ID NO: 144 is the nucleotide sequence of the DD43CR1 PCR amplicon.

SEQ ID NO: 145 is the nucleotide sequence of the DD43CR2 PCR amplicon.

SEQ ID NO: 146 is the nucleotide sequence of the DD43CR2 PCR amplicon.

5 SEQ ID NO: 147- 156 are the nucleotide sequence of mutations 1 to 10 for the DD20CR1 target site

SEQ ID NO: 157- 166 are the nucleotide sequence of mutations 1 to 10 for the DD20CR2 target site

10 SEQ ID NO: 167- 176 are the nucleotide sequence of mutations 1 to 10 for the DD43CR1 target site

SEQ ID NO: 177- 191 are the nucleotide sequence of mutations 1 to 10 for the DD43CR2 target site.

SEQ ID NO: 192 is the amino acid sequence of a maize optimized version of the Cas9 protein.

15 SEQ ID NO: 193 is the nucleotide sequence of the maize optimized version of the Cas9 gene of SEQ ID NO: 192.

SEQ ID NO: 194 is the DNA version of guide RNA (EPSPS sgRNA).

SEQ ID NO: 195 is the EPSPS polynucleotide modification template.

20 SEQ ID NO: 196 is a nucleotide fragment comprising the TIPS nucleotide modifications.

SEQ ID NO: 197-204 are primer sequences shown in Table 15.

SEQ ID NO: 205-208 are nucleotide fragments shown in Figure 14.

SEQ ID NO: 209 is an example of a TIPS edited *EPSPS* nucleotide sequence fragment shown in Figure 17.

25 SEQ ID NO: 210 is an example of a Wild-type *EPSPS* nucleotide sequence fragment shown in Figure 17.

SEQ ID NO: 211 is the nucleotide sequence of a maize enolpyruvylshikimate-3-phosphate synthase (*epsps*) locus

30 SEQ ID NO: 212 is the nucleotide sequence of a Cas9 endonuclease (genbank CS571758.1) from *S. thermophiles*.

SEQ ID NO: 213 is the nucleotide sequence of a Cas9 endonuclease (genbank CS571770.1) from *S. thermophiles*.

SEQ ID NO: 214 is the nucleotide sequence of a Cas9 endonuclease (genbank CS571785.1) from *S. agalactiae*.

SEQ ID NO: 215 is the nucleotide sequence of a Cas9 endonuclease, (genbank CS571790.1) from *S. agalactiae*.

5 SEQ ID NO: 216 is the nucleotide sequence of a Cas9 endonuclease (genbank CS571790.1) from *S. mutant*.

SEQ ID NOs: 217-228 are primer and probe nucleotide sequences described in Example 17.

10 SEQ ID NOs: 229 is the nucleotide sequence of the MHP14Cas1 target site.
 SEQ ID NOs: 230 is the nucleotide sequence of the MHP14Cas3 target site.
 SEQ ID NOs: 231 is the nucleotide sequence of the TS8Cas1 target site.
 SEQ ID NOs: 232 is the nucleotide sequence of the TS8Cas2 target site.
 SEQ ID NOs: 233 is the nucleotide sequence of the TS9Cas2 target site.
 SEQ ID NOs: 234 is the nucleotide sequence of the TS9Cas3 target site.
 15 SEQ ID NOs: 235 is the nucleotide sequence of the TS10Cas1 target site.
 SEQ ID NOs: 236 is the nucleotide sequence of the TS10Cas3 target site.
 SEQ ID NOs: 237-244 are the nucleotide sequences shown in Figure 19A-D.
 SEQ ID NOs: 245-252 are the nucleotide sequences of the guide RNA expression cassettes described in Example 18.

20 SEQ ID NOs: 253-260 are the nucleotide sequences of donor DNA expression cassettes described in Example 18.

SEQ ID NOs: 261-270 are the nucleotide sequences of the primers described in Example 18.

25 SEQ ID NOs: 271-294 are the nucleotide sequences of the primers and probes described in Example 18.

SEQ ID NO: 295 is the nucleotide sequence of GM-U6-13.1 PRO, a soybean U6 polymerase III promoter described herein,

SEQ ID NOs: 298, 300, 301 and 303 are the nucleotide sequences of the linked guideRNA/Cas9 expression cassettes.

30 SEQ ID NOs: 299 and 302 are the nucleotide sequences of the donor DNA expression cassettes.

SEQ ID NOs: 271-294 are the nucleotide sequences of the primers and probes described in Example 18.

SEQ ID NO: 304 is the nucleotide sequence of the DD20 qPCR amplicon.

SEQ ID NO: 305 is the nucleotide sequence of the DD43 qPCR amplicon.

5 SEQ ID NOs: 306-328 are the nucleotide sequences of the primers and probes described herein.

SEQ ID NOs: 329- 334 are the nucleotide sequences of PCR amplicons described herein.

10 SEQ ID NO: 335 is the nucleotide sequence of a soybean genomic region comprising the DD20CR1 target site.

SEQ ID NO: 364 is the nucleotide sequence of a soybean genomic region comprising the DD20CR2 target site.

SEQ ID NO: 386 is the nucleotide sequence of a soybean genomic region comprising the DD43CR1 target site.

15 SEQ ID NOs: 336-363, 365-385 and 387-414 are the nucleotide sequences of shown in Figure 26 A-C.

SEQ ID NOs: 415-444 are the nucleotide sequences of NHEJ mutations recovered based on the crRNA/tracrRNA/Cas endonuclease system shown in Figure 27A-C.

20 SEQ ID NO: 445-447 are the nucleotide sequence of the LIGCas-1, LIGCas2 and LIGCas3 crRNA expression cassettes, respectively.

SEQ ID NO: 448 is the nucleotide sequence of the tracrRNA expression cassette.

25 SEQ ID NO: 449 is the nucleotide sequence of LIGCas-2 forward primer for primary PCR

SEQ ID NO: 450 is the nucleotide sequence of LIGCas-3 forward primer for primary PCR.

SEQ ID NO: 451 is the nucleotide sequence of the maize genomic Cas9 endonuclease target site Zm-ARGOS8-CTS1.

30 SEQ ID NO: 452 is the nucleotide sequence of the maize genomic Cas9 endonuclease target site Zm-ARGOS8-CTS2.

SEQ ID NO: 453 is the nucleotide sequence of the maize genomic Cas9 endonuclease target site Zm-ARGOS8-CTS3

SEQ ID NOs: 454-458 are the nucleotide sequence of primers P1, P2, P3, P4, P5, respectively.

5 SEQ ID NO: 459 is the nucleotide sequence of a Primer Binding Site (PBS), a sequence to facilitate event screening.

SEQ ID NO: 460 is the nucleotide sequence of the Zm-GOS2 PRO-GOS2 INTRON, the maize GOS2 promoter and GOS2 intron1 including the promoter , 5'-UTR1, INTRON1 and 5'-UTR2.

10 SEQ ID NO:461 is the nucleotide sequence of the maize Zm-ARGOS8 promoter.

SEQ ID NO:462 is the nucleotide sequence of the maize Zm-ARGOS8 5'-UTR.

15 SEQ ID NO:463 is the nucleotide sequence of the maize Zm-ARGOS8 codon sequence

SEQ ID NO:464 is the nucleotide sequence of the maize Zm-GOS2 gene, including promoter, 5'-UTR, CDS, 3'-UTR and introns.

SEQ ID NO:465 is the nucleotide sequence of the maize Zm-GOS2 PRO promoter.

20 SEQ ID NO:466 is the nucleotide sequence of the maize GOS2 INTRON, maize GOS2 5'-UTR1 and intron1 and 5'-UTR2.

25 SEQ ID NOs: 467-468, 490-491, 503-504 are the nucleotide sequence of the soybean genomic Cas endonuclease target sequences soy EPSPS-CR1, soy EPSPS-CR2, soy EPSPS-CR4, soy EPSPS-CR5, soy EPSPS-CR6, soy EPSPS-CR7, respectively

SEQ ID NO:469 is the nucleotide sequence of the soybean U6 small nuclear RNA promoter GM-U6-13.1.

SEQ ID NOs:470, 471 are the nucleotide sequences of the QC868, QC879 plasmids, respectively.

30 SEQ ID NOs:472, 473, 492, 493, 494, 505, 506, 507 are the nucleotide sequences of the RTW1013A, RTW1012A, RTW1199, RTW1200, RTW1190A, RTW1201, RTW1202, RTW1192A respectively.

SEQ ID NOs:474-488 , 495-402 , 508-512 are the nucleotide sequences of primers and probes.

SEQ ID NO: 489 is the nucleotide sequence of the soybean codon optimized Cas9.

5 SEQ ID NO: 513 is the nucleotide sequence of the 35S enhancer.

SEQ ID NO: 514 is the nucleotide sequence of the 35S-CRTS for gRNA1 at 163-181 (including pam at 3'end).

SEQ ID NO: 515 is the nucleotide sequence of the 35S-CRTS for gRNA2 at 295-319 (including pam at 3'end).

10 SEQ ID NO: 516 is the nucleotide sequence of the 35S-CRT for gRNA3 at 331-350 (including pam at 3'end).

SEQ ID NO: 517 is the nucleotide sequence of the EPSPS-K90R template.

SEQ ID NO: 518 is the nucleotide sequence of the EPSPS-IME template. S

SEQ ID NO: 519 is the nucleotide sequence of the EPSPS-Tspliced template.

15 SEQ ID NO: 520 is the amino acid sequence of ZM-RAP2.7 peptide

SEQ ID NO: 521 is the nucleotide sequence ZM-RAP2.7 coding DNA sequence

SEQ ID NOs: 522 is the amino acid sequence of ZM-NPK1B peptide

20 SEQ ID NO: 523 is the nucleotide sequence of the ZM-NPK1B coding DNA sequence

SEQ ID NOs: 524 is the nucleotide sequence of the RAB17 promoter

SEQ ID NOs: 525 is the amino acid sequence of the Maize FTM1.

SEQ ID NO: 526 is the nucleotide sequence of the Maize FTM1 coding DNA sequence.

25 SEQ ID NOs: 527-532 are the nucleotide sequences shown in Figures 34, 35 and 37.

SEQ ID NOs: 533-534 are the nucleotide sequences of the Southern genomic probe and Southern MoPAT probe of Figure 38, respectively. SEQ ID

NOs: 535-541 are the nucleotide sequences of the RF-FPCas-1, RF-FPCas-2,

30 ALSCas-4, ALS modification repair template 804, ALS modification repair template 127, ALS Forward_primer and ALS Reverse_primer, respectively.

SEQ ID NOs: 542-549 are the nucleotide sequences of the soy ALS1-CR1, Cas9 target sequence, soy ALS2-CR2, Cas9 target sequence, QC880, QC881, RTW1026A, WOL900, Forward_primer, WOL578, Reverse_primer and WOL573, Forward_primer, respectively.

5 SEQ ID NO: 550 is the nucleotide sequence of a maize ALS protein.

SEQ ID NOs: 551-554 are the nucleotide sequences of the soy FAD2-1-CR1, Cas9 target sequence, soy FAD2-1-CR2, Cas9 target sequence, RTW1211 and RTW1212.

DETAILED DESCRIPTION

10 The present disclosure includes compositions and methods for genome modification of a target sequence in the genome of a plant or plant cell, for selecting plants, for gene editing, and for inserting a polynucleotide of interest into the genome of a plant without incorporating a selectable transgene marker. The methods employ a guide polynucleotide/Cas endonuclease system, wherein the

15 Cas endonuclease is guided by at least two guide polynucleotides to recognize and optionally introduce a double strand break at a specific first and second target site into the genome of a cell. The guide polynucleotide /Cas endonuclease system provides for an effective system for modifying target sites within the genome of a plant, plant cell or seed. Further provided are methods and compositions employing

20 a guide polynucleotide/Cas endonuclease for editing a nucleotide sequence in the genome of a cell. Once a genomic target site is identified, a variety of methods can be employed to further modify the target sites such that they contain a variety of polynucleotides of interest. Breeding methods utilizing a two component guide RNA/Cas endonuclease system are also disclosed. Compositions and methods are

25 also provided for editing a nucleotide sequence in the genome of a cell. The nucleotide sequence to be edited (the nucleotide sequence of interest) can be located within or outside a target site that is recognized by a Cas endonuclease.

CRISPR loci (Clustered Regularly Interspaced Short Palindromic Repeats) (also known as SPIDRs--SPacer Interspersed Direct Repeats) constitute a family of

30 recently described DNA loci. CRISPR loci consist of short and highly conserved DNA repeats (typically 24 to 40 bp, repeated from 1 to 140 times-also referred to as CRISPR-repeats) which are partially palindromic. The repeated sequences (usually

specific to a species) are interspaced by variable sequences of constant length (typically 20 to 58 by depending on the CRISPR locus (WO2007/025097 published March 1, 2007).

CRISPR loci were first recognized in *E. coli* (Ishino et al. (1987) *J. Bacterial.* 169:5429-5433; Nakata et al. (1989) *J. Bacterial.* 171:3553-3556). Similar interspersed short sequence repeats have been identified in *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena*, and *Mycobacterium tuberculosis* (Groenen et al. (1993) *Mol. Microbiol.* 10:1057-1065; Hoe et al. (1999) *Emerg. Infect. Dis.* 5:254-263; Masepohl et al. (1996) *Biochim. Biophys. Acta* 1307:26-30; Mojica et al. (1995) *Mol. Microbiol.* 17:85-93). The CRISPR loci differ from other SSRs by the structure of the repeats, which have been termed short regularly spaced repeats (SRSRs) (Janssen et al. (2002) *OMICS J. Integ. Biol.* 6:23-33; Mojica et al. (2000) *Mol. Microbiol.* 36:244-246). The repeats are short elements that occur in clusters, which are always regularly spaced by variable sequences of constant length (Mojica et al. (2000) *Mol. Microbiol.* 36:244-246).

Cas gene includes a gene that is generally coupled, associated or close to or in the vicinity of flanking CRISPR loci. The terms "Cas gene", "CRISPR-associated (Cas) gene" are used interchangeably herein. A comprehensive review of the Cas protein family is presented in Haft et al. (2005) *Computational Biology, PLoS Comput Biol* 1(6): e60. doi:10.1371/journal.pcbi.0010060.

As described therein, 41 CRISPR-associated (Cas) gene families are described, in addition to the four previously known gene families. It shows that CRISPR systems belong to different classes, with different repeat patterns, sets of genes, and species ranges. The number of Cas genes at a given CRISPR locus can vary between species.

Cas endonuclease relates to a Cas protein encoded by a Cas gene, wherein said Cas protein is capable of introducing a double strand break into a DNA target sequence. The Cas endonuclease is guided by the guide polynucleotide to recognize and optionally introduce a double strand break at a specific target site into the genome of a cell. As used herein, the term "guide polynucleotide/Cas endonuclease system" includes a complex of a Cas endonuclease and a guide polynucleotide that is capable of introducing a double strand break into a DNA target

sequence. The Cas endonuclease unwinds the DNA duplex in close proximity of the genomic target site and cleaves both DNA strands upon recognition of a target sequence by a guide RNA, but only if the correct protospacer-adjacent motif (PAM) is approximately oriented at the 3' end of the target sequence (Figure 2A, Figure 2B).

In one embodiment, the Cas endonuclease gene is a Cas9 endonuclease, such as but not limited to, Cas9 genes listed in SEQ ID NOs: 462, 474, 489, 494, 499, 505, and 518 of WO2007/025097 published March 1, 2007.

In another embodiment, the Cas endonuclease gene is plant, maize or soybean optimized Cas9 endonuclease (Figure 1 A).. In another embodiment, the Cas endonuclease gene is operably linked to a SV40 nuclear targeting signal upstream of the Cas codon region and a bipartite VirD2 nuclear localization signal (Tinland et al. (1992) Proc. Natl. Acad. Sci. USA 89:7442-6) downstream of the Cas codon region.

In one embodiment, the Cas endonuclease gene is a Cas9 endonuclease gene of SEQ ID NO:1, 124, 212, 213, 214, 215, 216, 193 or nucleotides 2037-6329 of SEQ ID NO:5, or any functional fragment or variant thereof.

The terms “functional fragment”, “fragment that is functionally equivalent” and “functionally equivalent fragment” are used interchangeably herein. These terms refer to a portion or subsequence of the Cas endonuclease sequence of the present disclosure in which the ability to create a double-strand break is retained.

The terms “functional variant”, “Variant that is functionally equivalent” and “functionally equivalent variant” are used interchangeably herein. These terms refer to a variant of the Cas endonuclease of the present disclosure in which the ability to create a double-strand break is retained. Fragments and variants can be obtained via methods such as site-directed mutagenesis and synthetic construction.

In one embodiment, the Cas endonuclease gene is a plant codon optimized *streptococcus pyogenes* Cas9 gene that can recognize any genomic sequence of the form N(12-30)NGG can in principle be targeted.

In one embodiment, the Cas endonuclease is introduced directly into a cell by any method known in the art, for example, but not limited to transient introduction methods, transfection and/or topical application.

Endonucleases are enzymes that cleave the phosphodiester bond within a polynucleotide chain, and include restriction endonucleases that cleave DNA at specific sites without damaging the bases. Restriction endonucleases include Type I, Type II, Type III, and Type IV endonucleases, which further include subtypes. In the Type I and Type III systems, both the methylase and restriction activities are contained in a single complex. Endonucleases also include meganucleases, also known as homing endonucleases (HEases), which like restriction endonucleases, bind and cut at a specific recognition site, however the recognition sites for meganucleases are typically longer, about 18 bp or more. (patent application WO-PCT PCT/US12/30061 filed on March 22, 2012) Meganucleases have been classified into four families based on conserved sequence motifs, the families are the LAGLIDADG, GIY-YIG, H-N-H, and His-Cys box families. These motifs participate in the coordination of metal ions and hydrolysis of phosphodiester bonds. HEases are notable for their long recognition sites, and for tolerating some sequence polymorphisms in their DNA substrates. The naming convention for meganuclease is similar to the convention for other restriction endonuclease. Meganucleases are also characterized by prefix F-, I-, or PI- for enzymes encoded by free-standing ORFs, introns, and inteins, respectively. One step in the recombination process involves polynucleotide cleavage at or near the recognition site. This cleaving activity can be used to produce a double-strand break. For reviews of site-specific recombinases and their recognition sites, see, Sauer (1994) *Curr Op Biotechnol* 5:521-7; and Sadowski (1993) *FASEB* 7:760-7. In some examples the recombinase is from the Integrase or Resolvase families.

TAL effector nucleases are a new class of sequence-specific nucleases that can be used to make double-strand breaks at specific target sequences in the genome of a plant or other organism. (Miller *et al.* (2011) *Nature Biotechnology* 29:143–148). Zinc finger nucleases (ZFNs) are engineered double-strand break inducing agents comprised of a zinc finger DNA binding domain and a double-strand-break-inducing agent domain. Recognition site specificity is conferred by the zinc finger domain, which typically comprising two, three, or four zinc fingers, for example having a C2H2 structure, however other zinc finger structures are known and have been engineered. Zinc finger domains are amenable for designing

polypeptides which specifically bind a selected polynucleotide recognition sequence. ZFNs include an engineered DNA-binding zinc finger domain linked to a non-specific endonuclease domain, for example nuclease domain from a Type IIs endonuclease such as FokI. Additional functionalities can be fused to the zinc-
 5 finger binding domain, including transcriptional activator domains, transcription repressor domains, and methylases. In some examples, dimerization of nuclease domain is required for cleavage activity. Each zinc finger recognizes three consecutive base pairs in the target DNA. For example, a 3 finger domain recognized a sequence of 9 contiguous nucleotides, with a dimerization requirement
 10 of the nuclease, two sets of zinc finger triplets are used to bind an 18 nucleotide recognition sequence.

Bacteria and archaea have evolved adaptive immune defenses termed clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems that use short RNA to direct degradation of foreign
 15 nucleic acids ((WO2007/025097 published March 1, 2007). The type II CRISPR/Cas system from bacteria employs a crRNA and tracrRNA to guide the Cas endonuclease to its DNA target. The crRNA (CRISPR RNA) contains the region complementary to one strand of the double strand DNA target and base pairs with the tracrRNA (trans-activating CRISPR RNA) forming a RNA duplex that directs the
 20 Cas endonuclease to cleave the DNA target (Figure 2 B).

As used herein, the term "guide RNA" includes a synthetic fusion of two RNA molecules, a crRNA (CRISPR RNA) comprising a variable targeting domain, and a tracrRNA (Figure 2 B). In one embodiment, the guide RNA comprises a variable
 25 targeting domain of 12 to 30 nucleotide sequences and a RNA fragment that can interact with a Cas endonuclease.

As used herein, the term "guide polynucleotide", includes a polynucleotide sequence that can form a complex with a Cas endonuclease and enables the Cas endonuclease to recognize and optionally cleave a DNA target site. The guide polynucleotide can be included a single molecule or a double molecule. The guide
 30 polynucleotide sequence can be a RNA sequence, a DNA sequence, or a combination thereof (a RNA-DNA combination sequence). Optionally, the guide polynucleotide can comprise at least one nucleotide, phosphodiester bond or

linkage modification such as, but not limited, to Locked Nucleic Acid (LNA), 5-methyl dC, 2,6-Diaminopurine, 2'-Fluoro A, 2'-Fluoro U, 2'-O-Methyl RNA, phosphorothioate bond, linkage to a cholesterol molecule, linkage to a polyethylene glycol molecule, linkage to a spacer 18 (hexaethylene glycol chain) molecule, or 5' to 3' covalent
 5 linkage resulting in circularization. A guide polynucleotide that solely comprises ribonucleic acids is also referred to as a "guide RNA".

The guide polynucleotide can be a double molecule (also referred to as duplex guide polynucleotide) comprising a first nucleotide sequence domain (referred to as Variable Targeting domain or VT domain) that is complementary to a
 10 nucleotide sequence in a target DNA and a second nucleotide sequence domain (referred to as Cas endonuclease recognition domain or CER domain) that interacts with a Cas endonuclease polypeptide. The CER domain of the double molecule guide polynucleotide comprises two separate molecules that are hybridized along a region of complementarity. The two separate molecules can be RNA, DNA, and/or
 15 RNA-DNA- combination sequences. In some embodiments, the first molecule of the duplex guide polynucleotide comprising a VT domain linked to a CER domain is referred to as "crDNA" (when composed of a contiguous stretch of DNA nucleotides) or "crRNA" (when composed of a contiguous stretch of RNA nucleotides), or "crDNA-RNA" (when composed of a combination of DNA and RNA nucleotides).
 20 The crNucleotide can comprise a fragment of the crRNA naturally occurring in Bacteria and Archaea. In one embodiment, the size of the fragment of the crRNA naturally occurring in Bacteria and Archaea that is present in a crNucleotide disclosed herein can range from, but is not limited to, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more nucleotides. In some embodiments the
 25 second molecule of the duplex guide polynucleotide comprising a CER domain is referred to as "tracrRNA" (when composed of a contiguous stretch of RNA nucleotides) or "tracrDNA" (when composed of a contiguous stretch of DNA nucleotides) or "tracrDNA-RNA" (when composed of a combination of DNA and RNA nucleotides). In one embodiment, the RNA that guides the RNA/ Cas9
 30 endonuclease complex, is a duplexed RNA comprising a duplex crRNA-tracrRNA..

The guide polynucleotide can also be a single molecule comprising a first nucleotide sequence domain (referred to as Variable Targeting domain or VT

domain) that is complementary to a nucleotide sequence in a target DNA and a second nucleotide domain (referred to as Cas endonuclease recognition domain or CER domain) that interacts with a Cas endonuclease polypeptide. By “domain” it is meant a contiguous stretch of nucleotides that can be RNA, DNA, and/or RNA-DNA-combination sequence. The VT domain and / or the CER domain of a single guide polynucleotide can comprise a RNA sequence, a DNA sequence, or a RNA-DNA-combination sequence. In some embodiments the single guide polynucleotide comprises a crNucleotide (comprising a VT domain linked to a CER domain) linked to a tracrNucleotide (comprising a CER domain), wherein the linkage is a nucleotide sequence comprising a RNA sequence, a DNA sequence, or a RNA-DNA combination sequence. The single guide polynucleotide being comprised of sequences from the crNucleotide and tracrNucleotide may be referred to as “single guide RNA” (when composed of a contiguous stretch of RNA nucleotides) or “single guide DNA” (when composed of a contiguous stretch of DNA nucleotides) or “single guide RNA-DNA” (when composed of a combination of RNA and DNA nucleotides). In one embodiment of the disclosure, the single guide RNA comprises a crRNA or crRNA fragment and a tracrRNA or tracrRNA fragment of the type II CRISPR/Cas system that can form a complex with a type II Cas endonuclease, wherein said guide RNA/Cas endonuclease complex can direct the Cas endonuclease to a plant genomic target site, enabling the Cas endonuclease to introduce a double strand break into the genomic target site. One aspect of using a single guide polynucleotide versus a duplex guide polynucleotide is that only one expression cassette needs to be made to express the single guide polynucleotide.

The term “variable targeting domain” or “VT domain” is used interchangeably herein and includes a nucleotide sequence that is complementary to one strand (nucleotide sequence) of a double strand DNA target site (Figure 2 A and 2 B). The % complementation between the first nucleotide sequence domain (VT domain) and the target sequence can be at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 63%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%. The variable target domain can be at least 12, 13, 14, 15, 16, 17, 18,

19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 nucleotides in length. In some embodiments, the variable targeting domain comprises a contiguous stretch of 12 to 30 nucleotides. The variable targeting domain can be composed of a DNA sequence, a RNA sequence, a modified DNA sequence, a modified RNA sequence, or any combination thereof.

The term “Cas endonuclease recognition domain” or “CER domain” of a guide polynucleotide is used interchangeably herein and includes a nucleotide sequence (such as a second nucleotide sequence domain of a guide polynucleotide), that interacts with a Cas endonuclease polypeptide. The CER domain can be composed of a DNA sequence, a RNA sequence, a modified DNA sequence, a modified RNA sequence (see for example modifications described herein), or any combination thereof.

The nucleotide sequence linking the crNucleotide and the tracrNucleotide of a single guide polynucleotide can comprise a RNA sequence, a DNA sequence, or a RNA-DNA combination sequence. In one embodiment, the nucleotide sequence linking the crNucleotide and the tracrNucleotide of a single guide polynucleotide can be at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100 nucleotides in length. In another embodiment, the nucleotide sequence linking the crNucleotide and the tracrNucleotide of a single guide polynucleotide can comprise a tetraloop sequence, such as, but not limiting to a GAAA tetraloop sequence.

Nucleotide sequence modification of the guide polynucleotide, VT domain and/or CER domain can be selected from, but not limited to , the group consisting of a 5' cap, a 3' polyadenylated tail, a riboswitch sequence, a stability control sequence, a sequence that forms a dsRNA duplex, a modification or sequence that targets the guide poly nucleotide to a subcellular location, a modification or sequence that provides for tracking , a modification or sequence that provides a binding site for proteins , a Locked Nucleic Acid (LNA), a 5-methyl dC nucleotide, a 2,6-Diaminopurine nucleotide, a 2'-Fluoro A nucleotide, a 2'-Fluoro U nucleotide; a

2'-O-Methyl RNA nucleotide, a phosphorothioate bond, linkage to a cholesterol molecule, linkage to a polyethylene glycol molecule, linkage to a spacer 18 molecule, a 5' to 3' covalent linkage, or any combination thereof. These modifications can result in at least one additional beneficial feature, wherein the additional beneficial feature is selected from the group of a modified or regulated stability, a subcellular targeting, tracking, a fluorescent label, a binding site for a protein or protein complex, modified binding affinity to complementary target sequence, modified resistance to cellular degradation, and increased cellular permeability.

In one embodiment, the guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at a DNA target site

In one embodiment of the disclosure the variable target domain is 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 nucleotides in length.

In one embodiment of the disclosure, the guide RNA comprises a crRNA (or crRNA fragment) and a tracrRNA (or tracrRNA fragment) of the type II CRISPR/Cas system that can form a complex with a type II Cas endonuclease, wherein said guide RNA/Cas endonuclease complex can direct the Cas endonuclease to a plant genomic target site, enabling the Cas endonuclease to introduce a double strand break into the genomic target site.

In one embodiment the guide RNA can be introduced into a plant or plant cell directly using any method known in the art such as, but not limited to, particle bombardment or topical applications.

In another embodiment the guide RNA can be introduced indirectly by introducing a recombinant DNA molecule comprising the corresponding guide DNA sequence operably linked to a plant specific promoter (as shown in Figure 1 B) that is capable of transcribing the guide RNA in said plant cell. The term "corresponding guide DNA" includes a DNA molecule that is identical to the RNA molecule but has a "T" substituted for each "U" of the RNA molecule.

In some embodiments, the guide RNA is introduced via particle bombardment or *Agrobacterium* transformation of a recombinant DNA construct

comprising the corresponding guide DNA operably linked to a plant U6 polymerase III promoter.

In one embodiment, the RNA that guides the RNA/ Cas9 endonuclease complex, is a duplexed RNA comprising a duplex crRNA-tracrRNA (as shown in Figure 2B). One advantage of using a guide RNA versus a duplexed crRNA-tracrRNA is that only one expression cassette needs to be made to express the fused guide RNA.

The terms “target site”, “target sequence”, “target DNA”, “target locus”, “genomic target site”, “genomic target sequence”, and “genomic target locus” are used interchangeably herein and refer to a polynucleotide sequence in the genome (including chloroplastic and mitochondrial DNA) of a plant cell at which a double-strand break is induced in the plant cell genome by a Cas endonuclease. The target site can be an endogenous site in the plant genome, or alternatively, the target site can be heterologous to the plant and thereby not be naturally occurring in the genome, or the target site can be found in a heterologous genomic location compared to where it occurs in nature. As used herein, terms “endogenous target sequence” and “native target sequence” are used interchangeably herein to refer to a target sequence that is endogenous or native to the genome of a plant and is at the endogenous or native position of that target sequence in the genome of the plant.

In one embodiment, the target site can be similar to a DNA recognition site or target site that is specifically recognized and/or bound by a double-strand break inducing agent such as a LIG3-4 endonuclease (US patent publication 2009-0133152 A1 (published May 21, 2009) or a MS26++ meganuclease (U.S. patent application 13/526912 filed June 19, 2012).

An “artificial target site” or “artificial target sequence” are used interchangeably herein and refer to a target sequence that has been introduced into the genome of a plant. Such an artificial target sequence can be identical in sequence to an endogenous or native target sequence in the genome of a plant but be located in a different position (*i.e.*, a non-endogenous or non-native position) in the genome of a plant.

An “altered target site”, “altered target sequence”, “modified target site”, “modified target sequence” are used interchangeably herein and refer to a target sequence as disclosed herein that comprises at least one alteration when compared to non-altered target sequence. Such “alterations” include, for example:

- 5 (i) replacement of at least one nucleotide, (ii) a deletion of at least one nucleotide, (iii) an insertion of at least one nucleotide, or (iv) any combination of (i) – (iii).

Methods for modifying a plant genomic target site are disclosed herein. In one embodiment, the method comprises a method for producing a genetic modification into a second gene of a plant genome without introducing an
10 exogenous selectable marker into said plant genome, the method comprising providing a first guide polynucleotide, a polynucleotide modification template, a second guide polynucleotide, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be modified to confer herbicide resistance, wherein said first guide polynucleotide and Cas endonuclease are capable of forming a first
15 complex that enables the Cas endonuclease to introduce a double strand break at a first target site, located in or near said first endogenous gene in the genome of said plant cell, wherein said second guide polynucleotide and Cas endonuclease are capable of forming a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant
20 cell, wherein said polynucleotide modification template comprises at least one nucleotide alteration when compared to the first endogenous gene.

Also provided is a method for modifying a target site in the genome of a plant cell, the method comprising introducing a guide RNA and a Cas endonuclease into said plant, wherein said guide RNA and Cas endonuclease are capable of forming a
25 complex that enables the Cas endonuclease to introduce a double strand break at said target site.

Further provided is a method for modifying a target site in the genome of a plant cell, the method comprising introducing a guide RNA and a donor DNA into a plant cell having a Cas endonuclease, wherein said guide RNA and Cas
30 endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at said target site, wherein said donor DNA comprises a polynucleotide of interest.

Further provided is a method for modifying a target site in the genome of a plant cell, the method comprising: a) introducing into a plant cell a guide RNA comprising a variable targeting domain and a Cas endonuclease, wherein said guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at said target site; and, b) identifying at least one plant cell that has a modification at said target, wherein the modification includes at least one deletion or substitution of one or more nucleotides in said target site.

Further provided, a method for modifying a target DNA sequence in the genome of a plant cell, the method comprising: a) introducing into a plant cell a first recombinant DNA construct capable of expressing a guide RNA and a second recombinant DNA construct capable of expressing a Cas endonuclease, wherein said guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at said target site; and, b) identifying at least one plant cell that has a modification at said target, wherein the modification includes at least one deletion or substitution of one or more nucleotides in said target site.

The length of the target site can vary, and includes, for example, target sites that are at least 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more nucleotides in length. It is further possible that the target site can be palindromic, that is, the sequence on one strand reads the same in the opposite direction on the complementary strand. The nick/cleavage site can be within the target sequence or the nick/cleavage site could be outside of the target sequence. In another variation, the cleavage could occur at nucleotide positions immediately opposite each other to produce a blunt end cut or, in other cases, the incisions could be staggered to produce single-stranded overhangs, also called "sticky ends", which can be either 5' overhangs, or 3' overhangs.

In some embodiment, the genomic target site capable of being cleaved by a Cas endonuclease comprises a 12 to 30 nucleotide fragment of a male fertility gene such as MS26 (see for example U.S. Patents 7,098,388, 7,517,975, 7,612,251), MS45 (see for example U.S. Patents 5,478,369, 6,265,640) or MSCA1 (see for example U.S. Patent 7,919,676), ALS or ESPS genes.

Active variants of genomic target sites can also be used. Such active variants can comprise at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to the given target site, wherein the active variants retain biological activity and hence are capable of being
 5 recognized and cleaved by an Cas endonuclease. Assays to measure the double-strand break of a target site by an endonuclease are known in the art and generally measure the overall activity and specificity of the agent on DNA substrates containing recognition sites.

Various methods and compositions can be employed to obtain a plant having
 10 a polynucleotide of interest inserted in a target site for a Cas endonuclease. Such methods can employ homologous recombination to provide integration of the polynucleotide of Interest at the target site. In one method provided, a polynucleotide of interest is provided to the plant cell in a donor DNA construct. As used herein, "donor DNA" is a DNA construct that comprises a polynucleotide of
 15 Interest to be inserted into the target site of a Cas endonuclease. The donor DNA construct further comprises a first and a second region of homology that flank the polynucleotide of Interest. The first and second regions of homology of the donor DNA share homology to a first and a second genomic region, respectively, present in or flanking the target site of the plant genome. By "homology" is meant DNA
 20 sequences that are similar. For example, a "region of homology to a genomic region" that is found on the donor DNA is a region of DNA that has a similar sequence to a given "genomic region" in the plant genome. A region of homology can be of any length that is sufficient to promote homologous recombination at the cleaved target site. For example, the region of homology can comprise at least 5-
 25 10, 5-15, 5-20, 5-25, 5-30, 5-35, 5-40, 5-45, 5- 50, 5-55, 5-60, 5-65, 5- 70, 5-75, 5-80, 5-85, 5-90, 5-95, 5-100, 5-200, 5-300, 5-400, 5-500, 5-600, 5-700, 5-800, 5-900, 5-1000, 5-1100, 5-1200, 5-1300, 5-1400, 5-1500, 5-1600, 5-1700, 5-1800, 5-1900, 5-2000, 5-2100, 5-2200, 5-2300, 5-2400, 5-2500, 5-2600, 5-2700, 5-2800, 5-2900, 5-3000, 5-3100 or more bases in length such that the region of homology has
 30 sufficient homology to undergo homologous recombination with the corresponding genomic region. "Sufficient homology" indicates that two polynucleotide sequences have sufficient structural similarity to act as substrates for a homologous

recombination reaction. The structural similarity includes overall length of each polynucleotide fragment, as well as the sequence similarity of the polynucleotides. Sequence similarity can be described by the percent sequence identity over the whole length of the sequences, and/or by conserved regions comprising localized
 5 similarities such as contiguous nucleotides having 100% sequence identity, and percent sequence identity over a portion of the length of the sequences.

The amount of homology or sequence identity shared by a target and a donor polynucleotide can vary and includes total lengths and/or regions having unit integral values in the ranges of about 1-20 bp, 20-50 bp, 50-100 bp, 75-150 bp, 100-
 10 250 bp, 150-300 bp, 200-400 bp, 250-500 bp, 300-600 bp, 350-750 bp, 400-800 bp, 450-900 bp, 500-1000 bp, 600-1250 bp, 700-1500 bp, 800-1750 bp, 900-2000 bp, 1-2.5 kb, 1.5-3 kb, 2-4 kb, 2.5-5 kb, 3-6 kb, 3.5-7 kb, 4-8 kb, 5-10 kb, or up to and including the total length of the target site. These ranges include every integer within the range, for example, the range of 1-20 bp includes 1, 2, 3, 4, 5, 6, 7, 8, 9,
 15 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20 bp. The amount of homology can also be described by percent sequence identity over the full aligned length of the two polynucleotides which includes percent sequence identity of about at least 50%, 55%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%,
 20 95%, 96%, 97%, 98%, 99% or 100%. Sufficient homology includes any combination of polynucleotide length, global percent sequence identity, and optionally conserved regions of contiguous nucleotides or local percent sequence identity, for example sufficient homology can be described as a region of 75-150 bp having at least 80% sequence identity to a region of the target locus. Sufficient homology can also be
 25 described by the predicted ability of two polynucleotides to specifically hybridize under high stringency conditions, see, for example, Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, NY); *Current Protocols in Molecular Biology*, Ausubel *et al.*, Eds (1994) *Current Protocols*, (Greene Publishing Associates, Inc. and John Wiley & Sons, Inc); and,
 30 Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, (Elsevier, New York).

As used herein, a “genomic region” is a segment of a chromosome in the genome of a plant cell that is present on either side of the target site or, alternatively, also comprises a portion of the target site. The genomic region can comprise at least 5-10, 5-15, 5-20, 5-25, 5-30, 5-35, 5-40, 5-45, 5- 50, 5-55, 5-60, 5-65, 5- 70, 5-75, 5-80, 5-85, 5-90, 5-95, 5-100, 5-200, 5-300, 5-400, 5-500, 5-600, 5-700, 5-800, 5-900, 5-1000, 5-1100, 5-1200, 5-1300, 5-1400, 5-1500, 5-1600, 5-1700, 5-1800, 5-1900, 5-2000, 5-2100, 5-2200, 5-2300, 5-2400, 5-2500, 5-2600, 5-2700, 5-2800, 5-2900, 5-3000, 5-3100 or more bases such that the genomic region has sufficient homology to undergo homologous recombination with the corresponding region of homology.

Polynucleotides of interest and/or traits can be stacked together in a complex trait locus as described in US-2013-0263324-A1, published 03 Oct 2013 and in PCT/US13/22891, published January 24, 2013.

The guide polynucleotide/Cas9 endonuclease system described herein provides for an efficient system to generate double strand breaks and allows for traits to be stacked in a complex trait locus.

In one embodiment, the guide polynucleotide/Cas endonuclease system is used for introducing one or more polynucleotides of interest or one or more traits of interest into one or more target sites by providing one or more guide polynucleotides, one Cas endonuclease, and optionally one or more donor DNAs to a plant cell without incorporating a selectable transgene marker. A fertile plant can be produced from that plant cell that comprises an alteration at said one or more target sites, wherein the alteration is selected from the group consisting of (i) replacement of at least one nucleotide, (ii) a deletion of at least one nucleotide, (iii) an insertion of at least one nucleotide, and (iv) any combination of (i) – (iii). Plants comprising these altered target sites can be crossed with plants comprising at least one gene or trait of interest in the same complex trait locus, thereby further stacking traits in said complex trait locus. (see also US-2013-0263324-A1, published 03 Oct 2013 and in PCT/US13/22891, published January 24, 2013).

In one embodiment, the method comprises a method for producing in a plant a complex trait locus without incorporating a selectable transgene marker comprising at least two altered target sequences in a genomic region of interest,

said method comprising: (a) selecting a genomic region in a plant, wherein the genomic region comprises a first target sequence and a second target sequence; (b) contacting at least one plant cell with at least a first guide polynucleotide, a second polynucleotide, and optionally at least one donor DNA, and a Cas endonuclease, wherein the first and second guide polynucleotide and the Cas endonuclease can form a complex that enables the Cas endonuclease to introduce a double strand break in at least a first and a second target sequence; (c) identifying a cell from (b) comprising a first alteration at the first target sequence and a second alteration at the second target sequence; and (d) recovering a first fertile plant from the cell of (c) said fertile plant comprising the first alteration and the second alteration, wherein the first alteration and the second alteration are physically linked.

In one embodiment, the method comprises a method for producing in a plant a complex trait locus without incorporating a selectable transgene marker comprising at least two altered target sequences in a genomic region of interest, said method comprising: (a) selecting a genomic region in a plant, wherein the genomic region comprises a first target sequence and a second target sequence; (b) contacting at least one plant cell with a first guide polynucleotide, a Cas endonuclease, and optionally a first donor DNA, wherein the first guide polynucleotide and the Cas endonuclease can form a complex that enables the Cas endonuclease to introduce a double strand break a first target sequence; (c) identifying a cell from (b) comprising a first alteration at the first target sequence; (d) recovering a first fertile plant from the cell of (c), said first fertile plant comprising the first alteration; (e) contacting at least one plant cell with a second guide polynucleotide, a Cas endonuclease and optionally a second Donor DNA; (f) identifying a cell from (e) comprising a second alteration at the second target sequence; (g) recovering a second fertile plant from the cell of (f), said second fertile plant comprising the second alteration; and, (h) obtaining a fertile progeny plant from the second fertile plant of (g), said fertile progeny plant comprising the first alteration and the second alteration, wherein the first alteration and the second alteration are physically linked.

The structural similarity between a given genomic region and the corresponding region of homology found on the donor DNA can be any degree of

sequence identity that allows for homologous recombination to occur. For example, the amount of homology or sequence identity shared by the "region of homology" of the donor DNA and the "genomic region" of the plant genome can be at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 5 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, such that the sequences undergo homologous recombination

The region of homology on the donor DNA can have homology to any sequence flanking the target site. While in some embodiments the regions of homology share significant sequence homology to the genomic sequence 10 immediately flanking the target site, it is recognized that the regions of homology can be designed to have sufficient homology to regions that may be further 5' or 3' to the target site. In still other embodiments, the regions of homology can also have homology with a fragment of the target site along with downstream genomic regions. In one embodiment, the first region of homology further comprises a first fragment of 15 the target site and the second region of homology comprises a second fragment of the target site, wherein the first and second fragments are dissimilar.

As used herein, "homologous recombination" includes the exchange of DNA fragments between two DNA molecules at the sites of homology. The frequency of homologous recombination is influenced by a number of factors. Different 20 organisms vary with respect to the amount of homologous recombination and the relative proportion of homologous to non-homologous recombination. Generally, the length of the region of homology affects the frequency of homologous recombination events: the longer the region of homology, the greater the frequency. The length of the homology region needed to observe homologous recombination is also species- 25 variable. In many cases, at least 5 kb of homology has been utilized, but homologous recombination has been observed with as little as 25-50 bp of homology. See, for example, Singer et al., (1982) Cell 31:25-33; Shen and Huang, (1986) Genetics 112:441-57; Watt et al., (1985) Proc. Natl. Acad. Sci. USA 82:4768-72, Sugawara and Haber, (1992) Mol Cell Biol 12:563-75, Rubnitz and Subramani, 30 (1984) Mol Cell Biol 4:2253-8; Ayares et al., (1986) Proc. Natl. Acad. Sci. USA 83:5199-203; Liskay et al., (1987) Genetics 115:161-7.

Homology-directed repair (HDR) is a mechanism in cells to repair double-stranded and single stranded DNA breaks. Homology-directed repair includes homologous recombination (HR) and single-strand annealing (SSA) (Lieber. 2010 Annu. Rev. Biochem. 79:181-211). The most common form of HDR is called

5 homologous recombination (HR), which has the longest sequence homology requirements between the donor and acceptor DNA. Other forms of HDR include single-stranded annealing (SSA) and breakage-induced replication, and these require shorter sequence homology relative to HR. Homology-directed repair at nicks (single-stranded breaks) can occur via a mechanism distinct from HDR at

10 double-strand breaks (Davis and Maizels. PNAS(0027-8424), 111 (10), p. E924-E932.

Alteration of the genome of a plant cell, for example, through homologous recombination (HR), is a powerful tool for genetic engineering. Despite the low frequency of homologous recombination in higher plants, there are a few examples

15 of successful homologous recombination of plant endogenous genes. The parameters for homologous recombination in plants have primarily been investigated by rescuing introduced truncated selectable marker genes. In these experiments, the homologous DNA fragments were typically between 0.3 kb to 2 kb. Observed frequencies for homologous recombination were on the order of 10^{-4} to

20 10^{-5} . See, for example, Halfter et al., (1992) Mol Gen Genet 231:186-93; Offringa et al., (1990) EMBO J 9:3077-84; Offringa et al., (1993) Proc. Natl. Acad. Sci. USA 90:7346-50; Paszkowski et al., (1988) EMBO J 7:4021-6; Hourda and Paszkowski, (1994) Mol Gen Genet 243:106-11; and Risseuw et al., (1995) Plant J 7:109-19.

Homologous recombination has been demonstrated in insects. In

25 *Drosophila*, Dray and Gloor found that as little as 3 kb of total template:target homology sufficed to copy a large non-homologous segment of DNA into the target with reasonable efficiency (Dray and Gloor, (1997) Genetics 147:689-99). Using FLP-mediated DNA integration at a target FRT in *Drosophila*, Golic et al., showed integration was approximately 10-fold more efficient when the donor and target

30 shared 4.1 kb of homology as compared to 1.1 kb of homology (Golic et al., (1997) Nucleic Acids Res 25:3665). Data from *Drosophila* indicates that 2-4 kb of homology is sufficient for efficient targeting, but there is some evidence that much

less homology may suffice, on the order of about 30 bp to about 100 bp (Nassif and Engels, (1993) *Proc. Natl. Acad. Sci. USA* 90:1262-6; Keeler and Gloor, (1997) *Mol Cell Biol* 17:627-34).

Homologous recombination has also been accomplished in other organisms.

- 5 For example, at least 150-200 bp of homology was required for homologous recombination in the parasitic protozoan *Leishmania* (Papadopoulou and Dumas, (1997) *Nucleic Acids Res* 25:4278-86). In the filamentous fungus *Aspergillus nidulans*, gene replacement has been accomplished with as little as 50 bp flanking homology (Chaverroche et al., (2000) *Nucleic Acids Res* 28:e97). Targeted gene
- 10 replacement has also been demonstrated in the ciliate *Tetrahymena thermophila* (Gaertig et al., (1994) *Nucleic Acids Res* 22:5391-8). In mammals, homologous recombination has been most successful in the mouse using pluripotent embryonic stem cell lines (ES) that can be grown in culture, transformed, selected and introduced into a mouse embryo. Embryos bearing inserted transgenic ES cells
- 15 develop as genetically offspring. By interbreeding siblings, homozygous mice carrying the selected genes can be obtained. An overview of the process is provided in Watson et al., (1992) *Recombinant DNA*, 2nd Ed., (Scientific American Books distributed by WH Freeman & Co.); Capecchi, (1989) *Trends Genet* 5:70-6; and Bronson, (1994) *J Biol Chem* 269:27155-8. Homologous recombination in
- 20 mammals other than mouse has been limited by the lack of stem cells capable of being transplanted to oocytes or developing embryos. However, McCreath et al., *Nature* 405:1066-9 (2000) reported successful homologous recombination in sheep by transformation and selection in primary embryo fibroblast cells.

- Error-prone DNA repair mechanisms can produce mutations at double-strand
- 25 break sites. The nonhomologous end-joining (NHEJ) pathways are the most common repair mechanism to bring the broken ends together (Bleuyard et al., (2006) *DNA Repair* 5:1-12). The structural integrity of chromosomes is typically preserved by the repair, but deletions, insertions, or other rearrangements are possible. The two ends of one double-strand break are the most prevalent
- 30 substrates of NHEJ (Kirik et al., (2000) *EMBO J* 19:5562-6), however if two different double-strand breaks occur, the free ends from different breaks can be ligated and result in chromosomal deletions (Siebert and Puchta, (2002) *Plant Cell* 14:1121-31),

or chromosomal translocations between different chromosomes (Pacher et al., (2007) Genetics 175:21-9).

Episomal DNA molecules can also be ligated into the double-strand break, for example, integration of T-DNAs into chromosomal double-strand breaks (Chilton and Que, (2003) Plant Physiol 133:956-65; Salomon and Puchta, (1998) EMBO J 17:6086-95). Once the sequence around the double-strand breaks is altered, for example, by exonuclease activities involved in the maturation of double-strand breaks, gene conversion pathways can restore the original structure if a homologous sequence is available, such as a homologous chromosome in non-dividing somatic cells, or a sister chromatid after DNA replication (Molinier et al., (2004) Plant Cell 16:342-52). Ectopic and/or epigenic DNA sequences may also serve as a DNA repair template for homologous recombination (Puchta, (1999) Genetics 152:1173-81).

Once a double-strand break is induced in the DNA, the cell's DNA repair mechanism is activated to repair the break. Error-prone DNA repair mechanisms can produce mutations at double-strand break sites. The most common repair mechanism to bring the broken ends together is the nonhomologous end-joining (NHEJ) pathway (Bleuyard et al., (2006) DNA Repair 5:1-12). The structural integrity of chromosomes is typically preserved by the repair, but deletions, insertions, or other rearrangements are possible (Siebert and Puchta, (2002) Plant Cell 14:1121-31; Pacher et al., (2007) Genetics 175:21-9).

Alternatively, the double-strand break can be repaired by homologous recombination between homologous DNA sequences. Once the sequence around the double-strand break is altered, for example, by exonuclease activities involved in the maturation of double-strand breaks, gene conversion pathways can restore the original structure if a homologous sequence is available, such as a homologous chromosome in non-dividing somatic cells, or a sister chromatid after DNA replication (Molinier et al., (2004) Plant Cell 16:342-52). Ectopic and/or epigenic DNA sequences may also serve as a DNA repair template for homologous recombination (Puchta, (1999) Genetics 152:1173-81).

DNA double-strand breaks appear to be an effective factor to stimulate homologous recombination pathways (Puchta et al., (1995) Plant Mol Biol 28:281-

92; Tzfira and White, (2005) Trends Biotechnol 23:567-9; Puchta, (2005) J Exp Bot 56:1-14). Using DNA-breaking agents, a two- to nine-fold increase of homologous recombination was observed between artificially constructed homologous DNA repeats in plants (Puchta et al., (1995) Plant Mol Biol 28:281-92). In maize
5 protoplasts, experiments with linear DNA molecules demonstrated enhanced homologous recombination between plasmids (Lyznik et al., (1991) Mol Gen Genet 230:209-18).

In one embodiment provided herein, the method comprises contacting a plant cell with the donor DNA and the endonuclease. Once a double-strand break is
10 introduced in the target site by the endonuclease, the first and second regions of homology of the donor DNA can undergo homologous recombination with their corresponding genomic regions of homology resulting in exchange of DNA between the donor and the genome. As such, the provided methods result in the integration of the polynucleotide of interest of the donor DNA into the double-strand break in the
15 target site in the plant genome, thereby altering the original target site and producing an altered genomic target site.

The donor DNA may be introduced by any means known in the art. For example, a plant having a target site is provided. The donor DNA may be provided by any transformation method known in the art including, for example,
20 Agrobacterium-mediated transformation or biolistic particle bombardment. The donor DNA may be present transiently in the cell or it could be introduced via a viral replicon. In the presence of the Cas endonuclease and the target site, the donor DNA is inserted into the transformed plant's genome.

Another approach uses protein engineering of existing homing
25 endonucleases to alter their target specificities. Homing endonucleases, such as I-SceI or I-CreI, bind to and cleave relatively long DNA recognition sequences (18 bp and 22 bp, respectively). These sequences are predicted to naturally occur infrequently in a genome, typically only 1 or 2 sites/genome. The cleavage specificity of a homing endonuclease can be changed by rational design of amino
30 acid substitutions at the DNA binding domain and/or combinatorial assembly and selection of mutated monomers (see, for example, Arnould et al., (2006) J Mol Biol 355:443-58; Ashworth et al., (2006) Nature 441:656-9; Doyon et al., (2006) J Am

Chem Soc 128:2477-84; Rosen et al., (2006) Nucleic Acids Res 34:4791-800; and Smith et al., (2006) Nucleic Acids Res 34:e149; Lyznik et al., (2009) U.S. Patent Application Publication No. 20090133152A1; Smith et al., (2007) U.S. Patent Application Publication No. 20070117128A1). Engineered meganucleases have
5 been demonstrated that can cleave cognate mutant sites without broadening their specificity. An artificial recognition site specific to the wild type yeast I-SceI homing nuclease was introduced in maize genome and mutations of the recognition sequence were detected in 1% of analyzed F1 plants when a transgenic I-SceI was introduced by crossing and activated by gene excision (Yang et al., (2009) Plant Mol
10 Biol 70:669-79). More practically, the maize liguleless locus was targeted using an engineered single-chain endonuclease designed based on the I-CreI meganuclease sequence. Mutations of the selected liguleless locus recognition sequence were detected in 3% of the T0 transgenic plants when the designed homing nuclease was introduced by Agrobacterium-mediated transformation of immature embryos (Gao et
15 al., (2010) Plant J 61:176-87).

Polynucleotides of interest are further described herein and are reflective of the commercial markets and interests of those involved in the development of the crop. Crops and markets of interest change, and as developing nations open up world markets, new crops and technologies will emerge also. In addition, as our
20 understanding of agronomic traits and characteristics such as yield and heterosis increase, the choice of genes for genetic engineering will change accordingly.

Genome editing using the guide RNA/Cas endonuclease system

As described herein, the guide RNA/Cas endonuclease system can be used in combination with a co-delivered polynucleotide modification template to allow for
25 editing of a genomic nucleotide sequence of interest. Also, as described herein, for each embodiment that uses a guide RNA/Cas endonuclease system, a similar guide polynucleotide/Cas endonuclease system can be deployed where the guide polynucleotide does not solely comprise ribonucleic acids but wherein the guide polynucleotide comprises a combination of RNA-DNA molecules or solely comprise
30 DNA molecules.

While numerous double-strand break-making systems exist, their practical applications for gene editing may be restricted due to the relatively low frequency of

induced double-strand breaks (DSBs). To date, many genome modification methods rely on the homologous recombination system. Homologous recombination (HR) can provide molecular means for finding genomic DNA sequences of interest and modifying them according to the experimental specifications. Homologous recombination takes place in plant somatic cells at low frequency. The process can be enhanced to a practical level for genome engineering by introducing double-strand breaks (DSBs) at selected endonuclease target sites. The challenge has been to efficiently make DSBs at genomic sites of interest since there is a bias in the directionality of information transfer between two interacting DNA molecules (the broken one acts as an acceptor of genetic information). Described herein is the use of a guide RNA/Cas system which provides flexible genome cleavage specificity and results in a high frequency of double-strand breaks at a DNA target site, thereby enabling efficient gene editing in a nucleotide sequence of interest, wherein the nucleotide sequence of interest to be edited can be located within or outside the target site recognized and cleaved by a Cas endonuclease.

A “modified nucleotide” or “edited nucleotide” refers to a nucleotide sequence of interest that comprises at least one alteration when compared to its non-modified nucleotide sequence. Such “alterations” include, for example: (i) replacement of at least one nucleotide, (ii) a deletion of at least one nucleotide, (iii) an insertion of at least one nucleotide, or (iv) any combination of (i) – (iii).

The term “polynucleotide modification template” includes a polynucleotide that comprises at least one nucleotide modification when compared to the nucleotide sequence to be edited. A nucleotide modification can be at least one nucleotide substitution, addition or deletion. Optionally, the polynucleotide modification template can further comprise homologous nucleotide sequences flanking the at least one nucleotide modification, wherein the flanking homologous nucleotide sequences provide sufficient homology to the desired nucleotide sequence to be edited.

In one embodiment, the disclosure describes a method for editing a nucleotide sequence in the genome of a cell without incorporating a selectable transgene marker, the method comprising providing at least one guide RNA, at least one polynucleotide modification template, and at least one Cas endonuclease to a

cell, wherein the Cas endonuclease is capable of introducing a double-strand break at a target sequence in the genome of said cell, wherein said polynucleotide modification template includes at least one nucleotide modification of said nucleotide sequence to be edited. Cells include, but are not limited to, human,
5 animal, bacterial, fungal, insect, and plant cells as well as plants and seeds produced by the methods described herein. The nucleotide to be edited can be located within or outside a target site recognized and cleaved by a Cas endonuclease. In one embodiment, the at least one nucleotide modification is not a modification at a target site recognized and cleaved by a Cas endonuclease. In
10 another embodiment, there are at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 40, 50, 100, 200, 300, 400, 500, 600, 700, 900 or 1000 nucleotides between the at least one nucleotide to be edited and the genomic target site.

In one embodiment, the method comprises a method for editing a second
15 gene of a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be modified to confer herbicide
20 resistance, wherein said first guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at a first target site (located in or near said first endogenous gene) in the genome of said plant cell, wherein said first polynucleotide modification template comprises at least one nucleotide modification of said first endogenous gene to
25 render said endogenous gene capable of conferring herbicide resistance to a plant cell, wherein said second guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at a second target site (located at a different locus than said first endogenous gene) in the genome of said plant cell, wherein said second polynucleotide modification
30 template comprises at least one nucleotide alteration when compared to the second gene to be edited.

In one embodiment of genome editing, editing of the endogenous enolpyruvylshikimate-3-phosphate synthase (*EPSPS*) gene is disclosed herein wherein, the polynucleotide modification template (*EPSPS* polynucleotide modification template) includes a partial fragment of the *EPSPS* gene (and therefore
5 does not encode a fully functional *EPSPS* polypeptide by itself). The *EPSPS* polynucleotide modification template contained three point mutations that were responsible for the creation of the T102I/P106S (TIPS) double mutant (Funke, T et al., J. Biol. Chem. 2009, 284:9854-9860), which provide glyphosate tolerance to transgenic plants expressing as *EPSPS* double mutant transgene.

10 As defined herein "Glyphosate" includes any herbicidally effective form of N-phosphonomethylglycine (including any salt thereof), other forms which result in the production of the glyphosate anion in plants and any other herbicides of the phosphonomethylglycine family.

Increased resistance to a herbicide is demonstrated when plants which
15 display the increased resistance to a herbicide are subjected to the herbicide and a dose/response curve is shifted to the right when compared with that provided by an appropriate control plant. Such dose/response curves have "dose" plotted on the x-axis and "percentage injury", "herbicidal effect" etc. plotted on the y-axis. Plants which are substantially resistant to the herbicide exhibit few, if any, bleached,
20 necrotic, lytic, chlorotic or other lesions and are not stunted, wilted or deformed when subjected to the herbicide at concentrations and rates which are typically employed by the agricultural community to kill weeds in the field. The terms resistance and tolerance may be used interchangeably.

Figure 12 shows a schematic representation of components used in the
25 genome editing procedure. A maize optimized Cas endonuclease, a guide RNA and a polynucleotide modification template were provided to a plant cell. For example, as shown in Figure 12, the polynucleotide modification template included three nucleotide modifications (indicated by arrows) when compared to the *EPSPS* genomic sequence to be edited. These three nucleotide modifications are referred
30 to as TIPS mutations as these nucleotide modifications result in the amino acid changes T-102 to I-102 and P-106 to S-106. The first point mutation results from the substitution of the C nucleotide in the codon sequence ACT with a T nucleotide, a

second mutation results from the substitution of the T nucleotide on the same codon sequence ACT with a C nucleotide to form the isoleucine codon ATC, the third point mutation results from the substitution of the first C nucleotide in the codon sequence CCA with a T nucleotide in order to form a serine codon TCA (Figure 12).

5 The nucleotide sequence to be edited can be a sequence that is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited. For example, the nucleotide sequence in the genome of a cell can be a native gene, a mutated gene, a non-native gene, a foreign gene, or a transgene that is stably incorporated into the genome of a cell. Editing of such nucleotide may result in a
10 further desired phenotype or genotype.

Regulatory sequence modifications using the guide polynucleotide/Cas endonuclease system

In one embodiment the nucleotide sequence to be modified can be a regulatory sequence such as a promoter wherein the editing of the promoter
15 comprises replacing the promoter (also referred to as a “promoter swap” or “promoter replacement”) or promoter fragment with a different promoter (also referred to as replacement promoter) or promoter fragment (also referred to as replacement promoter fragment), wherein the promoter replacement results in any one of the following or any one combination of the following: an increased promoter
20 activity, an increased promoter tissue specificity, a decreased promoter activity, a decreased promoter tissue specificity, a new promoter activity, an inducible promoter activity, an extended window of gene expression, a modification of the timing or developmental progress of gene expression in the same cell layer or other cell layer (such as but not limiting to extending the timing of gene expression in the
25 tapetum of maize anthers (US 5,837,850 issued November 17, 1998), a mutation of DNA binding elements and / or a deletion or addition of DNA binding elements. The promoter (or promoter fragment) to be modified can be a promoter (or promoter fragment) that is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited. The replacement promoter (or replacement promoter fragment) can be
30 a promoter (or promoter fragment) that is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited.

In one embodiment the nucleotide sequence can be a promoter wherein the editing of the promoter comprises replacing a native EPSPS1 promoter from with a plant ubiquitin promoter.

In one embodiment the nucleotide sequence can be a promoter wherein the promoter to be edited is selected from the group comprising Zea mays-PEPC1 promoter (Kausch et al, Plant Molecular Biology, 45: 1-15, 2001), Zea mays Ubiquitin promoter (UBI1ZM PRO, Christensen et al, plant Molecular Biology 18: 675-689, 1992), Zea mays-Rootmet2 promoter (US 7,214,855), Rice actin promoter (OS-ACTIN PRO, US5641876; McElroy et al, The Plant Cell, Vol 2, 163-171, Feb 1990), Sorghum RCC3 promoter (US 2012/0210463 filed on 13 Feb. 2012), Zea mays-GOS2 promoter (US 6,504,083), Zea mays-ACO2 promoter (US app 14/210,711 filed 14 March 2014) or Zea mays-oleosin promoter (US 8466341 B2).

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used in combination with a co-delivered polynucleotide modification template or donor DNA sequence to allow for the insertion of a promoter or promoter element into a genomic nucleotide sequence of interest without incorporating a selectable transgene marker, wherein the promoter insertion (or promoter element insertion) results in any one of the following or any one combination of the following: an increased promoter activity (increased promoter strength), an increased promoter tissue specificity, a decreased promoter activity, a decreased promoter tissue specificity, a new promoter activity, an inducible promoter activity, an extended window of gene expression, a modification of the timing or developmental progress of gene expression a mutation of DNA binding elements and / or an addition of DNA binding elements. Promoter elements to be inserted can be, but are not limited to, promoter core elements (such as, but not limited to, a CAAT box, a CCAAT box, a Pribnow box, a and / or TATA box, translational regulation sequences and / or a repressor system for inducible expression (such as TET operator repressor/operator/inducer elements, or sulphonylurea (Su) repressor/operator/inducer elements. The dehydration-responsive element (DRE) was first identified as a *cis*-acting promoter element in the promoter of the drought-responsive gene *rd29A*, which contains a 9 bp conserved core sequence, TACCGACAT (Yamaguchi-Shinozaki, K., and Shinozaki, K. (1994) Plant Cell 6,

251–264). Insertion of DRE into an endogenous promoter may confer a drought inducible expression of the downstream gene. Another example is ABA-responsive elements (ABREs) which contains a (C/T)ACGTGGC consensus sequence found to be present in numerous ABA and/or stress-regulated genes (Busk P. K., Pages
5 M.(1998) Plant Mol. Biol. 37:425–435). Insertion of 35S enhancer or MMV enhancer into an endogenous promoter region will increase gene expression (US patent 5196525). The promoter (or promoter element) to be inserted can be a promoter (or promoter element) that is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited.

10 In one embodiment, the guide polynucleotide/Cas endonuclease system can be used to insert an enhancer element, such as but not limited to a Cauliflower Mosaic Virus 35 S enhancer, in front of an endogenous FMT1 promoter to enhance expression of the FTM1.

15 In one embodiment, the guide polynucleotide/Cas endonuclease system can be used to insert a component of the TET operator repressor/operator/inducer system, or a component of the sulphonylurea (Su) repressor/operator/inducer system into plant genomes to generate or control inducible expression systems without incorporating a selectable transgene marker.

20 In another embodiment, the guide polynucleotide/Cas endonuclease system can be used to allow for the deletion of a promoter or promoter element, wherein the promoter deletion (or promoter element deletion) results in any one of the following or any one combination of the following: a permanently inactivated gene locus, an increased promoter activity (increased promoter strength), an increased promoter tissue specificity, a decreased promoter activity, a decreased promoter tissue
25 specificity, a new promoter activity, an inducible promoter activity, an extended window of gene expression, a modification of the timing or developmental progress of gene expression, a mutation of DNA binding elements and / or an addition of DNA binding elements. Promoter elements to be deleted can be, but are not limited to, promoter core elements, promoter enhancer elements or 35 S enhancer
30 elements (as described in Example 32) The promoter or promoter fragment to be deleted can be endogenous, artificial, pre-existing, or transgenic to the cell that is being edited.

Terminator modifications using the guide polynucleotide/Cas endonuclease system

In one embodiment the nucleotide sequence to be modified can be a terminator wherein the editing of the terminator comprises replacing the terminator (also referred to as a “terminator swap” or “terminator replacement”) or terminator
 5 fragment with a different terminator (also referred to as replacement terminator) or terminator fragment (also referred to as replacement terminator fragment), wherein the terminator replacement results in any one of the following or any one combination of the following: an increased terminator activity, an increased terminator tissue specificity, a decreased terminator activity, a decreased terminator
 10 tissue specificity, a mutation of DNA binding elements and / or a deletion or addition of DNA binding elements.” The terminator (or terminator fragment) to be modified can be a terminator (or terminator fragment) that is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited. The replacement terminator (or replacement terminator fragment) can be a terminator (or terminator fragment) that
 15 is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited.

In one embodiment the nucleotide sequence to be modified can be a terminator wherein the terminator to be edited is selected from the group comprising terminators from maize Argos 8 or SRTF18 genes, or other terminators, such as
 potato PinII terminator, sorghum actin terminator (SB-ACTIN TERM , WO
 20 2013/184537 A1 published Dec 2013) , sorghum SB-GKAF TERM (WO2013019461), rice T28 terminator (OS-T28 TERM ,WO 2013/012729 A2), AT-T9 TERM (WO 2013/012729 A2) or GZ-W64A TERM (US7053282).

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used in combination with a co-delivered polynucleotide modification template or
 25 donor DNA sequence to allow for the insertion of a terminator or terminator element into a genomic nucleotide sequence of interest, wherein the terminator insertion (or terminator element insertion) results in any one of the following or any one combination of the following: an increased terminator activity (increased terminator strength), an increased terminator tissue specificity, a decreased terminator activity,
 30 a decreased terminator tissue specificity, a mutation of DNA binding elements and / or an addition of DNA binding elements.

The terminator (or terminator element) to be inserted can be a terminator (or terminator element) that is endogenous, artificial, pre-existing, or transgenic to the cell that is being edited.

In another embodiment, the guide polynucleotide/Cas endonuclease system can be used to allow for the deletion of a terminator or terminator element, wherein the terminator deletion (or terminator element deletion) results in any one of the following or any one combination of the following: an increased terminator activity (increased terminator strength), an increased terminator tissue specificity, a decreased terminator activity, a decreased terminator tissue specificity, a mutation of DNA binding elements and / or an addition of DNA binding elements. The terminator or terminator fragment to be deleted can be endogenous, artificial, pre-existing, or transgenic to the cell that is being edited.

Additional regulatory sequence modifications using the guide polynucleotide/Cas endonuclease system

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used to modify or replace a regulatory sequence in the genome of a cell without incorporating a selectable transgene marker. A regulatory sequence is a segment of a nucleic acid molecule which is capable of increasing or decreasing the expression of specific genes within an organism and/or is capable of altering tissue specific expression of genes within an organism. Examples of regulatory sequences include, but are not limited to, 3' UTR (untranslated region) region, 5' UTR region, transcription activators, transcriptional enhancers transcriptions repressors, translational repressors, splicing factors, miRNAs, siRNA, artificial miRNAs, promoter elements, CAMV 35 S enhancer, MMV enhancer elements (PCT/US14/23451 filed March 11, 2013), SECIS elements , polyadenylation signals, and polyubiquitination sites. In some embodiments the editing (modification) or replacement of a regulatory element results in altered protein translation, RNA cleavage, RNA splicing, transcriptional termination or post translational modification. In one embodiment, regulatory elements can be identified within a promoter and these regulatory elements can be edited or modified do to optimize these regulatory elements for up or down regulation of the promoter.

In one embodiment, the genomic sequence of interest to be modified is a polyubiquitination site, wherein the modification of the polyubiquitination sites results in a modified rate of protein degradation. The ubiquitin tag condemns proteins to be degraded by proteasomes or autophagy. Proteasome inhibitors are known to cause
 5 a protein overproduction. Modifications made to a DNA sequence encoding a protein of interest can result in at least one amino acid modification of the protein of interest, wherein said modification allows for the polyubiquitination of the protein (a post translational modification) resulting in a modification of the protein degradation

In one embodiment, the genomic sequence of interest to be modified is a
 10 polyubiquitination site on a maize EPSPS gene, wherein the polyubiquitination site modified resulting in an increased protein content due to a slower rate of EPSPS protein degradation.

In one embodiment, the genomic sequence of interest to be modified is a an intron site, wherein the modification consist of inserting an intron enhancing motif
 15 into the intron which results in modulation of the transcriptional activity of the gene comprising said intron.

In one embodiment, the genomic sequence of interest to be modified is a an intron site, wherein the modification consist of replacing a soybean EPSP1 intron with a soybean ubiquitin intron 1 as described herein (Example 25)

20 In one embodiment, the genomic sequence of interest to be modified is a an intron or UTR site, wherein the modification consist of inserting at least one microRNA into said intron or UTR site, wherein expression of the gene comprising the intron or UTR site also results in expression of said microRNA, which in turn can silence any gene targeted by the microRNA without disrupting the gene expression
 25 of the native/transgene comprising said intron.

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used to allow for the deletion or mutation of a Zinc Finger transcription factor, wherein the deletion or mutation of the Zinc Finger transcription factor results in or allows for the creation of a dominant negative Zinc Finger transcription factor mutant
 30 (Li et al 2013 Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression PNAS 110:3167-3172). Insertion of a single base pair downstream zinc finger domain will result in a frame shift and produces a

new protein which still can bind to DNA without transcription activity. The mutant protein will compete to bind to cytokinin oxidase gene promoters and block the expression of cytokinin oxidase gene. Reduction of cytokinin oxidase gene expression will increase cytokinin level and promote panicle growth in rice and ear growth in maize, and increase yield under normal and stress conditions.

Modifications of splicing sites and/or introducing alternate splicing sites using the guide polynucleotide/Cas endonuclease system

Protein synthesis utilizes mRNA molecules that emerge from pre-mRNA molecules subjected to the maturation process. The pre-mRNA molecules are capped, spliced and stabilized by addition of polyA tails. Eukaryotic cells developed a complex process of splicing that result in alternative variants of the original pre-mRNA molecules. Some of them may not produce functional templates for protein synthesis. In maize cells, the splicing process is affected by splicing sites at the exon-intron junction sites. An example of a canonical splice site is AGGT. Gene coding sequences can contains a number of alternate splicing sites that may affect the overall efficiency of the pre-mRNA maturation process and as such may limit the protein accumulation in cells. The guide polynucleotide/Cas endonuclease system can be used in combination with a co-delivered polynucleotide modification template to edit a gene of interest to introduce a canonical splice site at a described junction or any variant of a splicing site that changes the splicing pattern of pre-mRNA molecules, without incorporating a selectable transgene marker.

In one embodiment, the nucleotide sequence of interest to be modified is a maize EPSPS gene, wherein the modification of the gene consists of modifying alternative splicing sites resulting in enhanced production of the functional gene transcripts and gene products (proteins).

In one embodiment, the nucleotide sequence of interest to be modified is a gene, wherein the modification of the gene consists of editing the intron borders of alternatively spliced genes to alter the accumulation of splice variants.

Modifications of nucleotide sequences encoding a protein of interest using the guide polynucleotide/ Cas endonuclease system

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used to modify or replace a coding sequence in the genome of a cell without

incorporating a selectable transgene marker , wherein the modification or replacement results in any one of the following, or any one combination of the following: an increased protein (enzyme) activity, an increased protein functionality, a decreased protein activity, a decreased protein functionality, a site specific
5 mutation, a protein domain swap, a protein knock-out, a new protein functionality, a modified protein functionality, .

In one embodiment the protein knockout is due to the introduction of a stop codon into the coding sequence of interest.

In one embodiment the protein knockout is due to the deletion of a start
10 codon into the coding sequence of interest.

Amino acid and/or protein fusions using the guide polynucleotide/ Cas endonuclease system

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used with or without a co-delivered polynucleotide sequence to fuse a first coding
15 sequence encoding a first protein to a second coding sequence encoding a second protein in the genome of a cell, without incorporating a selectable transgene marker, wherein the protein fusion results in any one of the following or any one combination of the following: an increased protein (enzyme) activity, an increased protein functionality, a decreased protein activity, a decreased protein functionality, a new
20 protein functionality, a modified protein functionality, a new protein localization, a new timing of protein expression, a modified protein expression pattern, a chimeric protein, or a modified protein with dominant phenotype functionality.

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used with or without a co-delivered polynucleotide sequence to fuse a first coding
25 sequence encoding a chloroplast localization signal to a second coding sequence encoding a protein of interest, wherein the protein fusion results in targeting the protein of interest to the chloroplast.

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used with or without a co-delivered polynucleotide sequence to fuse a first coding
30 sequence encoding a chloroplast localization signal to a second coding sequence encoding a protein of interest, wherein the protein fusion results in targeting the protein of interest to the chloroplast.

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used with or without a co-delivered polynucleotide sequence to fuse a first coding sequence encoding a chloroplast localization signal (e.g., a chloroplast transit peptide) to a second coding sequence, wherein the protein fusion results in a modified protein with dominant phenotype functionality

Gene silencing by expressing an inverted repeat into a gene of interest using the guide polynucleotide/Cas endonuclease system

In one embodiment, the guide polynucleotide/Cas endonuclease system can be used in combination with a co-delivered polynucleotide sequence to insert an inverted gene fragment into a gene of interest in the genome of an organism, without incorporating a selectable transgene marker, wherein the insertion of the inverted gene fragment can allow for an in-vivo creation of an inverted repeat (hairpin) and results in the silencing of said endogenous gene.

In one embodiment the insertion of the inverted gene fragment can result in the formation of an in-vivo created inverted repeat (hairpin) in a native (or modified) promoter of a gene and/or in a native 5' end of the native gene. The inverted gene fragment can further comprise an intron which can result in an enhanced silencing of the targeted gene.

Genome deletion for Trait Locus Characterization

Trait mapping in plant breeding often results in the detection of chromosomal regions housing one or more genes controlling expression of a trait of interest. For a qualitative trait, the guide polynucleotide/Cas endonuclease system can be used to eliminate candidate genes in the identified chromosomal regions to determine if deletion of the gene affects expression of the trait. For quantitative traits, expression of a trait of interest is governed by multiple quantitative trait loci (QTL) of varying effect-size, complexity, and statistical significance across one or more chromosomes. In cases of negative effect or deleterious QTL regions affecting a complex trait, the guide polynucleotide/Cas endonuclease system can be used to eliminate whole regions delimited by marker-assisted fine mapping, and to target specific regions for their selective elimination or rearrangement. Similarly, presence/absence variation (PAV) or copy number variation (CNV) can be

manipulated with selective genome deletion using the guide polynucleotide/Cas endonuclease system.

In one embodiment, the region of interest can be flanked by two independent guide polynucleotide/CAS endonuclease target sequences. Cutting would be done concurrently. The deletion event would be the repair of the two chromosomal ends without the region of interest. Alternative results would include inversions of the region of interest, mutations at the cut sites and duplication of the region of interest.

Methods for identifying at least one plant cell comprising in its genome a polynucleotide of Interest integrated at the target site.

Further provided are methods for identifying at least one plant cell comprising in its genome a polynucleotide of Interest integrated at the target site. A variety of methods are available for identifying those plant cells with insertion into the genome at or near to the target site without using a screenable marker phenotype. Such methods can be viewed as directly analyzing a target sequence to detect any change in the target sequence, including but not limited to PCR methods, sequencing methods, nuclease digestion, Southern blots, and any combination thereof. See, for example, US Patent Application 12/147,834.

The method also comprises recovering a plant from the plant cell comprising a polynucleotide of Interest integrated into its genome. The plant may be sterile or fertile. It is recognized that any polynucleotide of interest can be provided, integrated into the plant genome at the target site, and expressed in a plant.

Polynucleotides/polypeptides of interest include, but are not limited to, herbicide-resistance coding sequences, insecticidal coding sequences, nematocidal coding sequences, antimicrobial coding sequences, antifungal coding sequences, antiviral coding sequences, abiotic and biotic stress tolerance coding sequences, or sequences modifying plant traits such as yield, grain quality, nutrient content, starch quality and quantity, nitrogen fixation and/or utilization, fatty acids, and oil content and/or composition. More specific polynucleotides of interest include, but are not limited to, genes that improve crop yield, polypeptides that improve desirability of crops, genes encoding proteins conferring resistance to abiotic stress, such as drought, nitrogen, temperature, salinity, toxic metals or trace elements, or those

conferring resistance to toxins such as pesticides and herbicides, or to biotic stress, such as attacks by fungi, viruses, bacteria, insects, and nematodes, and development of diseases associated with these organisms. General categories of genes of interest include, for example, those genes involved in information, such as zinc fingers, those involved in communication, such as kinases, and those involved in housekeeping, such as heat shock proteins. More specific categories of transgenes, for example, include genes encoding important traits for agronomics, insect resistance, disease resistance, herbicide resistance, fertility or sterility, grain characteristics, and commercial products. Genes of interest include, generally, those involved in oil, starch, carbohydrate, or nutrient metabolism as well as those affecting kernel size, sucrose loading, and the like that can be stacked or used in combination with glyphosate resistance described herein.

The terms "FAD" and fatty acid desaturase are used interchangeably and refer to membrane bound microsomal oleoyl- and linoleoyl-phosphatidylcholine desaturases that convert oleic acid to linoleic acid and linoleic acid to linolenic acid, respectively, in reactions that reduce molecular oxygen to water and require the presence of NADH. The gene for microsomal delta 12 fatty acid desaturases described in WO 94/11516, can be used to make a high oleic acid soybean variety. The resulting high oleic acid soybean variety was one in which the polyunsaturated fatty acids were reduced from 70% of the total fatty acids to less than 5%.

Plant having a high oleic phenotype (a high oleic acid phenotype) and a herbicide tolerance phenotype conferred by suppression of a FAD2 gene in conjunction with the expression of a sequence that confers tolerance to inhibitors of ALS have been described in US patent no. 8,609,935, issued 2013/12/17)

Two soybean fatty acid desaturases, designated FAD2-1 and FAD2-2, are Δ 12 desaturases that introduce a second double bond into oleic acid to form linoleic acid, a polyunsaturated fatty acid. FAD2-1 is expressed only in the developing seed (Heppard et al. (1996) Plant Physiol. 110:311-319). The expression of this gene increases during the period of oil deposition, starting around 19 days after flowering, and its gene product is responsible for the synthesis of the polyunsaturated fatty acids found in soybean oil. GmFad 2-1 is described in detail by Okuley, J. et al. (1994) Plant Cell 6:147 158 and in WO94/11516. It is available from the ATCC in

the form of plasmid pSF2 169K (ATCC accession number 69092). FAD 2-2 is expressed in the seed, leaf, root and stem of the soy plant at a constant level and is the “housekeeping” 12-desaturase gene. The Fad 2-2 gene product is responsible for the synthesis of polyunsaturated fatty acids for cell membranes.

- 5 Since FAD2-1 is the major enzyme of this type in soybean seeds, reduction in the expression of FAD2-1 results in increased accumulation of oleic acid (18:1) and a corresponding decrease in polyunsaturated fatty acid content. Reduction of expression of FAD2-2 in combination with FAD2-1 leads to a greater accumulation of oleic acid and corresponding decrease in polyunsaturated fatty acid content.
- 10 FAD3 is a Δ 15 desaturase that introduces a third double bond into linoleic acid (18:2) to form linolenic acid (18:3). Reduction of expression of FAD3 in combination with reduction of FAD2-1 and FAD2-2 leads to a greater accumulation of oleic acid and corresponding decrease in polyunsaturated fatty acid content, especially linolenic acid.

- 15 Nucleic acid fragments encoding FAD2-1, FAD2-2, and FAD3 have been described in WO 94/11516 and WO 93/11245. Chimeric recombinant constructs comprising all or a part of these nucleic acid fragments or the reverse complements thereof operably linked to at least one suitable regulatory sequence can be constructed wherein expression of the chimeric gene results in an altered fatty acid
- 20 phenotype. A chimeric recombinant construct can be introduced into soybean plants via transformation techniques well known to those skilled in the art. Transgenic soybean plants resulting from a transformation with a recombinant DNA are assayed to select plants with altered fatty acid profiles. The recombinant construct may contain all or part of 1) the FAD2-1 gene or 2) the FAD2-2 gene or 3)
- 25 the FAD3 gene or 4) combinations of all or portions of the FAD2-1, Fad2-2, or FAD3 genes.

- Recombinant constructs comprising all or part of 1) the FAD2-1 gene with or without 2) all or part of the Fad2-2 gene with or without all or part of the FAD3 gene can be used in making a transgenic soybean plant having a high oleic phenotype.
- 30 An altered fatty acid profile, specifically an increase in the proportion of oleic acid and a decrease in the proportion of the polyunsaturated fatty acids, indicates that one or more of the soybean seed FAD genes (FAD2-1, Fad2-2, FAD3) have been

suppressed. Assays may be conducted on soybean somatic embryo cultures and seeds to determine suppression of FAD2-1, Fad2-2, or FAD3.

Agronomically important traits such as oil, starch, and protein content can be genetically altered in addition to using traditional breeding methods. Modifications include increasing content of oleic acid, saturated and unsaturated oils, increasing levels of lysine and sulfur, providing essential amino acids, and also modification of starch. Hordothionin protein modifications are described in U.S. Patent Nos. 5,703,049, 5,885,801, 5,885,802, and 5,990,389 .

Another example is lysine and/or sulfur rich seed protein encoded by the soybean 2S albumin described in U.S. Patent No. 5,850,016, and the chymotrypsin inhibitor from barley, described in Williamson *et al.* (1987) *Eur. J. Biochem.* 165:99-106 .

Commercial traits can also be encoded on a polynucleotide of interest that could increase for example, starch for ethanol production, or provide expression of proteins. Another important commercial use of transformed plants is the production of polymers and bioplastics such as described in U.S. Patent No. 5,602,321. Genes such as β -Ketothiolase, PHBase (polyhydroxybutyrate synthase), and acetoacetyl-CoA reductase (see Schubert *et al.* (1988) *J. Bacteriol.* 170:5837-5847) facilitate expression of polyhydroxyalkanoates (PHAs).

Derivatives of the coding sequences can be made by site-directed mutagenesis to increase the level of preselected amino acids in the encoded polypeptide. For example, the gene encoding the barley high lysine polypeptide (BHL) is derived from barley chymotrypsin inhibitor, U.S. Application Serial No. 08/740,682, filed November 1, 1996, and WO 98/20133.

Other proteins include methionine-rich plant proteins such as from sunflower seed (Lilley *et al.* (1989) *Proceedings of the World Congress on Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, ed. Applewhite (American Oil Chemists Society, Champaign, Illinois), pp. 497-502; herein incorporated by reference); corn (Pedersen *et al.* (1986) *J. Biol. Chem.* 261:6279; Kirihaara *et al.* (1988) *Gene* 71:359; and rice (Musumura *et al.* (1989) *Plant Mol. Biol.* 12:123.

Other agronomically important genes encode latex, Flourey 2, growth factors, seed storage factors, and transcription factors.

Polynucleotides that improve crop yield include dwarfing genes, such as Rht1 and Rht2 (Peng et al. (1999) Nature 400:256-261), and those that increase plant growth, such as ammonium-inducible glutamate dehydrogenase. Polynucleotides that improve desirability of crops include, for example, those that allow plants to have reduced saturated fat content, those that boost the nutritional value of plants, and those that increase grain protein. Polynucleotides that improve salt tolerance are those that increase or allow plant growth in an environment of higher salinity than the native environment of the plant into which the salt-tolerant gene(s) has been introduced.

Polynucleotides/polypeptides that influence amino acid biosynthesis include, for example, anthranilate synthase (AS; EC 4.1.3.27) which catalyzes the first reaction branching from the aromatic amino acid pathway to the biosynthesis of tryptophan in plants, fungi, and bacteria. In plants, the chemical processes for the biosynthesis of tryptophan are compartmentalized in the chloroplast. See, for example, US Pub. 20080050506, herein incorporated by reference. Additional sequences of interest include Chorismate Pyruvate Lyase (CPL) which includes a gene encoding an enzyme which catalyzes the conversion of chorismate to pyruvate and pHBA. The most well characterized CPL gene has been isolated from E. coli and bears the GenBank accession number M96268. See, US Patent No. 7,361,811.

Polynucleotide sequences of interest may encode proteins involved in providing disease or pest resistance. By "disease resistance" or "pest resistance" is intended that the plants avoid the harmful symptoms that are the outcome of the plant-pathogen interactions. Pest resistance genes may encode resistance to pests that have great yield drag such as rootworm, cutworm, European Corn Borer, and the like. Disease resistance and insect resistance genes such as lysozymes or cecropins for antibacterial protection, or proteins such as defensins, glucanases or chitinases for antifungal protection, or Bacillus thuringiensis endotoxins, protease inhibitors, collagenases, lectins, or glycosidases for controlling nematodes or insects are all examples of useful gene products. Genes encoding disease resistance traits

include detoxification genes, such as against fumonisin (U.S. Patent No. 5,792,931); avirulence (avr) and disease resistance (R) genes (Jones et al. (1994) Science 266:789; Martin et al. (1993) Science 262:1432; and Mindrinos et al. (1994) Cell 78:1089); and the like. Insect resistance genes may encode resistance to pests that have great yield drag such as rootworm, cutworm, European Corn Borer, and the like. Such genes include, for example, *Bacillus thuringiensis* toxic protein genes (U.S. Patent Nos. 5,366,892; 5,747,450; 5,736,514; 5,723,756; 5,593,881; and Geiser *et al.* (1986) *Gene* 48:109); and the like.

An "herbicide resistance protein" or a protein resulting from expression of an "herbicide resistance-encoding nucleic acid molecule" includes proteins that confer upon a cell the ability to tolerate a higher concentration of an herbicide than cells that do not express the protein, or to tolerate a certain concentration of an herbicide for a longer period of time than cells that do not express the protein. Herbicide resistance traits may be introduced into plants by genes coding for resistance to herbicides that act to inhibit the action of acetolactate synthase (ALS), in particular the sulfonylurea-type herbicides, genes coding for resistance to herbicides that act to inhibit the action of glutamine synthase, such as phosphinothricin or basta (e.g., the bar gene), glyphosate (e.g., the EPSP synthase gene and the GAT gene), HPPD inhibitors (e.g, the HPPD gene) or other such genes known in the art. See, for example, US Patent Nos. 7,626,077, 5,310,667, 5,866,775, 6,225,114, 6,248,876, 7,169,970, 6,867,293, and US Provisional Application No. 61/401,456.

The *bar* gene encodes resistance to the herbicide basta, the *nptII* gene encodes resistance to the antibiotics kanamycin and geneticin, and the ALS-gene mutants encode resistance to the herbicide chlorsulfuron.

Sterility genes can also be encoded in an expression cassette and provide an alternative to physical detasseling. Examples of genes used in such ways include male fertility genes such as MS26 (see for example U.S. Patents 7,098,388, 7,517,975, 7,612,251) , MS45 (see for example U.S. Patents 5,478,369, 6,265,640) or MSCA1 (see for example U.S. Patent 7,919,676). Maize plants (*Zea mays* L.) can be bred by both self-pollination and cross-pollination techniques. Maize has male flowers, located on the tassel, and female flowers, located on the ear, on the same

plant. It can self-pollinate ("selfing") or cross pollinate. Natural pollination occurs in maize when wind blows pollen from the tassels to the silks that protrude from the tops of the incipient ears. Pollination may be readily controlled by techniques known to those of skill in the art. The development of maize hybrids requires the

5 development of homozygous inbred lines, the crossing of these lines, and the evaluation of the crosses. Pedigree breeding and recurrent selections are two of the breeding methods used to develop inbred lines from populations. Breeding programs combine desirable traits from two or more inbred lines or various broad-based sources into breeding pools from which new inbred lines are developed by

10 selfing and selection of desired phenotypes. A hybrid maize variety is the cross of two such inbred lines, each of which may have one or more desirable characteristics lacked by the other or which complement the other. The new inbreds are crossed with other inbred lines and the hybrids from these crosses are evaluated to determine which have commercial potential. The hybrid progeny of the first

15 generation is designated F1. The F1 hybrid is more vigorous than its inbred parents. This hybrid vigor, or heterosis, can be manifested in many ways, including increased vegetative growth and increased yield.

Hybrid maize seed can be produced by a male sterility system incorporating manual detasseling. To produce hybrid seed, the male tassel is removed from the

20 growing female inbred parent, which can be planted in various alternating row patterns with the male inbred parent. Consequently, providing that there is sufficient isolation from sources of foreign maize pollen, the ears of the female inbred will be fertilized only with pollen from the male inbred. The resulting seed is therefore hybrid (F1) and will form hybrid plants.

25 Field variation impacting plant development can result in plants tasseling after manual detasseling of the female parent is completed. Or, a female inbred plant tassel may not be completely removed during the detasseling process. In any event, the result is that the female plant will successfully shed pollen and some female plants will be self-pollinated. This will result in seed of the female inbred

30 being harvested along with the hybrid seed which is normally produced. Female inbred seed does not exhibit heterosis and therefore is not as productive as F1

seed. In addition, the presence of female inbred seed can represent a germplasm security risk for the company producing the hybrid.

Alternatively, the female inbred can be mechanically detasseled by machine. Mechanical detasseling is approximately as reliable as hand detasseling, but is faster and less costly. However, most detasseling machines produce more damage to the plants than hand detasseling. Thus, no form of detasseling is presently entirely satisfactory, and a need continues to exist for alternatives which further reduce production costs and to eliminate self-pollination of the female parent in the production of hybrid seed.

Mutations that cause male sterility in plants have the potential to be useful in methods for hybrid seed production for crop plants such as maize and can lower production costs by eliminating the need for the labor-intensive removal of male flowers (also known as de-tasseling) from the maternal parent plants used as a hybrid parent. Mutations that cause male sterility in maize have been produced by a variety of methods such as X-rays or UV-irradiations, chemical treatments, or transposable element insertions (ms23, ms25, ms26, ms32) (Chaubal et al. (2000) Am J Bot 87:1193-1201). Conditional regulation of fertility genes through fertility/sterility "molecular switches" could enhance the options for designing new male-sterility systems for crop improvement (Unger et al. (2002) Transgenic Res 11:455-465).

Besides identification of novel genes impacting male fertility, there remains a need to provide a reliable system of producing genetic male sterility.

In U.S. Patent No. 5,478,369, a method is described by which the Ms45 male fertility gene was tagged and cloned on maize chromosome 9. Previously, there had been described a male fertility gene on chromosome 9, ms2, which had never been cloned and sequenced. It is not allelic to the gene referred to in the '369 patent. See Albertsen, M. and Phillips, R.L., "Developmental Cytology of 13 Genetic Male Sterile Loci in Maize" Canadian Journal of Genetics & Cytology 23:195-208 (Jan. 1981). The only fertility gene cloned before that had been the Arabidopsis gene described at Aarts, et al., supra.

Examples of genes that have been discovered subsequently that are important to male fertility are numerous and include the Arabidopsis ABORTED

MICROSPORES (AMS) gene, Sorensen et al., The Plant Journal (2003) 33(2):413-423); the Arabidopsis MS1 gene (Wilson et al., The Plant Journal (2001) 39(2):170-181); the NEF1 gene (Ariizumi et al., The Plant Journal (2004) 39(2):170-181); Arabidopsis AtGPAT1 gene (Zheng et al., The Plant Cell (2003) 15:1872-1887); the
5 Arabidopsis dde2-2 mutation was shown to be defective in the allene oxide synthase gene (Malek et al., Planta (2002) 216:187-192); the Arabidopsis faceless pollen-1 gene (flp1) (Ariizumi et al., Plant Mol. Biol. (2003) 53:107-116); the Arabidopsis MALE MEIOCYTE DEATH1 gene (Yang et al., The Plant Cell (2003) 15: 1281-1295); the tapetum-specific zinc finger gene, TAZ1 (Kapoor et al., The Plant Cell
10 (2002) 14:2353-2367); and the TAPETUM DETERMINANT1 gene (Lan et al., The Plant Cell (2003) 15:2792-2804).

Other known male fertility mutants or genes from Zea mays are listed in U.S. patent 7,919,676.

Other genes include kinases and those encoding compounds toxic to either
15 male or female gametophytic development.

Furthermore, it is recognized that the polynucleotide of interest may also comprise antisense sequences complementary to at least a portion of the messenger RNA (mRNA) for a targeted gene sequence of interest. Antisense nucleotides are constructed to hybridize with the corresponding mRNA.
20 Modifications of the antisense sequences may be made as long as the sequences hybridize to and interfere with expression of the corresponding mRNA. In this manner, antisense constructions having 70%, 80%, or 85% sequence identity to the corresponding antisense sequences may be used. Furthermore, portions of the antisense nucleotides may be used to disrupt the expression of the target gene.
25 Generally, sequences of at least 50 nucleotides, 100 nucleotides, 200 nucleotides, or greater may be used.

In addition, the polynucleotide of interest may also be used in the sense orientation to suppress the expression of endogenous genes in plants. Methods for suppressing gene expression in plants using polynucleotides in the sense
30 orientation are known in the art. The methods generally involve transforming plants with a DNA construct comprising a promoter that drives expression in a plant operably linked to at least a portion of a nucleotide sequence that corresponds to

the transcript of the endogenous gene. Typically, such a nucleotide sequence has substantial sequence identity to the sequence of the transcript of the endogenous gene, generally greater than about 65% sequence identity, about 85% sequence identity, or greater than about 95% sequence identity. See, U.S. Patent Nos.

5 5,283,184 and 5,034,323.

The polynucleotide of interest can also be a phenotypic marker. A phenotypic marker is screenable or a selectable marker that includes visual markers and selectable markers whether it is a positive or negative selectable marker. Any phenotypic marker can be used. Specifically, a selectable or screenable marker
10 comprises a DNA segment that allows one to identify, or select for or against a molecule or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to, production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like.

15 Examples of selectable markers include, but are not limited to, DNA segments that comprise restriction enzyme sites; DNA segments that encode products which provide resistance against otherwise toxic compounds including antibiotics, such as, spectinomycin, ampicillin, kanamycin, tetracycline, Basta, neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase
20 (HPT)); DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); DNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β -galactosidase, GUS; fluorescent proteins such as green fluorescent protein (GFP), cyan (CFP), yellow (YFP), red (RFP), and cell surface proteins); the generation of
25 new primer sites for PCR (e.g., the juxtaposition of two DNA sequence not previously juxtaposed), the inclusion of DNA sequences not acted upon or acted upon by a restriction endonuclease or other DNA modifying enzyme, chemical, etc.; and, the inclusion of a DNA sequences required for a specific modification (e.g., methylation) that allows its identification.

30 Additional selectable markers include genes that confer resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). See for example, Yarranton, (1992) Curr

Opin Biotech 3:506-11; Christopherson et al., (1992) Proc. Natl. Acad. Sci. USA 89:6314-8; Yao et al., (1992) Cell 71:63-72; Reznikoff, (1992) Mol Microbiol 6:2419-22; Hu et al., (1987) Cell 48:555-66; Brown et al., (1987) Cell 49:603-12; Figge et al., (1988) Cell 52:713-22; Deuschle et al., (1989) Proc. Natl. Acad. Sci. USA 86:5400-4; Fuerst et al., (1989) Proc. Natl. Acad. Sci. USA 86:2549-53; Deuschle et al., (1990) Science 248:480-3; Gossen, (1993) Ph.D. Thesis, University of Heidelberg; Reines et al., (1993) Proc. Natl. Acad. Sci. USA 90:1917-21; Labow et al., (1990) Mol Cell Biol 10:3343-56; Zambretti et al., (1992) Proc. Natl. Acad. Sci. USA 89:3952-6; Baim et al., (1991) Proc. Natl. Acad. Sci. USA 88:5072-6; Wyborski et al., (1991) Nucleic Acids Res 19:4647-53; Hillen and Wissman, (1989) Topics Mol Struc Biol 10:143-62; Degenkolb et al., (1991) Antimicrob Agents Chemother 35:1591-5; Kleinschmidt et al., (1988) Biochemistry 27:1094-104; Bonin, (1993) Ph.D. Thesis, University of Heidelberg; Gossen et al., (1992) Proc. Natl. Acad. Sci. USA 89:5547-51; Oliva et al., (1992) Antimicrob Agents Chemother 36:913-9; Hlavka et al., (1985) Handbook of Experimental Pharmacology, Vol. 78 (Springer-Verlag, Berlin); Gill et al., (1988) Nature 334:721-4. Commercial traits can also be encoded on a gene or genes that could increase for example, starch for ethanol production, or provide expression of proteins. Another important commercial use of transformed plants is the production of polymers and bioplastics such as described in U.S. Patent No. 5,602,321. Genes such as β -Ketothiolase, PHBase (polyhydroxybutyrate synthase), and acetoacetyl-CoA reductase (see Schubert *et al.* (1988) *J. Bacteriol.* 170:5837-5847) facilitate expression of polyhydroxyalkanoates (PHAs).

Exogenous products include plant enzymes and products as well as those from other sources including prokaryotes and other eukaryotes. Such products include enzymes, cofactors, hormones, and the like. The level of proteins, particularly modified proteins having improved amino acid distribution to improve the nutrient value of the plant, can be increased. This is achieved by the expression of such proteins having enhanced amino acid content.

The transgenes, recombinant DNA molecules, DNA sequences of interest, and polynucleotides of interest can be comprise one or more DNA sequences for gene silencing. Methods for gene silencing involving the expression of DNA

sequences in plant are known in the art include, but are not limited to, cosuppression, antisense suppression, double-stranded RNA (dsRNA) interference, hairpin RNA (hpRNA) interference, intron-containing hairpin RNA (ihpRNA) interference, transcriptional gene silencing, and micro RNA (miRNA) interference

5 As used herein, “nucleic acid” means a polynucleotide and includes a single or a double-stranded polymer of deoxyribonucleotide or ribonucleotide bases. Nucleic acids may also include fragments and modified nucleotides. Thus, the terms “polynucleotide”, “nucleic acid sequence”, “nucleotide sequence” and “nucleic acid fragment” are used interchangeably to denote a polymer of RNA and/or DNA
10 that is single- or double-stranded, optionally containing synthetic, non-natural, or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: “A” for adenosine or deoxyadenosine (for RNA or DNA, respectively), “C” for cytosine or deoxycytosine, “G” for guanosine or deoxyguanosine, “U” for uridine, “T” for
15 deoxythymidine, “R” for purines (A or G), “Y” for pyrimidines (C or T), “K” for G or T, “H” for A or C or T, “I” for inosine, and “N” for any nucleotide.

“Open reading frame” is abbreviated ORF.

The terms “subfragment that is functionally equivalent” and “functionally equivalent subfragment” are used interchangeably herein. These terms refer to a
20 portion or subsequence of an isolated nucleic acid fragment in which the ability to alter gene expression or produce a certain phenotype is retained whether or not the fragment or subfragment encodes an active enzyme. For example, the fragment or subfragment can be used in the design of genes to produce the desired phenotype in a transformed plant. genes can be designed for use in suppression by linking a
25 nucleic acid fragment or subfragment thereof, whether or not it encodes an active enzyme, in the sense or antisense orientation relative to a plant promoter sequence.

The term “conserved domain” or “motif” means a set of amino acids conserved at specific positions along an aligned sequence of evolutionarily related proteins. While amino acids at other positions can vary between homologous
30 proteins, amino acids that are highly conserved at specific positions indicate amino acids that are essential to the structure, the stability, or the activity of a protein. Because they are identified by their high degree of conservation in aligned

sequences of a family of protein homologues, they can be used as identifiers, or "signatures", to determine if a protein with a newly determined sequence belongs to a previously identified protein family.

Polynucleotide and polypeptide sequences, variants thereof, and the structural relationships of these sequences can be described by the terms "homology", "homologous", "substantially identical", "substantially similar" and "corresponding substantially" which are used interchangeably herein. These refer to polypeptide or nucleic acid fragments wherein changes in one or more amino acids or nucleotide bases do not affect the function of the molecule, such as the ability to mediate gene expression or to produce a certain phenotype. These terms also refer to modification(s) of nucleic acid fragments that do not substantially alter the functional properties of the resulting nucleic acid fragment relative to the initial, unmodified fragment. These modifications include deletion, substitution, and/or insertion of one or more nucleotides in the nucleic acid fragment.

Substantially similar nucleic acid sequences encompassed may be defined by their ability to hybridize (under moderately stringent conditions, e.g., 0.5X SSC, 0.1% SDS, 60°C) with the sequences exemplified herein, or to any portion of the nucleotide sequences disclosed herein and which are functionally equivalent to any of the nucleic acid sequences disclosed herein. Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions.

The term "selectively hybridizes" includes reference to hybridization, under stringent hybridization conditions, of a nucleic acid sequence to a specified nucleic acid target sequence to a detectably greater degree (e.g., at least 2-fold over background) than its hybridization to non-target nucleic acid sequences and to the substantial exclusion of non-target nucleic acids. Selectively hybridizing sequences typically have about at least 80% sequence identity, or 90% sequence identity, up to and including 100% sequence identity (i.e., fully complementary) with each other.

The term "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a probe will selectively hybridize to its target

sequence in an *in vitro* hybridization assay. Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous probing).

- 5 Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, optionally less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salt(s)) at pH 7.0 to 8.3, and at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

“Sequence identity” or “identity” in the context of nucleic acid or polypeptide sequences includes the nucleic acid bases or amino acid residues in two sequences that are the same when aligned for maximum correspondence over a specified comparison window.

The term “percentage of sequence identity” refers to the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in

both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the results by 100 to yield the percentage of sequence identity. Useful examples of percent sequence identities include, but are not limited to, 50%, 55%, 5 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95%, or any integer percentage from 50% to 100%. These identities can be determined using any of the programs described herein.

Sequence alignments and percent identity or similarity calculations may be determined using a variety of comparison methods designed to detect homologous 10 sequences including, but not limited to, the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the “default values” of the program referenced, unless otherwise specified. As used herein “default values” 15 will mean any set of values or parameters that originally load with the software when first initialized.

The “Clustal V method of alignment” corresponds to the alignment method labeled Clustal V (described by Higgins and Sharp, (1989) *CABIOS* 5:151-153; Higgins *et al.*, (1992) *Comput Appl Biosci* 8:189-191) and found in the MegAlign™ 20 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). For multiple alignments, the default values correspond to GAP PENALTY=10 and GAP LENGTH PENALTY=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS 25 SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a “percent identity” by viewing the “sequence distances” table in the same program.

The “Clustal W method of alignment” corresponds to the alignment method 30 labeled Clustal W (described by Higgins and Sharp, (1989) *CABIOS* 5:151-153; Higgins *et al.*, (1992) *Comput Appl Biosci* 8:189-191) and found in the MegAlign™ v6.1 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.,

Madison, WI). Default parameters for multiple alignment (GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergen Seqs (%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB). After alignment of the sequences using the Clustal W program, it is possible to obtain a “percent
5 identity” by viewing the “sequence distances” table in the same program.

Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using GAP Version 10 (GCG, Accelrys, San Diego, CA) using the following parameters: % identity and % similarity for a nucleotide sequence using a gap creation penalty weight of 50 and a gap length extension
10 penalty weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using a GAP creation penalty weight of 8 and a gap length extension penalty of 2, and the BLOSUM62 scoring matrix (Henikoff and Henikoff, (1989) *Proc. Natl. Acad. Sci. USA* 89:10915). GAP uses the algorithm of Needleman and Wunsch, (1970) *J Mol Biol* 48:443-53, to find an alignment of two
15 complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps, using a gap creation penalty and a gap extension penalty in units of matched bases.

20 “BLAST” is a searching algorithm provided by the National Center for Biotechnology Information (NCBI) used to find regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches to identify sequences having sufficient similarity to a query sequence such that the similarity
25 would not be predicted to have occurred randomly. BLAST reports the identified sequences and their local alignment to the query sequence.

It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides from other species or modified naturally or synthetically wherein such polypeptides have the same or similar
30 function or activity. Useful examples of percent identities include, but are not limited to, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95%, or any integer percentage from 50% to 100%. Indeed, any integer amino acid identity from 50% to

100% may be useful in describing the present disclosure, such as 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
5 96%, 97%, 98% or 99%.

“Gene” includes a nucleic acid fragment that expresses a functional molecule such as, but not limited to, a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its own
10 regulatory sequences.

A “mutated gene” is a gene that has been altered through human intervention. Such a “mutated gene” has a sequence that differs from the sequence of the corresponding non-mutated gene by at least one nucleotide addition, deletion, or substitution. In certain embodiments of the disclosure, the mutated gene
15 comprises an alteration that results from a guide polynucleotide/Cas endonuclease system as disclosed herein. A mutated plant is a plant comprising a mutated gene.

Targeted mutation includes a mutation in a native gene that was made by altering a target sequence within the native gene using a method involving a double-strand-break-inducing agent that is capable of inducing a double-strand break in the
20 DNA of the target sequence as disclosed herein or known in the art.

In one embodiment, the targeted mutation is the result of a guideRNA/Cas endonuclease induced gene editing, without incorporating a selectable transgene marker, as described herein. The guide RNA/Cas endonuclease induced targeted mutation can occur in a nucleotide sequence that is located within or outside a
25 genomic target site that is recognized and cleaved by a Cas endonuclease.

The term “genome” as it applies to a plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondria, or plastid) of the cell.

A “codon-modified gene” or “codon-preferred gene” or “codon-optimized
30 gene” is a gene having its frequency of codon usage designed to mimic the frequency of preferred codon usage of the host cell.

An "allele" is one of several alternative forms of a gene occupying a given locus on a chromosome. When all the alleles present at a given locus on a chromosome are the same, that plant is homozygous at that locus. If the alleles present at a given locus on a chromosome differ, that plant is heterozygous at that locus.

"Coding sequence" refers to a polynucleotide sequence which codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include, but are not limited to: promoters, translation leader sequences, 5' untranslated sequences, 3' untranslated sequences, introns, polyadenylation target sequences, RNA processing sites, effector binding sites, and stem-loop structures.

"A plant-optimized nucleotide sequence" is nucleotide sequence that has been optimized for increased expression in plants, particularly for increased expression in plants or in one or more plants of interest. For example, a plant-optimized nucleotide sequence can be synthesized by modifying a nucleotide sequence encoding a protein such as, for example, double-strand-break-inducing agent (e.g., an endonuclease) as disclosed herein, using one or more plant-preferred codons for improved expression. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of host-preferred codon usage.

Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498. Additional sequence modifications are known to enhance gene expression in a plant host. These include, for example, elimination of: one or more sequences encoding spurious polyadenylation signals, one or more exon-intron splice site signals, one or more transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the sequence may be adjusted to levels average for a given plant host, as calculated by reference to known genes expressed in the host plant cell. When possible, the sequence is

modified to avoid one or more predicted hairpin secondary mRNA structures. Thus, "a plant-optimized nucleotide sequence" of the present disclosure comprises one or more of such sequence modifications.

5 "Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. An "enhancer" is a DNA sequence that can stimulate promoter activity, and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived
10 in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, and/or comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. It is further
15 recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters".

It has been shown that certain promoters are able to direct RNA synthesis at
20 a higher rate than others. These are called "strong promoters". Certain other promoters have been shown to direct RNA synthesis at higher levels only in particular types of cells or tissues and are often referred to as "tissue specific promoters", or "tissue-preferred promoters" if the promoters direct RNA synthesis preferably in certain tissues but also in other tissues at reduced levels. Since
25 patterns of expression of a chimeric gene (or genes) introduced into a plant are controlled using promoters, there is an ongoing interest in the isolation of novel promoters which are capable of controlling the expression of a chimeric gene or (genes) at certain levels in specific tissue types or at specific plant developmental stages.

30 New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and

Goldberg, (1989) In *The Biochemistry of Plants*, Vol. 115, Stumpf and Conn, eds (New York, NY: Academic Press), pp. 1-82.

“Translation leader sequence” refers to a polynucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (e.g., Turner and Foster, (1995) *Mol Biotechnol* 3:225-236).

“3’ non-coding sequences”, “transcription terminator” or “termination sequences” refer to DNA sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3’ end of the mRNA precursor. The use of different 3’ non-coding sequences is exemplified by Ingelbrecht *et al.*, (1989) *Plant Cell* 1:671-680.

“RNA transcript” refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or pre-mRNA. A RNA transcript is referred to as the mature RNA or mRNA when it is a RNA sequence derived from post-transcriptional processing of the primary transcript pre mRNA. “Messenger RNA” or “mRNA” refers to the RNA that is without introns and that can be translated into protein by the cell. “cDNA” refers to a DNA that is complementary to, and synthesized from, a mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into double-stranded form using the Klenow fragment of DNA polymerase I. “Sense” RNA refers to RNA transcript that includes the mRNA and can be translated into protein within a cell or *in vitro*. “Antisense RNA” refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA, and that blocks the expression of a target gene (see, e.g., U.S. Patent No. 5,107,065). The complementarity of an antisense RNA may be with any part of the specific gene

transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" includes antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes. The terms "complement" and "reverse complement" are used interchangeably
5 herein with respect to mRNA transcripts, and are meant to define the antisense RNA of the message.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a coding sequence when it is
10 capable of regulating the expression of that coding sequence (i.e., the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in a sense or antisense orientation. In another example, the complementary RNA regions can be operably linked, either directly or indirectly, 5' to the target mRNA, or 3' to the target mRNA, or within the
15 target mRNA, or a first complementary region is 5' and its complement is 3' to the target mRNA.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor,
20 NY (1989). Transformation methods are well known to those skilled in the art and are described *infra*.

"PCR" or "polymerase chain reaction" is a technique for the synthesis of specific DNA segments and consists of a series of repetitive denaturation, annealing, and extension cycles. Typically, a double-stranded DNA is heat
25 denatured, and two primers complementary to the 3' boundaries of the target segment are annealed to the DNA at low temperature, and then extended at an intermediate temperature. One set of these three consecutive steps is referred to as a "cycle".

The term "recombinant" refers to an artificial combination of two otherwise
30 separated segments of sequence, e.g., by chemical synthesis, or manipulation of isolated segments of nucleic acids by genetic engineering techniques.

The terms "plasmid", "vector" and "cassette" refer to an extra chromosomal element often carrying genes that are not part of the central metabolism of the cell, and usually in the form of double-stranded DNA. Such elements may be autonomously replicating sequences, genome integrating sequences, phage, or nucleotide sequences, in linear or circular form, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a polynucleotide of interest into a cell. "Transformation cassette" refers to a specific vector containing a gene and having elements in addition to the gene that facilitates transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a gene and having elements in addition to the gene that allow for expression of that gene in a host.

The terms "recombinant DNA molecule", "recombinant construct", "expression construct", "construct", "construct", and "recombinant DNA construct" are used interchangeably herein. A recombinant construct comprises an artificial combination of nucleic acid fragments, e.g., regulatory and coding sequences that are not all found together in nature. For example, a construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. Such a construct may be used by itself or may be used in conjunction with a vector. If a vector is used, then the choice of vector is dependent upon the method that will be used to transform host cells as is well known to those skilled in the art. For example, a plasmid vector can be used. The skilled artisan is well aware of the genetic elements that must be present on the vector in order to successfully transform, select and propagate host cells. The skilled artisan will also recognize that different independent transformation events may result in different levels and patterns of expression (Jones *et al.*, (1985) *EMBO J* 4:2411-2418; De Almeida *et al.*, (1989) *Mol Gen Genetics* 218:78-86), and thus that multiple events are typically screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished standard molecular biological, biochemical, and other assays including Southern analysis of DNA, Northern analysis of mRNA expression, PCR,

real time quantitative PCR (qPCR), reverse transcription PCR (RT-PCR), immunoblotting analysis of protein expression, enzyme or activity assays, and/or phenotypic analysis.

5 The term “expression”, as used herein, refers to the production of a functional end-product (e.g., an mRNA, guide RNA, or a protein) in either precursor or mature form.

10 The term “introduced” means providing a nucleic acid (e.g., expression construct) or protein into a cell. Introduced includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell, and includes reference to the transient provision of a nucleic acid or protein to the cell. Introduced includes reference to stable or transient transformation methods, as well as sexually crossing. Thus, “introduced” in the context of inserting a nucleic acid fragment (e.g., a recombinant DNA construct/expression construct) into a cell, means “transfection” or
15 “transformation” or “transduction” and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid, or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

20 “Mature” protein refers to a post-translationally processed polypeptide (i.e., one from which any pre- or propeptides present in the primary translation product have been removed). “Precursor” protein refers to the primary product of translation of mRNA (i.e., with pre- and propeptides still present). Pre- and propeptides may be but are not limited to intracellular localization signals.

25 “Stable transformation” refers to the transfer of a nucleic acid fragment into a genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance. In contrast, “transient transformation” refers to the transfer of a nucleic acid fragment into the nucleus, or other DNA-containing organelle, of a host organism resulting in gene expression without
30 integration or stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as “transgenic” organisms.

The commercial development of genetically improved germplasm has also advanced to the stage of introducing multiple traits into crop plants, often referred to as a gene stacking approach. In this approach, multiple genes conferring different characteristics of interest can be introduced into a plant. Gene stacking can be
5 accomplished by many means including but not limited to co-transformation, retransformation, and crossing lines with different genes of interest.

The term "plant" refers to whole plants, plant organs, plant tissues, seeds, plant cells, seeds and progeny of the same. Plant cells include, without limitation, cells from seeds, suspension cultures, embryos, meristematic regions, callus tissue,
10 leaves, roots, shoots, gametophytes, sporophytes, pollen and microspores. Plant parts include differentiated and undifferentiated tissues including, but not limited to roots, stems, shoots, leaves, pollens, seeds, tumor tissue and various forms of cells and culture (e.g., single cells, protoplasts, embryos, and callus tissue). The plant tissue may be in plant or in a plant organ, tissue or cell culture. The term "plant
15 organ" refers to plant tissue or a group of tissues that constitute a morphologically and functionally distinct part of a plant. The term "genome" refers to the entire complement of genetic material (genes and non-coding sequences) that is present in each cell of an organism, or virus or organelle; and/or a complete set of chromosomes inherited as a (haploid) unit from one parent. "Progeny" comprises
20 any subsequent generation of a plant.

A transgenic plant includes, for example, a plant which comprises within its genome a heterologous polynucleotide introduced by a transformation step. The heterologous polynucleotide can be stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous
25 polynucleotide may be integrated into the genome alone or as part of a recombinant DNA construct. A transgenic plant can also comprise more than one heterologous polynucleotide within its genome. Each heterologous polynucleotide may confer a different trait to the transgenic plant. A heterologous polynucleotide can include a sequence that originates from a foreign species, or, if from the same species, can
30 be substantially modified from its native form. Transgenic can include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenics initially so altered

as well as those created by sexual crosses or asexual propagation from the initial transgenic. The alterations of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods, by the genome editing procedure described herein that does not result in an insertion of a foreign polynucleotide, or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation are not intended to be regarded as transgenic.

In certain embodiments of the disclosure, a fertile plant is a plant that produces viable male and female gametes and is self-fertile. Such a self-fertile plant can produce a progeny plant without the contribution from any other plant of a gamete and the genetic material contained therein. Other embodiments of the disclosure can involve the use of a plant that is not self-fertile because the plant does not produce male gametes, or female gametes, or both, that are viable or otherwise capable of fertilization. As used herein, a "male sterile plant" is a plant that does not produce male gametes that are viable or otherwise capable of fertilization. As used herein, a "female sterile plant" is a plant that does not produce female gametes that are viable or otherwise capable of fertilization. It is recognized that male-sterile and female-sterile plants can be female-fertile and male-fertile, respectively. It is further recognized that a male fertile (but female sterile) plant can produce viable progeny when crossed with a female fertile plant and that a female fertile (but male sterile) plant can produce viable progeny when crossed with a male fertile plant.

A "centimorgan" (cM) or "map unit" is the distance between two linked genes, markers, target sites, loci, or any pair thereof, wherein 1% of the products of meiosis are recombinant. Thus, a centimorgan is equivalent to a distance equal to an 1% average recombination frequency between the two linked genes, markers, target sites, loci, or any pair thereof.

Breeding methods and methods for selecting plants utilizing a two component RNA guide and Cas endonuclease system

The present disclosure finds use in the breeding of plants comprising one or more transgenic traits. Most commonly, transgenic traits are randomly inserted throughout the plant genome as a consequence of transformation systems based on

Agrobacterium, biolistics, or other commonly used procedures. More recently, gene targeting protocols have been developed that enable directed transgene insertion. One important technology, site-specific integration (SSI) enables the targeting of a transgene to the same chromosomal location as a previously inserted transgene.

- 5 Custom-designed meganucleases and custom-designed zinc finger meganucleases allow researchers to design nucleases to target specific chromosomal locations, and these reagents allow the targeting of transgenes at the chromosomal site cleaved by these nucleases.

The currently used systems for precision genetic engineering of eukaryotic
10 genomes, e.g. plant genomes, rely upon homing endonucleases, meganucleases, zinc finger nucleases, and transcription activator–like effector nucleases (TALENs), which require de novo protein engineering for every new target locus. The highly specific, RNA-directed DNA nuclease, guide RNA/ Cas9 endonuclease system described herein, is more easily customizable and therefore more useful when
15 modification of many different target sequences is the goal. This disclosure takes further advantage of the two component nature of the guide RNA/ Cas system, with its constant protein component, the Cas endonuclease, and its variable and easily reprogrammable targeting component, the guide RNA or the crRNA.

The guide RNA/Cas system described herein is especially useful for genome
20 engineering, especially plant genome engineering, in circumstances where nuclease off-target cutting can be toxic to the targeted cells. In one embodiment of the guide RNA/Cas system described herein, the constant component, in the form of an expression-optimized Cas9 gene, is stably integrated into the target genome, e.g. plant genome. Expression of the Cas9 gene is under control of a promoter, e.g.
25 plant promoter, which can be a constitutive promoter, tissue-specific promoter or inducible promoter, e.g. temperature-inducible, stress-inducible, developmental stage inducible, or chemically inducible promoter. In the absence of the variable component, i.e. the guide RNA or crRNA, the Cas9 protein is not able to cut DNA and therefore its presence in the plant cell should have little or no consequence.
30 Hence a key advantage of the guide RNA/Cas system described herein is the ability to create and maintain a cell line or transgenic organism capable of efficient expression of the Cas9 protein with little or no consequence to cell viability. In order

to induce cutting at desired genomic sites to achieve targeted genetic modifications, guide RNAs or crRNAs can be introduced by a variety of methods into cells containing the stably-integrated and expressed cas9 gene. For example, guide RNAs or crRNAs can be chemically or enzymatically synthesized, and introduced
5 into the Cas9 expressing cells via direct delivery methods such a particle bombardment or electroporation.

Alternatively, genes capable of efficiently expressing guide RNAs or crRNAs in the target cells can be synthesized chemically, enzymatically or in a biological system, and these genes can be introduced into the Cas9 expressing cells via direct
10 delivery methods such a particle bombardment, electroporation or biological delivery methods such as Agrobacterium mediated DNA delivery.

In one embodiment, the method comprises a method for introducing a polynucleotide of interest into a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first
15 guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be modified to confer herbicide resistance, wherein said first guide RNA and Cas endonuclease are capable of forming a first complex that enables the Cas endonuclease to introduce a double
20 strand break at a first target site located in or near said first endogenous gene in the genome of said plant cell, wherein said first polynucleotide modification template comprises at least one nucleotide modification of said first endogenous gene to render said endogenous gene capable of conferring herbicide resistance to a plant cell, wherein said second guide RNA and Cas endonuclease are capable of
25 forming a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant cell, wherein said second polynucleotide modification template comprises at least one polynucleotide of interest to be introduced into said plant genome.

As disclosed herein, a guide RNA/Cas system mediating gene targeting
30 without incorporating a selectable transgene marker can be used in methods for directing transgene insertion and / or for producing complex transgenic trait loci comprising multiple transgenes in a fashion similar as disclosed in

WO2013/0198888 (published August 1, 2013) where instead of using a double strand break inducing agent to introduce a gene of interest, a guide RNA/Cas system or a guide polynucleotide/Cas system as disclosed herein is used. In one embodiment, a complex transgenic trait locus is a genomic locus that has multiple transgenes genetically linked to each other. By inserting independent transgenes within 0.1, 0.2, 0.3, 0.4, 0.5, 1, 2, or even 5 centimorgans (cM) from each other, the transgenes can be bred as a single genetic locus (see, for example, U.S. patent application 13/427,138) or PCT application PCT/US2012/030061. After selecting a plant comprising a transgene, plants containing (at least) one transgene can be crossed to form an F1 that contains both transgenes. In progeny from these F1 (F2 or BC1) 1/500 progeny would have the two different transgenes recombined onto the same chromosome. The complex locus can then be bred as single genetic locus with both transgene traits. This process can be repeated to stack as many traits as desired.

Chromosomal intervals that correlate with a phenotype or trait of interest can be identified. A variety of methods well known in the art are available for identifying chromosomal intervals. The boundaries of such chromosomal intervals are drawn to encompass markers that will be linked to the gene controlling the trait of interest. In other words, the chromosomal interval is drawn such that any marker that lies within that interval (including the terminal markers that define the boundaries of the interval) can be used as a marker for northern leaf blight resistance. In one embodiment, the chromosomal interval comprises at least one QTL, and furthermore, may indeed comprise more than one QTL. Close proximity of multiple QTLs in the same interval may obfuscate the correlation of a particular marker with a particular QTL, as one marker may demonstrate linkage to more than one QTL. Conversely, e.g., if two markers in close proximity show co-segregation with the desired phenotypic trait, it is sometimes unclear if each of those markers identify the same QTL or two different QTL. The term "quantitative trait locus" or "QTL" refers to a region of DNA that is associated with the differential expression of a quantitative phenotypic trait in at least one genetic background, e.g., in at least one breeding population. The region of the QTL encompasses or is closely linked to the gene or genes that affect the trait in question. An "allele of a QTL" can comprise multiple

genes or other genetic factors within a contiguous genomic region or linkage group, such as a haplotype. An allele of a QTL can denote a haplotype within a specified window wherein said window is a contiguous genomic region that can be defined, and tracked, with a set of one or more polymorphic markers. A haplotype can be defined by the unique fingerprint of alleles at each marker within the specified window.

A variety of methods are available to identify those cells having an altered genome at or near a target site without using a screenable marker phenotype. Such methods can be viewed as directly analyzing a target sequence to detect any change in the target sequence, including but not limited to PCR methods, sequencing methods, nuclease digestion, Southern blots, and any combination thereof.

Proteins may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known. For example, amino acid sequence variants of the protein(s) can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations include, for example, Kunkel, (1985) *Proc. Natl. Acad. Sci. USA* 82:488-92; Kunkel *et al.*, (1987) *Meth Enzymol* 154:367-82; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein. Guidance regarding amino acid substitutions not likely to affect biological activity of the protein is found, for example, in the model of Dayhoff *et al.*, (1978) *Atlas of Protein Sequence and Structure* (Natl Biomed Res Found, Washington, D.C.). Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be preferable. Conservative deletions, insertions, and amino acid substitutions are not expected to produce radical changes in the characteristics of the protein, and the effect of any substitution, deletion, insertion, or combination thereof can be evaluated by routine screening assays. Assays for double-strand-break-inducing activity are known and generally measure the overall activity and specificity of the agent on DNA substrates containing target sites.

A variety of methods are known for the introduction of nucleotide sequences and polypeptides into an organism, including, for example, transformation, sexual crossing, and the introduction of the polypeptide, DNA, or mRNA into the cell.

Methods for contacting, providing, and/or introducing a composition into
 5 various organisms are known and include but are not limited to, stable transformation methods, transient transformation methods, virus-mediated methods, and sexual breeding. Stable transformation indicates that the introduced polynucleotide integrates into the genome of the organism and is capable of being inherited by progeny thereof. Transient transformation indicates that the introduced
 10 composition is only temporarily expressed or present in the organism.

Protocols for introducing polynucleotides and polypeptides into plants may vary depending on the type of plant or plant cell targeted for transformation, such as monocot or dicot. Suitable methods of introducing polynucleotides and polypeptides into plant cells and subsequent insertion into the plant genome include
 15 microinjection (Crossway *et al.*, (1986) *Biotechniques* 4:320-34 and U.S. Patent No. 6,300,543), meristem transformation (U.S. Patent No. 5,736,369), electroporation (Riggs *et al.*, (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-6, *Agrobacterium*-mediated transformation (U.S. Patent Nos. 5,563,055 and 5,981,840), direct gene transfer (Paszkowski *et al.*, (1984) *EMBO J* 3:2717-22), and ballistic particle acceleration
 20 (U.S. Patent Nos. 4,945,050; 5,879,918; 5,886,244; 5,932,782; Tomes *et al.*, (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment" in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg & Phillips (Springer-Verlag, Berlin); McCabe *et al.*, (1988) *Biotechnology* 6:923-6; Weissinger *et al.*, (1988) *Ann Rev Genet* 22:421-77; Sanford *et al.*, (1987)
 25 *Particulate Science and Technology* 5:27-37 (onion); Christou *et al.*, (1988) *Plant Physiol* 87:671-4 (soybean); Finer and McMullen, (1991) *In Vitro Cell Dev Biol* 27P:175-82 (soybean); Singh *et al.*, (1998) *Theor Appl Genet* 96:319-24 (soybean); Datta *et al.*, (1990) *Biotechnology* 8:736-40 (rice); Klein *et al.*, (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-9 (maize); Klein *et al.*, (1988) *Biotechnology* 6:559-63
 30 (maize); U.S. Patent Nos. 5,240,855; 5,322,783 and 5,324,646; Klein *et al.*, (1988) *Plant Physiol* 91:440-4 (maize); Fromm *et al.*, (1990) *Biotechnology* 8:833-9 (maize); Hooykaas-Van Slogteren *et al.*, (1984) *Nature* 311:763-4; U.S. Patent No.

5,736,369 (cereals); Bytebier *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-9 (*Liliaceae*); De Wet *et al.*, (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.*, (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.*, (1990) *Plant Cell Rep* 9:415-8) and Kaeppler *et al.*, (1992) *Theor Appl Genet* 5 84:560-6 (whisker-mediated transformation); D'Halluin *et al.*, (1992) *Plant Cell* 4:1495-505 (electroporation); Li *et al.*, (1993) *Plant Cell Rep* 12:250-5; Christou and Ford (1995) *Annals Botany* 75:407-13 (rice) and Osjoda *et al.*, (1996) *Nat Biotechnol* 14:745-50 (maize via *Agrobacterium tumefaciens*).

Alternatively, polynucleotides may be introduced into plants by contacting
 10 plants with a virus or viral nucleic acids. Generally, such methods involve incorporating a polynucleotide within a viral DNA or RNA molecule. In some examples a polypeptide of interest may be initially synthesized as part of a viral polyprotein, which is later processed by proteolysis *in vivo* or *in vitro* to produce the desired recombinant protein. Methods for introducing polynucleotides into plants
 15 and expressing a protein encoded therein, involving viral DNA or RNA molecules, are known, see, for example, U.S. Patent Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367 and 5,316,931. Transient transformation methods include, but are not limited to, the introduction of polypeptides, such as a double-strand break inducing agent, directly into the organism, the introduction of polynucleotides such as DNA
 20 and/or RNA polynucleotides, and the introduction of the RNA transcript, such as an mRNA encoding a double-strand break inducing agent, into the organism. Such methods include, for example, microinjection or particle bombardment. See, for example Crossway *et al.*, (1986) *Mol Gen Genet* 202:179-85; Nomura *et al.*, (1986) *Plant Sci* 44:53-8; Hepler *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91:2176-80; and,
 25 Hush *et al.*, (1994) *J Cell Sci* 107:775-84.

The term "dicot" refers to the subclass of angiosperm plants also known as "dicotyledoneae" and includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds, plant cells, and progeny of the same. Plant cell, as used
 30 herein includes, without limitation, seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores.

The term “crossed” or “cross” or “crossing” in the context of this disclosure means the fusion of gametes via pollination to produce progeny (i.e., cells, seeds, or plants). The term encompasses both sexual crosses (the pollination of one plant by another) and selfing (self-pollination, i.e., when the pollen and ovule (or microspores and megaspores) are from the same plant or genetically identical plants).

The term “introgression” refers to the transmission of a desired allele of a genetic locus from one genetic background to another. For example, introgression of a desired allele at a specified locus can be transmitted to at least one progeny plant via a sexual cross between two parent plants, where at least one of the parent plants has the desired allele within its genome. Alternatively, for example, transmission of an allele can occur by recombination between two donor genomes, e.g., in a fused protoplast, where at least one of the donor protoplasts has the desired allele in its genome. The desired allele can be, e.g., a transgene, a modified (mutated or edited) native allele, or a selected allele of a marker or QTL.

Standard DNA isolation, purification, molecular cloning, vector construction, and verification/characterization methods are well established, see, for example Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, NY). Vectors and constructs include circular plasmids, and linear polynucleotides, comprising a polynucleotide of interest and optionally other components including linkers, adapters, regulatory regions, introns, restriction sites, enhancers, insulators, selectable markers, nucleotide sequences of interest, promoters, and/or other sites that aid in vector construction or analysis. In some examples a recognition site and/or target site can be contained within an intron, coding sequence, 5' UTRs, 3' UTRs, and/or regulatory regions.

The present disclosure further provides expression constructs for expressing in a plant, plant cell, or plant part a guide RNA/Cas system that is capable of binding to and creating a double strand break in a target site. In one embodiment, the expression constructs of the disclosure comprise a promoter operably linked to a nucleotide sequence encoding a Cas gene and a promoter operably linked to a guide RNA of the present disclosure. The promoter is capable of driving expression of an operably linked nucleotide sequence in a plant cell.

A promoter is a region of DNA involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A plant promoter is a promoter capable of initiating transcription in a plant cell, for a review of plant promoters, see, Potenza *et al.*, (2004) *In Vitro Cell Dev Biol* 40:1-22. Constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell *et al.*, (1985) *Nature* 313:810-2); rice actin (McElroy *et al.*, (1990) *Plant Cell* 2:163-71); ubiquitin (Christensen *et al.*, (1989) *Plant Mol Biol* 12:619-32; Christensen *et al.*, (1992) *Plant Mol Biol* 18:675-89); pEMU (Last *et al.*, (1991) *Theor Appl Genet* 81:581-8); MAS (Velten *et al.*, (1984) *EMBO J* 3:2723-30); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters are described in, for example, U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142 and 6,177,611. In some examples an inducible promoter may be used. Pathogen-inducible promoters induced following infection by a pathogen include, but are not limited to those regulating expression of PR proteins, SAR proteins, beta-1,3-glucanase, chitinase, *etc.*

Chemical-regulated promoters can be used to modulate the expression of a gene in a plant through the application of an exogenous chemical regulator. The promoter may be a chemical-inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters include, but are not limited to, the maize In2-2 promoter, activated by benzene sulfonamide herbicide safeners (De Veylder *et al.*, (1997) *Plant Cell Physiol* 38:568-77), the maize GST promoter (GST-II-27, WO93/01294), activated by hydrophobic electrophilic compounds used as pre-emergent herbicides, and the tobacco PR-1a promoter (Ono *et al.*, (2004) *Biosci Biotechnol Biochem* 68:803-7) activated by salicylic acid. Other chemical-regulated promoters include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter (Schena *et al.*, (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-5; McNellis *et al.*, (1998) *Plant J* 14:247-257); tetracycline-inducible and tetracycline-repressible promoters (Gatz *et al.*, (1991) *Mol Gen Genet* 227:229-37; U.S. Patent Nos. 5,814,618 and 5,789,156).

Tissue-preferred promoters can be utilized to target enhanced expression within a particular plant tissue. Tissue-preferred promoters include, for example, Kawamata *et al.*, (1997) *Plant Cell Physiol* 38:792-803; Hansen *et al.*, (1997) *Mol Gen Genet* 254:337-43; Russell *et al.*, (1997) *Transgenic Res* 6:157-68; Rinehart *et al.*, (1996) *Plant Physiol* 112:1331-41; Van Camp *et al.*, (1996) *Plant Physiol* 112:525-35; Canevascini *et al.*, (1996) *Plant Physiol* 112:513-524; Lam, (1994) *Results Probl Cell Differ* 20:181-96; and Guevara-Garcia *et al.*, (1993) *Plant J* 4:495-505. Leaf-preferred promoters include, for example, Yamamoto *et al.*, (1997) *Plant J* 12:255-65; Kwon *et al.*, (1994) *Plant Physiol* 105:357-67; Yamamoto *et al.*, (1994) *Plant Cell Physiol* 35:773-8; Gotor *et al.*, (1993) *Plant J* 3:509-18; Orozco *et al.*, (1993) *Plant Mol Biol* 23:1129-38; Matsuoka *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90:9586-90; Simpson *et al.*, (1958) *EMBO J* 4:2723-9; Timko *et al.*, (1988) *Nature* 318:57-8. Root-preferred promoters include, for example, Hire *et al.*, (1992) *Plant Mol Biol* 20:207-18 (soybean root-specific glutamine synthase gene); Miao *et al.*, (1991) *Plant Cell* 3:11-22 (cytosolic glutamine synthase (GS)); Keller and Baumgartner, (1991) *Plant Cell* 3:1051-61 (root-specific control element in the GRP 1.8 gene of French bean); Sanger *et al.*, (1990) *Plant Mol Biol* 14:433-43 (root-specific promoter of *A. tumefaciens* mannopine synthase (MAS)); Bogusz *et al.*, (1990) *Plant Cell* 2:633-41 (root-specific promoters isolated from *Parasponia andersonii* and *Trema tomentosa*); Leach and Aoyagi, (1991) *Plant Sci* 79:69-76 (*A. rhizogenes* rolC and rolD root-inducing genes); Teeri *et al.*, (1989) *EMBO J* 8:343-50 (*Agrobacterium* wound-induced TR1' and TR2' genes); VtENOD-GRP3 gene promoter (Kuster *et al.*, (1995) *Plant Mol Biol* 29:759-72); and rolB promoter (Capana *et al.*, (1994) *Plant Mol Biol* 25:681-91; phaseolin gene (Murai *et al.*, (1983) *Science* 23:476-82; Sengopta-Gopalen *et al.*, (1988) *Proc. Natl. Acad. Sci. USA* 82:3320-4). See also, U.S. Patent Nos. 5,837,876; 5,750,386; 5,633,363; 5,459,252; 5,401,836; 5,110,732 and 5,023,179.

Seed-preferred promoters include both seed-specific promoters active during seed development, as well as seed-germinating promoters active during seed germination. See, Thompson *et al.*, (1989) *BioEssays* 10:108. Seed-preferred promoters include, but are not limited to, Cim1 (cytokinin-induced message); cZ19B1 (maize 19 kDa zein); and milps (myo-inositol-1-phosphate synthase);

(WO00/11177; and U.S. Patent 6,225,529). For dicots, seed-preferred promoters include, but are not limited to, bean β -phaseolin, napin, β -conglycinin, soybean lectin, cruciferin, and the like. For monocots, seed-preferred promoters include, but are not limited to, maize 15 kDa zein, 22 kDa zein, 27 kDa gamma zein, waxy, shrunk 1, shrunk 2, globulin 1, oleosin, and nuc1. See also, WO00/12733, where seed-preferred promoters from *END1* and *END2* genes are disclosed.

A phenotypic marker is a screenable or selectable marker that includes visual markers and selectable markers whether it is a positive or negative selectable marker. Any phenotypic marker can be used. Specifically, a selectable or screenable marker comprises a DNA segment that allows one to identify, or select for or against a molecule or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to, production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like.

Examples of selectable markers include, but are not limited to, DNA segments that comprise restriction enzyme sites; DNA segments that encode products which provide resistance against otherwise toxic compounds including antibiotics, such as, spectinomycin, ampicillin, kanamycin, tetracycline, Basta, neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT); DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); DNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β -galactosidase, GUS; fluorescent proteins such as green fluorescent protein (GFP), cyan (CFP), yellow (YFP), red (RFP), and cell surface proteins); the generation of new primer sites for PCR (e.g., the juxtaposition of two DNA sequence not previously juxtaposed), the inclusion of DNA sequences not acted upon or acted upon by a restriction endonuclease or other DNA modifying enzyme, chemical, etc.; and, the inclusion of a DNA sequences required for a specific modification (e.g., methylation) that allows its identification.

Additional selectable markers include genes that confer resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). See for example, Yarranton, (1992) *Curr*

- Opin Biotech* 3:506-11; Christopherson *et al.*, (1992) *Proc. Natl. Acad. Sci. USA* 89:6314-8; Yao *et al.*, (1992) *Cell* 71:63-72; Reznikoff, (1992) *Mol Microbiol* 6:2419-22; Hu *et al.*, (1987) *Cell* 48:555-66; Brown *et al.*, (1987) *Cell* 49:603-12; Figge *et al.*, (1988) *Cell* 52:713-22; Deuschle *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:5400-4; Fuerst *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:2549-53; Deuschle *et al.*, (1990) *Science* 248:480-3; Gossen, (1993) Ph.D. Thesis, University of Heidelberg; Reines *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90:1917-21; Labow *et al.*, (1990) *Mol Cell Biol* 10:3343-56; Zambretti *et al.*, (1992) *Proc. Natl. Acad. Sci. USA* 89:3952-6; Baim *et al.*, (1991) *Proc. Natl. Acad. Sci. USA* 88:5072-6; Wyborski *et al.*, (1991) *Nucleic Acids Res* 19:4647-53; Hillen and Wissman, (1989) *Topics Mol Struc Biol* 10:143-62; Degenkolb *et al.*, (1991) *Antimicrob Agents Chemother* 35:1591-5; Kleinschmidt *et al.*, (1988) *Biochemistry* 27:1094-104; Bonin, (1993) Ph.D. Thesis, University of Heidelberg; Gossen *et al.*, (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-51; Oliva *et al.*, (1992) *Antimicrob Agents Chemother* 36:913-9; Hlavka *et al.*, (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.*, (1988) *Nature* 334:721-4.

The cells having the introduced sequence may be grown or regenerated into plants using conventional conditions, see for example, McCormick *et al.*, (1986) *Plant Cell Rep* 5:81-4. These plants may then be grown, and either pollinated with the same transformed strain or with a different transformed or untransformed strain, and the resulting progeny having the desired characteristic and/or comprising the introduced polynucleotide or polypeptide identified. Two or more generations may be grown to ensure that the polynucleotide is stably maintained and inherited, and seeds harvested.

Any plant can be used, including monocot and dicot plants. Examples of monocot plants that can be used include, but are not limited to, corn (*Zea mays*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*)), wheat (*Triticum aestivum*), sugarcane (*Saccharum spp.*), oats (*Avena*), barley (*Hordeum*), switchgrass (*Panicum virgatum*), pineapple (*Ananas comosus*), banana (*Musa spp.*), palm, ornamentals, turfgrasses, and other grasses. Examples of dicot plants

that can be used include, but are not limited to, soybean (*Glycine max*), canola (*Brassica napus* and *B. campestris*), alfalfa (*Medicago sativa*), tobacco (*Nicotiana tabacum*), *Arabidopsis* (*Arabidopsis thaliana*), sunflower (*Helianthus annuus*), cotton (*Gossypium arboreum*), and peanut (*Arachis hypogaea*), tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*) etc.

The transgenes, recombinant DNA molecules, DNA sequences of interest, and polynucleotides of interest can comprise one or more genes of interest. Such genes of interest can encode, for example, a protein that provides agronomic advantage to the plant.

10 *Marker Assisted Selection and Breeding of Plants*

A primary motivation for development of molecular markers in crop species is the potential for increased efficiency in plant breeding through marker assisted selection (MAS). Genetic marker alleles, or alternatively, quantitative trait loci (QTL alleles, are used to identify plants that contain a desired genotype at one or more loci, and that are expected to transfer the desired genotype, along with a desired phenotype to their progeny. Genetic marker alleles (or QTL alleles) can be used to identify plants that contain a desired genotype at one locus, or at several unlinked or linked loci (e.g., a haplotype), and that would be expected to transfer the desired genotype, along with a desired phenotype to their progeny. It will be appreciated that for the purposes of MAS, the term marker can encompass both marker and QTL loci.

After a desired phenotype and a polymorphic chromosomal locus, e.g., a marker locus or QTL, are determined to segregate together, it is possible to use those polymorphic loci to select for alleles corresponding to the desired phenotype - a process called marker-assisted selection (MAS). In brief, a nucleic acid corresponding to the marker nucleic acid is detected in a biological sample from a plant to be selected. This detection can take the form of hybridization of a probe nucleic acid to a marker, e.g., using allele-specific hybridization, southern blot analysis, northern blot analysis, in situ hybridization, hybridization of primers followed by PCR amplification of a region of the marker or the like. A variety of procedures for detecting markers are well known in the art. After the presence (or

absence) of a particular marker in the biological sample is verified, the plant is selected, i.e., used to make progeny plants by selective breeding.

Plant breeders need to combine traits of interest with genes for high yield and other desirable traits to develop improved plant varieties. Screening for large numbers of samples can be expensive, time consuming, and unreliable. Use of markers, and/or genetically-linked nucleic acids is an effective method for selecting plant having the desired traits in breeding programs. For example, one advantage of marker-assisted selection over field evaluations is that MAS can be done at any time of year regardless of the growing season. Moreover, environmental effects are irrelevant to marker-assisted selection.

When a population is segregating for multiple loci affecting one or multiple traits, the efficiency of MAS compared to phenotypic screening becomes even greater because all the loci can be processed in the lab together from a single sample of DNA.

The DNA repair mechanisms of cells are the basis to introduce extraneous DNA or induce mutations on endogenous genes. DNA homologous recombination is a specialized way of DNA repair that the cells repair DNA damages using a homologous sequence. In plants, DNA homologous recombination happens at frequencies too low to be routinely used in gene targeting or gene editing until it has been found that the process can be stimulated by DNA double-strand breaks (Bibikova et al., (2001) Mol. Cell Biol. 21:289-297; Puchta and Baltimore, (2003) Science 300:763; Wright et al., (2005) Plant J. 44:693-705).

The meaning of abbreviations is as follows: “sec” means second(s), “min” means minute(s), “h” means hour(s), “d” means day(s), “μL” means microliter(s), “mL” means milliliter(s), “L” means liter(s), “μM” means micromolar, “mM” means millimolar, “M” means molar, “mmol” means millimole(s), “μmole” mean micromole(s), “g” means gram(s), “μg” means microgram(s), “ng” means nanogram(s), “U” means unit(s), “bp” means base pair(s) and “kb” means kilobase(s).

Also, as described herein, for each example or embodiment that cites a guide RNA, a similar guide *polynucleotide* can be designed wherein the guide polynucleotide does not solely comprise ribonucleic acids but wherein the guide

polynucleotide comprises a combination of RNA-DNA molecules or solely comprises DNA molecules.

Non-limiting examples of compositions and methods disclosed herein are as follows:

- 5 1. A method for producing a genetic modification into a second gene of a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide polynucleotide, a polynucleotide modification template, a second guide polynucleotide, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be
 - 10 modified to confer herbicide resistance, wherein said first guide polynucleotide and Cas endonuclease are capable of forming a first complex that enables the Cas endonuclease to introduce a double strand break at a first target site, located in or near said first endogenous gene in the genome of said plant cell, wherein said second guide polynucleotide and Cas endonuclease are capable of
 - 15 forming a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant cell, wherein said polynucleotide modification template comprises at least one nucleotide alteration when compared to the first endogenous gene.
 2. The method of embodiment 1, wherein said at least one nucleotide alteration
 - 20 encodes for an amino acid change in first endogenous gene.
 3. The method of embodiment 1, wherein said first endogenous gene is modified to confer herbicide resistance to a plant cell.
 4. The method of embodiment 1, wherein said first endogenous gene is selected from the group consisting of an acetolactate synthase (ALS) and an
 - 25 enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene.
 5. The method of embodiment 1, wherein the first target site and the second target site are located at two different genomic loci.
 6. The method of embodiment 4, wherein the endogenous ALS gene is modified to confer sulfonylurea resistance.
 - 30 7. The method of embodiment 4, wherein the endogenous EPSPS gene is modified to confer glyphosate resistance.

8. The method of embodiment 1, wherein Cas endonuclease is a Cas9 endonuclease
9. The method of embodiment 1, further comprising selecting at least one plant cell that has resistance to an herbicide and comprises a modification into said second gene, wherein said modification includes at least one deletion, insertion or substitution of one or more nucleotides in said second gene of the plant genome.
10. The method of embodiment 9, wherein said herbicide is a sulphonylurea, such as chlorsulfuron or an imidazolinone herbicide.
11. The method of embodiment 9, wherein said modification into said second gene includes at least one deletion, insertion, or substitution of one or more nucleotides in a FAD2-1 gene.
12. A plant grown or cultured from the plant cell of embodiment 9, a seed thereof, or progeny thereof
13. The plant of embodiment 12, wherein the plant is a monocot or a dicot.
14. The plant of embodiment 13, wherein the monocot is selected from the group consisting of maize, rice, sorghum, rye, barley, wheat, millet, oats, sugarcane, turfgrass, or switchgrass.
15. The plant of embodiment 13, wherein the dicot is selected from the group consisting of soybean, canola, alfalfa, sunflower, cotton, tobacco, peanut, potato, tobacco, Arabidopsis, or safflower.
16. An herbicide resistant plant grown or cultured from the plant cell of embodiment 9, a seed thereof, or progeny thereof having herbicide resistance and a high oleic phenotype.
17. A method for introducing a polynucleotide of interest into a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be modified to confer herbicide resistance, wherein said first guide RNA and Cas endonuclease are capable of forming a first complex that enables the Cas endonuclease to introduce a double strand break

- at a first target site located in or near said first endogenous gene in the genome of said plant cell, wherein said first polynucleotide modification template comprises at least one nucleotide modification of said first endogenous gene to render said endogenous gene capable of conferring herbicide resistance to a plant cell, wherein said second guide RNA and Cas endonuclease are capable of forming a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant cell, wherein said second polynucleotide modification template comprises at least one polynucleotide of interest to be introduced into said plant genome.
18. The method of embodiment 17, wherein the first target site and the second target site are located at two different genomic loci.
 19. The method of embodiment 17, wherein said first endogenous gene is selected from the group consisting of a plant acetolactate synthase (ALS) and a plant enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene.
 20. The method of embodiment 17, wherein the endogenous ALS gene is modified to confer sulfonylurea resistance.
 21. The method of embodiment 17, wherein the endogenous EPSPS gene is modified to confer glyphosate resistance.
 22. The method of embodiment 17, wherein Cas endonuclease is a Cas9 endonuclease
 23. The method of embodiment 17, further comprising selecting at least one plant cell that has resistance to an herbicide and comprises said at least one polynucleotide of interest in said plant genome.
 24. The method of embodiment 17, wherein the at least one polynucleotide of interest is selected from the group consisting of herbicide-tolerance coding sequences, insecticidal coding sequences, nematocidal coding sequences, antimicrobial coding sequences, antifungal coding sequences, antiviral coding sequences, abiotic and biotic stress tolerance coding sequences, or sequences modifying plant traits such as yield, grain quality, nutrient content, starch quality and quantity, nitrogen fixation and/or utilization, and oil content and/or composition.

25. The method of embodiment 23, wherein said herbicide is a sulphonylurea, such as chlorsulfuron or an imidazolinone herbicide.
26. A plant grown or cultured from the plant cell of embodiment 23, a seed thereof, or progeny thereof
- 5 27. The plant of embodiment 26, wherein the plant is a monocot or a dicot.
28. An herbicide resistant plant grown or cultured from the plant cell of embodiment 23, a seed thereof, or progeny thereof having herbicide resistance and stably inherited the at least one polynucleotide of interest in said plant genome.
- 10 29. A method for editing a second gene of a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and a Cas endonuclease to a plant cell comprising a first endogenous gene that can be
- 15 modified to confer herbicide resistance, wherein said first guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at a first target site (located in or near said first endogenous gene) in the genome of said plant cell, wherein said first polynucleotide modification template comprises at least one nucleotide
- 20 modification of said first endogenous gene to render said endogenous gene capable of conferring herbicide resistance to a plant cell, wherein said second guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to introduce a double strand break at a second target site (located at a different locus than said first endogenous gene) in the
- 25 genome of said plant cell, wherein said second polynucleotide modification template comprises at least one nucleotide alteration when compared to the second gene to be edited.
30. The method of embodiment 29, wherein said first endogenous gene is an acetolactate synthase (ALS) gene.
- 30 31. The method of embodiment 29, wherein the endogenous ALS gene is modified to confer sulfonylurea resistance.

32. The method of embodiment 29, wherein the second gene is an EPSPS gene modified to confer glyphosate resistance.
33. The method of embodiment 29, wherein Cas endonuclease is a Cas9 endonuclease
- 5 34. The method of embodiment 29, further comprising selecting at least one plant cell that has resistance to an herbicide and comprises a modified EPSPS gene that confers glyphosate resistance.
35. The method of embodiment 34, wherein said herbicide is a sulphonylurea, such as chlorsulfuron or an imidazolinone herbicide.
- 10 36. A plant grown or cultured from the plant cell of embodiment 34, a seed thereof, or progeny thereof
37. The plant of embodiment 36, wherein the plant is a monocot or a dicot.
38. A method of generating a glyphosate resistant soybean plant without introducing an exogenous selectable marker into said plant genome the method
- 15 comprising providing a maize plant cell wherein its endogenous chromosomal ALS gene and endogenous chromosomal EPSPS gene have been modified through a guide RNA/Cas endonuclease system to produce a glyphosate resistant EPSPS protein and growing a maize plant from said maize plant cell, wherein said plant is resistant to glyphosate.

EXAMPLES

In the following Examples, unless otherwise stated, parts and percentages are by weight and degrees are Celsius. It should be understood that these Examples, while indicating embodiments of the disclosure, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can make various changes and modifications of the disclosure to adapt it to various usages and conditions. Such modifications are also intended to fall within the scope of the appended claims.

EXAMPLE 1

10 Maize optimized expression cassettes for guide RNA/Cas endonuclease based genome modification in maize plants

For genome engineering applications, the type II CRISPR/Cas system minimally requires the Cas9 protein and a duplexed crRNA/tracrRNA molecule or a synthetically fused crRNA and tracrRNA (guide RNA) molecule for DNA target site recognition and cleavage (Gasiunas *et al.* (2012) *Proc. Natl. Acad. Sci. USA* 109:E2579-86, Jinek *et al.* (2012) *Science* 337:816-21, Mali *et al.* (2013) *Science* 339:823-26, and Cong *et al.* (2013) *Science* 339:819-23). Described herein is a guideRNA/Cas endonuclease system that is based on the type II CRISPR/Cas system and consists of a Cas endonuclease and a guide RNA (or duplexed crRNA and tracrRNA) that together can form a complex that recognizes a genomic target site in a plant and introduces a double-strand break into said target site.

To test the guide RNA/Cas endonuclease system in maize, the Cas9 gene from *Streptococcus pyogenes* M1 GAS (SF370) (SEQ ID NO: 1) was maize codon optimized per standard techniques known in the art and the potato ST-LS1 intron (SEQ ID NO: 2) was introduced in order to eliminate its expression in *E.coli* and *Agrobacterium* (Figure 1 A). To facilitate nuclear localization of the Cas9 protein in maize cells, *Simian virus 40* (SV40) monopartite amino terminal nuclear localization signal (MAPKKKRKV, SEQ ID NO: 3) and *Agrobacterium tumefaciens* bipartite VirD2 T-DNA border endonuclease carboxyl terminal nuclear localization signal (KRPRDRHDGELGGRKRAR, SEQ ID NO: 4) were incorporated at the amino and carboxyl-termini of the Cas9 open reading frame (Figure 1 A), respectively. The maize optimized Cas9 gene was operably linked to a maize constitutive or regulated

promoter by standard molecular biological techniques. An example of the maize optimized Cas9 expression cassette (SEQ ID NO: 5) is illustrated in Figure 1 A. Figure 1A shows a maize optimized Cas9 gene containing the ST-LS1 intron, SV40 amino terminal nuclear localization signal (NLS) and VirD2 carboxyl terminal NLS driven by a plant Ubiquitin promoter.

The second component necessary to form a functional guide RNA/Cas endonuclease system for genome engineering applications is a duplex of the crRNA and tracrRNA molecules or a synthetic fusing of the crRNA and tracrRNA molecules, a guide RNA. To confer efficient guide RNA expression (or expression of the duplexed crRNA and tracrRNA) in maize, the maize U6 polymerase III promoter (SEQ ID NO: 9) and maize U6 polymerase III terminator (first 8 bases of SEQ ID NO: 10) residing on chromosome 8 were isolated and operably fused to the termini of a guide RNA (Figure 1 B) using standard molecular biology techniques. Two different guide RNA configurations were developed for testing in maize, a short guide RNA (SEQ ID NO: 11) based on Jinek *et al.* (2012) *Science* 337:816-21 and a long guide RNA (SEQ ID NO: 8) based on Mali *et al.* (2013) *Science* 339:823-26. An example expression cassette (SEQ ID NO: 12) is shown in Figure 1 B which illustrates a maize U6 polymerase III promoter driving expression of a long guide RNA terminated with a U6 polymerase III terminator.

As shown in Figures 2 A and 2B, the guide RNA or crRNA molecule contains a region complementary to one strand of the double strand DNA target (referred to as the variable targeting domain) that is approximately 12-30 nucleotides in length and upstream of a PAM sequence (5'NGG3' on antisense strand of Figure 2A-2B, corresponding to 5'CCN3' on sense strand of Figure 2A-2B) for target site recognition and cleavage (Gasiunas *et al.* (2012) *Proc. Natl. Acad. Sci. USA* 109:E2579-86, Jinek *et al.* (2012) *Science* 337:816-21, Mali *et al.* (2013) *Science* 339:823-26, and Cong *et al.* (2013) *Science* 339:819-23). To facilitate the rapid introduction of maize genomic DNA target sequences into the crRNA or guide RNA expression constructs, two Type IIS BbsI restriction endonuclease target sites were introduced in an inverted tandem orientation with cleavage orientated in an outward direction as described in Cong *et al.* (2013) *Science* 339:819-23. Upon cleavage, the Type IIS restriction endonuclease excises its target sites from the crRNA or

guide RNA expression plasmid, generating overhangs allowing for the in-frame directional cloning of duplexed oligos containing the desired maize genomic DNA target site into the variable targeting domain. In this example, only target sequences starting with a G nucleotide were used to promote favorable polymerase

5 III expression of the guide RNA or crRNA,.

Expression of both the Cas endonuclease gene and the guide RNA then allows for the formation of the guide RNA/Cas complex depicted in Figure 2 B (SEQ ID NO: 8). Alternatively, expression of the Cas endonucleases gene, crRNA, and tracrRNA allow for the formation of the crRNA/tracrRNA/Cas complex as depicted in

10 Figure 2 A, (SEQ ID NOs: 6-7).

EXAMPLE 2

The guide RNA/Cas endonuclease system cleaves chromosomal DNA in maize and introduces mutations by imperfect non-homologous end-joining

To test whether the maize optimized guide RNA/Cas endonuclease

15 described in example 1 could recognize, cleave, and mutate maize chromosomal DNA through imprecise non-homologous end-joining (NHEJ) repair pathways, three different genomic target sequences in 5 maize loci were targeted for cleavage (see Table 1) and examined by deep sequencing for the presence of NHEJ mutations.

20 Table 1. Maize genomic target sites targeted by a guideRNA/Cas endonuclease system.

<u>Locus</u>	<u>Location</u>	<u>Guide RNA Used</u>	<u>Target Site Designation</u>	<u>Maize Genomic Target Site Sequence</u>	<u>PAM Sequence</u>	<u>SEQ ID NO:</u>
MS26	Chr. 1: 51.81cM	Long	MS26Cas-1	GTACTCCATCCGCCCCATCGA GTA	GGG	13
		Long	MS26Cas-2	GCACGTACGTCACCATCCCGC	CGG	14
		Long	MS26Cas-3	GACGTACGTGCCCTACTCGAT	GGG	15
LIG	Chr. 2: 28.45cM	Long	LIGCas-1	GTACCGTACGTGCCCCGGCGG	AGG	16
		Long	LIGCas-2	GGAATTGTACCGTACGTGCCC	CGG	17
		Long	LIGCas-3	GCGTACGCGTACGTGTG	AGG	18
MS45	Chr. 9: 119.15cM	Long	MS45Cas-1	GCTGGCCGAGGTGCGACTAC	CGG	19
		Long	MS45Cas-2	GGCCGAGGTGCGACTACCGGC	CGG	20
		Long	MS45Cas-3	GGCGCGAGCTCGTGCTTCAC	CGG	21
ALS	Chr. 4:	Long	ALSCas-1	GGTGCCAATCATGCGTCG	CGG	22

	107.73cM and Chr. 5: 115.49cM	Long	ALSCas-2	GGTCGCCATCACGGGAC	AGG	23
		Long	ALSCas-3	GTCGCGGCACCTGTCCCGTGA	TGG	24
	EPSP S	Long	EPSPSCas-1	GGAATGCTGGAAGTCAATG	CGG	25
		Long	EPSPSCas-2	GCAGCTCTTCTTGGGGAATGC	TGG	26
		Long	EPSPSCas-3	GCAGTAACAGCTGCTGTCAA	TGG	27

MS26=Male Sterility Gene 26, LIG=Liguleless 1 Gene Promoter, MS45=Male Sterility Gene 45,
ALS=Acetolactate Synthase Gene, EPSPS=Enolpyruvylshikimate Phosphate Synthase Gene

The maize optimized Cas9 endonuclease and long guide RNA expression cassettes containing the specific maize variable targeting domains were co-

5 delivered to 60-90 Hi-II immature maize embryos by particle-mediated delivery (see Example 10) in the presence of BBM and WUS2 genes (see Example 11). Hi-II maize embryos transformed with either the LIG3-4 or MS26++ homing endonucleases (see Example 9) targeting the same maize genomic loci as the LIGCas or MS26Cas target sites served as a positive control and embryos

10 transformed with only the Cas9 or guide RNA expression cassette served as negative controls. After 7 days, the 20-30 most uniformly transformed embryos from each treatment were pooled and total genomic DNA was extracted. The region surrounding the intended target site was PCR amplified with Phusion® High Fidelity PCR Master Mix (New England Biolabs, M0531L) adding on the sequences

15 necessary for amplicon-specific barcodes and Illumina sequencing using “tailed” primers through two rounds of PCR. The primers used in the primary PCR reaction are shown in Table 2 and the primers used in the secondary PCR reaction were AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACG (forward, SEQ ID NO: 53) and CAAGCAGAAGACGGCATA (reverse, SEQ ID NO: 54).

20

Table 2. PCR primer sequences

<u>Target Site</u>	<u>Primer Orientation</u>	<u>Primary PCR Primer Sequence</u>	<u>SEQ ID NO:</u>
MS26Cas-1	Forward	CTAACTCTTTCCCTACACGACGCTCTTCCGATCTA GGACCGGAAGCTCGCCGCGT	28
MS26Cas-1	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTTC	29

		CTGGAGGACGACGTGCTG	
MS26Cas-2	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTAA GGTCCTGGAGGACGACGTGCTG	30
MS26Cas-2	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTCC GGAAGCTCGCCGCGT	31
MS26Cas-3	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTTC CTCCGGAAGCTCGCCGCGT	32
MS26Cas-3	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTTC CTGGAGGACGACGTGCTG	29
MS26 Meganucleas e	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTTT CCTCCTGGAGGACGACGTGCTG	33
MS26 Meganucleas e	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTCC GGAAGCTCGCCGCGT	31
LIGCas-1	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTA GGACTGTAACGATTTACGCACCTGCTG	34
LIGCas-1	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTGC AAATGAGTAGCAGCGCACGTAT	35
LIGCas-2	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTTC CTCTGTAACGATTTACGCACCTGCTG	36
LIGCas-2	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTGC AAATGAGTAGCAGCGCACGTAT	35
LIGCas-3	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTAA GGCGCAAATGAGTAGCAGCGCAC	37
LIGCas-3	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTCA CCTGCTGGGAATTGTACCGTA	38
LIG3-4 Meganucleas e	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTC CTTCGCAAATGAGTAGCAGCGCAC	39
LIG3-4 Meganucleas e	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTCA CCTGCTGGGAATTGTACCGTA	38
MS45Cas-1	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTA GGAGGACCCGTTCCGGCCTCAGT	40
MS45Cas-1	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTGC CGGCTGGCATTGTCTCTG	41
MS45Cas-2	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTTC CTGGACCCGTTCCGGCCTCAGT	42
MS45Cas-2	Reverse	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTGC CGGCTGGCATTGTCTCTG	41
MS45Cas-3	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTG	43

		AAGGGACCCGTTCCGGCCTCAGT	
MS45Cas-3	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTGC CGGCTGGCATTGTCTCTG	41
ALSCas-1	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTAA GGCGACGATGGGCGTCTCCTG	44
ALSCas-1	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTGC GTCTGCATCGCCACCTC	45
ALSCas-2	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTTT CCCGACGATGGGCGTCTCCTG	46
ALSCas-2	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTGC GTCTGCATCGCCACCTC	45
ALSCas-3	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTG GAACGACGATGGGCGTCTCCTG	47
ALSCas-3	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTGC GTCTGCATCGCCACCTC	45
EPSPSCas-1	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTG GAAGAGGAAACATACGTTGCATTTCCTCA	48
EPSPSCas-1	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTG GTGGAAAGTTCCAGTTGAGGA	49
EPSPSCas-2	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTAA GCGGTGGAAAGTTCCAGTTGAGGA	50
EPSPSCas-2	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTGA GGAAACATACGTTGCATTTCCTCA	51
EPSPSCas-3	Forward	CTACACTCTTTCCCTACACGACGCTCTTCCGATCTC CTTGAGGAAACATACGTTGCATTTCCTCA	52
EPSPSCas-3	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCCGATCTG GTGGAAAGTTCCAGTTGAGGA	49

The resulting PCR amplifications were purified with a Qiagen PCR purification spin column, concentration measured with a Hoechst dye-based fluorometric assay, combined in an equimolar ratio, and single read 100 nucleotide-length deep sequencing was performed on Illumina's MiSeq Personal Sequencer with a 30-40% (v/v) spike of PhiX control v3 (Illumina, FC-110-3001) to off-set sequence bias. Only those reads with a ≥ 1 nucleotide indel arising within the 10 nucleotide window centered over the expected site of cleavage and not found in a similar level in the negative control were classified as NHEJ mutations. NHEJ

mutant reads with the same mutation were counted and collapsed into a single read and the top 10 most prevalent mutations were visually confirmed as arising within the expected site of cleavage. The total numbers of visually confirmed NHEJ mutations were then used to calculate the % mutant reads based on the total
5 number of reads of an appropriate length containing a perfect match to the barcode and forward primer.

The frequency of NHEJ mutations recovered by deep sequencing for the guide RNA/Cas endonuclease system targeting the three LIGCas targets (SEQ ID NOS: 16, 17, 18) compared to the LIG3-4 homing endonuclease targeting the same
10 locus is shown in Table 3. The ten most prevalent types of NHEJ mutations recovered based on the guide RNA/Cas endonuclease system compared to the LIG3-4 homing endonuclease are shown in Figure 3 A (corresponding to SEQ ID NOS: 55-75) and Figure 3 B (corresponding to SEQ ID NOS: 76-96). Approximately, 12-23 fold higher frequencies of NHEJ mutations were observed when using a guide
15 RNA/Cas system to introduce a double strand break at a maize genomic target site (Cas target sites), relative to the LIG3-4 homing endonuclease control. As shown in Table 4, a similar difference between the guide RNA/Cas system and meganuclease double-strand break technologies was observed at the MS26 locus with approximately 14-25 fold higher frequencies of NHEJ mutations when a guide
20 RNA/Cas endonuclease system was used. High frequencies of NHEJ mutations were also recovered at the MS45, ALS and EPSPS Cas targets (see Table 5) when using a guide RNA/Cas endonuclease system. This data indicates that the guide RNA/Cas9 endonuclease system described herein can be effectively used to introduce an alteration at genomic sites of interest such as those related to male
25 fertility, wherein an alteration results in the creation of a male sterile gene locus and male sterile plants. Altering the EPSPS target can result in the production of plants that are tolerant and/or resistant against glyphosate based herbicides. Altering the acetolactate synthase (ALS) gene target site can result in the production of plants that are tolerant and/or resistant to imidazolinone and sulphonylurea herbicides.

Table 3. Percent (%) mutant reads at maize Liguleless 1 target locus produced by a guide RNA/Cas system versus a homing endonuclease system.

<u>System</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
Cas9 Only Control	640,063	1	0.00%
guide RNA Only Control	646,774	1	0.00%
LIG3-4 Homing Endonuclease	616,536	1,211	0.20%
LIGCas-1 guide /Cas9	716,854	33,050	4.61%
LIGCas-2 guide /Cas9	711,047	16,675	2.35%
LIGCas-3 guide /Cas9	713,183	27,959	3.92%

5 Table 4. Percent (%) mutant reads at maize Male Sterility 26 target locus produced by a guide RNA/Cas system versus a homing endonuclease.

<u>System</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
Cas9 Only Control	403,123	15	0.00%
MS26++ Homing Endonuclease	512,784	642	0.13%
MS26Cas-1 guide /Cas9	575,671	10,073	1.75%
MS26Cas-2 guide /Cas9	543,856	16,930	3.11%
MS26Cas-3 guide /Cas9	538,141	13,879	2.58%

Table 5. Percent (%) mutant reads at maize Male Sterility 45, Acetolactate Synthase and Enolpyruvylshikimate Phosphate Synthase target loci produced by the guide RNA/Cas system.

<u>System</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
Cas9 Only Control (MS45)	899,500	27	0.00%
MS45Cas- 1 guide /Cas9	812,644	3,795	0.47%
MS45Cas- 2 guide /Cas9	785,183	14,704	1.87%
MS45Cas- 3 guide /Cas9	728,023	9,203	1.26%
Cas9 Only	534,764	19	0.00%

Control (ALS)			
ALSCas- 1 guide /Cas9	434,452	9,669	2.23%
ALSCas- 2 guide /Cas9	472,351	6,352	1.345%
ALSCas- 3 guide /Cas9	497,786	8,535	1.715%
Cas9 Only Control (EPSPS)	1,347,086	6	0.00%
EPSPSCas- 1 guide /Cas9	1,420,274	13,051	0.92%
EPSPSCas- 2 guide /Cas9	1,225,082	26,340	2.15%
EPSPSCas- 3 guide /Cas9	1,406,905	53,603	3.81%

Taken together, our data indicate that the maize optimized guide RNA/Cas endonuclease system described herein using a long guide RNA expression cassette efficiently cleaves maize chromosomal DNA and generates imperfect NHEJ mutations at frequencies greater than the engineered LIG3-4 and MS26++ homing endonucleases.

EXAMPLE 3

Long guide RNA of the maize optimized guide RNA/Cas endonuclease system cleaves maize chromosomal DNA more efficiently than the short guide RNA

To determine the most effective guide RNA (comprising a fusion of the crRNA and tracrRNA) for use in maize, the recovery of NHEJ mutations using a short guide RNA (SEQ ID NO: 11) based on Jinek *et al.* (2012) *Science* 337:816-21 and a long guide RNA (SEQ ID NO: 8) based on Mali *et al.* (2013) *Science* 339:823-26 was examined.

The variable targeting domains of the guide RNA targeting the maize genomic target sites at the LIG locus (LIGCas-1, LIGCas-2 and LIGCas-3, SEQ ID NOs: 16, 17 and 18, Table1) were introduced into both the maize optimized long and short guide RNA expression cassettes as described in Example 1 and co-transformed along with the maize optimized Cas9 endonuclease expression cassette into immature maize embryos and deep sequenced for NHEJ mutations as described in Example 2. Embryos transformed with only the Cas9 endonuclease expression cassette served as a negative control.

As shown in Table 6 below, the frequency of NHEJ mutations recovered with the long guide RNA far exceeded those obtained with the short guide RNA. This data indicates that the long guide RNA paired with the maize optimized Cas9 endonuclease gene described herein more efficiently cleaves maize chromosomal DNA.

Table 6. Percent (%) mutant reads at the maize Liguleless 1 target locus produced by a guide RNA/Cas system with a long versus a short guide RNA.

<u>System</u>	<u>guide RNA Used</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
Cas9 Only	N/A	640,063	1	0.00%
LIGCas-1 guide/Cas9	Short	676,870	43	0.01%
LIGCas-2 guide/Cas9	Short	747,945	91	0.01%
LIGCas-3 guide/Cas9	Short	655,157	10	0.00%
LIGCas-1 guide/ Cas9	Long	716,854	33,050	4.61%
LIGCas-2 guide/cas9	Long	711,047	16,675	2.35%
LIGCas-3 guide/Cas9	Long	713,183	27,959	3.92%

10

EXAMPLE 4

The guide RNA/Cas endonuclease system may be multiplexed to simultaneously target multiple chromosomal loci in maize for mutagenesis by imperfect non-homologous end-joining

To test if multiple chromosomal loci may be simultaneously mutagenized with the guide RNA/maize optimized Cas endonuclease system described herein, the long guide RNA expression cassettes targeting the MS26Cas-2 target site (SEQ ID NO: 14), the LIGCas-3 target site (SEQ ID NO: 18) and the MS45Cas-2 target site (SEQ ID NO: 20), were co-transformed into maize embryos either in duplex or in triplex along with the Cas9 endonuclease expression cassette and examined by deep sequencing for the presence of imprecise NHEJ mutations as described in Example 2.

Hi-II maize embryos co-transformed with the Cas9 expression cassette and the corresponding guide RNA expression cassette singly served as a positive

control and embryos transformed with only the Cas9 expression cassette served as a negative control.

- As shown in Table 7 below, mutations resulting from imprecise NHEJ were recovered at all relevant loci when multiple guide RNA expression cassettes were simultaneously introduced either in duplex or triplex with frequencies of mutant reads near those of the positive control. Thus, demonstrating that the maize optimized guide RNA/Cas endonuclease system described herein may be used to simultaneously introduce imprecise NHEJ mutations at multiple loci in maize.

10 Table 7. Percent (%) mutant reads at maize target loci produced by a multiplexed guide RNA/Cas system.

<u>Target Site Examined for NHEJ Mutations</u>	<u>guide RNAs Co-transformed Individually, in Duplex, or in Triplex with Cas9</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
LIGCas-3, MS26Cas-2, MS45Cas-2	None (Cas9 Only control)	527,691	9	0.00%
LIGCas-3	LIGCas-3	645,107	12,631	1.96%
	LIGCas-3 MS26Cas-2	579,992	10,348	1.78%
	LIGCas-3 MS26Cas-2 MS45Cas-2	648,901	12,094	1.86%
MS26Cas-2	MS26 Cas 2	699,154	17,247	2.47%
	LIGCas-3 MS26Cas-2	717,158	10,256	1.43%
	MS26Cas-2 MS45Cas-2	613,431	9,931	1.62%
	LIGCas-3 MS26Cas-2 MS45Cas-2	471,890	7,311	1.55%
MS45Cas-2	MS45Cas-2	503,423	10,034	1.99%
	MS26Cas-2 MS45Cas-2	480,178	8,008	1.67%
	LIGCas-3 MS26Cas-2 MS45Cas-2	416,711	7,190	1.73%

EXAMPLE 5

Guide RNA/Cas endonuclease mediated DNA cleavage in maize chromosomal loci can stimulate homologous recombination repair-mediated transgene insertion

To test the utility of the maize optimized guide RNA/Cas system described
5 herein to cleave maize chromosomal loci and stimulate homologous recombination
(HR) repair pathways to site-specifically insert a transgene, a HR repair DNA vector
(also referred to as a donor DNA) (SEQ ID NO: 97) was constructed as illustrated in
Figure 4 using standard molecular biology techniques and co-transformed with a
10 long guide RNA expression cassette, comprising a variable targeting domain
corresponding to the LIGCas-3 genomic target site, and a Cas9 endonuclease
expression cassette into immature maize embryos as described in Example 2.

Maize embryos co-transformed with the HR repair DNA vector and LIG3-4
homing endonuclease (see Example 9) targeting the same genomic target site as
LIGCas-3 served as a positive control. Since successful delivery of the HR repair
15 DNA vector confers bialaphos herbicide resistance, callus events containing
putative HR-mediated transgenic insertions were selected by placing the callus on
herbicide containing media. After selection, stable callus events were sampled, total
genomic DNA extracted, and using the primer pairs shown in Figure 5
(corresponding to SEQ ID NOs: 98-101), PCR amplification was carried out at both
20 possible transgene genomic DNA junctions to identify putative HR-mediated
transgenic insertions. The resulting amplifications were sequenced for confirmation.

Sequence confirmed PCR amplifications indicating site-specific transgene
insertion for the guide RNA/Cas system were detected for 37 out of 384 stable
transformants with 15 containing amplifications across both transgene genomic
25 DNA junctions indicating near perfect site-specific transgene insertion. The LIG3-4
homing endonuclease positive control yielded PCR amplifications indicating site-
specific transgene insertion for 3 out of 192 stable transformants with 1 containing
amplifications across both transgene genomic DNA junctions. The data clearly
demonstrates that maize chromosomal loci cleaved with the maize optimized guide
30 RNA/Cas system described herein can be used to stimulate HR repair pathways to
site-specifically insert transgenes at frequencies greater than the LIG3-4 homing
endonuclease.

EXAMPLE 6

Guide RNA/Cas endonuclease system transformed together on a single vector results in greater recovery of imperfect non-homologous end-joining mutations

5 To evaluate different delivery methods for the maize optimized guide RNA/Cas endonuclease system described herein, the recovery of NHEJ mutations when the guide RNA/Cas expression cassettes were either co-transformed as separate DNA vectors as in Examples 2, 3, 4 and 5 or transformed as a single vector DNA (comprising both guide RNA and Cas endonuclease expression
10 cassettes, as shown in Figure 1C) was examined.

The long guide RNA expression cassette for LIGCas-3 and the Cas9 expression cassette were consolidated onto a single vector DNA (Figure 1 C, SEQ ID NO: 102) by standard molecular biology techniques and transformed into immature Hi-II maize embryos as described in Examples 10 and 11 by particle-mediated delivery. Hi-II embryos co-transformed with the Cas9 and LIGCas-3 long
15 guide RNA expression cassettes served as a positive control while embryos transformed with only the Cas9 expression cassette served as a negative control. Deep sequencing for NHEJ mutations was performed as described in Example 2.

As shown in Table 8 below, the frequency of NHEJ mutations recovered
20 when the Cas endonuclease and long guide RNA expression cassettes were delivered together as a single vector DNA was approximately 2-fold greater than that observed from the equivalent co-transformation experiment. This indicates that delivery of the guide RNA/Cas system expression cassettes together on a single vector DNA results in a greater recovery of imperfect non-homologous end-joining
25 mutations.

Table 8. Percent (%) mutant reads at the maize Liguleless 1 target locus produced by a guide RNA/Cas system with Cas9 and guide RNA expression cassettes combined into one DNA vector versus two separate DNA vectors.

<u>System</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
Cas9 Only Control	1,519,162	97	0.01%
LIGCas-3 guide/Cas9 (Two	1,515,0607	36,346	2.40%

vector DNAs)			
LIGCas-3 guide/Cas9 (Single vector DNA)	1,860,031	105,854	5.69%

EXAMPLE 7

Delivery methods for plant genome editing using the guide RNA/Cas endonuclease system

5 This example describes methods to deliver or maintain and express the Cas9
endonuclease and guide RNA (or individual crRNA and tracrRNAs) into, or within
plants, respectively, to enable directed DNA modification or gene insertion via
homologous recombination. More specifically this example describes a variety of
methods which include, but are not limited to, delivery of the Cas9 endonuclease as
10 a DNA, RNA (5'-capped and polyadenylated) or protein molecule. In addition, the
guide RNA may be delivered as a DNA or RNA molecule.

Shown in Example 2, a high mutation frequency was observed when Cas9
endonuclease and guide RNA were delivered as DNA vectors by biolistic
transformation of immature corn embryos. Other embodiments of this disclosure
15 can be to deliver the Cas9 endonuclease as a DNA, RNA or protein and the guide
RNA as a DNA or RNA molecule or as a duplex crRNA/tracrRNA molecule as RNA
or DNA or a combination. Various combinations of Cas9 endonuclease, guide RNA
and crRNA/tracrRNA delivery methods can be, but are not limited to, the methods
shown in Table 9.

20

Table 9: Various combinations of delivery of the cas9 endonuclease, guide RNA or
crRNA+tracrRNA.

combination	Components delivered. (Delivery method is shown between brackets)
1	Cas9 (DNA vector), guide RNA (DNA vector)
2	Cas9 (DNA vector), guide RNA (RNA)
3	Cas9 (RNA), guide RNA (DNA)
4	Cas9 (RNA), guide RNA (RNA)
5	Cas9 (Protein), guide RNA (DNA)
6	Cas9 (Protein), guide RNA (RNA)
7	Cas9 (DNA vector), crRNA (DNA), tracrRNA (DNA)

8	Cas9 (DNA vector), crRNA (RNA), tracrRNA (DNA)
9	Cas9 (DNA vector), crRNA (RNA), tracrRNA (RNA)
10	Cas9 (DNA vector) crRNA (DNA), tracrRNA (RNA)
11	Cas9 (RNA), crRNA (DNA), tracrRNA (DNA)
12	Cas9 (RNA), crRNA (RNA), tracrRNA (DNA)
13	Cas9 (RNA), crRNA (RNA), tracrRNA (RNA)
14	Cas9 (RNA), crRNA (DNA), tracrRNA (RNA)
15	Cas9 (Protein), crRNA (DNA), tracrRNA (DNA)
16	Cas9 (Protein), crRNA (RNA), tracrRNA (DNA)
17	Cas9 (Protein), crRNA (RNA), tracrRNA 18(RNA)
18	Cas9 (Protein), crRNA (DNA), tracrRNA (RNA)

Delivery of the Cas9 (as DNA vector) and guide RNA (as DNA vector) example (Table 9, combination1) can also be accomplished by co-delivering these DNA cassettes on a single or multiple *Agrobacterium* vectors and transforming plant tissues by *Agrobacterium* mediated transformation. In addition, a vector containing a constitutive, tissue-specific or conditionally regulated Cas9 gene can be first delivered to plant cells to allow for stable integration into the plant genome to establish a plant line that contains only the Cas9 gene in the plant genome. In this example, single or multiple guide RNAs, or single or multiple crRNA and a tracrRNA can be delivered as either DNA or RNA, or combination, to the plant line containing the genome-integrated version of the Cas9 gene for the purpose of generating mutations or promoting homologous recombination when HR repair DNA vectors for targeted integration are co-delivered with the guide RNAs. As extension of this example, plant line containing the genome-integrated version of the Cas9 gene and a tracrRNA as a DNA molecule can also be established. In this example single or multiple crRNA molecules can be delivered as RNA or DNA to promote the generation of mutations or to promote homologous recombination when HR repair DNA vectors for targeted integration are co-delivered with crRNA molecule(s) enabling the targeted mutagenesis or homologous recombination at single or multiple sites in the plant genome.

EXAMPLE 8**Components of the guide RNA/Cas endonuclease system delivered directly as RNA in plants**

This example illustrates the use of the methods as described in Table 9 configuration of Example 7 [Cas9 (DNA vector), guide RNA (RNA)] for modification or mutagenesis of chromosomal loci in plants. The maize optimized Cas9 endonuclease expression cassette described in Example 1 was co-delivered by particle gun as described in Example 2 along with single stranded RNA molecules (synthesized by Integrated DNA Technologies, Inc.) constituting a short guide RNA targeting the maize locus and sequence shown in Table 10. Embryos transformed with only the Cas9 expression cassette or short guide RNA molecules served as negative controls. Seven days post-bombardment, the immature embryos were harvested and analyzed by deep sequencing for NHEJ mutations as described in Example 2. Mutations not present in the negative controls were found at the site (Figure 6, corresponding to SEQ ID NOs: 104-110). These mutations were similar to those found in Examples 2, 3, 4 and 6. This data indicates that component(s) of the maize optimized guide RNA/Cas endonuclease system described herein may be delivered directly as RNA.

Table 10. Maize genomic target site and location for short guide RNA delivered as RNA.

<u>Locu</u> <u>s</u>	<u>Locatio</u> <u>n</u>	<u>Guid</u> <u>e</u> <u>RNA</u> <u>Used</u>	<u>Designation</u>	<u>Maize Target Site</u>	<u>PAM</u> <u>Sequenc</u> <u>e</u>	<u>SE</u> <u>Q ID</u> <u>NO</u>
55	Chr. 1: 51.78 cM	Short	55CasRNA -1	TGGGCAGGTCTCACGACGG T	TGG	103

EXAMPLE 9**Creation of rare cutting engineered meganucleases****LIG3-4 meganuclease and LIG3-4 intended recognition sequence**

An endogenous maize genomic target site comprising the LIG3-4 intended recognition sequence (SEQ ID NO: 111) was selected for design of a rare-cutting double-strand break inducing agent (SEQ ID NO: 112) as described in US patent

publication 2009-0133152 A1 (published May 21, 2009). The LIG3-4 intended recognition sequence is a 22 bp polynucleotide having the following sequence: ATATACCTCACACGTACGCGTA (SEQ ID NO: 111).

MS 26 ++ meganuclease

- 5 An endogenous maize genomic target site designated "TS-MS26" (SEQ ID NO: 113) was selected for design of a custom double-strand break inducing agent MS26++ as described in U.S. patent application 13/526912 filed June 19, 2012). The TS-MS26 target site is a 22 bp polynucleotide positioned 62 bps from the 5' end of the fifth exon of the maize *MS26* gene and having the following
- 10 sequence: gatggtgacgtac[^]gtgccctac (SEQ ID NO: 113). The double strand break site and overhang region is underlined, the enzyme cuts after C13, as indicated by the [^]. Plant optimized nucleotide sequences for an engineered endonuclease (SEQ ID NO: 114) encoding an engineered MS26++ endonuclease were designed to bind and make double-strand breaks at the selected TS-MS26 target site.

15 EXAMPLE 10

Transformation of Maize Immature Embryos

Transformation can be accomplished by various methods known to be effective in plants, including particle-mediated delivery, *Agrobacterium*-mediated transformation, PEG-mediated delivery, and electroporation.

20 a. Particle-mediated delivery

Transformation of maize immature embryos using particle delivery is performed as follows. Media recipes follow below.

- The ears are husked and surface sterilized in 30% Clorox bleach plus 0.5% Micro detergent for 20 minutes, and rinsed two times with sterile water. The
- 25 immature embryos are isolated and placed embryo axis side down (scutellum side up), 25 embryos per plate, on 560Y medium for 4 hours and then aligned within the 2.5-cm target zone in preparation for bombardment. Alternatively, isolated embryos are placed on 560L (Initiation medium) and placed in the dark at temperatures ranging from 26°C to 37°C for 8 to 24 hours prior to placing on 560Y for 4 hours at
- 30 26°C prior to bombardment as described above.

Plasmids containing the double strand brake inducing agent and donor DNA are constructed using standard molecular biology techniques and co-bombarded

with plasmids containing the developmental genes ODP2 (AP2 domain transcription factor ODP2 (Ovule development protein 2); US20090328252 A1) and Wushel (US2011/0167516).

The plasmids and DNA of interest are precipitated onto 0.6 μm (average
5 diameter) gold pellets using a water-soluble cationic lipid transfection reagent as follows. DNA solution is prepared on ice using 1 μg of plasmid DNA and optionally other constructs for co-bombardment such as 50 ng (0.5 μl) of each plasmid containing the developmental genes ODP2 (AP2 domain transcription factor ODP2 (Ovule development protein 2); US20090328252 A1) and Wushel. To the pre-mixed
10 DNA, 20 μl of prepared gold particles (15 mg/ml) and 5 μl of the a water-soluble cationic lipid transfection reagent is added in water and mixed carefully. Gold particles are pelleted in a microfuge at 10,000 rpm for 1 min and supernatant is removed. The resulting pellet is carefully rinsed with 100 ml of 100% EtOH without resuspending the pellet and the EtOH rinse is carefully removed. 105 μl of 100%
15 EtOH is added and the particles are resuspended by brief sonication. Then, 10 μl is spotted onto the center of each macrocarrier and allowed to dry about 2 minutes before bombardment.

Alternatively, the plasmids and DNA of interest are precipitated onto 1.1 μm (average diameter) tungsten pellets using a calcium chloride (CaCl_2) precipitation
20 procedure by mixing 100 μl prepared tungsten particles in water, 10 μl (1 μg) DNA in Tris EDTA buffer (1 μg total DNA), 100 μl 2.5 M CaCl_2 , and 10 μl 0.1 M spermidine. Each reagent is added sequentially to the tungsten particle suspension, with mixing. The final mixture is sonicated briefly and allowed to incubate under constant vortexing for 10 minutes. After the precipitation period, the tubes are centrifuged
25 briefly, liquid is removed, and the particles are washed with 500 ml 100% ethanol, followed by a 30 second centrifugation. Again, the liquid is removed, and 105 μl 100% ethanol is added to the final tungsten particle pellet. For particle gun bombardment, the tungsten/DNA particles are briefly sonicated. 10 μl of the tungsten/DNA particles is spotted onto the center of each macrocarrier, after which
30 the spotted particles are allowed to dry about 2 minutes before bombardment.

The sample plates are bombarded at level #4 with a Biorad Helium Gun. All samples receive a single shot at 450 PSI, with a total of ten aliquots taken from each tube of prepared particles/DNA.

Following bombardment, the embryos are incubated on 560P (maintenance
5 medium) for 12 to 48 hours at temperatures ranging from 26C to 37C, and then placed at 26C. After 5 to 7 days the embryos are transferred to 560R selection medium containing 3 mg/liter Bialaphos, and subcultured every 2 weeks at 26C. After approximately 10 weeks of selection, selection-resistant callus clones are transferred to 288J medium to initiate plant regeneration. Following somatic embryo
10 maturation (2-4 weeks), well-developed somatic embryos are transferred to medium for germination and transferred to a lighted culture room. Approximately 7-10 days later, developing plantlets are transferred to 272V hormone-free medium in tubes for 7-10 days until plantlets are well established. Plants are then transferred to inserts in flats (equivalent to a 2.5" pot) containing potting soil and grown for 1 week in a
15 growth chamber, subsequently grown an additional 1-2 weeks in the greenhouse, then transferred to Classic 600 pots (1.6 gallon) and grown to maturity. Plants are monitored and scored for transformation efficiency, and/or modification of regenerative capabilities.

Initiation medium (560L) comprises 4.0 g/l N6 basal salts (SIGMA C-1416),
20 1.0 ml/l Eriksson's Vitamin Mix (1000X SIGMA-1511), 0.5 mg/l thiamine HCl, 20.0 g/l sucrose, 1.0 mg/l 2,4-D, and 2.88 g/l L-proline (brought to volume with D-I H₂O following adjustment to pH 5.8 with KOH); 2.0 g/l Gelrite (added after bringing to volume with D-I H₂O); and 8.5 mg/l silver nitrate (added after sterilizing the medium and cooling to room temperature).

25 Maintenance medium (560P) comprises 4.0 g/l N6 basal salts (SIGMA C-1416), 1.0 ml/l Eriksson's Vitamin Mix (1000X SIGMA-1511), 0.5 mg/l thiamine HCl, 30.0 g/l sucrose, 2.0 mg/l 2,4-D, and 0.69 g/l L-proline (brought to volume with D-I H₂O following adjustment to pH 5.8 with KOH); 3.0 g/l Gelrite (added after bringing to volume with D-I H₂O); and 0.85 mg/l silver nitrate (added after sterilizing the
30 medium and cooling to room temperature).

Bombardment medium (560Y) comprises 4.0 g/l N6 basal salts (SIGMA C-1416), 1.0 ml/l Eriksson's Vitamin Mix (1000X SIGMA-1511), 0.5 mg/l thiamine HCl,

120.0 g/l sucrose, 1.0 mg/l 2,4-D, and 2.88 g/l L-proline (brought to volume with D-I H₂O following adjustment to pH 5.8 with KOH); 2.0 g/l Gelrite (added after bringing to volume with D-I H₂O); and 8.5 mg/l silver nitrate (added after sterilizing the medium and cooling to room temperature).

- 5 Selection medium (560R) comprises 4.0 g/l N6 basal salts (SIGMA C-1416), 1.0 ml/l Eriksson's Vitamin Mix (1000X SIGMA-1511), 0.5 mg/l thiamine HCl, 30.0 g/l sucrose, and 2.0 mg/l 2,4-D (brought to volume with D-I H₂O following adjustment to pH 5.8 with KOH); 3.0 g/l Gelrite (added after bringing to volume with D-I H₂O); and 0.85 mg/l silver nitrate and 3.0 mg/l bialaphos (both added after sterilizing the
10 medium and cooling to room temperature).

- Plant regeneration medium (288J) comprises 4.3 g/l MS salts (GIBCO 11117-074), 5.0 ml/l MS vitamins stock solution (0.100 g nicotinic acid, 0.02 g/l thiamine HCL, 0.10 g/l pyridoxine HCL, and 0.40 g/l glycine brought to volume with polished D-I H₂O) (Murashige and Skoog (1962) *Physiol. Plant.* 15:473), 100 mg/l myo-
15 inositol, 0.5 mg/l zeatin, 60 g/l sucrose, and 1.0 ml/l of 0.1 mM abscisic acid (brought to volume with polished D-I H₂O after adjusting to pH 5.6); 3.0 g/l Gelrite (added after bringing to volume with D-I H₂O); and 1.0 mg/l indoleacetic acid and 3.0 mg/l bialaphos (added after sterilizing the medium and cooling to 60°C).

- Hormone-free medium (272V) comprises 4.3 g/l MS salts (GIBCO 11117-074), 5.0
20 ml/l MS vitamins stock solution (0.100 g/l nicotinic acid, 0.02 g/l thiamine HCL, 0.10 g/l pyridoxine HCL, and 0.40 g/l glycine brought to volume with polished D-I H₂O), 0.1 g/l myo-inositol, and 40.0 g/l sucrose (brought to volume with polished D-I H₂O after adjusting pH to 5.6); and 6 g/l bacto-agar (added after bringing to volume with polished D-I H₂O), sterilized and cooled to 60°C.

- 25 b. *Agrobacterium*-mediated transformation

- Agrobacterium*-mediated transformation was performed essentially as described in Djukanovic *et al.* (2006) *Plant Biotech J* 4:345-57. Briefly, 10-12 day old immature embryos (0.8 -2.5 mm in size) were dissected from sterilized kernels and placed into liquid medium (4.0 g/L N6 Basal Salts (Sigma C-1416), 1.0 ml/L
30 Eriksson's Vitamin Mix (Sigma E-1511), 1.0 mg/L thiamine HCl, 1.5 mg/L 2, 4-D, 0.690 g/L L-proline, 68.5 g/L sucrose, 36.0 g/L glucose, pH 5.2). After embryo collection, the medium was replaced with 1 ml *Agrobacterium* at a concentration of

0.35-0.45 OD₅₅₀. Maize embryos were incubated with *Agrobacterium* for 5 min at room temperature, then the mixture was poured onto a media plate containing 4.0 g/L N6 Basal Salts (Sigma C-1416), 1.0 ml/L Eriksson's Vitamin Mix (Sigma E-1511), 1.0 mg/L thiamine HCl, 1.5 mg/L 2, 4-D, 0.690 g/L L-proline, 30.0 g/L sucrose, 0.85 mg/L silver nitrate, 0.1 nM acetosyringone, and 3.0 g/L Gelrite, pH 5.8. Embryos were incubated axis down, in the dark for 3 days at 20°C, then incubated 4 days in the dark at 28°C, then transferred onto new media plates containing 4.0 g/L N6 Basal Salts (Sigma C-1416), 1.0 ml/L Eriksson's Vitamin Mix (Sigma E-1511), 1.0 mg/L thiamine HCl, 1.5 mg/L 2, 4-D, 0.69 g/L L-proline, 30.0 g/L sucrose, 0.5 g/L MES buffer, 0.85 mg/L silver nitrate, 3.0 mg/L Bialaphos, 100 mg/L carbenicillin, and 6.0 g/L agar, pH 5.8. Embryos were subcultured every three weeks until transgenic events were identified. Somatic embryogenesis was induced by transferring a small amount of tissue onto regeneration medium (4.3 g/L MS salts (Gibco 11117), 5.0 ml/L MS Vitamins Stock Solution, 100 mg/L myo-inositol, 0.1 µM ABA, 1 mg/L IAA, 0.5 mg/L zeatin, 60.0 g/L sucrose, 1.5 mg/L Bialaphos, 100 mg/L carbenicillin, 3.0 g/L Gelrite, pH 5.6) and incubation in the dark for two weeks at 28°C. All material with visible shoots and roots were transferred onto media containing 4.3 g/L MS salts (Gibco 11117), 5.0 ml/L MS Vitamins Stock Solution, 100 mg/L myo-inositol, 40.0 g/L sucrose, 1.5 g/L Gelrite, pH 5.6, and incubated under artificial light at 28°C. One week later, plantlets were moved into glass tubes containing the same medium and grown until they were sampled and/or transplanted into soil.

EXAMPLE 11

Transient Expression of BBM Enhances Transformation

Parameters of the transformation protocol can be modified to ensure that the BBM activity is transient. One such method involves precipitating the BBM-containing plasmid in a manner that allows for transcription and expression, but precludes subsequent release of the DNA, for example, by using the chemical PEI. In one example, the BBM plasmid is precipitated onto gold particles with PEI, while the transgenic expression cassette (UBI::moPAT~GFPm::PinII; moPAT is the maize optimized PAT gene) to be integrated is precipitated onto gold particles using the standard calcium chloride method.

Briefly, gold particles were coated with PEI as follows. First, the gold particles were washed. Thirty-five mg of gold particles, 1.0 in average diameter (A.S.I. #162-0010), were weighed out in a microcentrifuge tube, and 1.2 ml absolute EtOH was added and vortexed for one minute. The tube was incubated for 15 minutes at room temperature and then centrifuged at high speed using a microfuge for 15 minutes at 4°C. The supernatant was discarded and a fresh 1.2 ml aliquot of ethanol (EtOH) was added, vortexed for one minute, centrifuged for one minute, and the supernatant again discarded (this is repeated twice). A fresh 1.2 ml aliquot of EtOH was added, and this suspension (gold particles in EtOH) was stored at -20°C for weeks. To coat particles with polyethylimine (PEI; Sigma #P3143), 250 µl of the washed gold particle/EtOH mix was centrifuged and the EtOH discarded. The particles were washed once in 100 µl ddH₂O to remove residual ethanol, 250 µl of 0.25 mM PEI was added, followed by a pulse-sonication to suspend the particles and then the tube was plunged into a dry ice/EtOH bath to flash-freeze the suspension, which was then lyophilized overnight. At this point, dry, coated particles could be stored at -80°C for at least 3 weeks. Before use, the particles were rinsed 3 times with 250 µl aliquots of 2.5 mM HEPES buffer, pH 7.1, with 1x pulse-sonication, and then a quick vortex before each centrifugation. The particles were then suspended in a final volume of 250 µl HEPES buffer. A 25 µl aliquot of the particles was added to fresh tubes before attaching DNA. To attach uncoated DNA, the particles were pulse-sonicated, then 1 µg of DNA (in 5 µl water) was added, followed by mixing by pipetting up and down a few times with a Pipetteman and incubated for 10 minutes. The particles were spun briefly (i.e. 10 seconds), the supernatant removed, and 60 µl EtOH added. The particles with PEI-precipitated DNA-1 were washed twice in 60 µl of EtOH. The particles were centrifuged, the supernatant discarded, and the particles were resuspended in 45 µl water. To attach the second DNA (DNA-2), precipitation using a water-soluble cationic lipid transfection reagent was used. The 45 µl of particles/DNA-1 suspension was briefly sonicated, and then 5 µl of 100 ng/µl of DNA-2 and 2.5 µl of a water-soluble cationic lipid transfection reagent were added. The solution was placed on a rotary shaker for 10 minutes, centrifuged at 10,000g for 1 minute. The supernatant was removed, and the particles resuspended in 60 µl of EtOH. The solution was spotted onto

macrocarriers and the gold particles onto which DNA-1 and DNA-2 had been sequentially attached were delivered into scutellar cells of 10 DAP Hi-II immature embryos using a standard protocol for the PDS-1000. For this experiment, the DNA-1 plasmid contained a UBI::RFP::pinII expression cassette, and DNA-2

5 contained a UBI::CFP::pinII expression cassette. Two days after bombardment, transient expression of both the CFP and RFP fluorescent markers was observed as numerous red & blue cells on the surface of the immature embryo. The embryos were then placed on non-selective culture medium and allowed to grow for 3 weeks before scoring for stable colonies. After this 3-week period, 10 multicellular, stably-

10 expressing blue colonies were observed, in comparison to only one red colony. This demonstrated that PEI-precipitation could be used to effectively introduce DNA for transient expression while dramatically reducing integration of the PEI-introduced DNA and thus reducing the recovery of RFP-expressing transgenic events. In this manner, PEI-precipitation can be used to deliver transient expression of BBM and/or

15 WUS2.

For example, the particles are first coated with UBI::BBM::pinII using PEI, then coated with UBI::moPAT~YFP using a water-soluble cationic lipid transfection reagent, and then bombarded into scutellar cells on the surface of immature embryos. PEI-mediated precipitation results in a high frequency of transiently

20 expressing cells on the surface of the immature embryo and extremely low frequencies of recovery of stable transformants. Thus, it is expected that the PEI-precipitated BBM cassette expresses transiently and stimulates a burst of embryogenic growth on the bombarded surface of the tissue (i.e. the scutellar surface), but this plasmid will not integrate. The PAT~GFP plasmid released from

25 the Ca⁺⁺/gold particles is expected to integrate and express the selectable marker at a frequency that results in substantially improved recovery of transgenic events. As a control treatment, PEI-precipitated particles containing a UBI::GUS::pinII (instead of BBM) are mixed with the PAT~GFP/Ca⁺⁺ particles. Immature embryos from both treatments are moved onto culture medium containing 3mg/l bialaphos.

30 After 6-8 weeks, it is expected that GFP⁺, bialaphos-resistant calli will be observed in the PEI/BBM treatment at a much higher frequency relative to the control treatment (PEI/GUS).

As an alternative method, the BBM plasmid is precipitated onto gold particles with PEI, and then introduced into scutellar cells on the surface of immature embryos, and subsequent transient expression of the BBM gene elicits a rapid proliferation of embryogenic growth. During this period of induced growth, the explants are treated with *Agrobacterium* using standard methods for maize (see Example 1), with T-DNA delivery into the cell introducing a transgenic expression cassette such as UBI::moPAT~GFPm::pinII. After co-cultivation, explants are allowed to recover on normal culture medium, and then are moved onto culture medium containing 3 mg/l bialaphos. After 6-8 weeks, it is expected that GFP+, bialaphos-resistant calli will be observed in the PEI/BBM treatment at a much higher frequency relative to the control treatment (PEI/GUS).

It may be desirable to “kick start” callus growth by transiently expressing the BBM and/or WUS2 polynucleotide products. This can be done by delivering BBM and WUS2 5'-capped polyadenylated RNA, expression cassettes containing BBM and WUS2 DNA, or BBM and/or WUS2 proteins. All of these molecules can be delivered using a biolistics particle gun. For example 5'-capped polyadenylated BBM and/or WUS2 RNA can easily be made in vitro using Ambion's mMessage mMachine kit. RNA is co-delivered along with DNA containing a polynucleotide of interest and a marker used for selection/screening such as Ubi::moPAT~GFPm::PinII. It is expected that the cells receiving the RNA will immediately begin dividing more rapidly and a large portion of these will have integrated the agronomic gene. These events can further be validated as being transgenic clonal colonies because they will also express the PAT~GFP fusion protein (and thus will display green fluorescence under appropriate illumination). Plants regenerated from these embryos can then be screened for the presence of the polynucleotide of interest.

EXAMPLE 12

DNA constructs to test the guide RNA/Cas endonuclease system for soybean genome modifications

To test if a guide RNA/Cas endonuclease system, similar to that described in Example 1 for maize, is functional in a dicot such as soybean, a Cas9 (SO) gene (SEQ ID NO:115) soybean codon optimized from *Streptococcus pyogenes* M1 GAS

(SF370) was expressed with a strong soybean constitutive promoter GM-EF1A2 (US patent application 20090133159 (SEQ ID NO: 116). A simian vacuolating virus 40 (SV40) large T-antigen nuclear localization signal (SEQ ID NO:117), representing the amino acid molecules of PKKKRKV (with a linker SRAD (SRADPKKKRKV), was added to the carboxyl terminus of the codon optimized Cas9 to facilitate transporting the codon optimized Cas9 protein (SEQ ID NO:118) to the nucleus. The codon optimized Cas9 gene was synthesized as two pieces by GenScript USA Inc. (Piscataway, NJ) and cloned in frame downstream of the GM-EF1A2 promoter to make DNA construct QC782 shown in Figure 7 (SEQ ID NO:119).

Plant U6 RNA polymerase III promoters have been cloned and characterized from such as *Arabidopsis* and *Medicago truncatula* (Waibel and Filipowicz, NAR 18:3451-3458 (1990); Li et al., J. Integrat. Plant Biol. 49:222-229 (2007); Kim and Nam, Plant Mol. Biol. Rep. 31:581-593 (2013); Wang et al., RNA 14:903-913 (2008)). Soybean U6 small nuclear RNA (snRNA) genes were identified herein by searching public soybean variety Williams82 genomic sequence using *Arabidopsis* U6 gene coding sequence. Approximately 0.5 kb genomic DNA sequence upstream of the first G nucleotide of a U6 gene was selected to be used as a RNA polymerase III promoter for example, GM-U6-13.1 promoter (SEQ ID NO:120), to express guide RNA to direct Cas9 nuclease to designated genomic site. The guide RNA coding sequence was 76 bp long (Figure 8B) and comprised a 20 bp variable targeting domain from a chosen soybean genomic target site on the 5' end and a tract of 4 or more T residues as a transcription terminator on the 3' end. (SEQ ID NO:121, Figure 8 B). The first nucleotide of the 20 bp variable targeting domain was a G residue to be used by RNA polymerase III for transcription. The U6 gene promoter and the complete guide RNA was synthesized and then cloned into an appropriate vector to make, for example, DNA construct QC783 shown in Figure 8 A (SEQ ID NO:122). Other soybean U6 homologous genes promoters were similarly cloned and used for small RNA expression.

Since the Cas9 endonuclease and the guide RNA need to form a protein/RNA complex to mediate site-specific DNA double strand cleavage, the Cas9 endonuclease and guide RNA must be expressed in same cells. To improve

their co-expression and presence, the Cas9 endonuclease and guide RNA expression cassettes were linked into a single DNA construct, for example, QC815 in Figure 9 A (SEQ ID NO:123), which was then used to transform soybean cells to test the soybean optimized guide RNA/Cas system for genome modification. Similar

5 DNA constructs were made to target different genomic sites using guide RNAs containing different target sequences.

EXAMPLE 13

Selection of soybean genomic sites to be cleaved by the guide RNA/Cas endonuclease system

10 A region of the soybean chromosome 4 (Gm04) was selected to test if the soybean optimized guide RNA/Cas endonuclease system could recognize, cleave, and mutate soybean chromosomal DNA through imprecise non-homologous end-joining (NHEJ) repair. Two genomic target sites were selected one close to a predicted gene Glyma04g39780.1 at 114.13 cM herein named DD20 locus (Figure

15 10A) and another close to Glyma04g39550.1 at 111.95 cM herein named DD43 locus (Figure 10B). Each of the 20 bp variable targeting domain of the guide RNA started with a G residue required by RNA polymerase III and was followed in the soybean genome by a 3 bp PAM motif (Table 11). The chromosome positions of the soybean genomic targets sites in close proximity to the PAM sequences were

20 determined by blast searching the public soybean variety Williams82 genomic sequence. The soybean genomic target sites DD20CR1(SEQ ID NO: 125), DD20CR2 (SEQ ID NO: 126), and DD43CR1 (SEQ ID NO: 127) were identified as all unique in soybean genome while a second identical 23 bp genomic target site DD43CR2 (SEQ ID NO: 128) was found at Gm06:12072339-12072361 so there are

25 two potential cleavage sites targeted by DD43CR2 guide RNA. Both DD43CR1 and DD43CR2 are complementary strand sequences indicated by “c” after the positions.

Table 11. Soybean genomic target sites for a guide RNA/Cas endonuclease system.

Chromosome	Positions	Designation	Genomic Target Sites	PAM
Gm04, 114.13 cM	45936311-45936333	DD20CR1	GGAAGTACACACGACATGA	TGG
	45936324-45936346	DD20CR2	GACATGATGGAACGTGACTA	AGG
Gm04,	45731921-45731943c	DD43CR1	GTCCCTTGTAAGTTGTACGTA	CGG

111.95 cM	45731895-45731917c	DD43CR2	GTATTCTAGAAAAGAGGAAT	TGG
-----------	--------------------	---------	----------------------	-----

Guide RNA expression cassette comprising a variable targeting domain targeting one of DD20CR1, DD20CR2, DD43CR2 genomic target sites were similarly constructed and linked to the soybean Cas9 expression cassette to make DNA constructs QC817, QC818, and QC816 that are similar to QC815 in Figure 9 A (SEQ ID NO:123) except for the 20 bp variable targeting domain of the guide RNA

Since up to six continuous mismatches in the 5' regions of the genomic target site (protospacer) with the 20 bp variable targeting domain can be tolerated, i.e., a continuous stretch of 14 base pairs between the variable targeting domain and the crRNA sequence proximate to the PAM is necessarily enough for efficient targets cleavage any 23 bp genomic DNA sequence following the pattern N(20)NGG can be selected as a target site for the guide RNA/Cas endonuclease system. The last NGG is the PAM sequence that should not be included in the 20 bp variable targeting domain of the guide RNA. If the first N is not endogenously a G residue it must be replaced with a G residue in guide RNA target sequence to accommodate RNA polymerase III, which should not sacrifice recognition specificity of the target site by the guide RNA.

EXAMPLE 14

Delivery of the guide RNA/Cas endonuclease system DNA to soybean by transient transformation

The soybean optimized Cas9 endonuclease and guide RNA expression cassettes were delivered to young soybean somatic embryos in the form of embryogenic suspension cultures by particle gun bombardment. Soybean embryogenic suspension cultures were induced as follows. Cotyledons (~3 mm in length) were dissected from surface sterilized, immature seeds and were cultured for 6-10 weeks in the light at 26 °C on a Murashige and Skoog (MS) media containing 0.7% agar and supplemented with 10 mg/ml 2,4-D (2,4-Dichlorophenoxyacetic acid). Globular stage somatic embryos, which produced secondary embryos, were then excised and placed into flasks containing liquid MS medium supplemented with 2,4-D (10 mg/ml) and cultured in the light on a rotary shaker. After repeated selection for clusters of somatic embryos that multiplied as

early, globular staged embryos, the soybean embryogenic suspension cultures were maintained in 35 ml liquid media on a rotary shaker, 150 rpm, at 26 °C with fluorescent lights on a 16:8 hour day/night schedule. Cultures were subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 ml of the same fresh liquid MS medium.

Soybean embryogenic suspension cultures were then transformed by the method of particle gun bombardment using a DuPont Biolistic™ PDS1000/HE instrument (Bio-Rad Laboratories, Hercules, CA). To 50 µl of a 60 mg/ml 1.0 mm gold particle suspension were added (in order): 30 µl of 30 ng/µl QC815 DNA fragment U6-13.1:DD43CR1+EF1A2:CAS9 as an example, 20 µl of 0.1 M spermidine, and 25 µl of 5 M CaCl₂. The particle preparation was then agitated for 3 minutes, spun in a centrifuge for 10 seconds and the supernatant removed. The DNA-coated particles were then washed once in 400 µl 100% ethanol and resuspended in 45 µl of 100% ethanol. The DNA/particle suspension was sonicated three times for one second each. Then 5 µl of the DNA-coated gold particles was loaded on each macro carrier disk.

Approximately 100 mg of a two-week-old suspension cultures were placed in an empty 60x15 mm Petri dish and the residual liquid removed from the tissue with a pipette. Membrane rupture pressure was set at 1100 psi and the chamber was evacuated to a vacuum of 28 inches mercury. The tissue was placed approximately 3.5 inches away from the retaining screen and bombarded once. The tissue clumps were rearranged and bombarded another time. Minimum amount of liquid MS media without 2,4-D supplement was added to the tissue to prevent the cultures from drying or overgrowing. The 60x15 mm Petri dish was sealed in a 100x25 mm Petri dish containing agar solid MS media to as another measure to keep the tissues from drying up. The tissues were harvested seven days after and genomic DNA was extracted for PCR analysis.

EXAMPLE 15

Analysis of guide RNA/Cas endonuclease system mediated site-specific NHEJ by deep sequencing

To evaluate DNA double strand cleavage at a soybean genomic target site mediated by the guide RNA/Cas endonuclease system, a region of approximately

100 bp genomic DNA surrounding the target site was amplified by PCR and the PCR product was then sequenced to check mutations at the target site as results of NHEJs. The region was first amplified by 20 cycles of PCR with Phusion High Fidelity mastermix (New England Biolabs) from 100 ng genomic DNA using gene-specific primers that also contain adaptors and amplicon-specific barcode sequences needed for a second round PCR and subsequence sequence analysis. For examples, the first PCR for the four experiments listed in Table 2 were done using primers DD20-S3 (SEQ ID NO:133)/DD20-A (SEQ ID NO:134), DD20-S4 (SEQ ID NO:135)/DD20-A, DD43-S3 (SEQ ID NO:136)/DD43-A (SEQ ID NO:137) and DD43-S4 (SEQ ID NO:138)/DD43-A, . One micro liter of the first round PCR products was further amplified by another 20 cycles of PCR using universal primers (SEQ ID NOs:140, 141) with Phusion High Fidelity mastermix. The resulting PCR products were separated on 1.5% agarose gel and the specific DNA bands were purified with Qiagen gel purification spin columns. DNA concentrations were measured with a DNA Bioanalyzer (Agilent) and equal molar amounts of DNA for up to 12 different samples each with specific barcode were mixed as one sample for Illumina deep sequencing analysis. Single read 100 nucleotide-length deep sequencing was performed at a DuPont core facility on a Illumnia's MiSeq Personal Sequencer with a 40% (v/v) spike of PhiX control v3 (Illumina, FC-110-3001) to offset sequence bias.

Since the genomic target site is located in the middle of the ~100 bp long PCR amplicon (SEQ ID NOs: 142, 143, 144, 145), the 100 nucleotide-length deep sequencing is sufficient to cover the targets site region. A window of 10 nucleotides centered over the expected cleavage site, i.e., 3 bp upstream of the PAM, was selected for sequence analysis. Only those reads with one or more nucleotide indel arising within the 10 nucleotide window and not found in a similar level in negative controls were classified as NHEJ mutations. NHEJ mutant reads of different lengths but with the same mutation were counted into a single read and up to 10 most prevalent mutations were visually confirmed to be specific mutations before they were then used to calculate the % mutant reads based on the total analyzed reads containing specific barcode and forward primer.

The frequencies of NHEJ mutations revealed by deep sequencing for four target sites DD20CR1, DD20CR2, DD43CR1, DD43CR2 with one RNA polymerase III promoter GM-U6-13.1 are shown in Table 2. The visually confirmed most prevalent NHEJ mutations are shown in FIG. 11A-11D. The mutant sequences in Figure 11A-11E are listed as SEQ ID NOs:147-201. The top row is the original reference sequence with the target site sequence underlined. Deletions in the mutated sequences are indicated by "---" while additions and replacements are indicated by bold letters. Total count of each mutation of different reads is given in the last column. Cas9 nuclease construct only, guide RNA construct only, and no DNA bombardment negative controls were similarly performed and analyzed but data not shown since no-specific mutations were detected. Other targets sites and guide RNAs were also tested with similar positive results and data not shown.

Table 12. Target site-specific mutations introduced by guide RNA/Cas endonuclease mediated NHEJ.

Experiment	DNA	Mutant reads	Total reads	% Mutants
U6-13.1:DD20CR1+EF1A2:CAS9	QC817	339	710,339	0.048%
U6-13.1:DD20CR2+EF1A2:CAS9	QC818	419	693,483	0.060%
U6-13.1:DD43CR1+EF1A2:CAS9	QC815	489	682,207	0.072%
U6-13.1:DD43CR2+EF1A2:CAS9	QC816	917**	539,681	0.170%

**At least the top 15 reads are specific mutations but only the top 10 are counted in the table to be consistent with other experiments. If all top 15 mutations are counted, the total Mutant reads is 1080 and the % Mutants is 0.200%.

In conclusion, our data indicate that the soybean optimized guide RNA/Cas endonuclease system is able to effectively cleave soybean endogenous genomic DNA and create imperfect NHEJ mutations at the specified genomic target sites.

EXAMPLE 16

The guide RNA/Cas endonuclease system delivers double-strand breaks (DBSs) to the maize *epsps* locus resulting in desired point mutations

Two maize optimized Cas9 endonucleases were developed and evaluated for their ability to introduce a double-strand break at a genomic target sequence. A first Cas9 endonuclease was as described in Figure 1A (Example 2 and expression cassette SEQ ID NO:5). A second maize optimized Cas9 endonuclease (moCas9 endonuclease; SEQ ID NO:192) was supplemented with the SV40 nuclear localization signal by adding the signal coding sequence to the 5' end of the *moCas9* coding sequence (Figure 13). The plant moCas9 expression cassette was subsequently modified by the insertion of the ST-LS1 intron into the *moCas9* coding sequence in order to enhance its expression in maize cells and to eliminate its expression in *E.coli* and *Agrobacterium*. The maize ubiquitin promoter and the potato proteinase inhibitor II gene terminator sequences complemented the *moCas9* endonuclease gene designs. The structural elements of the moCas9 expression cassette are shown in Figure 13 and its amino acid and nucleotide sequences are listed as SEQ ID Nos: 192 and 193.

A single guide RNA (sgRNA) expression cassette was essentially as described in Example 1 and shown in Figure 1B. It consists of the U6 polymerase III maize promoter (SEQ ID NO: 9) and its cognate U6 polymerase III termination sequences (TTTTTTTT). The guide RNA (SEQ ID NO: 194) comprised a 20 nucleotide variable targeting domain (nucleotide1-20 of SEQ ID NO: 194) followed by a RNA sequence capable of interacting with the double strand break inducing endonuclease.

A maize optimized Cas9 endonuclease target sequence (moCas9 target sequence) within the EPSPS codon sequence was complementary to the 20 nucleotide variable sequence of the guide sgRNA determined the site of the Cas9 endonuclease cleavage within the EPSPS coding sequence.

The moCAS9 target sequence (nucleotides 25-44 of SEQ ID NO:209) was synthesized and cloned into the guide RNA-Cas9 expression vector designed for delivery of the components of the guide RNA-Cas9 system to the BMS (Black Mexican Sweet) cells through *Agrobacterium*-mediated transformation.

Agrobacterium T-DNA delivered also the yeast FLP site-specific recombinase and the WDV (wheat dwarf virus) replication-associated protein (replicase). Since the moCas9 target sequences were flanked by the FLP recombination targets (*FRT*), they were excised by FLP in maize cells forming episomal (chromosome-like) structures. Such circular DNA fragments were replicated by the WDV replicase (the origin of replication was embedded into the WDV promoter) allowing their recovery in *E.coli* cells. If the maize optimizedCas9 endonuclease made a double-strand break at the moCas9 target sequence, its repair might produce mutations. The procedure is described in detail in: Lyznik, L.A., Djukanovic, V., Yang, M. and Jones, S. (2012) Double-strand break-induced targeted mutagenesis in plants. In: *Transgenic plants: Methods and Protocols* (Dunwell, J.M. and Wetten, A.C. eds). New York Heidelberg Dordrecht London: Springer, pp. 399-416.

The guideRNA/Cas endonuclease systems using either one of the maize optimized Cas9 endonucleases described herein, generated double-strand breaks in the moCas9 target sequence (Table 13). Table 13 shows the percent of the moCas9 target sequences mutagenized in the maize BMS cells using the moCas9 endonuclease of SEQ ID NO: 192 or the maize optimized cas9 endonuclease described in Figure 1A and expressed by the expression cassette of SEQ ID NO:5. Both guideRNA/Cas endonuclease systems generated double-strand breaks (as judged by the number of targeted mutagenesis events) ranging from 67 to 84% of the moCas9 target sequences available on episomal DNA molecules in maize BMS cells. A sample of mutagenized EPSPS target sequences is shown in Figure 14. This observation indicates that the maize optimized Cas9 endonuclease described herein is functional in maize cells and efficiently generates double-strand breaks at the moCas9 target sequence.

TABLE 13. Percent of the moCas9 target sequences mutagenized in the maize BMS cells by maize optimized Cas9 endonucleases.

Cas9 endonuclease version	# of moCas9 target sequences analyzed	# of intact moCas9 target sequences recovered	# of mutagenized moCas9 target sequences found	Percent mutagenesis (%)
SEQ ID	81	13	68	84%

NO:193 (FIG.13)				
SEQ ID NO: 5 (FIG. 1A)	93	31	62	67%

In order to accomplish targeted genome editing of the maize chromosomal EPSPS gene, a polynucleotide modification template which provided genetic information for editing the *EPSPS* coding sequence was created (SEQ ID NO:195) and co-delivered with the guide RNA/Cas9 system components..

As shown in Figure 12, the polynucleotide modification template comprised three nucleotide modifications (indicated by arrows) when compared to the EPSPS genomic sequence to be edited. These three nucleotide modifications are referred to as TIPS mutations as these nucleotide modifications result in the amino acid changes T-102 to I-102 and P-106 to S-106. The first point mutation results from the substitution of the C nucleotide in the codon sequence ACT with a T nucleotide, a second mutation results from the substitution of the T nucleotide on the same codon sequence ACT with a C nucleotide to form the isoleucine codon (ATC), the third point mutation results from the substitution of the first C nucleotide in the codon sequence CCA with a T nucleotide in order to form a serine codon, TCA.(Figure 12). Both codon sequences were located within 9 nucleotides of each other as shown in SEQ ID NO: 196: **atcgcaatg**cggtca. The three nucleotide modifications are shown in bold. The nucleotides between the two codon sequences were homologous to the non-edited *EPSPS* gene on the *epsps* locus. The polynucleotide modification template further comprised DNA fragments of maize *EPSPS* genomic sequence that were used as homologous sequence for the *EPSPS* gene editing. The short arm of homologous sequence (HR1-Figure 12) was 810 base pairs long and the long arm of homologous sequence (HR2-Figure 12) was 2,883 base pairs long (SEQ ID NO: 195).

In this example, the EPSPS polynucleotide modification template was co-delivered using particle gun bombardment as a plasmid (see template vector 1, Figure 15) together with the guide sgRNA expression cassette and a maize optimizedCas9 endonuclease expression vector which contained the maize optimized Cas9 endonuclease expression cassette described in Figure 1A

(Example 1, SEQ ID NO:5) and also contained a *moPAT* selectable marker gene. Ten to eleven day-old immature embryos were placed, embryo-axis down, onto plates containing the N6 medium (Table 14) and incubated at 28 °C for 4-6 hours before bombardment. The plates were placed on the third shelf from the bottom in the PDS-1000 apparatus and bombarded at 200 psi. Post-bombardment, embryos were incubated in the dark overnight at 28°C and then transferred to plates containing the N6-2 media for 6-8 days at 28°C. The embryos were then transferred to plates containing the N6-3 media for three weeks, followed by transferring the responding callus to plates containing the N6-4 media for an additional three-week selection. After six total weeks of selection at 28°C, a small amount of selected tissue was transferred onto the MS regeneration medium and incubated for three weeks in the dark at 28°C.

Table 14. Composition of Culture Media.

Culture medium	Composition
N6	4.0 g/L N ₆ Basal Salts (Sigma C-1416; Sigma-Aldrich Co., St. Louis, MO, USA), 1.0 ml/L Ericksson's Vitamin Mix (Sigma E-1511), 0.5 mg/L thiamine HCl, 190 g/L sucrose, 1.0 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D), 2.88 g/L L-proline, 8.5 mg/L silver nitrate, 25 mg/L cefotaxime, and 6.36 g/L Sigma agar at pH 5.8
N6-2	4.0 g/L N ₆ Basal Salts (Sigma C-1416), 1.0 ml/L Ericksson's Vitamin Mix (Sigma E-1511), 0.5 mg/L thiamine HCl, 20 g/L sucrose, 1.0 mg/L 2,4-D, 2.88 g/L L-proline, 8.5 mg/L silver nitrate, 25 mg/L cefotaxime, and 8.5 g/L Sigma agar at pH 5.8
N6-3	4.0 g/L N ₆ Basal Salts (Sigma C-1416), 1.0 ml/L Ericksson's Vitamin Mix (Sigma E-1511), 0.5 mg/L thiamine HCl, 30 g/L sucrose, 1.5 mg/L 2,4-D, 0.69 g/L L-proline, 0.5 g/L 2-(N-morpholino)ethanesulphonic acid (MES) buffer, 0.85 mg/L silver nitrate, 5 mg/L glufosinate NH ₄ , and 8.0 g/L Sigma agar at pH 5.8
N6-4	4.0 g/L N ₆ Basal Salts (Sigma C-1416), 1.0 ml/L Ericksson's Vitamin Mix (Sigma E-1511), 0.5 mg/L thiamine HCl, 30 g/L

	sucrose, 1.5 mg/L 2,4-D, 0.69 g/L L-proline, 0.5 g/L MES buffer, 0.85 mg/L silver nitrate, 3 mg/L bialophos, and 8.0 g/L Sigma agar at pH 5.8
MS	4.3 g/L Murashige and Skoog (MS) salts (Gibco 11117; Gibco, Grand Island, NY), 5.0 ml/L MS Vitamins Stock Solution (Sigma M3900), 100 mg/L myo-inositol, 0.1 μ mol abscisic acid (ABA), 1 mg/L indoleacetic acid (IAA), 0.5 mg/L zeatin, 60.0 g/L sucrose, 3.0 mg/L Bialaphos, and 8.0 g/L Sigma agar at pH 5.6

DNA was extracted by placing callus cell samples, two stainless-steel beads, and 450 μ l of extraction buffer (250 mM NaCl, 200 mM Tris-HCl pH 7.4, 25 mM EDTA, 4.2 M Guanidine HCl) into each tube of a Mega titer rack. The rack was shaken in the Genogrinder at 1650 r.p.m. for 60 seconds and centrifuged at 3000 x g for 20 min at 4°C. Three hundred μ l of supernatant was transferred to the wells of the Unifilter 96-well DNA Binding GF/F Microplate (770-2810, WhatmanTM, GE Healthcare). The plate was placed on the top of a Multi-well plate vacuum manifold (5017, Pall Life Sciences). A vacuum pressure was applied until the wells were completely dried. The vacuum filtration procedure was repeated one time with 100 μ l extraction buffer and two times with 250 μ l washing buffer (50 mM Tris-HCl pH 7.4, 200 mM NaCl, 70% ethanol). The residual ethanol was removed by placing the GF/F filter plate on an empty waste collection plate and centrifuged for 10 min at 3000 x g. The DNA was eluted in 100 μ l Elution Buffer (10 mM Tris-HCl, pH 8.3) and centrifuged at 3000 x g for 1 min. For each sample, four PCR reactions were run. They included approximately 40 ng genomic DNA, 10 μ l REDExtract-N-Amp PCR ReadyMix (R4775, Sigma-Aldrich Co.), and 5 picomoles of each primer in a total volume of 20 μ l. Primer combinations for each PCR reaction are listed in the Table 15.

Table 15. Primer combinations for PCR reactions.

PCR reaction	Primer sequence	SEQ ID NO:	PCR product
F-E2	CCGAGGAGATCGTGCTGCA	197	Template randomly integrated or gene editing event
	CAATGGCCGCATTGCAGTTC	198	

F-T	CCGAGGAGATCGTGCTGCA	199	Wild-type EPSPS allele
	TGACCGCATTGCGATTCCAG	200	
H-T	TCCAAGTCGCTTTCCAACAGGATC	201	TIPS editing event
	TGACCGCATTGCGATTCCAG	202	
F-E3	CCGAGGAGATCGTGCTGCA	203	A fragment of the epsps locus for cloning and sequencing
	ACCAAGCTGCTTCAATCCGACAAC	204	

The same PCR reactions were done on five samples of genomic DNA obtained from untransformed maize inbred plantlets. After an initial denaturation at 95°C for 5 minutes, each PCR amplification was carried out over 35 cycles using DNA Engine Tetrad2 Thermal Cycler (BioRad Laboratories, Hercules, CA) at 94°C for 30 sec denaturation, 68°C for 30 sec annealing, and 72°C for 1 min extension. PCR products F-E2, F-T and H-T were separated in 1% agarose gel at 100 Volts for 45 minutes, with 100bp DNA Ladder (N0467S, NewEngland Biolabs). For sequencing, the F-F3 PCR amplified fragments from selected calli were cloned into pCR 2.1-TOPO vectors using the TOPO TA Cloning Kit (Invitrogen Corp, Carlsbad, CA). DNA sequencing was done with BigDye Terminator chemistry on ABI 3700 capillary sequencing machines (Applied Biosystems, Foster City, CA). Each sample contained about 0.5 ug Topo plasmid DNA and 6.4 pmole primer E3-EPex3 Rev (ACCAAGCTGCTTCAATCCGACAAC, SEQ ID NO: 204). Sequences were analyzed using the Sequencer program.

A sample of thirty one callus events selected on media containing bialaphos (the *moPAT* selectable marker gene was part of the guide RNA- moCas9 expression vector) were screened for the presence of the TIPS point mutations. Twenty four events contained the TIPS point mutations integrated into genomic DNA (Figure 16, the F-E2 treatment). Among them, six events showed the PCR amplification product of the chromosomal EPSPS gene with TIPS mutations (Figure 16, the H-T treatment). The pair of PCR primers (one that can hybridize to the genomic *epsps* sequence not present in the *EPSPS* polynucleotide modification template and the other one binding to the edited *EPSPS* sequence present in the EPSPS polynucleotide modification template) distinguished the EPSPS-TIPS editing products from the wild-type *epsps* alleles or random insertions of the TIPS mutations. If one *EPSPS* allele was edited to contain the TIPS substitutions, it should be detected as a DNA fragment originating from the genomic *epsps* locus,

regardless whether the TIPS substitutions were selected for during the PCR amplification process. The TIPS primer was replaced with the wild-type EPSPS primer (Table 15, the F-E3 pair of primers) and the PCR amplification products were cloned into the TOPO cloning vectors and sequenced. The sequencing data
5 represented a random sample of the genomic *epsps* locus sequences in one of the selected events (Figure 17, callus A12 3360.92). Figure 17 shows that the method disclosed herein resulted in the successful nucleotide editing of three nucleotides (Figure 17 bold) responsible for the TIPS mutations without altering any of the other *epsps* nucleotides, while the moCas9 target sequence (the site of guide RNA
10 binding underlined in Figure 17) was not mutagenized.

Also, the other EPSPS allele was not edited indicating that only one EPSPS allele was edited in this particular event (Figure 17, lower section).

This data further shows that the present disclosure of the use of the guide RNA/Cas system for the gene editing demonstrates the ability to recover gene
15 editing events at a high efficiency of 1 out of fewer than 10 selected events.

EXAMPLE 17

The guide RNA/Cas endonuclease system delivers double-strand breaks to the maize *epsps* locus resulting in maize plants containing an *EPSPS-TIPS* edited gene.

20 The *EPSPS* gene edited events were produced and selected as described in the Example 16. In short, the EPSPS polynucleotide modification template was co-delivered using particle gun bombardment as a plasmid (see template vector 1, Figure 15) together with the guide RNA expression cassette and a maize optimized Cas9 endonuclease expression vector which contained the maize optimized Cas9
25 endonuclease expression cassette described in Figure 1A (Example 1, SEQ ID NO:5) and also contained a *moPAT* selectable marker gene.

After six weeks of selection at 28°C, a small amount of selected tissue was transferred onto the MS regeneration medium and incubated for three weeks in the dark at 28°C. After the three week incubation visible shoots were transferred to
30 plates containing the MS-1 medium and incubated at 26°C in the light for 1-2 weeks until they were ready to be sent to a greenhouse and transferred into soil flats. The Ms-1 medium contained: 4.3 g/L MS salts (Gibco 11117), 5.0 ml/L MS Vitamins

Stock Solution (Sigma M3900), 100 mg/L myo-inositol, 40.0 g/L sucrose, and 6.0 g/L Bacto-Agar at pH 5.6.

Using the procedures described above, 390 T0 maize plants were produced originating from 3282 embryos, resulting in an overall transformation efficiency of 12%, further indicating that the guide RNA/Cas system used herein results in low or no toxicity (Table 16).

Table 16. Transformation efficiency of the EPSPS editing.

Treatment	# Embryos	# Calli selected	Selection efficiency	T0 plants to GH	Overall Efficiency
Particle bombardment	3282	489	15%	390	12%

DNA was extracted from each T0 plantlet 7-10 days after transfer to the greenhouse and PCR procedures were conducted as described in the Example 16 to screen the T0 plants for mutations at the *epsps* locus.

Seventy two percent of analyzed T0 plants (270/375, Table 17) contained mutagenized *EPSPS* alleles as determined by the end-point PCR procedure described in the Example 16. Most of the mutations (230/375 or 89%) were produced as a result of error-prone non-homologous end joining (NHEJ) while forty T0 plants (40/ 375 or 11%) contained the TIPS edited *EPSPS* alleles indicating the involvement of a templated double-strand break repair mechanism (Table 17).

Table 17. Mutations at the *epsps* locus.

Transformation	TO Plants Analyzed	Mutations at the <i>epsps</i> locus	Mutation rate	TIPS editing	Gene Editing Rate (TIPS)
Particle bombardment	375	270	72%	40	11 %

A pair of primers (Table 15, the F-E3 pair of primers) was used to amplify a native, endogenous fragment of the *epsps* locus containing the moCas6 target sequence and the EPSPS editing site from the genomic DNA of selected T0 plants.

The PCR amplification products were cloned into the TOPO cloning vectors and

sequenced as described in Example 16. The sequencing data represent a random sample of the genomic *epsps* locus sequences from a particular T0 plant (Table 18) and indicate the genotype of the selected T0 plants. The list of the EPSPS-TIPS allele-containing T0 plants transferred to the pots is presented in Table 18 (a
 5 selected set of T0 plants from the original 40 TIPS-containing events).

Table 18. The *epsps* locus genotypes observed in T0 plants. TIPS refers to a clone comprising the TIPS edited *EPSPS* sequence. NHEJ refers to the presence of a NHEJ mutation and WT refers to the presence of a wild-type *EPSPS* sequence
 10 amplified from the native *epsps* locus.

Event (T0 plant)	Observed Sequences found at the <i>epsps</i> locus
E1	16 TIPS, 13 NHEJ
E2	28 TIPS, 0 NHEJ
E3	2 TIPS, 20 WT
E4	1 TIPS, 28 NHEJ
E5	2 TIPS, 2 NHEJ, 9 WT
E6	10 TIPS, 17 NHEJ
E7	12 TIPS, 17 NHEJ
E8	11 TIPS, 15 NHEJ
E9	17 TIPS, 10 NHEJ

As presented in Table 18, the selected plants of E1 and E3 to E9 contained the *EPSPS-TIPS* edited version of the *EPSPS* gene either accompanied by a wild-type *EPSPS* allele (WT) or a NHEJ mutagenized *EPSPS* allele (NHEJ). The numbers
 15 before TIPS, WT, NHEJ in Table 18 indicate the frequency at which a particular version of the *EPSPS* allele was identified. If all clones contained the TIPS-edited *EPSPS* sequence, the analyzed plant was likely to be homozygous for the *EPSPS-TIPS* allele (see for example E2). If only about 50% of clones contained a TIPS-edited *EPSPS* sequence, the analyzed plant was likely to be hemizygous for the
 20 *EPSPS-TIPS* allele (see for example E1). Other plants, such as E3 or E4, were

likely to be chimeric for TIPS. In one event, E2, the T0 plant contained only TIPS-edited sequence at the *epsps* locus indicating that the guide RNA/Cas endonuclease system disclosed herein resulted in the successful nucleotide editing of three nucleotides (Figure 17 bold) responsible for the two *EPSPS-TIPS* alleles at the *epsps* locus in maize plants.

A qPCR analysis was performed on the selected T0 plants to estimate the copy number of the wild-type *EPSPS* genes and the moCas9 endonuclease sequences. Multiplex qPCR amplifications of the maize *EPSPS* gene and the *ADH* housekeeping gene were carried out on the DNA samples from T0 plants. The primers and probes used in the PCR reaction are shown in Table 19.

Table 19. Primers used in qPCR analysis of T0 plants.

Primer / probe	Primary PCR Primer Sequence	SEQ ID NO:
primer qADH F	5'-CAAGTCGCGGTTTTCAATCA-3	SEQ ID NO:217
Primer qADH R	5'-TGAAGGTGGAAGTCCCAACAA-3'	SEQ ID NO:218
probe ADH-VIC	VIC-TGGGAAGCCTATCTACCAC	SEQ ID NO:219
Probe wtEPSPS	6FAM-CGGCCATTGACAGCA-MGB-NFQ	SEQ ID NO:220
Forward primer qEPSPS F	5'-TCTTGGGGAATGCTGGAAC-3'	SEQ ID NO:221
reverse primer qEPSPSR	5'-CACCAGCAGCAGTAACAGCTG-3'	SEQ ID NO:222
FAM-wtEPSPS R probe	6FAM-TGCTGTCAATGGCCGCA	SEQ ID NO:223
forward primer qEPSPS F	5'-TCTTGGGGAATGCTGGAAC-3'	SEQ ID NO:224
reverse primer q wtEPSPS RA	5'-CCACCAGCAGCAGTAACAGC-3	SEQ ID NO:225)

All analyses were conducted using the LightCycler 480 Real-Time PCR System (Roche Diagnostics). A threshold value for the wtEPSPS genotype was set at 1.76. Every sample showing less than 1.76 copies of *EPSPS*, with the end-point fluorescence measurements up to two times lower than the wild-type control, was

categorized as the One Allele EPSPS genotype (hemizygous for the wild-type *EPSPS* allele).

A qPCR method was used to estimate the TIPS sequence copy number. The primers and probes used in the qPCR reaction are shown in Table 20.

5

Table 20. Primers used in qPCR analysis to estimate the TIPS sequence copy number.

<u>Primer / probe</u>	<u>Primary PCR Primer Sequence</u>	<u>SEQ ID NO:</u>
forward primer q epTIPS F	5'- GGAAGTGCAGCTCTTCTTGGG-3'	SEQ ID NO:226
reverse primer q epTIPS R	5'- AGCTGCTGTCAATGACCGC-3'	SEQ ID NO:227
TIPS probe	6FAM-AATGCTGGAATCGCA	SEQ ID NO:228)

A comparative Ct method with Delta Ct values normalized to the average Delta
10 Ct from the bi-allelic TIPS genotypes provided a copy number estimation for the TIPS sequence detected in the analyzed plant samples.

Table 21. qPCR genotyping and copy number of selected T0 plants.

Event name	TIPS EPSPS allele	Wild-type EPSPS allele #	TIPS copy #	moCas9 coding sequence
E1	positive	Null	5	positive
E2	positive	Null	2	positive
E7	positive	Null	6	positive
E8	positive	Null	1	positive
E9	positive	Null	3	positive

15 The qPCR genotyping indicated that no wild-type *EPSPS* alleles were detected in the selected T0 plants of Events E1, E2, E7, E8 and E9 (Table 21). Both, the TIPS template sequences and the moCas9 coding sequence were found in the selected T0 plants, presumably, as a result of random insertions associated

with the transformation process (Table 21: for the TIPS template sequences E1, E7, and E9 T0 plants). Both genetic elements (the randomly inserted TIPS templates and the moCas9 expression cassette) can be segregated out by standard breeding procedures in the T1 progeny generation, if not linked to the edited *EPSPS-TIPS* gene.

T0 plants grew well in the greenhouse and were fertile. A sample of T0 plants was sprayed with a 1X dose of glyphosate (Roundup Powermax) at V3 growth stage using the spray booth setting of 20 gallons per acre. The 1X dose of glyphosate was prepared as follow: 2.55 ml Powermax in 300 ml water (active ingredient: glyphosate, N-(phosphonomethyl) glycine, in the form of its potassium salt at 48.7%). Seven days after glyphosate application, no leaf tissue damage was observed in some of the T0 plants. These plantlets were hemizygous for the *EPSPS-TIPS* alleles, while other plantlets were severely damaged. One plant showing no damage to the leaf tissue 14 days after herbicide application contained 21 *EPSPS-TIPS* alleles among 44 genomic clones of the *epsps* locus (cloned and sequenced as described in the Example 16).

These data indicate that a guide RNA/Cas system can be used to create a TIPS-edited *EPSPS* allele in maize. Maize plants homozygous at the *epsps-tips* locus (two *EPSPS* alleles edited) with no additional insertion of the TIPS template (plant E2) were obtained. Furthermore, some *EPSPS-TIPS* edited maize plants did show some level of tolerance against a 1X dose of glyphosate.

EXAMPLE 18

Guide RNA/Cas endonuclease mediated DNA cleavage in maize chromosomal loci enables transgene insertion in an elite maize line

To test whether a maize optimized guide RNA/Cas system can cleave an maize chromosomal locus and enable homologous recombination (HR) mediated pathways to site-specifically insert a transgene in an elite maize line, 4 loci were selected on the maize chromosome 1 located between 51.54cM to 54.56 cM (Figure 18). Two target sites for a Cas endonuclease were identified at each of the four loci and are referred to as MHP14Cas-1, MHP14Cas-3, TS8Cas-1, TS8Cas2, TS9Cas-2, TS9Cas-3, TS10Cas-1 and TS10Cas-3 (Figure 19, Table 22, SEQ ID NOs:229-236).

Table 22Maize genomic target sites targeted by a guide RNA/Cas endonuclease.

<u>Locus</u>	<u>Location</u>	<u>Target Site</u>	<u>Maize Genomic Target Site Sequence</u>	<u>PAM</u>	<u>SEQ ID NO:</u>
MHP14	Chr. 1: 51.54cM	MHP14Cas-1	gttaaactcgacgtgaatctgtt	TGG	229
		MHP14Cas-3	acaaacattgaagcgacatag	TGG	230
TS8	Chr. 1: 52.56cM	TS8Cas-1	gtacgtaacgtgcagtac	TGG	231
		TS8Cas-2	gctcatcagtgatcagctgg	TGG	232
TS9	Chr. 1: 53.56cM	TS9Cas-2	ggctgtttgcggcctcg	AGG	233
		TS9Cas-3	gcctcgaggtgcacgcacgt	CGG	234
TS10	Chr.1: 54.56cM	TS10Cas-1	gcctcgccttcgctagtaa	GGG	235
		TS10Cas-3	gctcgtgttgagataca	GGG	236

The maize optimized Cas endonuclease cassette (SEQ ID NO: 5 was as

5 prepared as describe in Example 1. Long guide RNA expression cassettes comprising a variable targeting domain targeting one of the 8 genomic target sites, driven by a maize U6 polymerase III promoter, and terminated by a maize U6 polymerase III terminator were designed as described in Example 1 and 3 and listed in Table 23. A donor DNA (HR repair DNA) containing a selectable marker (a

10 phosphomannose-isomerase (PMI) expression cassette) flanked by two homologous regions was constructed using standard molecular biology techniques (Figure 20).

Table 23.List of guide RNA (gRNA) and Donor DNA expression cassettes

<u>Locus</u>	<u>Target Site</u>	<u>gRNA (SEQ ID NO:)</u>	<u>Donor DNA (SEQ ID NO:)</u>
MHP14	MHP14Cas-1	245	253
	MHP14Cas-3	246	254
TS8	TS8Cas-1	247	255
	TS8Cas-2	248	256
TS9	TS9Cas-2	249	257
	TS9Cas-3	250	258

TS10	TS10Cas-1	251	259
	TS10Cas-3	252	260

A vector containing the maize optimized Cas9 endonuclease of SEQ ID NO: 5, a vector containing one of eight long guide RNA expression cassettes of SEQ ID NOs: 245-252, and a vector containing one of eight donor DNAs of SEQ ID NOs: 253-260 were co-delivered to maize elite line immature embryos by particle-mediated delivery as described in Example 10. About 1000 embryos were bombarded for each target site. Since the donor DNA contained a selectable marker, PMI, successful delivery of the donor DNA allowed for callus growth on mannose media. Putative HR-mediated transgenic insertions were selected by placing the callus on mannose containing media. After selection, stable shoots on maturation plates were sampled, total genomic DNA extracted, and using the primer pairs shown in Table 24 (corresponding to SEQ ID NOs: 261-270), PCR amplification was carried out at both possible transgene genomic DNA junctions to identify putative HR-mediated transgenic insertions.

15

Table 24. Primer sequences used for integration event screening at each target site

<u>Locus</u>	<u>Target Site</u>	<u>Junction</u>	<u>Primer</u>	<u>SEQ ID NO:</u>
UBIR	donor	1	CCATGTCTAACTGTTCAATTTATATGAT TCTCT	261
PSBF	donor	2	GCTCGTGTCCAAGCGTCACTTACGAT TAGCT	262
MHP14	MHP14Cas-1 MHP14Cas-3	14-1HR1f	CTCACATGAGGCTCTTCTTTGCTTGC T	263
		14-1HR2r	AGGATCCTATTCCCCAATTTGTAGAT	264
CHR1-8	TS8Cas-1 TS8Cas-2	8HR1f	CAGTCCGTGGATTGAAGCCAT	265
		8HR2r	CTCTGTCTCCGAGACGTGCTTA	266
CHR1-9	TS9Cas-2 TS9Cas-3	9HR1f	GGAGCAAATGTTTTAGGTATGAAATG	267
		9HR2r	CGGATTCTAAAGATCATACGTAAATG AA	268
CHR1-10	TS10Cas-1 TS10Cas-3	10HR1f	TGGCTTGTCTATGCGCATCTC	269
		10HR2r	CCAGACCCAAACAGCAGGTT	270

The same genomic primers were used for each of the two target sites at one locus. The resulting amplifications were sequenced to determine if these sites were mutated or contained a transgene insertion.

The “Event Recovery frequency” was calculated using the number of events recovered divided by the total number of embryos bombarded, and may indicate if an endonuclease has some toxic effect or not (Table 26). Hence, if 1000 embryos were bombarded and 240 were recovered, the Event Recovery frequency is 24%. Table 26 indicates that for all target sites analyzed the Event Recovery frequency ranged between 17 and 28%, indicating that the guide RNA/Cas system used herein results in low or no toxicity. Cas endonuclease activity was measured *in-planta* by determining the “Target Site Mutation frequency” (Table 26) is defined as: (number of events with target site modification / total number recovered events)*100 %. Hence, if 240 events were recovered and 180 events showed a mutation, the Target Site Mutation frequency is 75%. The target site mutation frequency was measured using target site allele copy number as described in Example 9 of US application 13/886317, filed on May 3, 2013. The primers and probes for obtaining the target site copy number using qPCR at each site were as listed in Table 25 (SEQ ID NO: 271-294).

Table 25

Primer and probe sequences used to assess DNA cleavage at 8 maize genomic target sites

<u>Target Site Designation</u>	<u>Probe primers</u>	<u>Primer sequence</u>	<u>SEQ ID NO:</u>
MHP14Cas-1	probe	CAGATTCACGTCAGATTT	271
	forward	CATAGTGGTGTATGAAAGGAAGCACTT	272
	reverse	CATTTTGGATTGTAATATGTGTACCTCATA	273
MHP14Cas-3	probe	CACCACTATGTCGCTTC	274
	forward	CGGATGCACGAAAATTGTAGGA	275
	reverse	CTGACGTGAATCTGTTTGAATTG	276
TS8Cas-1	probe	TACGTAACGTGCAGTACT	277
	forward	ACGGACGGACCATACGTTATG	278
	reverse	TCAGCTGGTGGAGTATATTAGTTTCGT	279
TS8Cas-2	probe	CCAGCTGATCACTGATGA	280
	forward	ACGGACGGACCATACGTTATG	281
	reverse	CGCACATGTTATAAATTACAATGCAT	282
TS9Cas-2	probe	CTGTTTGCGGCCTC	283
	forward	CTGCGGAGCTGCTGGCGAT	284
	reverse	CTTGCTGGCTTCGTCTGTCA	285

TS9Cas-3	probe	CCGACGTGCGTGCAA	286
	forward	CTGCGGAGCTGCTGGCGAT	287
	reverse	CTTGCTGGCTTCGTCTGTCA	288
TS10Cas-1	probe	TCGCCTTCGCTAGTTAA	289
	forward	AAGACCTGGCCGGTTTTCCA	290
	reverse	TAGCGGCCATTGCCATCA	291
TS10Cas-3	probe	CTGTATCTCCAACACGAGC	292
	forward	AAGACCTGGCCGGTTTTCCA	293
	reverse	TAGCGGCCATTGCCATCA	294

As shown in Table 26, all 8 guide RNA/Cas9 systems were very efficient in cleaving their target DNA and inducing mutations (by non-homologous end joining (NHEJ) as is evidenced by a mutation frequency ranging from 33-90%.

- 5 All events were also screened for the presence of an inserted transgene. The insertion event screening for each target site is illustrated in Figure 21. The primers used for insertion PCR analysis at each site are listed in Table 24. Figure 22 shows one example of an insertion event screening PCR result. The frequency of transgene insertion was determined by calculating the "Insertion frequency" which is defined as: (number of events with target site insertion / total number recovered events)*100 %. Hence, if 240 events were recovered and 21 events showed a transgene insertion, the Insertion frequency was 9%.

15 Table 26. Activity of the guide RNA/Cas 9 system at 8 target sites as determined by target site mutation frequency and transgene insertion frequency at the desired target site in maize plant tissue

Target Site	Event Recovery (%)	Target Site Mutation (%)	Insertion frequency (%)
TS10Cas-1	24%	75%	9% (7*)
TS10Cas-3	22%	83%	16%(20*)
TS8Cas-1	17%	90%	14%(9*)
TS8Cas-2	27%	84%	8%(10*)
MHP14Cas-1	17%	33%	2%(2*)

MHP14Cas-3	28%	68%	4%(1*)
TS9Cas-2	23%	62%	8%**
TS9Cas-3	28%	84%	8%**

* Number of events with HR1 and HR2 both junctions positive

** only HR2 junction available

Sequence –confirmed-PCR amplifications indicated a site-specific transgene
 5 insertion for each of the 8 target sites as shown in Table 26 (column Insertion frequency). A transgene cassette was inserted at all 8 target sites with high efficiency (2-16%). The number of events containing amplifications across both transgene genomic DNA junctions, indicating near perfect site-specific transgene insertion, are show in brackets in Table 26.

10 Taken together, these data demonstrates that maize chromosomal loci cleaved with the maize optimized guide RNA/Cas system described herein can be used to insert transgenes at high frequencies in maize elite inbred line.

EXAMPLE 19

15 Delivery of the guide RNA/Cas9 endonuclease system DNA to soybean by stable transformation

A soybean U6 small nuclear RNA promoter (GM-U6-9.1; SEQ ID NO: 295) was identified in a similar manner as the soybean promoter GM-U6-13.1 (SEQ ID NO:120) described in Example 12. The GM-U6-9.1 promoter was used to express guide RNA to direct Cas9 nuclease to designated genomic target site.

20 A soybean codon optimized Cas9 endonuclease expression cassette (such as for example EF1A2:CAS9, SEQ ID NO: 296) and a guide RNA expression cassette (such as for example U6-9.1:DD20CR1; SEQ ID NO: 297) were linked (such as U6-9.1: DD20CR1 + EF1A2:CAS9; SEQ ID NO: 298, Figure 23A) and integrated into a DNA plasmid that was co-delivered with another plasmid
 25 comprising a donor DNA (repair DNA) cassette (such as DD20HR1-SAMS:HPT-DD20HR2; SEQ ID NO: 299) to young soybean somatic embryos in the form of embryogenic suspension cultures by particle gun bombardment (Fig. 23A and 23B). Other guide RNA/Cas9 DNA constructs targeting various soybean genomic sites and donor DNA constructs for site-specific transgene integration through

homologous recombination were similarly configured and are listed in Table 27. The four gRNA/Cas9 constructs differed only in the 20 bp guide RNA targeting domain (variable targeting domain) targeting the soybean genomic target sites DD20CR1 (SEQ ID NO: 125), DD20CR2 (SEQ ID NO: 126), DD43CR1 (SEQ ID NO: 127), or DD43CR2 (SEQ ID NO: 128). The two donor DNA constructs differed only in the homologous regions such as DD20HR1 and DD20HR (Figure 23B), or DD43HR1 and DD43HR2. These guide RNA/Cas9 DNA constructs and donor DNAs were co-delivered to an elite (93B86) or a non-elite (Jack) soybean genome by the stable transformation procedure described below.

Table 27. Guide RNA/Cas9 Mediated Soybean Stable Transformation .

Experiment	Guide RNA/Cas9	Donor DNA	SEQ ID NOs:
U6-9.1DD20CR1	U6-9.1:DD20CR1+EF1A2:CAS9	DD20HR1-SAMS:HPT-DD20HR2	298, 299
U6-9.1DD20CR2	U6-9.1:DD20CR2+EF1A2:CAS9	DD20HR1-SAMS:HPT-DD20HR2	300, 299
U6-9.1DD43CR1	U6-9.1:DD43CR1+EF1A2:CAS9	DD43HR1-SAMS:HPT-DD43HR2	301, 302
U6-9.1DD43CR2	U6-9.1:DD43CR2+EF1A2:CAS9	DD43HR1-SAMS:HPT-DD43HR2	303,302

Soybean somatic embryogenic suspension cultures were induced from a DuPont Pioneer proprietary elite cultivar 93B86 as follows. Cotyledons (~3 mm in length) were dissected from surface sterilized, immature seeds and were cultured for 6-10 weeks in the light at 26 °C on a Murashige and Skoog (MS) media containing 0.7% agar and supplemented with 10 mg/ml 2,4-D (2,4-Dichlorophenoxyacetic acid). Globular stage somatic embryos, which produced secondary embryos, were then excised and placed into flasks containing liquid MS medium supplemented with 2,4-D (10 mg/ml) and cultured in light on a rotary shaker. After repeated selection for clusters of somatic embryos that multiplied as early, globular staged embryos, the soybean embryogenic suspension cultures were maintained in 35 ml liquid media on a rotary shaker, 150 rpm, at 26 °C with fluorescent lights on a 16:8 hour day/night schedule. Cultures were subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 ml of the same fresh liquid MS medium.

Soybean embryogenic suspension cultures were then transformed by the method of particle gun bombardment using a DuPont Biolistic™ PDS1000/HE instrument (Bio-Rad Laboratories, Hercules, CA). To 50 µl of a 60 mg/ml 1.0 mm gold particle suspension were added in order: 30 µl of equal amount (30 ng/µl) plasmid DNA comprising, for example, U6-9.1:DD20CR1+EF1A2:CAS9 (SEQ ID NO:298) and plasmid DNA comprising, for example, (DD20HR1-SAMS:HPT-DD20HR2, SEQ ID NO: 299) (Experiment U6-9.1DD20CR1 listed in Table 27) 20 µl of 0.1 M spermidine, and 25 µl of 5 M CaCl₂. The particle preparation was then agitated for 3 minutes, spun in a centrifuge for 10 seconds and the supernatant removed. The DNA-coated particles were then washed once in 400 µl 100% ethanol and resuspended in 45 µl of 100% ethanol. The DNA/particle suspension was sonicated three times for one second each. Then 5 µl of the DNA-coated gold particles was loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture was placed in an empty 60x15 mm Petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5 to 10 plates of tissue were bombarded. Membrane rupture pressure was set at 1100 psi and the chamber was evacuated to a vacuum of 28 inches mercury. The tissue was placed approximately 3.5 inches away from the retaining screen and bombarded once. Following bombardment, the tissue was divided in half and placed back into liquid media and cultured as described above.

Five to seven days post bombardment, the liquid media was exchanged with fresh media containing 30 mg/ml hygromycin as selection agent. This selective media was refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue was observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue was removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each clonally propagated culture was treated as an independent transformation event and subcultured in the same liquid MS media supplemented with 2,4-D (10 mg/ml) and 30 ng/ml hygromycin selection agent to increase mass. The embryogenic suspension cultures were then transferred to agar solid MS media plates without 2,4-D supplement to allow somatic embryos to

develop. A sample of each event was collected at this stage for quantitative PCR analysis.

Cotyledon stage somatic embryos were dried-down (by transferring them into an empty small Petri dish that was seated on top of a 10 cm Petri dish containing some agar gel to allow slow dry down) to mimic the last stages of soybean seed development. Dried-down embryos were placed on germination solid media and transgenic soybean plantlets were regenerated. The transgenic plants were then transferred to soil and maintained in growth chambers for seed production. Transgenic events were sampled at somatic embryo stage or T0 leaf stage for molecular analysis.

Similar transformation experiments (U6-9.1DD20CR2, U6-9.1DD43CR1, U6-9.1DD43CR2) with the components listed in Table 27 and using the elite cultivar 93B86 were performed as described above.

Two transformation experiments, U6-9.1DD20CR1 and U6-9.1DD43CR1 listed in Table 27, were also performed in a non-elite soybean cultivar "Jack" to test the gRNA/Cas9 system performance in different soybean genotypes.

EXAMPLE 20

Detection of Site-specific NHEJ Mediated by the Guide RNA/Cas9 System in Stably Transformed Soybean

Genomic DNA was extracted from somatic embryo samples and analyzed by quantitative PCR using a 7500 real time PCR system (Applied Biosystems, Foster City, CA) with target site-specific primers and FAM-labeled fluorescence probe to check copy number changes of the target site DD20 or DD43 (Figure 24 A-C). The qPCR analysis was done in duplex reactions with a heat shock protein (HSP) gene as the endogenous controls and a wild type 93B86 genomic DNA sample that contains one copy of the target site with 2 alleles, as the single copy calibrator. The HSP endogenous control qPCR employed primer probe set HSP-F/HSP-T/HSP-R. The DD20-CR1 (SEQ ID NO:306) and DD20-CR2 (SEQ ID NO:307) specific qPCR employed primer probe set DD20-F (SEQ ID NO:308)/DD20-T(SEQ ID NO:309)/DD20-R(SEQ ID NO:310). The DD43-CR1 (SEQ ID NO:311) specific qPCR employed primer probe set DD43-F(SEQ ID NO:313)/DD43-T(SEQ ID NO:315)/DD43-R (SEQ ID NO:316) while the DD43-CR2 (SEQ ID NO:312) specific

qPCR employed primer probe set DD43-F2(SEQ ID NO:314)/DD43-T/DD43-R. The guide RNA/Cas9 DNA (SEQ ID NOs: 298, 300, 301, and 303) specific qPCR employed primer probe set Cas9-F (SEQ ID NO:317/Cas9-T(SEQ ID NO:318)/Cas-9-R(SEQ ID NO:319). The donor DNA (SEQ ID NOS: 299, and 302) specific qPCR employed primer probe set Sams-76F (SEQ ID NO:320)/ FRT1I63-T (SEQ ID NO:321)/ FRT1I-41F (SEQ ID NO:322). The endogenous control probe HSP-T was labeled with VIC and the gene-specific probes DD20-T, DD43-T, Cas9-T, and FRT1I63-T were labeled with FAM for the simultaneous detection of both fluorescent probes (Applied Biosystems). PCR reaction data were captured and analyzed using the sequence detection software provided with the 7500 real time PCR system and the gene copy numbers were calculated using the relative quantification methodology (Applied Biosystems).

Since the wild type 93B86 genomic DNA with two alleles of the target site was used as the single copy calibrator, events without any change of the target site would be detected as one copy herein termed Wt-Homo (qPCR value ≥ 0.7), events with one allele changed, which is no longer detectable by the target site-specific qPCR, would be detected as half copy herein termed NHEJ-Hemi (qPCR value between 0.1 and 0.7), while events with both alleles changed would be detected as null herein termed NHEJ-Null (qPCR value ≤ 0.1). The wide range of the qPCR values suggested that most of the events contained mixed mutant and wild type sequences of the target site. High percentage of NHEJ-Hemi (ranging from 10.1 to 33.5%, Table 28) and NHEJ-Null (ranging from 32.3 to 46.4%, Table 21) were detected in all four experiments with combined NHEJ average frequencies of more than 60% (Table 28).

Table 28. Target Site Mutations and Site Specific Gene Integration Induced by the Guide RNA/Cas9 system in elite soybean germplasm. Numbers indicate no. of events (numbers in parentheses are %). NA= not analyzed.

Project	Total event	Wt-Homo (%)	NHEJ-Hemi (%)	NHEJ-Null (%)	Insertion Frequency(%)
U6-9.1DD20CR1	239	85 (35.6%)	77 (32.2%)	77 (32.2%)	11 (4.6%)

U6-9.1DD20CR2	79	43 (54.4%)	8 (10.1%)	28 (35.4%)	NA
U6-9.1DD43CR1	263	53 (20.2%)	88 (33.5%)	122 (46.4%)	10 (3.8%)

Table 29. Target Site Mutations and Site Specific Gene Integration Induced by the Guide RNA/Cas9 system in non-elite soybean germplasm. Numbers indicate no. of events (numbers in parentheses are % of the total analyzed events).

Project	Total event	Wt-Homo (%)	NHEJ-Hemi (%)	NHEJ-Null (%)	Insertion frequency (%)
U6-9.1DD20CR1-Jack	149	99 (66.4%)	34 (22.8%)	16 (10.7%)	0 (0%)
U6-9.1DD43CR1-Jack	141	84 (59.6%)	27 (19.1%)	30 (21.3%)	1 (0.7%)

5

Both NHEJ-Hemi and NHEJ-Null were detected in the two experiments U6-9.1DD20CR1-Jack and U6-9.1DD43CR1-Jack repeated in “Jack” genotype though at lower frequencies (Table 29). The differences between NHEJ frequencies were likely caused by variations between transformation experiments.

10 The target region of NHEJ-Null events were amplified by regular PCR from the same genomic DNA samples using DD20-LB (SEQ ID NO: 323) and DD20-RB (SEQ ID NO: 326) primers specific respectively to DD20-HR1 and DD20-HR2 for DD20 target site specific HR1-HR2 PCR amplicon (Figure 25 A-C; SEQ ID NO: 329), or DD43-LB (SEQ ID NO: 327) and DD43-RB (SEQ ID NO: 328) primers
15 specific respectively to DD43-HR1 and DD43-HR2 for DD43 target site specific HR1-HR2 PCR amplicon (SEQ ID NO: 332). The PCR bands were cloned into pCR2.1 vector using a TOPO-TA cloning kit (Invitrogen) and multiple clones were sequenced to check for target site sequence changes as the results of NHEJ. Various small deletions at the Cas9 cleavage site, 3 bp upstream of the PAM, were
20 revealed at all four tested target sites (Figure 26 A-C). Small insertions were also detected in some sequences. Different mutated sequences were identified from some of the same events indicating the chimeric nature of these events. Some of the same mutated sequences were also identified from different events suggesting that the same mutations could have happened independently or some of the events
25 could be clonal events. These sequence analysis confirmed the occurrence of NHEJ mediated by the guide RNA/Cas9 system at the specific Cas9 target sites.

EXAMPLE 21

Identification of Site-Specific Gene Integration via Homologous Recombination Mediated by the Guide RNA/Cas9 System in Stably Transformed Soybean

Site-specific gene integration via guide RNA/Cas9 system mediated DNA

5 homologous recombination was determined by border-specific PCR analysis. The 5' end borders of DD20CR1 and DD20CR2 events were amplified as a 1204 bp DD20 HR1-SAMS PCR amplicon (SEQ ID NO: 330) by PCR with primers DD20-LB (SEQ ID NO: 323) and Sams-A1 (SEQ ID NO: 324) while the 3' borders of the same events were amplified as a 1459 bp DD20 NOS-HR2 PCR amplicon (SEQ ID NO:

10 331) with primers QC498A-S1 and DD20-RB (Figure 25 A-C). Any events with both the 5' border and 3' border-specific bands amplified are considered as site-specific integration events through homologous recombination containing the transgene from the donor DNA fragment DD20HR1-SAMS:HPT-DD20HR2or its circular form (FIG. 23). The 5' end borders of DD43CR1 and DD43CR2 events were amplified as

15 a 1202 bp DD43 HR1-SAMS PCR amplicon (SEQ ID NO: 333) by PCR with primers DD43-LB and Sams-A1 while the 3' borders of the same events were amplified as a 1454 bp DD43 NOS-HR2 PCR amplicon (SEQ ID NO: 334) with primers QC498A-S1 (SEQ ID NO: 325) and DD43-RB (SEQ ID NO: 328). Any events with both the 5' border and 3' border-specific bands amplified are considered

20 as site-specific integration events through homologous recombination containing the transgene from repair DNA fragment DD43HR1-SAMS:HPT-DD43HR2or its circular form. Some of the border-specific PCR fragments were sequenced and were all confirmed to be recombined sequences as expected from homologous recombination. On average, gene integration through the guide RNA/Cas9 mediated

25 homologous recombination occurred at approximately 4% of the total transgenic events (Insertion frequency, Table 28 and Table 29). One homologous recombination event was identified from experiment U6-9.1DD43CR1-Jack repeated in "Jack" genotype (Table 29).

EXAMPLE 22

The crRNA/tracrRNA/Cas endonuclease system cleaves chromosomal DNA in maize and introduces mutations by imperfect non-homologous end-joining

To test whether the maize optimized crRNA/tracrRNA/Cas endonuclease system described in Example 1 could recognize, cleave, and mutate maize chromosomal DNA through imprecise non-homologous end-joining (NHEJ) repair pathways, three different genomic target sequences were targeted for cleavage (see Table 30) and examined by deep sequencing for the presence of NHEJ mutations.

10 Table 30. Maize genomic target sequences targeted by a crRNA/tracrRNA/Cas endonuclease system.

<u>Locus</u>	<u>Location</u>	<u>Cas RNA System Used</u>	<u>Target Site Designation</u>	<u>Maize Genomic Target Site Sequence</u>	<u>PAM Sequence</u>	<u>SEQ ID NO:</u>
LIG	Chr. 2: 28.45cM	crRNA/tracrRNA	LIGCas-1	GTACCGTACGTGCCCCGGCGG	AGG	16
		crRNA/tracrRNA	LIGCas-2	GGAATTGTACCGTACGTGCCC	CGG	17
		crRNA/tracrRNA	LIGCas-3	GCGTACGCGTACGTGTG	AGG	18

LIG=Liguleless 1 Gene Promoter

The maize optimized Cas9 endonuclease expression cassette, crRNA expression cassettes containing the specific maize variable targeting domains (SEQ ID NOs: 445-447) complementary to the antisense strand of the maize genomic target sequences listed in Table 30 and tracrRNA expression cassette (SEQ ID NO: 448) were co-delivered to 60-90 Hi-II immature maize embryos by particle-mediated delivery (see Example 5) in the presence of BBM and WUS2 genes (see Example 6). Hi-II maize embryos transformed with the Cas9 and long guide RNA expression cassettes targeting the LIGCas-3 genomic target site (SEQ ID NO: 18) for cleavage served as a positive control and embryos transformed with only the Cas9 expression cassette served as a negative control. After 7 days, the 20-30 most uniformly transformed embryos from each treatment were pooled and total genomic DNA was extracted. The region surrounding the intended target site was PCR amplified with Phusion® High Fidelity PCR Master Mix (New England Biolabs, M0531L) adding on the sequences necessary for amplicon-specific barcodes and Illumina sequencing using “tailed” primers through two rounds of PCR. The primers

used in the primary PCR reaction are shown in Table 31 and the primers used in the secondary PCR reaction were

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACG (forward, SEQ ID NO: 53) and CAAGCAGAAGACGGCATA (reverse, SEQ ID NO: 54).

5

Table 31. PCR primer sequences

<u>Target Site</u>	<u>Cas RNA System Used</u>	<u>Primer Orientation</u>	<u>Primary PCR Primer Sequence</u>	<u>SEQ ID NO:</u>
LIGCas-1	crRNA/ tracrRNA	Forward	CTACACTCTTTCCCTACACGACGCTCTTCC GATCTTCCTCTGTAACGATTTACGCACCTG CTG	36
LIGCas-1	crRNA/ tracrRNA	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCC GATCTGCAAATGAGTAGCAGCGCACGTAT	35
LIGCas-2	crRNA/ tracrRNA	Forward	CTACACTCTTTCCCTACACGACGCTCTTCC GATCTGAAGCTGTAACGATTTACGCACCTG CTG	449
LIGCas-2	crRNA/ tracrRNA	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCC GATCTGCAAATGAGTAGCAGCGCACGTAT	35
LIGCas-3	crRNA/ tracrRNA	Forward	CTACACTCTTTCCCTACACGACGCTCTTCC GATCTAAGGCGCAAATGAGTAGCAGCGCA C	37
LIGCas-3	crRNA/ tracrRNA	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCC GATCTCACCTGCTGGGAATTGTACCGTA	38
LIGCas-3	Long guide RNA	Forward	CTACACTCTTTCCCTACACGACGCTCTTCC GATCTTTCCCGCAAATGAGTAGCAGCGCA C	450
LIGCas-3	Long guide RNA	Reverse	CAAGCAGAAGACGGCATAACGAGCTCTTCC GATCTCACCTGCTGGGAATTGTACCGTA	38

The resulting PCR amplifications were purified with a Qiagen PCR purification spin column, concentration measured with a Hoechst dye-based fluorometric assay, combined in an equimolar ratio, and single read 100 nucleotide-length deep sequencing was performed on Illumina's MiSeq Personal Sequencer with a 30-40% (v/v) spike of PhiX control v3 (Illumina, FC-110-3001) to off-set sequence bias. Only those reads with a ≥ 1 nucleotide indel arising within the 10 nucleotide window centered over the expected site of cleavage and not found in a similar level in the negative control were classified as NHEJ mutations. NHEJ mutant reads with the same mutation were counted and collapsed into a single read

and the top 10 most prevalent mutations were visually confirmed as arising within the expected site of cleavage. The total numbers of visually confirmed NHEJ mutations were then used to calculate the % mutant reads based on the total number of reads of an appropriate length containing a perfect match to the barcode and forward primer.

The frequency of NHEJ mutations recovered by deep sequencing for the crRNA/tracrRNA/Cas endonuclease system targeting the three LIGCas targets (SEQ ID NOS: 16, 17, 18) compared to the long guide RNA/Cas endonuclease system targeting the same locus is shown in Table 32.

10

Table 32. Percent (%) mutant reads at maize Liguleless 1 target locus produced by crRNA/tracrRNA/Cas endonuclease system compared to the long guide RNA/Cas endonuclease system

<u>System</u>	<u>Total Number of Reads</u>	<u>Number of Mutant Reads</u>	<u>% Mutant Reads</u>
Cas9 Only Control	1,744,427	0	0.00%
LIGCas-3 long guide RNA	1,596,955	35,300	2.21%
LIGCas-1 crRNA/tracrRNA	1,803,163	4,331	0.24%
LIGCas-2 crRNA/tracrRNA	1,648,743	3,290	0.20%
LIGCas-3 crRNA/tracrRNA	1,681,130	2,409	0.14%

The ten most prevalent types of NHEJ mutations recovered based on the crRNA/tracrRNA/Cas endonuclease system are shown in Figure 27A (for LIGCas-1 target site, corresponding to SEQ ID NOS:415-424), Figure 27B (for LIGCas-2 target site corresponding to SEQ ID NOS: 425-434) and Figure 27C (for LIGCas-3 target site corresponding to SEQ ID NOS:435-444). Approximately, 9-16 fold lower frequencies of NHEJ mutations were observed when using a crRNA/tracrRNA/Cas endonuclease system to introduce a double strand break at a maize genomic target site, relative to the long guide RNA/Cas endonuclease system control.

Taken together, our data indicate that the maize optimized crRNA/tracrRNA/Cas endonuclease system described herein cleaves maize chromosomal DNA and generates imperfect NHEJ mutations.

25

EXAMPLE 23

Modifying the ARGOS8 gene to improve drought tolerance and nitrogen use efficiency in maize plants

ARGOS is a negative regulator for ethylene responses in plants (WO 2013/066805 A1, published 10 May 2013). ARGOS proteins target the ethylene signal transduction pathway. When over-expressed in maize plants, ARGOS reduces plant sensitivity to ethylene and promotes organ growth, leading to increased drought tolerance (DRT) and improved nitrogen use efficiency (NUE) ((WO 2013/066805 A1, published 10 May 2013). To achieve optimal ethylene sensitivity, promoters have been tested for driving Zm-ARGOS8 over-expression in transgenic maize plants. Field trials showed that a maize promoter, Zm-GOS2 PRO:GOS2 INTRON (SEQ ID NO:460, US 6,504,083 patent issued on January 7, 2003; Zm-GOS2 is a maize homologous gene of rice GOS2. Rice GOS2 stands for Gene from Oryza Sativa 2), provided a favorable expression level and tissue coverage for Zm-ARGOS8 and the transgenic plants have a higher grain yield than non-transgenic controls under drought stress and low nitrogen conditions (WO 2013/066805 A1, published 10 May 2013). However, these transgenic plants contain two ARGOS8 genes, the endogenous gene and the transgene. ARGOS8 protein levels, therefore, are determined by these two genes. Because the endogenous ARGOS8 gene varies in sequence and the expression level among different inbred lines, the ARGOS8 protein level will be different when the transgene is integrated into different inbreds. Here we present a mutagenization (gene editing) method to modify the promoter region of the endogenous ARGOS8 gene to attain desired expression patterns and eliminate the need for a transgene.

The promoter Zm-GOS2 PRO:GOS2 INTRON (SEQ ID NO:460; US 6,504,083 patent issued on January 7, 2003) was inserted into the 5'-UTR of Zm-ARGOS8 (SEQ ID NO:462) by using a guideRNA/Cas9 system. The Zm-GOS2 PRO:GOS2 INTRON fragment also included a primer binding site (SEQ ID NO:459) at its 5' end to facilitate event screening with PCR. We also substituted the native promoter of Zm-ARGOS8 (SEQ ID NO:461) with Zm-GOS2 PRO::GOS2 INTRON (SEQ ID NO:460). Resulted maize lines carry a new ARGOS8 allele whose expression levels and tissue specificity will differ from the native form. We expect

that these lines will recapitulate the phenotype of increased drought tolerance and improved NUE as observed in the Zm-GOS2 PRO:Zm-ARGOS8 transgenic plants (WO 2013/066805 A1, published 10 May 2013). These maize lines are different from those conventional transgenic events: (1) there is only one ARGOS8 gene in the genome; (2) this modified version of Zm-ARGOS8 resides at its native locus; (3) the ARGOS8 protein level and the tissue specificity of gene expression are entirely controlled by the edited allele. The DNA reagents used during the mutagenization, such as guideRNA, Cas9endonuclease, transformation selection marker and other DNA fragments are not required for function of the newly generated ARGOS8 allele and can be eliminated from the genome by segregation through standard breeding methods. Because the promoter Zm-GOS2 PRO:GOS2 INTRON was copied from maize GOS2 gene (SEQ ID NO:464) and inserted into the ARGOS8 locus through homologous recombination, this ARGOS8 allele is indistinguishable from natural mutant alleles.

A. Insertion of Zea mays-GOS2 PRO:GOS2 INTRON into maize-ARGOS 8 Promoter

To insert Zm-GOS2 PRO:GOS2 INTRON into the 5'-UTR of maize ARGOS8 gene, a guideRNA construct, gRNA1, was made using maize U6 promoter and terminator as described herein. The 5'-end of the guide RNA contained a 19-bp variable targeting domain targeting the genomic target sequence 1 (CTS1; SEQ ID NO; 451) in the 5'-UTR of Zm-ARGOS8 (Figure 28). A polynucleotide modification template containing the Zm-GOS2 PRO:GOS2 INTRON that was flanked by two genomic DNA fragments (HR1 and HR2, 370 and 430-bp in length, respectively) derived from the upstream and downstream region of the CTS1 (Figure 28). The gRNA1 construct, the polynucleotide modification template, a Cas9 cassette and transformation selection marker phosphomannose isomerase (PMI) were introduced into maize immature embryo cells by using a particle bombardment method. PMI-resistant calli were screened with PCR for Zm-GOS2 PRO:GOS2 INTRON insertion (Figure 29A and 29B). Multiple callus events were identified and plants were regenerated. The insertion events were confirmed by amplifying the Zm-ARGOS8 region in T0 plants with PCR (Figure 29C) and sequencing the PCR products.

B. Replacement of Zm-ARGOS 8 Promoter with Zm-GOS2 PRO:GOS2 INTRON promoter (Promoter swap).

To substitute (replace) the native promoter of Zm-ARGOS8 with Zm-GOS2 PRO:GOS2 INTRON, a guide RNA construct, gRNA3, was made for targeting the genomic target site CTS3 (SEQ ID NO:453), located 710-bp upstream of the Zm-ARGOS8 start codon (Figure 30). Another guide RNA, gRNA2, was designed to target the genomic target site CTS2 (SEQ ID NO:452) located in the 5'-UTR of Zm-ARGOS8 (Figure 30). The polynucleotide modification template contained a 400-bp genomic DNA fragment derived from the upstream region of CTS3, Zm-GOS2 PRO:GOS2 INTRON and a 360-bp genomic DNA fragment derived from the downstream region of CTS2 (Figure 30). The gRNA3 and gRNA2, the Cas9 cassette, the polynucleotide modification template and the PMI selection marker were used to transform immature embryo cells. Multiple promoter swap (promoter replacement) events were identified by PCR screening of the PMI-resistance calli (Figure 31A, 31B & 31C) and plants were regenerated. The swap events were confirmed by PCR analysis of the Zm-ARGOS8 region in T0 plants (Figure 31D).

C. Deletion of Zm-ARGOS 8 Promoter

To delete the promoter of Zm-ARGOS8, we screened the PMI-resistance calli obtained from the above gRNA3/gRNA2 experiment to look for events that produce a 1.1-kb PCR product (Figure 32A). Multiple deletion events were identified (Figure 32B) and plants were regenerated. The deletion events were confirmed by amplifying the Zm-ARGOS8 region in T0 plants with PCR and sequencing of the PCR products.

EXAMPLE 24

Gene editing of the soybean EPSPS1 gene using the guide RNA/Cas endonuclease system

A. guideRNA/Cas9 endonuclease target site design on the Soybean EPSPS genes.

Two guideRNA/ Cas9 endonuclease target sites (soy EPSPS-CR1 and soy EPSPS-CR2) were identified in the Exon2 of the soybean EPSPS1 gene Glyma01g33660 (Table 33).

Table 33. Guide RNA/Cas9 endonuclease target sites on soybean EPSPS1 gene

Name of gRNA-Cas9 endonuclease target site	Cas endonuclease target sequence (SEQ ID NO:)	Physical location
soy EPSPS-CR1	467	Gm01:45865337..45865315
soy EPSPS-CR2	468	Gm01: 45865311..45865333

B. Guide-RNA expression cassettes, Cas9 endonuclease expression cassettes and polynucleotide modification templates for introduction of specific amino acid changes in the soybean EPSPS1 gene

The soybean U6 small nuclear RNA promoter, GM-U6-13.1 (SEQ ID. NO: 469), was used to express guide RNAs to direct Cas9 nuclease to designated genomic target sites (Table 34). A soybean codon optimized Cas9 endonuclease (SEQ ID NO: 489) expression cassette and a guide RNA expression cassette were linked in a first plasmid that was co-delivered with a polynucleotide modification template. The polynucleotide modification template contained specific nucleotide changes that encoded for amino acid changes in the EPSPS1 polypeptide (Glyma01g33660), such as the T183I and P187S (TIPS) in the Exon2. Other amino acid changes in the EPSPS1 polypeptide can also be obtained using the guide RNA/Cas endonuclease system described herein . Specific amino acid modifications can be achieved by homologous recombination between the genomic DNA and the polynucleotide modification template facilitated by the guideRNA/Cas endonuclease system.

Table 34. Guide RNA/Cas9 expression cassettes and polynucleotide modification templates used in soybean stable transformation for the specific amino acid modifications of the EPSPS1 gene.

Experiment	Guide RNA/Cas9 (plasmid name)	SEQ ID NO:	polynucleotide modification template	SEQ ID NO:
soy EPSPS-CR1	U6-13.1:EPSPS CR1 + EF1A2:CAS9 (QC878)	470	RTW1013A	472
soy EPSPS-	U6-13.1:EPSPS CR2 + EF1A2:CAS9	471	RTW1012A	473

CR2	(QC879)			
-----	---------	--	--	--

C. Detection of site-specific non-homologous-end-joining (NHEJ) mediated by the guide RNA/Cas9 system in stably transformed soybean

- Genomic DNA was extracted from somatic embryo samples and analyzed by quantitative PCR using a 7500 real time PCR system (Applied Biosystems, Foster City, CA) with target site-specific primers and FAM-labeled fluorescence probe to check copy number changes of the double strand break target sites. The qPCR analysis was done in duplex reactions with a syringolide induced protein (SIP) as the endogenous controls and a wild type 93B86 genomic DNA sample that contains one copy of the target site with 2 alleles, as the single copy calibrator. The presence or absence of the guide RNA –Cas9 expression cassette in the transgenic events was also analyzed with the qPCR primer/probes for guideRNA/Cas9 (SEQ IDs: 477-479) and for PinII (SEQ ID: 480-482).he qPCR primers/probes are listed in Table 35.

15 Table 35. Primers/Probes used in qPCR analyses of transgenic soybean events.

Target Site	Primer/Probe Name	Sequences	SEQ ID NOs:
EPSPS-CR1 & EPSPS-CR2	Soy1-F1	CCACTAGTAAGGAATCTAAAGATGAAATCA	474
	Soy1-R2	CCTGCAGCAACCACAGCTGCTGTC	475
	Soy1-T1 (FAM-MGB)	CTGCAATGCGTCCTT	476
gRNA/CAS9	Cas9-F	CCTTCTTCCACCGCCTTGA	477
	Cas9-R	TGGGTGTCTCTCGTGCTTTTT	478
	Cas9-T (FAM-MGB)	AATCATTCCTGGTGGAGGA	479
pINII	pINII-99F	TGATGCCCCACATTATAGTGATTAGC	480
	pINII-13R	CATCTTCTGGATTGGCCAACTT	481
	pINII-69T (FAM-MGB)	ACTATGTGTGCATCCTT	482
SIP	SIP-130F	TTCAAGTTGGGCTTTTTCAGAAG	483
	SIP-198R	TCTCCTTGGTGCTCTCATCACA	484
	SIP-170T (VIC-MGB)	CTGCAGCAGAACCAA	485

The endogenous control probe SIP-T was labeled with VIC and the gene-specific probes for all the target sites were labeled with FAM for the simultaneous detection of both fluorescent probes (Applied Biosystems). PCR reaction data were captured and analyzed using the sequence detection software provided with the 7500 real time PCR system and the gene copy numbers were calculated using the relative quantification methodology (Applied Biosystems).

Since the wild type 93B86 genomic DNA with two alleles of the double strand break target site was used as the single copy calibrator, events without any change of the target site would be detected as one copy herein termed Wt-Homo (qPCR value ≥ 0.7), events with one allele changed, which is no longer detectable by the target site-specific qPCR, would be detected as half copy herein termed NHEJ-Hemi (qPCR value between 0.1 and 0.7), while events with both alleles changed would be detected as null herein termed NHEJ-Null (qPCR value ≤ 0.1). As shown in Table 36, both guideRNA/Cas endonuclease systems targeting the soy EPSPS-CR1 and EPSPS-CR2 sites can introduce efficient Double Strand Break (DSB) efficiency at their designed target sites. Both NHEJ-Hemi and NHEJ-Null were detected in the 93B86 genotype. NHEJ (Non-Homologous-End-Joining) mutations mediated by the guide RNA/Cas9 system at the specific Cas9 target sites were confirmed by PCR/topo cloning/sequencing.

Table 36. Target Site Double Strand Break Rate Mutations Induced by the Guide RNA/Cas9 system on soybean EPSPS1 gene. Numbers indicate no. of events (numbers in parentheses are %).

Project	Total event	Wt-Homo (%)	NHEJ-Hemi (%)	NHEJ-Null (%)
U6-13.1 EPSPS-CR1	168	63 (38%)	66 (39%)	39 (23%)
U6-13.1 EPSPS-CR2	111	50 (45%)	21 (19%)	40 (36%)

D. Detection of the TIPS mutation in the soybean EPSPS gene

In order to edit specific amino acids at the native EPSPS gene (such as those resulting in a TIPS modification), a polynucleotide modification template, such as RTW1013A or RTW1012A (Table 34), was co-delivered with the guideRNA/Cas9 expression cassettes into soybean cells.

The modification of the native EPSPS1 gene via guide RNA/Cas9 system mediated DNA homologous recombination was determined by specific PCR analysis. A specific PCR assay with primer pair WOL569 (SEQ ID NO: 486) and WOL876 (SEQ ID NO: 487) was used to detect perfect TIPS modification at the native EPSPS1 gene. A second primer pair WOL569 (SEQ ID NO: 486) and WOL570 (SEQ ID NO: 488) was used to amplify both TIPS modified EPSPS1 allele and WT (wild type)/NHEJ mutated allele. Topo cloning/sequencing was used to verify the sequences.

EXAMPLE 25

Intron replacement of soybean genes using the guideRNA/Cas endonuclease system

A. guideRNA/Cas9 endonuclease target site design.

Four guideRNA/ Cas9 endonuclease target sites were identified in the soybean EPSPS1 gene Glyma01g33660 (Table 37). Two of the target sites (soy EPSPS-CR1 and soy EPSPS-CR2) were identified to target the Exon2 of the soybean EPSPS gene as described in Example 24. Another two target sites (soy EPSPS-CR4 and soy EPSPS-CR5) were designed near the 5' end of the intron1 of the soybean EPSPS gene.

Table 37. Guide RNA/Cas9 endonuclease target sites on soybean EPSPS1 gene.

Name of gRNA-Cas9 endonuclease target site	Cas endonuclease target sequence (SEQ ID NO:)	Physical location
soy EPSPS-CR1	467	Gm01:45865337..45865315
soy EPSPS-CR2	468	Gm01: 45865311..45865333
soy EPSPS-CR4	490	Gm01: 45866302..45866280
soy EPSPS-CR5	491	Gm01:45866295..45866274

B. Guide RNA/Cas9 endonuclease expression cassettes and polynucleotide modification templates used in soybean stable transformation for the replacement of the Intron1 of the soybean EPSPS1 gene with the soybean ubiquitin (UBQ) intron1

The soybean U6 small nuclear RNA promoter GM-U6-13.1 (SEQ ID. NO: 469) was used to express two guide RNAs (soy-EPSPS-CR1 and soy-EPSPS-CR4, or soy-EPSPS-CR1 and soy-EPSPS-CR5) to direct Cas9 endonuclease to designated genomic target sites (Table 38). One of the target sites (soy-EPSPS-CR1) was located in the exon2, as described in Example 24, and a second target site (soy-EPSPS-CR4 or soy-EPSPS-CR5) was located near the 5' end of intron1 of the native EPSPS1 gene. A soybean codon optimized Cas9 endonuclease expression cassette and a guide RNA expression cassette were linked in the expression plasmids QC878/RTW1199 (SEQ ID NO:470/492) or QC878/RTW1200 (SEQ ID NO:470/493) that was co-delivered with a polynucleotide modification template. The polynucleotide modification template, RTW1190A (SEQ ID NO:494), contained 532bp intron1 of the soybean UBQ gene and the TIPS modified Exon2. Soybean EPSPS1 intron 1 replacement with the soybean UBQ intron1 can be achieved with the guide RNA/Cas system by homologous recombination between the genomic DNA and the polynucleotide modification template, resulting in enhancement of the native or modified soy EPSPS1 gene expression.

Table 38. Guide RNA/Cas9 endonuclease expression cassettes and polynucleotide modification templates used in soybean stable transformation for the replacement of the Intron1 of the soybean EPSPS1 gene with the soybean ubiquitin (UBQ) intron1

Experiment	Guide RNA/Cas9	SEQ ID NO:	polynucleotide modification template	SEQ ID NO:
soy EPSPS-CR1 and soy EPSPS-CR4	U6-13.1:EPSPS CR1 + CR4+ EF1A2:CAS9 (QC878/RTW1199)	470/492	RTW1190A	494
soy EPSPS-CR1 and soy EPSPS-CR5	U6-13.1:EPSPS CR1 + CR5+ EF1A2:CAS9 (QC878/RTW1200)	470/493	RTW1190A	494

C. Detection of site-specific NHEJ mediated by the guide RNA/Cas9 system in stably transformed soybean

Site-specific NHEJ was detected as described in Example 24 C, using the qPCR primers/probes listed in Table 39.

5 Table 39. Primers/Probes used in qPCR analyses of transgenic soybean events.

Target Site	Primer/Probe Name	Sequences	SEQ ID NOs:
EPSPS-CR1 & EPSPS-CR2	Soy1-F1	CCACTAGTAAGGAATCTAAAGATGA AATCA	474
	Soy1-R2	CCTGCAGCAACCACAGCTGCTGTC	475
	Soy1-T1 (FAM-MGB)	CTGCAATGCGTCCTT	476
EPSPS-CR4	Soy1-F3	GTTTGTGTTGTTGTTGGGTGTGGG	495
	Soy1-R3	GACATGATGCTTCATTTTCACAGAA	496
	Soy-T2 (FAM-MGB)	TGTGTAGAGTGGATTTTG	497
EPSPS-CR5	Soy1-F2	TGTTGTTGGGTGTGGGAATAGG	498
	Soy1-R3	GACATGATGCTTCATTTTCACAGAA	496
	Soy1-T2 (FAM-MGB)	TGTGTAGAGTGGATTTTG	497
gRNA/CAS9	Cas9-F	CCTTCTTCCACCGCCTTGA	477
	Cas9-R	TGGGTGTCTCTCGTGCTTTTT	478
	Cas9-T (FAM-MGB)	AATCATTCCTGGTGGAGGA	479
pINII	pINII-99F	TGATGCCACATTATAGTGATTAGC	480
	pINII-13R	CATCTTCTGGATTGGCCAACTT	481
	pINII-69T (FAM-MGB)	ACTATGTGTGCATCCTT	482
SIP	SIP-130F	TTCAAGTTGGGCTTTTTCAGAAG	483
	SIP-198R	TCTCCTTGGTGCTCTCATCACA	484
	SIP-170T (VIC-MGB)	CTGCAGCAGAACCAA	485

D. Detection of the replacement of the soybean EPSPS1 intron1 with the soybean UBQ intron1 using the guide RNA/Cas9 endonuclease system.

In order to replace the soybean EPSPS1 intron1 with the soybean UBQ intron1 at the native EPSPS1 gene, two guideRNA expression vectors were used as shown in Table 38. The QC878 vector (SEQ ID NO: 470) was targeting the exon2 and the RTW1199 (SEQ ID NO:492) or RTW1200 (SEQ ID NO:493) was targeting the 5' end of the intron1. The double cleavage of soybean EPSPS gene with the two guide RNA/Cas systems resulted in the removal of the native EPSPS1 intron1/partial Exon2 fragment. At the same time, a polynucleotide modification template RTW1190A (SEQ ID NO:494) was co-delivered into soybean cells and homologous recombination between the polynucleotide modification template and the genomic DNA resulted in the replacement of EPSPS1 intron1 with the soybean UBQ intron1 and the desired amino acid modifications in exon2 as evidenced by PCR analysis. PCR assays with primer WOL1001/WOL1002 pair (SEQ ID NO: 499 and 500) and WOL1003/WOL1004 pair (SEQ ID NO: 501 and 502) were used to detect the intron replacement events.

EXAMPLE 26

Promoter replacement (promoter swap) of soybean genes using the guideRNA/Cas endonuclease system

A. guideRNA/Cas9 endonuclease target site design.

Four guideRNA/ Cas9 endonuclease target sites were identified in the soybean EPSPS1 gene Glyma01g33660 (Table 40). Two of the target sites (soy EPSPS-CR1 and soy EPSPS-CR2) were identified to target the Exon2 of the soybean EPSPS gene as described in Example 24. The soy EPSPS-CR6 and soy EPSPS-CR7 were identified near the 5' end of the -798 bp of the native EPSPS promoter.

Table 40. Guide RNA/Cas9 endonuclease target sites on soybean EPSPS1 gene.

Name of gRNA-Cas9 endonuclease target site	Cas endonuclease target sequence (SEQ ID NO:)	Physical location
soy EPSPS-CR1	467	Gm01:45865337..45865315
soy EPSPS-CR2	468	Gm01: 45865311..45865333

soy EPSPS-CR6	503	Gm01:45867471..45867493
soy EPSPS-CR7	504	Gm01:45867459..45867481

B. Guide RNA / Cas9 endonuclease expression cassettes and polynucleotide modification templates used in soybean stable transformation for the replacement of the -798bp soybean EPSPS1 promoter with the soybean UBQ promoter.

The soybean U6 small nuclear RNA promoter GM-U6-13.1 (SEQ ID. NO: 469) was used to express two guide RNAs (soyEPSPS-CR1 and soyEPSPS-CR6, or soyEPSPS-CR1 and soyEPSPS-CR7) to direct Cas9 nuclease to designated genomic target sites (Table 41). One of the target sites (soy-EPSPS-CR1) was located in the exon2 as described in Example 24 and a second target site (soy-EPSPS-CR6 or soy-EPSPS-CR7) was located near 5' end of the -798bp of the native EPSPS1 promoter. A soybean codon optimized Cas9 endonuclease expression cassette and a guide RNA expression cassette were linked in the expression plasmids QC878/RTW1201 (SEQ ID NO:470/505) or QC878/RTW1202 (SEQ ID NO:470/506) that was co-delivered with a polynucleotide modification template, RTW1192A (SEQ ID NO:507). The polynucleotide modification template contained 1369 bp of the soybean UBQ gene promoter, 47bp 5UTR and 532bp UBQ intron1. Specific soybean EPSPS1 promoter replacement with the soybean UBQ promoter can be achieved with the guide RNA/Cas system by homologous recombination between the genomic DNA and the polynucleotide modification template, resulting enhancement of the native or modified soy EPSPS1 gene expression

Table 41. Guide RNA / Cas9 endonuclease expression cassettes and polynucleotide modification templates used in soybean stable transformation for the replacement of the -798bp soybean EPSPS1 promoter with the soybean UBQ promoter

Experiment	Guide RNA/Cas9	SEQ ID NO:	polynucleotide modification template	SEQ ID NO:
soy EPSPS-CR1	U6-13.1:EPSPS CR1 +	470,	RTW1192A	507

and soy EPSPS-CR6	CR6+ EF1A2:CAS9 (QC878/RTW1201)	505		
soy EPSPS-CR1 and soy EPSPS-CR7	U6-13.1:EPSPS CR1 + CR7+ EF1A2:CAS9 (QC878/RTW1202)	470, 506	RTW1192A	507

C. Detection of site-specific NHEJ mediated by the guide RNA/Cas9 system in stably transformed soybean

Site-specific NHEJ was detected as described in Example 24 C, using the
5 qPCR primers/probes listed in Table 42.

Table 42. Primers/Probes used in qPCR analyses of transgenic soybean events

Target Site	Primer/Probe Name	Sequences	SEQ ID NOs:
EPSPS-CR1 & EPSPS-CR12	Soy1-F1	CCACTAGTAAGGAATCTAAAGATGAA ATCA	474
	Soy1-R2	CCTGCAGCAACCACAGCTGCTGTC	475
	Soy1-T1 (FAM-MGB)	CTGCAATGCGTCCTT	476
EPSPS-CR6 & EPSPS-CR7	Soy1-F4	TCAATAATACTACTCTCTTAGACACCA AACAA	508
	Soy1-R4	CAAGGAAAATGAATGATGGCTTT	509
	Soy1-T3 (FAM-MGB)	CCTTCCCAAACCTATAATC	510
gRNA/CAS9	Cas9-F	CCTTCTTCCACCGCCTTGA	477
	Cas9-R	TGGGTGTCTCTCGTGCTTTTT	478
	Cas9-T (FAM-MGB)	AATCATTCCTGGTGGAGGA	479
pNII	pNII-99F	TGATGCCCACATTATAGTGATTAGC	480
	pNII-13R	CATCTTCTGGATTGGCCAACTT	481
	pNII-69T (FAM-MGB)	ACTATGTGTGCATCCTT	482
SIP	SIP-130F	TTCAAGTTGGGCTTTTTCAGAAG	483

	SIP-198R	TCTCCTTGGTGCTCTCATCACA	484
	SIP-170T (VIC-MGB)	CTGCAGCAGAACCAA	485

D. Detection of the promoter replacement of the soybean EPSPS1 promoter with the soybean UBQ promoter using the guide RNA/Cas9 endonuclease system.

In order to replace the soybean EPSPS1 promoter with the soybean UBQ promoter at the native EPSPS1 gene, two guideRNA expression vectors were used in each soybean transformation experiment as shown in Table 41. The QC878 (SEQ ID NO: 470) was targeting the exon2 and the RTW1201(SEQ ID NO: 505) or RTW1202 (SEQ ID NO: 506) was targeting the 5' end of the soybean -798bp promoter. The double cleavage of the soybean EPSPS1 gene with the two guide RNA/Cas systems resulted in removal of the native EPSPS1 promoter/5'UTR-Exon1/Intron1/partial Exon2 fragment at the native EPSPS gene. At the same time, a polynucleotide modification template RTW1192A (SEQ ID NO: 507) was co-delivered into soybean cells. This RTW1192A DNA contained 1369 bp soybean UBQ promoter, its 47bp 5-UTR and 532bp UBQ intron1 in front of the EPSPS1 exon1-Intron1-modified Exon2. Homologous recombination between the polynucleotide modification template and the genomic DNA resulted in the replacement of EPSPS1 promoter/5'UTR with the soybean UBQ promoter/5'UTR/Intron1 and the desired amino acid modifications evidenced by PCR analysis. PCR assays with primer WOL1005/WOL1006 pair (SEQ ID NO: 511 and 512) and WOL1003/WOL1004 pair (SEQ ID NO: 501 and 502) were used to detect the promoter replacement events.

EXAMPLE 27

Enhancer element deletions using the guideRNA/Cas endonuclease system

The guide RNA/Cas endonuclease system described herein can be used to allow for the deletion of a promoter element from either a transgenic (pre-existing, artificial) or endogenous gene. Promoter elements, such enhancer elements, or often introduced in promoters driving gene expression cassettes in multiple copies (3X= 3 copies of enhancer element, Figure 33) for trait gene testing or to produce transgenic plants expressing specific trait. Enhancer elements can be, but are not limited to, a 35S enhancer element (Benfey et al, EMBO J, Aug 1989; 8(8): 2195-

2202, SEQ ID NO:513). In some plants (events), the enhancer elements can cause an unwanted phenotype, a yield drag, or a change in expression pattern of the trait of interest that is not desired. For example, as shown in Figure 33, a plant comprising multiple enhancer elements (3 copies, 3X) in its genomic DNA located between two trait cassettes (Trait A en Trait B) was characterized to show an unwanted phenotype. It is desired to remove the extra copies of the enhancer element while keeping the trait gene cassettes intact at their integrated genomic location. The guide RNA/Cas endonuclease system described herein can be used to removing the unwanted enhancing element from the plant genome. A guide RNA can be designed to contain a variable targeting region targeting a target site sequence of 12-30 bps adjacent to a NGG (PAM) in the enhancer. If a Cas endonuclease target site sequence is present in all copies of the enhancer elements (such as the three Cas endonuclease target sites 35S-CRTS1 (SEQ ID NO:514), 35S-CRTS2 (SEQ ID NO:515), 35S-CRTS3 (SEQ ID NO:516)), only one guide RNA is needed to guide the Cas endonuclease to the target sites and induce a double strand break in all the enhancer elements at once. The Cas endonuclease can make cleavage to remove one or multiple enhancers. The guideRNA/Cas endonuclease system can introduced by either agrobacterium or particle gun bombardment. Alternatively, two different guide RNAs (targeting two different genomic target sites) can be used to remove all 3X enhancer elements from the genome of an organism, in a manner similar to the removal of a (transgenic or endogenous) promoter described herein.

EXAMPLE 28

Regulatory sequence modifications using the guide RNA/Cas endonuclease system

A. Modification of polyubiquitination sites

There are defined ubiquitination sites on proteins to be degraded and they were found within the maize EPSPS protein by using dedicated computer programs (for example, the CKSAAP_UbSite (Ziding Zhang's Laboratory of Protein Bioinformatics College of Biological Sciences, China Agricultural University, 100193 Beijing, China). One of the selected polyubiquitination site within the maize EPSPS coding sequence is shown in Figure 34A and its amino acid signature sequence is

compared to the equivalent EPSPS sites from the other plants (Figure 34A). The lysine amino acid (K) at position 90 (highly conserved in other plant species) was selected as a potential site of the EPSPS protein polyubiquitination. The polynucleotide modification template (referred to as EPSPS polynucleotide maize K90R template) used to edit the *epsps* locus is listed as SEQ ID NO: 517. This template allowed for editing the *epsps* locus to contain the lysine (K) to arginine (R) substitution at position 90 (K90R) and two additional TIPS substitutions at positions 102 and 106 (Figure 34B and 34C). Maize genomic DNA was edited using the guideRNA/Cas endonuclease system described herein and T0 plants were produced as described herein. The T0 plants that contained the nucleotide modifications, as specified by the information provided on the K90R template (Figure 34C), were selected by the genotyping methods described herein. F1 EPSPS-K90R plants can be selected for elevated protein content due to a slower rate of the EPSPS protein degradation.

15

B. Editing intron elements to introduce intron mediated enhancer elements (IMEs)

Transcriptional activity of the native EPSPS gene can be modulated by transcriptional enhancers positioned in the vicinity of other transcription controlling elements. Introns are known to contain enhancer elements affecting the overall rate of transcription from native promoters including the EPSPS promoter. For example, the first intron of the maize ubiquitin 5'UTR confers a high level of expression in monocot plants as specified in the WO 2011/156535 A1 patent application. An intron enhancing motif CATATCTG (Figure 35 A) , also referred to as a intron-mediated enhancer element, (IME) was identified by proprietary analysis (WO2011/156535 A1, published on December 15, 2011) and appropriate nucleotide sites at the 5' end of the EPSPS first intron were selected for editing in order to introduce the intron-mediated enhancer elements (IMEs) (Figure 35B-35C). The polynucleotide modification template (referred to as EPSPS polynucleotide maize IME template) is listed as SEQ ID No: 518. The polynucleotide modification template allows for editing of the *epsps* locus to contain three IMEs (two on one strand of the DNA, one on the reverse strand) in the first EPSPS intron and the TIPS substitutions at positions 102 and 106. The genomic DNA of maize plants was

30

edited using the guideRNA/Cas endonuclease system described herein. Maize plants containing the IME edited EPSPS coding sequence can be selected by genotyping the T0 plants and can be further evaluated for elevated EPSPS-TIPS protein content due to the enhanced transcription rate of the native *EPSPS* gene.

5

EXAMPLE 29

Modifications of splicing sites and/or introducing alternate splicing sites using the guide RNA/ Cas endonuclease system

In maize cells, the splicing process is affected by splicing sites at the exon-intron junction sites as illustrated in the EPSPS mRNA production (Figure 36A-36B). Fig. 36A shows analysis of EPSPS amplified pre-mRNA (cDNA panel on left). Lane 14 in Fig. 36A shows amplification of the EPSPS pre-mRNA containing the 3rd intron unspliced, resulting in a 804bp diagnostic fragment indicative for an alternate splicing event. Lanes E3 and F8 show the EPSPS PCR amplified fragments resulting from regular spliced introns. Diagnostic fragments such as the 804 bp fragment of lane 14 are not amplified unless cDNA is synthesized (as is evident by the absence of bands in lanes E3, 14, and F8 comprising total RNA (shown in the total RNA panel on right of Fig. 36A). The canonical splice site in the maize *EPSPS* gene and genes from other species is AGGT, while other (alternative) variants of the splice sites may lead to the aberrant processing of pre-mRNA molecules. The EPSPS coding sequence contains a number of alternate splicing sites that may affect the overall efficiency of the pre-mRNA maturation process and as such may limit the EPSPS protein accumulation in maize cells.

In order to limit the occurrence of alternate splicing events during *EPSPS* gene expression, a guideRNA/Cas endonuclease system as described herein can be used to edit splicing sites. The splicing site at the junction of the second native *EPSPS* intron and the third exon is AGTT and can be edited in order to introduce the canonical AGGT splice site at this junction (Figure 37). The T>G substitution does not affect the native EPSPS open reading frame and it does not change the EPSPS amino acid sequence. The polynucleotide modification template (referred to as EPSPS polynucleotide maize Tspliced template) is listed as SEQ ID NO: 519. This polynucleotide modification template allows for editing of the *epsps* locus to contain the canonical AGGT splice site at the 2nd intron-3rd exon junction site and

the TIPS substitutions at positions 102 and 106. Maize plants are edited using the procedures described herein. F1 EPSPS-Tspliced maize plants can be evaluated for increased protein content due to the enhanced production of functional EPSPS mRNA messages.

5

EXAMPLE 30

Shortening Maturity via Manipulation of Early Flowering Phenotype with ZmRap2.7

Down-Regulation using the guide RNA/Cas endonuclease system

Overall plant maturity can be shortened by modulating the flowering time phenotype of plants through modulation of a maize ZmRap2.7 gene. Shortening of plant maturity can be obtained by an early flowering phenotype.

RAP2.7 is an acronym for Related to APETALA 2.7. RAPL means RAP2.7 LIKE and RAP2.7 functions as an AP2-family transcription factor that suppresses floral transition (SEQ ID NOs:520 and 521). Transgenic phenotype upon silencing or knock-down of Rap2.7 resulted in early flowering, reduced plant height, but surprisingly developed normal ear and tassel as compared the wild-type plants (PCT/US14/26279 application, filed March 13, 2014). The guide RNA/Cas endonuclease system described herein can be used to target and induce a double strand break at a Cas endonuclease target site located within the RAP2.7 gene. Plants comprising NHEJ within the RAP2.7 gene can be selected and evaluated for the presence of a shortened maturity phenotype.

20

EXAMPLE 31

Modulating expression of a maize NPK1B gene for Engineering Frost Tolerance in

Maize Using a guide RNA/Cas endonuclease system

Nicotiana Protein Kinase1 (NPK1) is a mitogen activated protein kinase kinase kinase that is involved in cytokinesis regulation and oxidative stress signal transduction. The ZM-NPK1B (SEQ ID NO: 522 and SEQ ID NO: 523) which has about 70% amino acid similarity to rice NPKL3 has been tested for frost tolerance in maize seedlings and reproductive stages (PCT/US14/26279 application, filed March 13, 2014) . Transgenic seedlings and plants comprising a ZM-NPK1B driven by an inducible promoter Rab17, had significantly higher frost tolerance than control seedlings and control plants. The gene seemed induced after cold acclimation and

30

during -3°C treatment period in most of the events but at low levels.
(PCT/US14/26279 application, filed March 13, 2014).

A guide RNA/Cas endonuclease system described herein can be used to replace the endogenous promoter of NPK1 gene, with a stress-inducible promoter
5 such as the maize RAB17 promoter stages (SEQ ID NO: 524; PCT/US14/26279 application, filed March 13, 2014), thus modulate NPK1B expression in a stress-responsive manner and provide frost tolerance to the modulated maize plants.

EXAMPLE 32

Shortening Maturity via Manipulation of Early Flowering Phenotype with 10 FTM1 Expression using a guide RNA/ Cas endonuclease systems

Overall plant maturity can be shortened by modulating the flowering time phenotype of plants through expressing a transgene. Such a phenotype modification can also be achieved with additional transgenes or through a breeding approach.

FTM1 stands for Floral Transition MADS 1 transcription factor (SEQ ID NOs:
15 525 and 526). It is a MADS Box transcriptional factor and induces floral transition. Upon expression of FTM1 under a constitutive promoter, transgenic plants exhibited early flowering and shortened maturity, but surprisingly ear and tassel developed normally as compared to the wild-type plants (PCT/US14/26279 application, filed March 13, 2014).

20 FTM1-expressing maize plants demonstrated that by manipulating a floral transition gene, time to flowering can be reduced significantly, leading to a shortened maturity for the plant. As maturity can be generally described as time from seeding to harvest, a shorter maturity is desired for ensuring that a crop can finish in the northern continental dry climatic environment (PCT/US14/26279
25 application, filed March 13, 2014).

A guide RNA/ Cas endonuclease system described herein can be used to introduce enhancer elements such as the CaMV35S enhancers (Benfey et al, EMBO J, Aug 1989; 8(8): 2195-2202, SEQ ID NO:512), specifically targeted in front of the endogenous promoter of FTM1, in order to enhance the expression of FTM1
30 while preserving most of the tissue and temporal specificities of native expression, providing shortened maturity to the modulated plants.

EXAMPLE 33**Inserting inducible responsive elements in plant genomes**

Inducible expression systems controlled by an external stimulus are desirable for functional analysis of cellular proteins as well as trait development as changes in the expression level of the gene of interest can lead to an accompanying phenotype modification. Ideally such a system would not only mediate an "on/off" status for gene expression but would also permit limited expression of a gene at a defined level.

The guide RNA/Cas endonuclease system described herein can be used to introduce components of repressor/operator/inducer systems to regulate gene expression of an organism. Repressor/operator/inducer systems and their components are well known in the art (US 2003/0186281 published Oct. 2, 2003; US 6,271,348). For example, not limited to, components of the tetracycline (Tc) resistance system of E. coli have been found to function in eukaryotic cells and have been used to regulate gene expression (US 6,271,348). Nucleotide sequences of tet operators of different classes are known in the art see for example: classA, classB, classC, classD, classE TET operator sequences listed as SEQ ID NOs:11-15 of US 6,271,348.

Components of a sulfonylurea-responsive repressor system (as described in US 8,257,956, issued on September 4, 2012) can also be introduced into plant genomes to generate a repressor/operator/inducer systems into said plant where polypeptides can specifically bind to an operator, wherein the specific binding is regulated by a sulfonylurea compound.

EXAMPLE 34**Genome deletion for Trait Locus Characterization**

Trait mapping in plant breeding often results in the detection of chromosomal regions housing one or more genes controlling expression of a trait of interest. For quantitative traits, expression of a trait of interest is governed by multiple quantitative trait loci (QTL) of varying effect-size, complexity, and statistical significance across one or more chromosomes. A QTL or haplotype that is associated with suppression of kernel-row number in the maize ear can be found to be endemic in elite breeding germplasm. The negative effect of this QTL for kernel

row number can be fine-mapped to an acceptable resolution to desire selective elimination of this negative QTL segment within specific recipient germplasm. Two flanking cut sites for the guide polynucleotide/ Cas endonuclease system are designed via haplotype, marker, and/or DNA sequence context at the targeted QTL region, and the two guide polynucleotide/Cas endonuclease systems are deployed simultaneously or sequentially to produce the desired end product of two independent double strand breaks (cuts) that liberate the intervening region from the chromosome. Individuals harboring the desired deletion event would result by the NHEJ repair of the two chromosomal ends and eliminating the intervening DNA region. Assays to identify these individuals is based on the presence of flanking DNA marker regions, but absence of intervening DNA markers. A proprietary haplotype for kernel-row-number is created that is not extant in the previously defined elite breeding germplasm pool.

An alternative approach would be to delete a region containing a fluorescent gene. Recovery of plants with, and without, fluorescence would give an approximate indication of the efficiency of the deletion process.

EXAMPLE 35

Engineering drought tolerance and nitrogen use efficiency into maize via gene silencing by expressing an inverted repeat into an ACS6 gene using the guide

RNA/Cas endonuclease system

ACC (1-aminocyclopropane-1-carboxylic acid) synthase (ACS) genes encode enzymes that catalyze the rate limiting step in ethylene biosynthesis. A construct containing one of the maize ACS genes, ZM-ACS6, in an inverted repeat configuration, has been extensively tested for improved abiotic stress tolerance in maize (PCT/US2010/051358, filed October 4, 2010; PCT/US2010/031008, filed April 14, 2010). Multiple transgenic maize events containing a ZM-ACS6 RNAi sequence driven by a ubiquitin constitutive promoter had reduced ethylene emission, and a concomitant increase in grain yield relative to controls under both drought and low nitrogen field conditions (Plant Biotechnology Journal: 12 MAR 2014, DOI: 10.1111/pbi.12172).

In an embodiment the insertion of the inverted gene fragment can result in the formation of an in-vivo created inverted repeat (hairpin) in a native (or modified)

promoter of an ACS6 gene and/or in a native 5' end of the native ACS6 gene. The inverted gene fragment can further comprise an intron which can result in an enhanced silencing of the targeted ethylene biosynthetic gene.

EXAMPLE 36

- 5 T0 plants from the multiplexed guide RNA/Cas experiment carried high frequency of
bi-allelic mutations and demonstrated proper inheritance of
mutagenized alleles in the T1 population.

This example demonstrates the high efficiency of the guide RNA/Cas endonuclease system in generating maize plants with multiple mutagenized loci and
 10 their inheritance in the consecutive generation(s).

Mutated events generated in the multiplexed experiment described in Example 4 were used to regenerate T0 plants with mutations at 3 different target sites: MS26Cas-2 target site (SEQ ID NO: 14), LIGCas-3 target site (SEQ ID NO: 18) and MS45Cas-2 target site (SEQ ID NO: 20).

- 15 For further analysis, total genomic DNA was extracted from leaf tissue of individual T0 plants. Fragments spanning all 3 target sites were PCR amplified using primer pairs for the corresponding target sites, cloned into the pCR2.1-TOPO cloning vector (Invitrogen), and sequenced. Table 43 shows examples of mutations detected in four T0 plants resulting from imprecise NHEJ at all relevant loci when
 20 multiple guide RNA expression cassettes were simultaneously introduced either in duplex (see TS = Lig34/MS26) or triplex (see TS = Lig34/MS26/MS45), respectively.

Table 43. Examples of mutations at maize target loci produced by a multiplexed guide RNA/Cas system

Target sites (TS)	T0 plant	qPCR data	Sequencing data		
			Lig3/4 TS	Ms26 TS	Ms45 TS
Lig34 / MS26	1	NULL / NULL*	1bp ins / 2bp del+1 bp ins	1bp ins / 19bp del	
	2	NULL / NULL	1bp ins / 1bp del	1bp ins / 1bp ins	
Lig34 / MS26 / MS45	1	NULL / NULL / NULL	1bp ins / large del	1bp ins / 1bp del	15bp del / large del
	2	INDEL** / NULL	1bp ins / WT	1bp (T) ins /	1bp ins / large

		/ NULL		1bp (C) ins	del
--	--	--------	--	-------------	-----

* NULL indicates that both alleles are mutated ** INDEL indicates mutation in one of the two alleles. del = deletion, ins = insertion, bp = base pair

All T0 plants were crossed with wild type maize plants to produce T1 seeds.

5 T1 progeny plants (32 plants) of the second T0 plant from the triplex experiment (see Table 43, Lig34 / MS26 / MS45) were analyzed by sequencing to evaluate segregation frequencies of the mutated alleles. Our results demonstrated proper inheritance and expected (1:1) segregation of the mutated alleles as well as between mutated and wild type alleles at all three target sites.

10 The data clearly demonstrate that the guide RNA/maize optimized Cas endonuclease system described herein, can be used to simultaneously mutagenize multiple chromosomal loci and produce progeny plants containing the stably inherited multiple gene knock-outs.

EXAMPLE 37

15 Guide RNA/Cas endonuclease mediated DNA cleavage in maize chromosomal loci can stimulate homologous recombination repair-mediated transgene insertion and resulting T1 progeny plants demonstrated proper inheritance of the modified alleles.

Maize events generated in the experiment described in Example 5 were used to regenerate T0 plants. T0 plants were regenerated from 7 independent callus
20 events with correct amplifications across both transgene genomic DNA junctions and analyzed. Leaf tissue was sampled, total genomic DNA extracted, and PCR amplification at both transgene genomic DNA junctions was carried out using the primer pairs (corresponding to SEQ ID NOs: 98-101). The resulting amplification products were sequenced for confirmation. Plants with confirmed junctions at both
25 ends were further analyzed by Southern hybridization (Figure 38) using two probes, genomic (outside HR1 region, SEQ ID: 533) and transgenic (within MoPAT gene, SEQ ID: 534). PCR, sequencing and Southern hybridization data demonstrated that plants regenerated from two of the 7 events (events 1 and 2) demonstrated perfect, clean, single copy transgene integration at the expected target site via homologous
30 recombination. Plants regenerated from the remaining 5 events contained either additional, randomly integrated copies of the transgene (events 4, 5, and 6) or rearranged copies of the transgene integrated into the target site (events 3 and 7).

T0 plants from events 1 and 2 were crossed with wild type maize plants to produce T1 seeds. Ninety-six T1 plants from events 1 and 2 were analyzed by Southern hybridization (using the same probes as above) to evaluate segregation frequencies of the transgene locus. Southern results demonstrated proper inheritance and expected (1:1) segregation of the transgene and wild type loci.

The data clearly demonstrate that maize chromosomal loci cleaved with the maize optimized guide RNA/Cas system described herein can be used to stimulate HR repair pathways to site-specifically insert transgenes and produce progeny plants that have the inserted transgene stably inherited.

EXAMPLE 38

Production of maize transgenic lines with pre-integrated Cas9 for transient delivery of guide RNA

This example describes the rationale, production, and testing of maize transgenic lines with an integrated Cas9 gene under constitutive and temperature inducible promoters.

As demonstrated in Example 2, a high mutation frequency was observed when Cas9 endonuclease and guide RNA were delivered as DNA vectors by biolistic transformation to immature corn embryo cells. When Cas9 endonuclease was delivered as a DNA vector and guide RNA as RNA molecules, a reduced mutation frequency was observed (Table 44).

Table 44. Mutant reads at LigCas-3 target site produced by transiently delivered guide RNA.

Target Site Examined for Mutations	Transient Delivery	Expression Cassette	Mutant Reads	Total Reads
LIGCas-3	-	Cas9	24.2	1,599,492
LIGCas-3	-	Cas9/guide RNA	44170	1,674,825
LIGCas-3	35ng guide RNA	Cas9	418	1,622,180
LIGCas-3	70ng guide RNA	Cas9	667	1,791,388
LIGCas-3	140ng guide RNA	Cas9	239	1,632,137

Increased efficiency (increased mutant reads) may occur when the Cas9 protein and guide RNA are present in the cell at the same time. To facilitate the presence of both Cas9 endonuclease and guide RNA in the same cell, a vector containing a constitutive and conditionally regulated Cas9 gene can be first delivered to plant cells to allow for stable integration into the plant genome to establish a plant line that contains only the Cas9 gene in the plant genome. Then, single or multiple guide RNAs can be delivered as either DNA or RNA, or combination, to the embryo cells of the plant line containing the genome-integrated version of the Cas9 gene.

Transgenic maize (genotype Hi-II) lines with an integrated Cas9 gene driven by either a constitutive (Ubi) or an inducible (CAS) promoter were generated via *Agrobacterium*-mediated transformation. Besides the Cas9 gene, the Agro vector also contained a visible marker (END2:Cyan) and a Red Fluorescent Protein sequence interrupted with a 318 bp long linker (H2B:RF-FP). The linker sequence was flanked with 370 bp long direct repeats to promote recombination and restoration of a functional RFP gene sequence upon double strand break within the linker.

Lines with single copies of the transgene were identified and used for further experiments. Two guide RNA constructs targeting 2 different sites (Table 45 in the linker sequence, were delivered into immature embryo cells via particle bombardment. Meganuclease variant LIG3-4 B65 with very high cutting activity previously used in similar experiments was used as the positive control.

Table 45. Target sites in the RF-FP linker for guideRNA/Cas endonuclease system.

Locus	Guide RNA Used	Target Site Designation	Target Site Sequence	PAM Sequence	SEQ ID NO:
RF-FP linker	Long	RF-FPCas-1	GCAGGTCTCACGACGGT	TGG	535
	Long	RF-FPCas-2	GTAAAGTACGCGTACGTGTG	AGG	536

After transformation, embryos with Cas9 gene under Ubiquitin promoter were incubated at 28°C while embryos with Cas9 gene under temperature inducible CAS

promoter were first incubated at 37°C for 15-20 hours and then transferred to 28°C. Embryos were examined 3-5 days after bombardment under luminescent microscope. Expression and activity of the pre-integrated Cas9 protein was visually evaluated based on the number of embryo cells with RFP protein expression. In most lines, the guide RNA/Cas endonuclease system demonstrated similar or higher frequency of RFP repair than LIG3-4 B65 meganuclease indicating high level of Cas9 protein expression and activity in the generated transgenic lines.

This example describes the production of transgenic lines with a pre-integrated Cas9 gene that can be used in further experiments to evaluate efficiency of mutagenesis at a target site upon transient delivery of guide RNA in the form of RNA molecules.

EXAMPLE 39

The guide RNA/Cas endonuclease system delivers double-strand breaks to the maize ALS locus and facilitates editing of the ALS gene

This example demonstrates that the guide RNA/Cas endonuclease system can be efficiently used to introduce specific changes into the nucleotide sequence of the maize ALS gene resulting in resistance to sulfonylurea class herbicides, specifically, chlorsulfuron.

Endogenous ALS protein is the target site of ALS inhibitor sulfonylurea class herbicides. Expression of the herbicide tolerant version of ALS protein in crops confers tolerance to this class of herbicides. The ALS protein contains N-terminal transit peptides, and the mature protein is formed following transport into the chloroplast and subsequent cleavage of the transit peptide. The mature protein starts at residue S41, resulting in a mature protein of 598 amino acids with a predicted molecular weight of 65 kDa (SEQ ID NO: 550).

Table 46: Deduced Amino Acid Sequence of the Full-Length ZM-ALS Protein (SEQ ID no: 550)

1	MATAAAASTA	LTGATTAAPK	ARRRAHLLAT	RRALAAPIRC	SAASPAMPMA
51	PPATPLRPWG	PTPRKGADI	LVESLERCGV	RDVFAYPGGA	SMEIHQALTR
101	SPVIANHLFR	HEQGEAFAAS	GYARSSGRVG	VCIATSGPGA	TNLVSALADA
151	LLDSVPMVAI	TGQVPRRMIG	TDAFQETPIV	EVTRSITKHN	YLVLDVDDIP
201	RVVQEAFFLA	SSGRPGPVLV	DIPKDIQQQM	AVPVWDKPMs	LPGYIARLPK
251	PPATELLEQV	LRLVGESRRP	VLYVGGGCAA	SGEELRRFVE	LTGIPVTTTL
301	MGLGNFPSDD	PLSLRMLGMH	GTVYANYAVD	KADLLLALGV	RFDDRVTGKI

351 EAFASRAKIV HVDIDPAEIG KKNQPHVSIC ADVKLALQGM NALLEGSTSK
 401 KSFDGFSWND ELDQQKREFP LGYKTSNEEI QPQYAIQVLD ELTKGEAIIG
 451 TGVGQHQMWA AQYYTYKRPR QWLSSAGLGA MGFGLPAAAG ASVANPGVTV
 501 VDIGDGSFL MNVQELAMIR IENLPVKVFV LNNQHLGMVV QWEDRFYKAN
 5 551 RAHTYLGNP NESEIYPDFV TIAKGFNIPA VRVTKKNEVR AAIKKMLETF
 601 GPYLLDIIVP HQEHVLPMP SGGAFKDMIL DGDGRTVY

Modification of a single amino acid residue (P165A or P165S, shown in bold) from the endogenous maize acetoacetate synthase protein provides resistance to herbicides in maize.

There are two ALS genes in maize, ALS1 and ALS2, located on chromosomes 5 and 4, respectively. As described in Example 2, guide RNA expressing constructs for 3 different target sites within the ALS genes were tested. Based on polymorphism between ALS1 and ALS2 nucleotide sequences, ALS1-specific and ALSCas-4 target site were identified and tested. ALSCas-1 guide RNA expressing construct targeting both ALS1 and ALS2 genes was used as control (Table 47)

Table 47. Maize ALS genomic target sites tested.

Locu s	Location	Guide RNA	Target Site Designation	Maize Genomic Target Site Sequence	PAM Sequence	SEQ ID NO:
ALS	Chr. 4: 107.73cM and Chr. 5: 115.49cM	Long	ALSCas-1	GGTGCCAATCATGCGTCG	CGG	22
		Long	ALSCas-4	GCTGCTCGATTCCGTC A	TGG*	537

* Target site in the ALS1 gene; bolded nucleotides are different in the ALS2 gene.

The experiment was conducted and mutation frequency determined as described in Example 2 and results are shown in Table 48.

Table 48. Frequencies of NHEJ mutations at the two ALS target sites recovered by deep sequencing.

TS	Total Reads	Mutant reads (ALS1)	Mutant reads (ALS2)
ALSCas-1	204,230	5072 (2.5%)	2704 (1.3%)
ALSCas-4	120,766	3294 (2.7%)	40 (0.03%)

The results demonstrated that ALSCas-4 guide RNA/Cas9 system mutates the ALS1 gene with approximately 90 times higher efficiency than the ALS2 gene.

Therefore, the ALSCas-4 target site and the corresponding guide RNA were selected for the ALS gene editing experiment.

To produce edited events, the ALS polynucleotide modification repair template was co-delivered using particle bombardment as a plasmid with an 804 bp long homologous region (SEQ ID NO: 538) or as a single-stranded 127 bp DNA fragment (SEQ ID NO: 539), the maize optimized Cas9 endonuclease expression vector described in Example 1, the guide RNA expression cassette (targeting ALSCas-4 site), a moPAT-DsRed fusion as selectable and visible markers, and developmental genes (ODP-2 and WUS). Approximately 1000 Hi-II immature embryos were bombarded with each of the two repair templates described above. Forty days after bombardment, 600 young callus events (300 for each repair template) were collected and transferred to the media with bialaphos selection. The embryos with remaining events were transferred to the media with 100 ppm of chlorsulfuron for selection. A month later, events that continued growing under chlorsulfuron selection were collected and used for analysis.

A small amount of callus tissue from each selected event was used for total DNA extraction. A pair of genomic primers outside the repair/donor DNA fragment (SEQ ID NO:540 and SEQ ID NO:541) was used to amplify an endogenous fragment of the ALS1 locus containing the ALSCas4 target sequence. The PCR amplification products were gel purified, cloned into the pCR2.1 TOPO cloning vector (Invitrogen) and sequenced. A total of 6 events demonstrated the presence of the specifically edited ALS1 allele as well as either a wild type or a mutagenized second allele.

These data indicate that a guide RNA/Cas system can be successfully used to create edited ALS allele in maize. The data further demonstrates that the guide RNA/maize optimized Cas endonuclease system described herein, can be used to produce progeny plants containing gene edits that are stably inherited.

EXAMPLE 40Gene editing of the soybean ALS1 gene and use as a transformation selectable marker for soybean transformation with the guide RNA/Cas endonuclease system*A. guideRNA/Cas9 endonuclease target site design on the Soybean ALS1 gene.*

- 5 There are four ALS genes in soybean (Glyma04g37270, Glyma06g17790, Glyma13g31470 and Glyma15g07860). Two guideRNA/ Cas9 endonuclease target sites (soy ALS1-CR1 and soy ALS1-CR2) were designed near the Proline 178 of the soybean ALS1 gene Glyma04g37270 (Table 49).

10 Table 49. Guide RNA/Cas9 endonuclease target sites on soybean ALS1 gene

Name of gRNA-Cas9 endonuclease target site	Cas endonuclease target sequence (SEQ ID NO:)	Physical location
soy ALS1-CR1	542	Gm04:43645633..43645612
soy ALS1-CR2	543	Gm04: 43645594..43645615

- B. Guide-RNA expression cassettes, Cas9 endonuclease expression cassettes, polynucleotide modification templates for introduction of specific amino acid changes and use the P178S modified ALS1 allele as a soybean transformation*
- 15 *selectable marker*

- The soybean U6 small nuclear RNA promoter, GM-U6-13.1 (SEQ ID. NO: 469), was used to express guide RNAs to direct Cas9 nuclease to designated genomic target sites (Table 50). A soybean codon optimized Cas9 endonuclease (SEQ ID NO:489) expression cassette and a guide RNA expression cassette were
- 20 linked in a first plasmid that was co-delivered with a polynucleotide modification template. The polynucleotide modification template contained specific nucleotide changes that encoded for amino acid changes in the soy ALS1 polypeptide (Glyma04g37270), such as the P178S. Other amino acid changes in the ALS1 polypeptide can also be obtained using the guide RNA/Cas endonuclease system
- 25 described herein. Specific amino acid modifications can be achieved by homologous recombination between the genomic DNA and the polynucleotide modification template facilitated by the guideRNA/Cas endonuclease system.

Table 50. Guide RNA/Cas9 expression cassettes and polynucleotide modification templates used in soybean stable transformation for the specific amino acid modifications of the soy ALS1 gene.

Experiment	Guide RNA/Cas9 (plasmid name)	SEQ ID NO:	polynucleotide modification template	SEQ ID NO:
soy ALS1-CR1	U6-13.1:ALS1-CR1 + EF1A2:CAS9 (QC880)	544	RTW1026A	546
soy ALS-CR2	U6-13.1:ALS1-CR2 + EF1A2:CAS9 (QC881)	545	RTW1026A	546

5 *C. Detection of the P178S mutation in the soybean ALS1 gene in the event selected by chlorsulfuron*

In order to edit specific amino acids at the native ALS1 gene (such as the P178S modification), a polynucleotide modification template such as RTW1026A (Table 50), was co-delivered with the guideRNA/Cas9 expression cassettes into soybean cells. Chlorsulfuron (100 ppb) was used to select the P178S ALS1 gene editing events in soybean transformation process.

The modification of the native ALS1 gene via guide RNA/Cas9 system mediated DNA homologous recombination was determined by specific PCR analysis. A specific PCR assay with primer pair WOL900 (SEQ ID NO: 547) and WOL578 (SEQ ID NO: 548) was used to detect perfect P178S modification at the native ALS1 gene. A second primer pair WOL573 (SEQ ID NO: 549) and WOL578 (SEQ ID NO: 548) was used to amplify both a P178S modified Soy ALS1 allele and a NHEJ mutated allele. A chlorsulfuron tolerant event (MSE3772 -18) was generated from the soy ALS1-CR2 experiment. The event contained a perfect P178S modified allele and a 2nd allele with a 5bp deletion at the soyALS1-CR2 cleavage site. Topo cloning/sequencing was used to verify the sequences. Our results demonstrated one P178S modified ALS1 allele is sufficient to provide Chlorsulfuron selection in soybean transformation process.

EXAMPLE 41Soybean FAD2-1 gene knockout in combination with P178S ALS1 modification as a selectable marker using the guideRNA/Cas endonuclease system

This example describes a method for gene modification in a plant genome using the guide RNA/Cas system without the insertion of an exogenous selectable marker into said plant genome. Instead, an herbicide resistance trait is generated by editing a first gene located in a plant genome while simultaneously knocking out a second gene located at a different locus than the first gene. This is accomplished by using two guides in combination with the Cas endonuclease as described below.

A. guideRNA/Cas9 endonuclease target site design on the Soybean FAD2-1 genes.

There are two FAD2-1 genes in soybean (FAD2-1A for Glyma10g42470 and FAD2-1B for Glyma20g24530). Two guideRNA/ Cas9 endonuclease target sites (soy FAD2-1-CR1 and soy FAD2-1-CR2) were designed to target both FAD2-1 genes for each gRNA (Table 51).

Table 51. Guide RNA/Cas9 endonuclease target sites on soybean FAD2-1 gene

Name of gRNA-Cas9 endonuclease target site	Cas endonuclease target sequence (SEQ ID NO:)	Physical location
soy FAD2-1-CR1	551	Gm10:49417140..49417118 Gm20:34178367..34178345
soy FAD2-1-CR2	552	Gm10:49417140..49417161 Gm20:34178367..34178388

B. Guide-RNA expression cassettes, Cas9 endonuclease expression cassettes and knockout of the soybean FAD2-1 genes with P178S ALS1 gene editing as the transformation selectable marker.

The soybean U6 small nuclear RNA promoter, GM-U6-13.1 (SEQ ID. NO: 469), was used to express guide RNAs to direct Cas9 nuclease to designated genomic target sites (Table 52). A soybean codon optimized Cas9 endonuclease (SEQ ID NO: 489) expression cassette and a guide RNA expression cassette for the ALS1 gene were linked in a first plasmid (QC881) that was co-delivered with an ALS1 polynucleotide modification template (RTW1026A). With a second gRNA plasmid (RTW1211 or RTW1212) targeting the FAD2-1 genes simultaneously, the

FAD2-1 knockout events can be identified in the chlorsulfuron tolerant events by qPCR and PCR assays.

Table 52. Guide RNA/Cas9 expression cassettes and polynucleotide modification

5 templates used in soybean stable transformation for the specific amino acid modifications of the soy ALS1 gene and FAD2-1 knock-out.

Experiment	Guide RNA/Cas9 (plasmid name)	SEQ ID NO:	polynucleotide modification template	SEQ ID NO:
soy ALS1-CR2	QC881	545	RTW1026A	546
soy FAD2-1-CR1	RTW1211	553		
soy ALS-CR2	QC881	545	RTW1026A	546
soy FAD2-1 CR2	RTW1212	554		

EXAMPLE 42

Soybean EPSPS1 gene editing in combination with P178S ALS1 modification as a

10 selectable marker using the guide RNA/Cas endonuclease system

This example describes a method for gene modification in a plant genome using the guide RNA/Cas system without the insertion of an exogenous selectable marker into said plant genome. Instead, an herbicide resistance trait is generated by editing a first gene located in a plant genome while simultaneously inserting a
15 polynucleotide of interest located at a different locus in the plant genome. This is accomplished by using two guides in combination with the Cas endonuclease as described below.

The EPSPS gene editing as described in Example 24, 25 and 26 can also be carried out in combination with the P178S ALS1 modification by using multiple
20 gRNAs targeting both the soybean EPSPS gene and soy ALS1 gene. The soybean EPSPS1 gene editing events can be identified in the chlorsulfuron tolerant events.

EXAMPLE 43Targeted gene integration in combination with P178S ALS1 modification as a selectable marker using the guide RNA/Cas endonuclease system

5 This example describes a method for gene modification in a plant genome using the guide RNA/Cas system without the insertion of an exogenous selectable marker into said plant genome. Instead, an herbicide resistance trait is generated by editing a first gene located in a plant genome while simultaneously editing a second gene located at a different locus in the plant genome. This is accomplished by using two guides in combination with the Cas endonuclease as described below.

10 The targeted gene integration in soybean as described in Example 19, 20 and 21 can also be carried out in combination with the P178S ALS1 modification by using multiple gRNAs targeting both the intended gene integration sites and soy ALS1 gene. The perfect gene integration events can be identified in the chlorsulfuron tolerant events.

THAT WHICH IS CLAIMED:

1. A method for producing a genetic modification into a target site of a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide polynucleotide, a
5 polynucleotide modification template, a second guide polynucleotide, and a Cas endonuclease to a plant cell comprising a first endogenous gene selected from the group consisting of an acetolactate synthase (ALS) gene, a glutamine synthase gene and an enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene, wherein said first guide polynucleotide and Cas endonuclease form a first complex that
10 enables the Cas endonuclease to introduce a double strand break at a first target site, located in or sufficiently near said first endogenous gene in the genome of said plant cell to affect its function, wherein said second guide polynucleotide and Cas endonuclease form a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant
15 cell, wherein said polynucleotide modification template comprises at least one nucleotide alteration when compared to the first endogenous gene and said first endogenous gene is modified to confer herbicide resistance to a plant cell, wherein the first target site and the second target site are located at two different genomic loci.
20
2. A method for introducing a polynucleotide of interest into a plant genome without introducing an exogenous selectable marker into said plant genome, the method comprising providing a first guide RNA, a first polynucleotide modification template, a second guide RNA, a second polynucleotide modification template, and
25 a Cas endonuclease to a plant cell comprising a first endogenous gene selected from the group consisting of an acetolactate synthase (ALS) gene, a glutamine synthase gene and an enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene, wherein said first guide RNA and Cas endonuclease form a first complex that enables the Cas endonuclease to introduce a double strand break at a first target
30 site located in or sufficiently near said first endogenous gene in the genome of said plant cell to affect its function, wherein said first polynucleotide modification template

comprises at least one nucleotide modification of said first endogenous gene such that said endogenous gene confers herbicide resistance to a plant cell, wherein said second guide RNA and Cas endonuclease form a second complex that enables the Cas endonuclease to introduce a double strand break at a second target site in the genome of said plant cell, wherein said second polynucleotide modification template comprises at least one polynucleotide of interest to be introduced into said plant genome, wherein the first target site and the second target site are located at two different genomic loci; wherein the polynucleotide of interest does not encode an exogenous selectable marker.

10

3. The method of claim 1, wherein the genetic modification into the target site comprises editing said second target site, wherein said first guide polynucleotide is RNA, said second guide polynucleotide is RNA, wherein said method further comprises providing a second polynucleotide modification template, wherein said second polynucleotide modification template comprises at least one nucleotide alteration when compared to the second target site to be edited.

15

4. The method of any one of claims 1 to 3, wherein Cas endonuclease is a Cas9 endonuclease.

20

5. The method of claim 1, further comprising selecting at least one plant cell that has resistance to a herbicide and comprises a modification into said second target site, wherein said modification includes at least one deletion, insertion or substitution of one or more nucleotides in said second target site of the plant genome.

25

6. The method of claim 2, further comprising selecting at least one plant cell that has resistance to a herbicide and comprises said at least one polynucleotide of interest in said plant genome.

30

7. The method of claim 3, further comprising selecting at least one plant cell that has resistance to a herbicide and comprises a modified EPSPS gene that confers glyphosate resistance.
- 5 8. The method of any one of claims 5 to 7, wherein said herbicide is a sulphonylurea or an imidazolinone herbicide.
9. The method of claim 8, wherein the sulphonylurea herbicide is chlorsulfuron.
- 10 10. The method of claim 5, wherein said modification into said target site includes at least one deletion, insertion, or substitution of one or more nucleotides in a FAD2-1 gene.
11. The method of claim 2 wherein the at least one at least one polynucleotide of
15 interest is selected from the group consisting of insecticidal coding sequences, nematocidal coding sequences, antimicrobial coding sequences, antifungal coding sequences, antiviral coding sequences, abiotic and biotic stress tolerance coding sequences, and sequences modifying plant traits selected from the group consisting of yield, grain quality, nutrient content, starch quality and quantity, nitrogen fixation,
20 nitrogen utilization, oil content and oil composition.
12. A cell of a herbicide resistant plant grown or cultured from the plant cell obtained by the method of claim 5, wherein the cell of the herbicide resistant plant comprises a mutation of at least one amino acid in the first endogenous gene
25 selected from the group consisting of an acetolactate synthase (ALS) gene, a glutamine synthase gene and an enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene that confers herbicide resistance to the cell and wherein the cell further comprises at least one deletion, insertion or substitution of one or more nucleotides in said second target site.

30

13. The cell of claim 12, wherein the at least one deletion, insertion or substitution of one or more nucleotides is in a FAD2-1 gene, and wherein the plant has a high oleic phenotype.

5 14. A cell of a herbicide resistant plant grown or cultured from the plant cell obtained by the method of claim 6, wherein the cell of the herbicide resistant plant comprises a mutation of at least one amino acid in the first endogenous gene selected from the group consisting of an acetolactate synthase (ALS) gene, a glutamine synthase gene and an enolpyruvylshikimate-3-phosphate synthase
10 (EPSPS) gene that confers herbicide resistance to the cell and wherein the cell further comprises at least one polynucleotide of interest stably inherited in said plant cell genome.

15 15. The cell of claim 14, wherein the at least one polynucleotide of interest is selected from the group consisting of insecticidal coding sequences, nematocidal coding sequences, antimicrobial coding sequences, antifungal coding sequences, antiviral coding sequences, abiotic and biotic stress tolerance coding sequences, and sequences modifying plant traits selected from the group consisting of yield, grain quality, nutrient content, starch quality and quantity, nitrogen fixation, nitrogen
20 utilization, oil content and oil composition.

16. The cell of any one of claims 12-15, wherein the plant is a monocot or a dicot.

25 17. The cell of claim 16, wherein:
(i) the monocot is selected from the group consisting of maize, rice, sorghum, rye, barley, wheat, millet, oats, sugarcane, turfgrass, and switchgrass; or
(ii) the dicot is selected from the group consisting of soybean, canola, alfalfa, sunflower, cotton, tobacco, peanut, potato, Arabidopsis, and safflower.

30

FIGURE 1 A

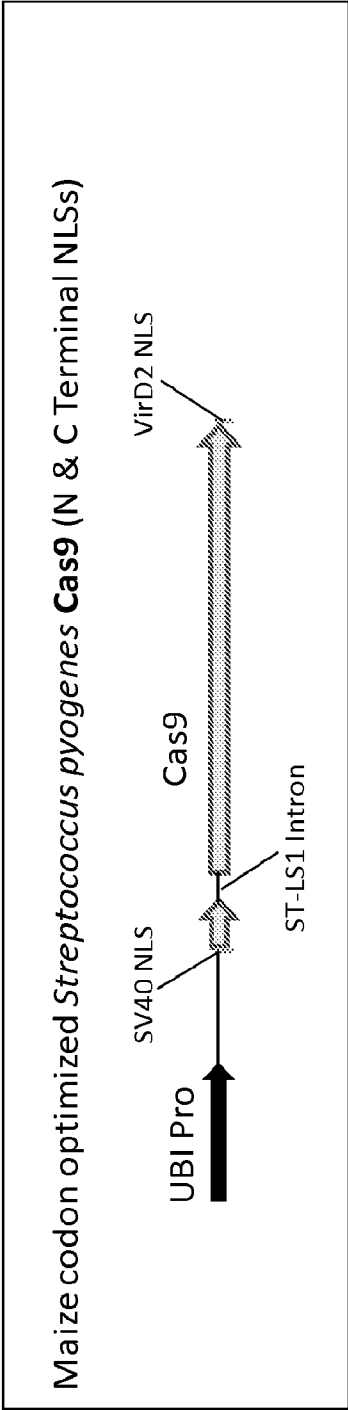


FIGURE 1 B

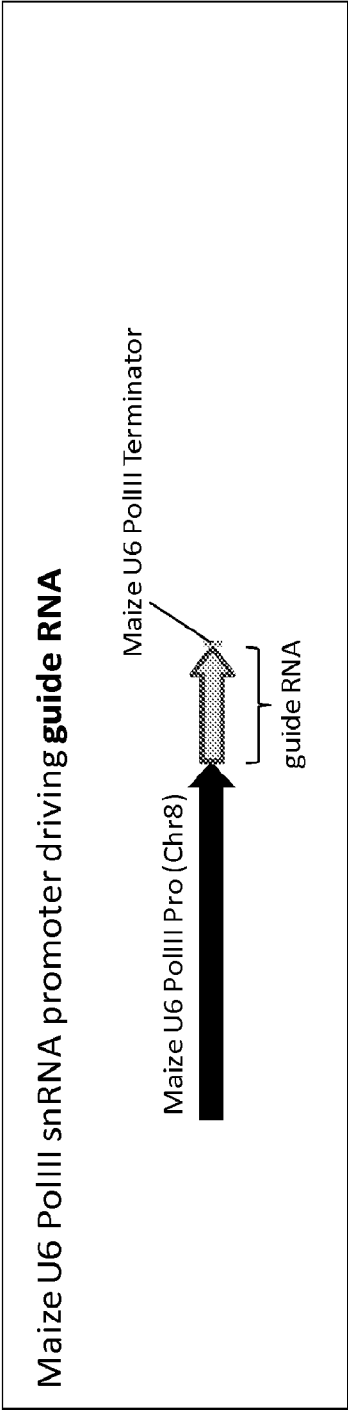
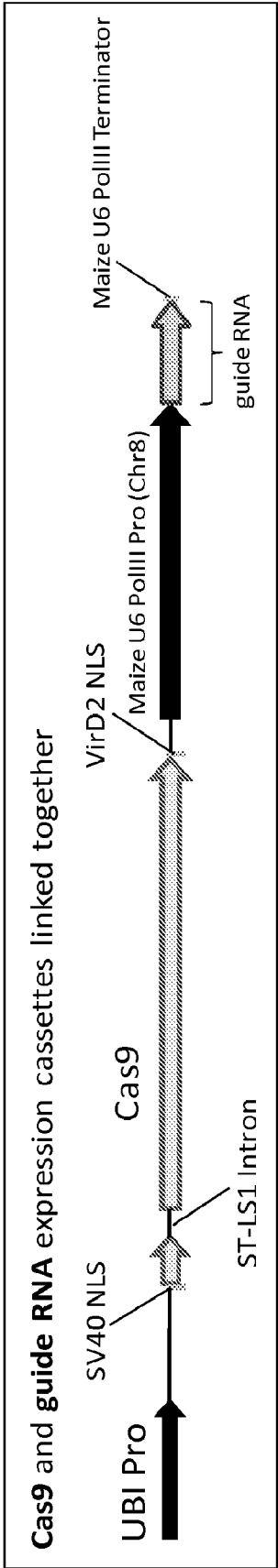


FIGURE 1 C



[illegible]

Long guide RNA (Fused crRNA and tracrRNA)

3' UCGUGGCGUGAGCCACGGUGCAGAAAAGUUUCANCUAUUUGCCUGAUGGGAUAUAAAUUGAGACGUAUAGAUUGGAGCAUUUUG 5'

3' GUGUGCAUGCGCAUGC 5'

Region Complementing the DNA Target
(Variable Targeting Domain)

Cas9

5' GATATATATA^C

3' CTATATATAT^G

PAM^A GTGTGCATGCCGATGCG^C

TATATACGTG 3'

ATATATGCAC 5'

FIGURE 3 A

Expected Site of Cleavage

PAM

LIGCas-1			Count	SEQ ID NO:
Reference	CTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		14488	55
Mutation 1	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGTCCGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		7746	56
Mutation 2	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		5028	57
Mutation 3	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		1425	58
Mutation 4	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		1056	59
Mutation 5	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		963	60
Mutation 6	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		732	61
Mutation 7	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		730	62
Mutation 8	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		492	63
Mutation 9	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		390	64
Mutation 10	AGGACTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC			65

Expected Site of Cleavage

PAM

LIGCas-2			Count	SEQ ID NO:
Reference	CTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		4221	55
Mutation 1	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		3452	66
Mutation 2	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		3395	67
Mutation 3	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		1870	68
Mutation 4	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		1344	69
Mutation 5	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		876	70
Mutation 6	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		507	71
Mutation 7	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		364	72
Mutation 8	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC		331	73
Mutation 9	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC			74
Mutation 10	TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATATACCTCACACGTACCGGTACCGGTATATATAC			75

Expected Site of Cleavage

Reference	LIGas-3	Count	SEQ ID NO:
Mutation 1	CGCAAAATGAGTAGCAGCGCACGTATATATACGCGTACGCGTACGTGTGAGGTATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	76	
Mutation 2	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATACGCGTACGCGTACGTGTGAGGTATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	77	
Mutation 3	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATACGCGTACGCGTACGTGTGAGGTATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	78	
Mutation 4	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATACGCGTACGCGTACGTGTGAGGTATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	79	
Mutation 5	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATACGCGTACGCGTACGTGTGAGGTATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	80	
Mutation 6	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATATATATATATATATATATATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	81	
Mutation 7	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATATATATATATATATATATATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	82	
Mutation 8	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATATATATATATATATATATATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	83	
Mutation 9	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATATATATATATATATATATATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	84	
Mutation 10	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATATATATATATATATATATATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	85	
Mutation 11	AAGGCGCAAAATGAGTAGCAGCGCACGTATATATATATATATATATATATATATATATATCTCCGCCGGGGCACGTACGGTACAATTTCCCG	86	

Expected Site of Cleavage

Reference	LIQ3-4 HOMING ENDONUCLEASE	Count	SEQ ID NO:
Mutation 1	CGCAAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGTAATAATCCTCCGCCGGGCACGTACGGTACAAATCCCAG	76	
Mutation 2	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGTG--AGGTATATAATCCTCCGCCGGGCACGTACGGTACAAATCCCAG	358	
Mutation 3	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGTAATA-----TCCTCCGCCGGGCACGTACGGTACAAATCCCAG	241	
Mutation 4	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGTAATA-----ACGTACGGTACAAATCCCAG	150	
Mutation 5	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGTAATAACG-----CGTACGGTACAAATCCCAG	143	
Mutation 6	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGTAATAACGCGT-----CGCCGGGGCACGTACGGTACAAATCCCAG	97	
Mutation 7	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGCGTACGTAATAATCCTCCGCCGGGCACGTACGGTACAAATCCCAG	52	
Mutation 8	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGCGTACGCGTACGTAATAATCCTCCGCCGGGCACGTACGGTACAAATCCCAG	50	
Mutation 9	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGCGTACGCGTACGTAATA-----CCTCCGCCGGGCACGTACGGTACAAATCCCAG	46	
Mutation 10	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGCGTACGCGTACGCGTACGTAATAATCCTCCGCCGGGCACGTACGGTACAAATCCCAG	42	
Mutation 11	CCCTCGCAAATGAGTAGCAGGCACGTAATAIAACGCGTACGCGTACGCGTACGCGTACGCGTACGCGTACGTAATAATCCTCCGCCGGGCACGTACGGTACAAATCCCAG	32	

FIGURE 4

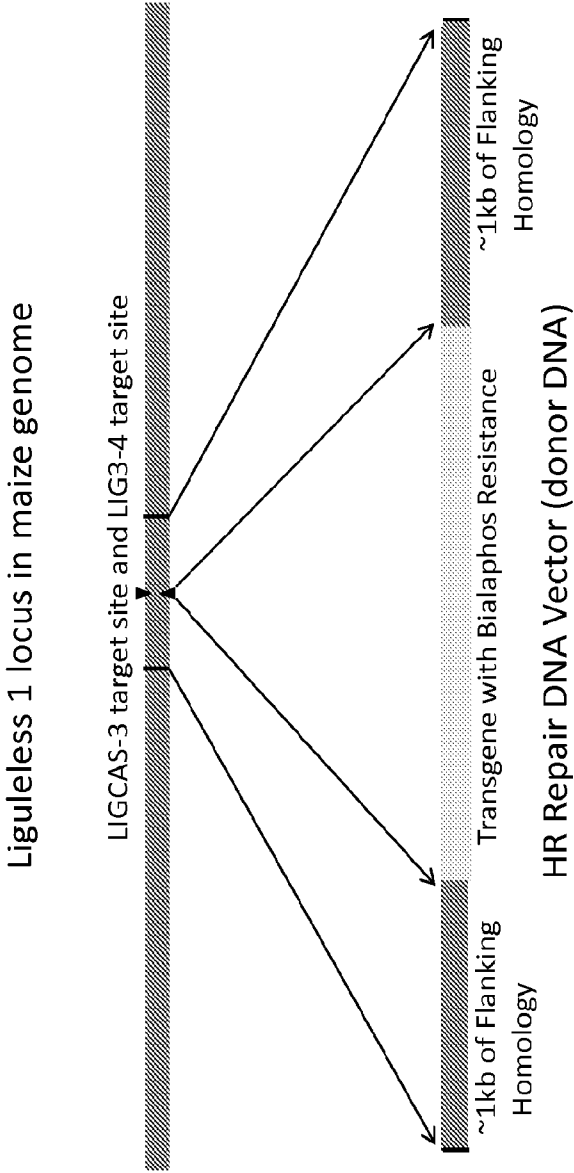


FIGURE 5

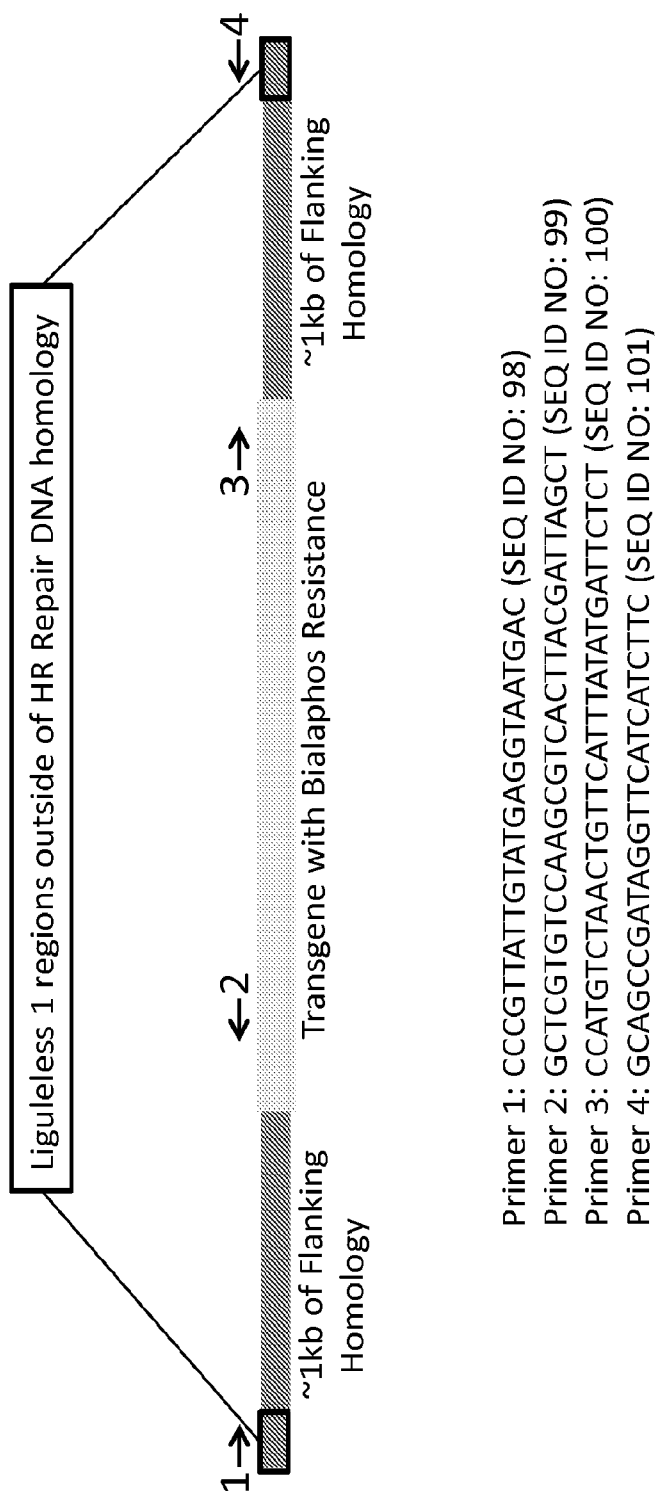


FIGURE 6

Expected Site of Cleavage

	55CasRNA-1	SEQ ID NO:
Reference	CCGGTTTCGCGTGCTCTGGCTTTACATTACATGGGCAGGTCTCAGCAGGTTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	104
Mutation 1	CCGGTTTCGCGTGCTCTGGCTTTACATTACATGGGCAGGTCTCAGCA-GGTTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	105
Mutation 2	CCGGTTTCGCGTGCTCTGGCTTTACATTACATGGGCAGGTCTCAC-ACGGTTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	106
Mutation 3	CCGGTTTCGCGTGCTCTGGCTTTACATTACATGGGCAGGTCTCAGCAGGTTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	107
Mutation 4	CCGGTTTCGCGTGCTCTGGCTTTACATTGCATGAGCAGGTCGT--GACGGTTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	108
Mutation 5	GGGCAGGTC'T--CGACGGTTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	109
Mutation 6	CCGGTTTCGCGTGCTC-----TTGGGCTGGAGAGCCCGGCTGGTAGGGAGGACCTCAACGGC	110

PAM

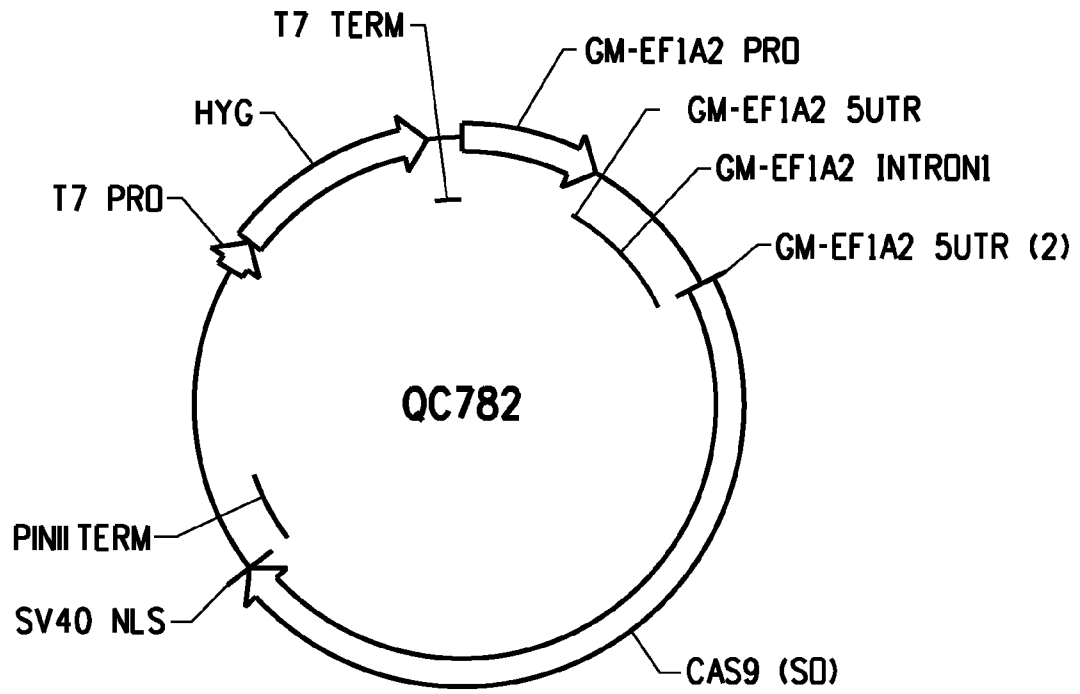
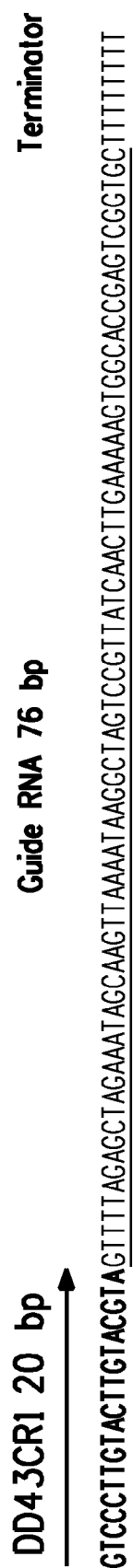


FIG. 7

FIG. 8B



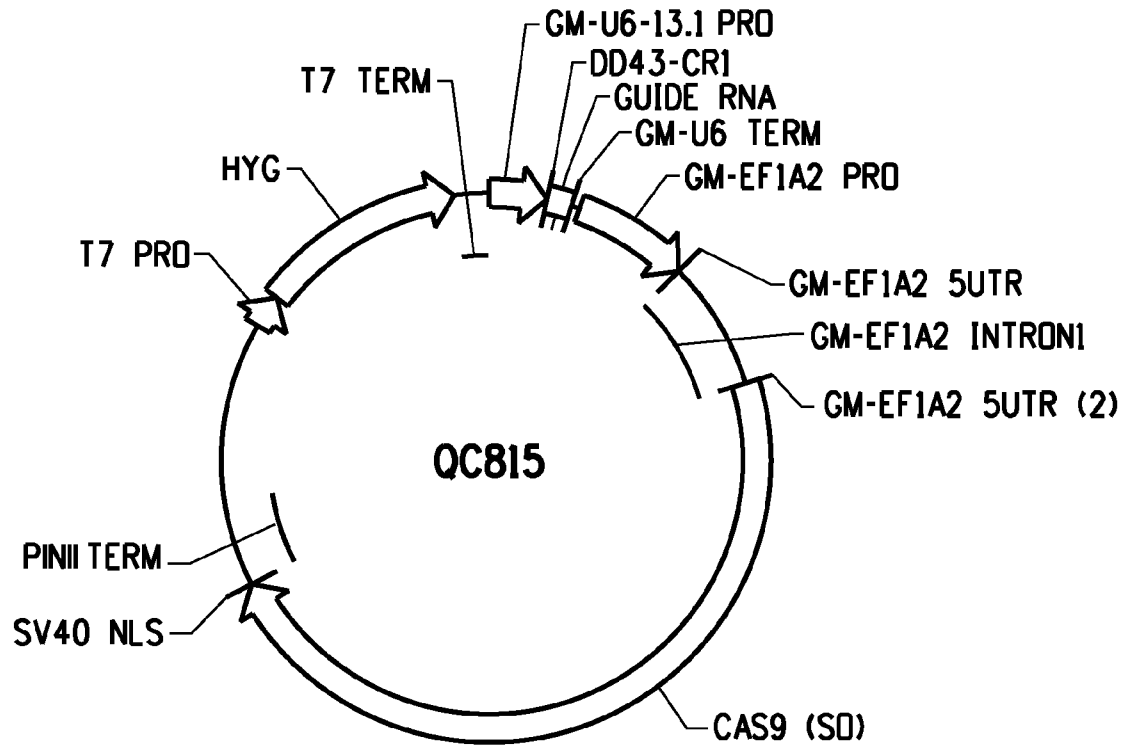


FIG. 9

FIGURE 10 A



FIGURE 10 B

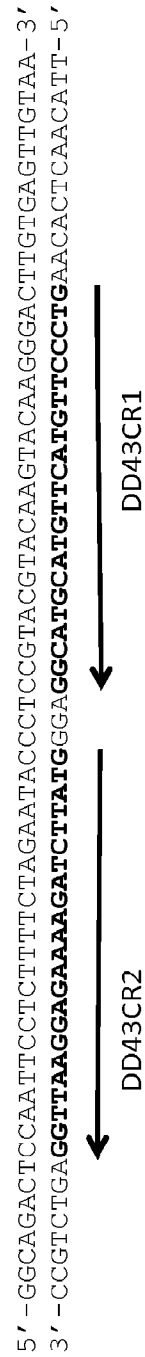


FIGURE 11B

U6-13.1:DDZ0CR2÷EFLA2:CA59

[illegible]

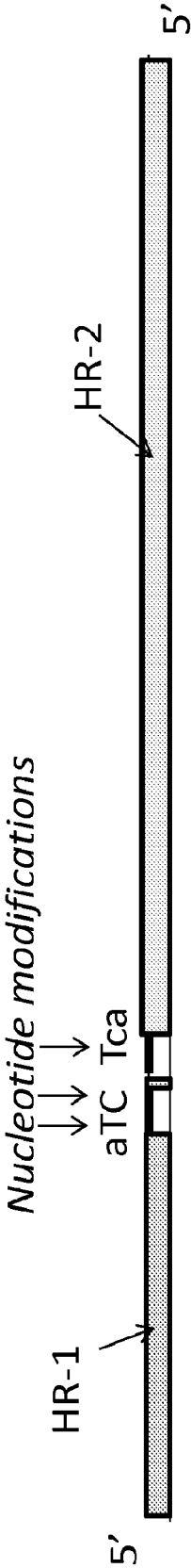
[illegible]

FIGURE 11 D

[illegible]

FIGURE 12

A : polynucleotide modification template (EPSPS template)



B: nucleotide sequence to be edited (wild type epsps locus)

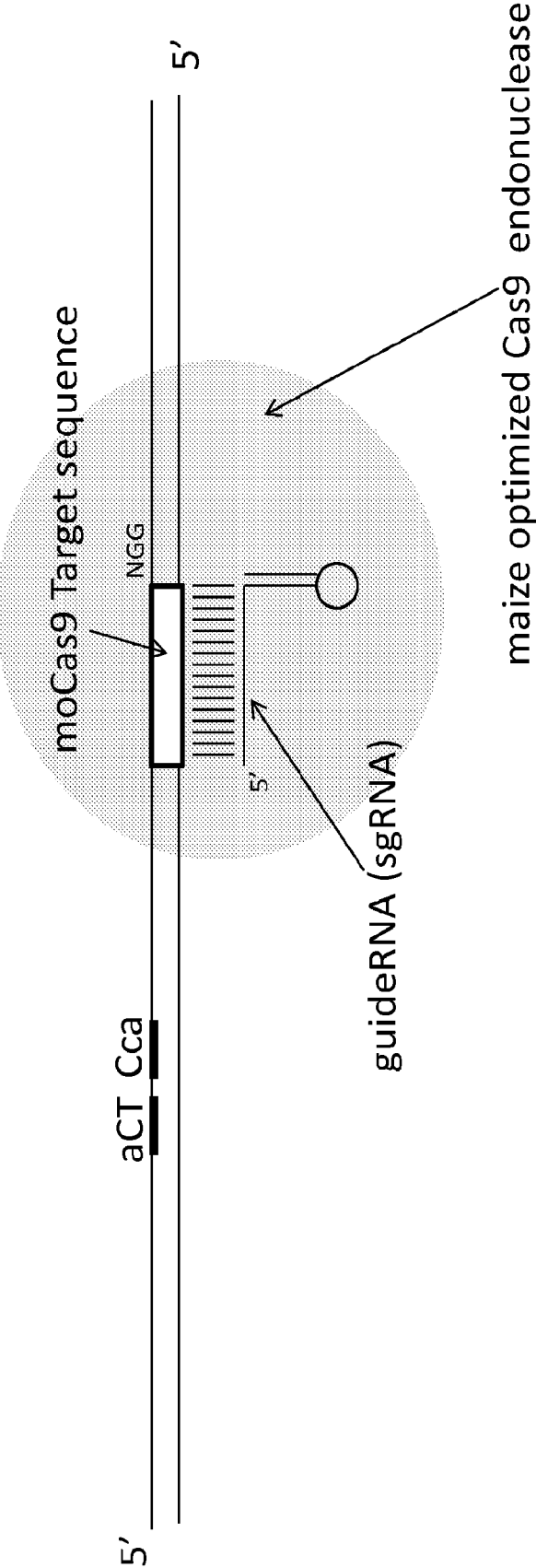


FIGURE 13

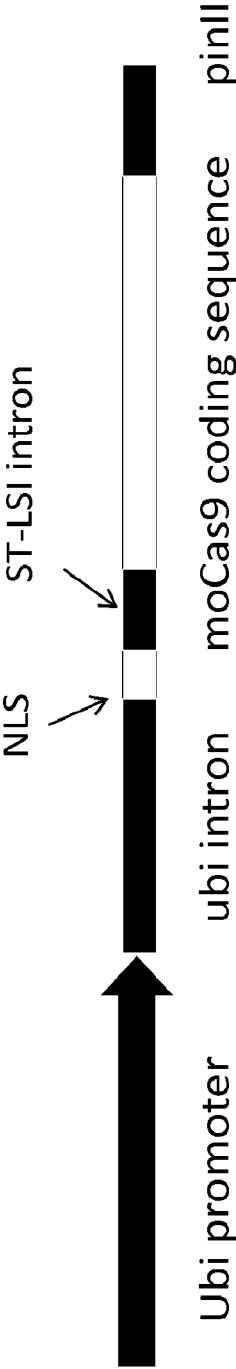


FIGURE 14

↓

Events with Intact moCas target sequence (underlined)

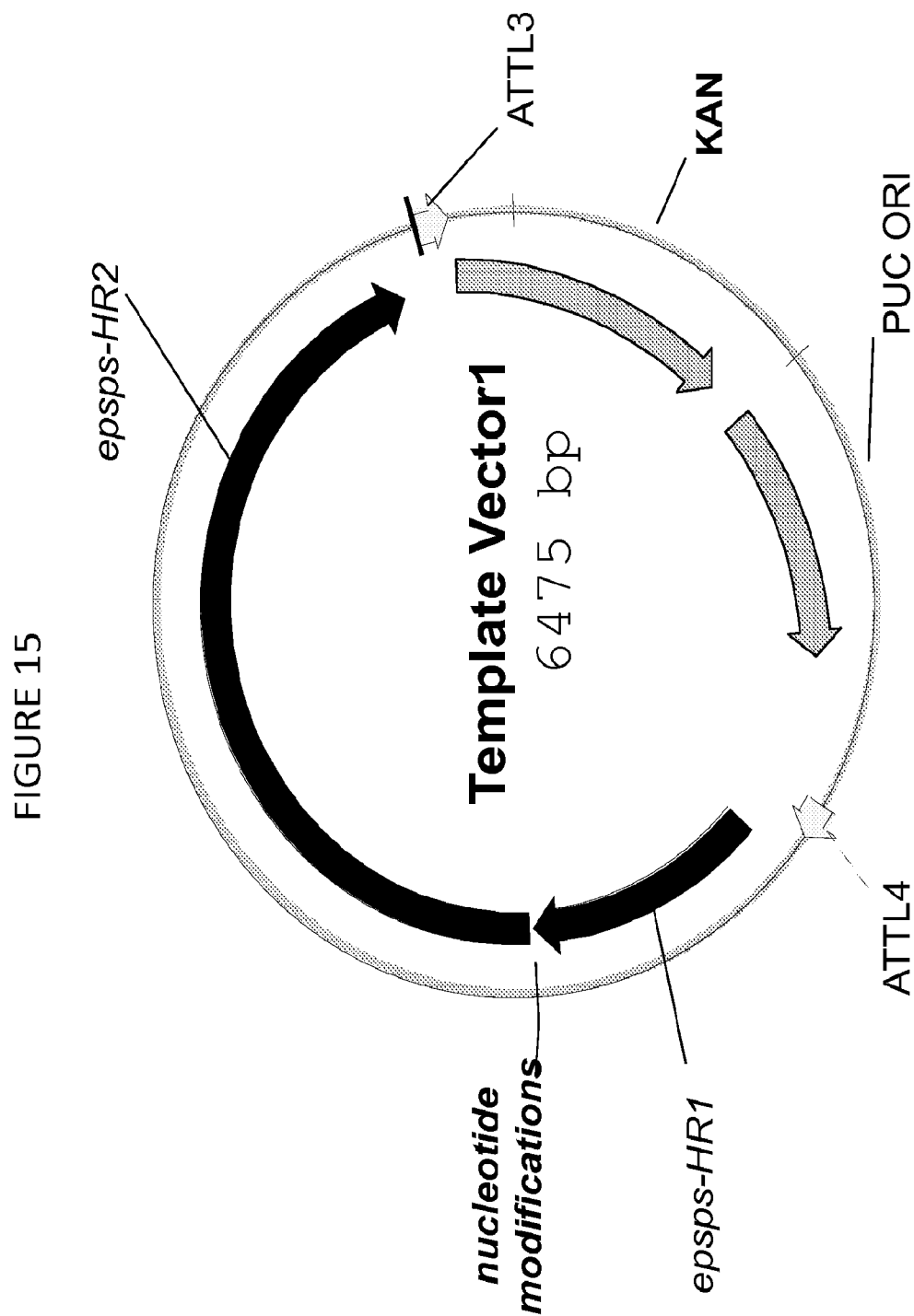
SEQ ID NO: 205 GGGGAA TGCTGGAACTGCAATGCCGCCA TTGACAGCAGCTGTACTGCTGCTGGTGGAAATGC

Events with mutagenized moCas target sequences (underlined)

SEQ ID NO: 206 GGGGAA TGCTGGAACTGCAATGCCGCCA TTG---GCAGCTGTACTGCTGCTGGTGGAAATGC

SEQ ID NO: 207 GGGGAA TGCTGGAACTGCA-----CAGCAGCTGTACTGCTGCTGGTGGAAATGC

SEQ ID NO: 208 GGGGAA TGCTG-----TTACTGCTGCTGGTGGAAATGC



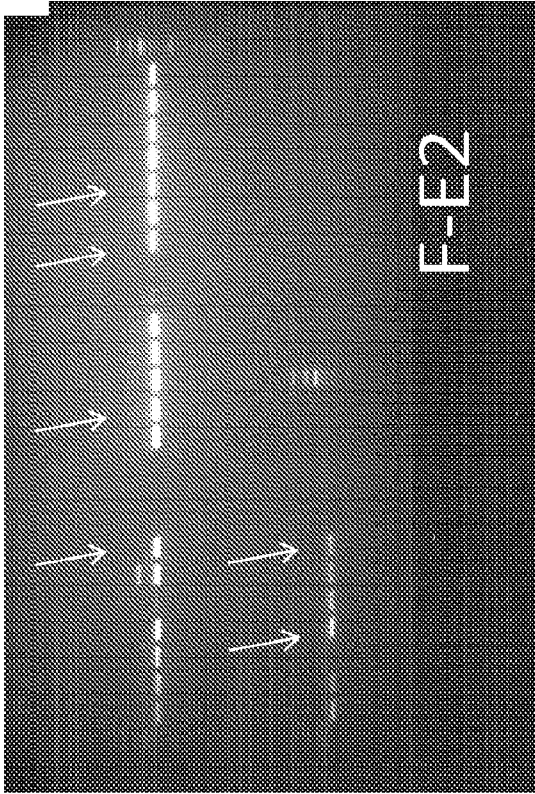
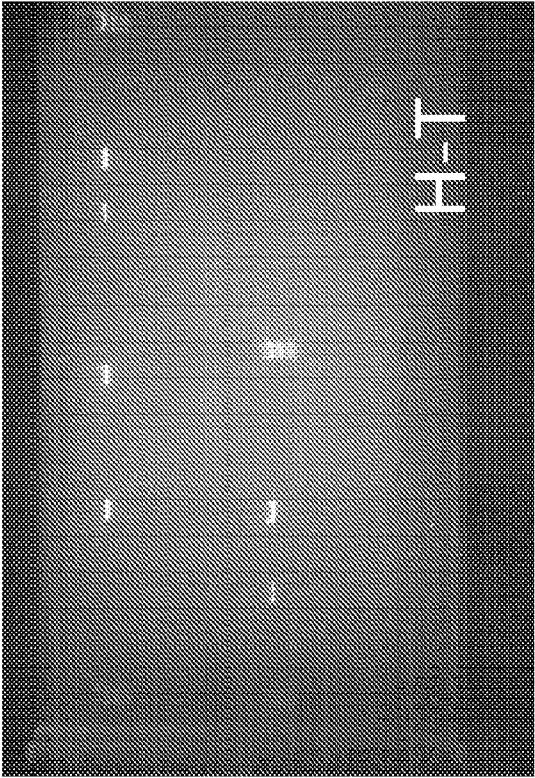
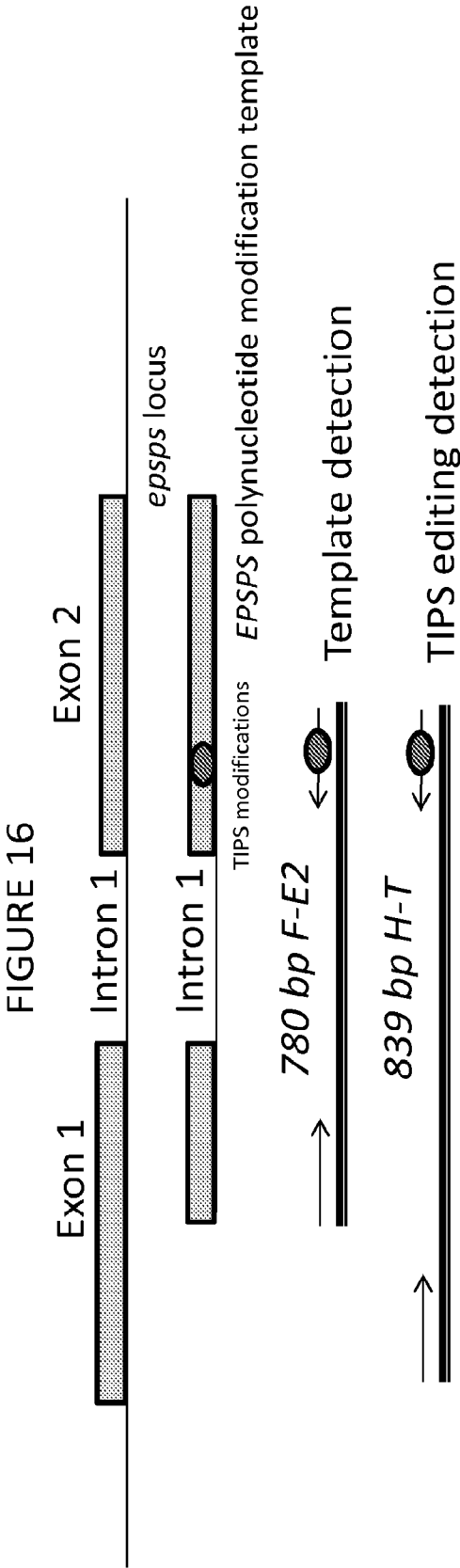
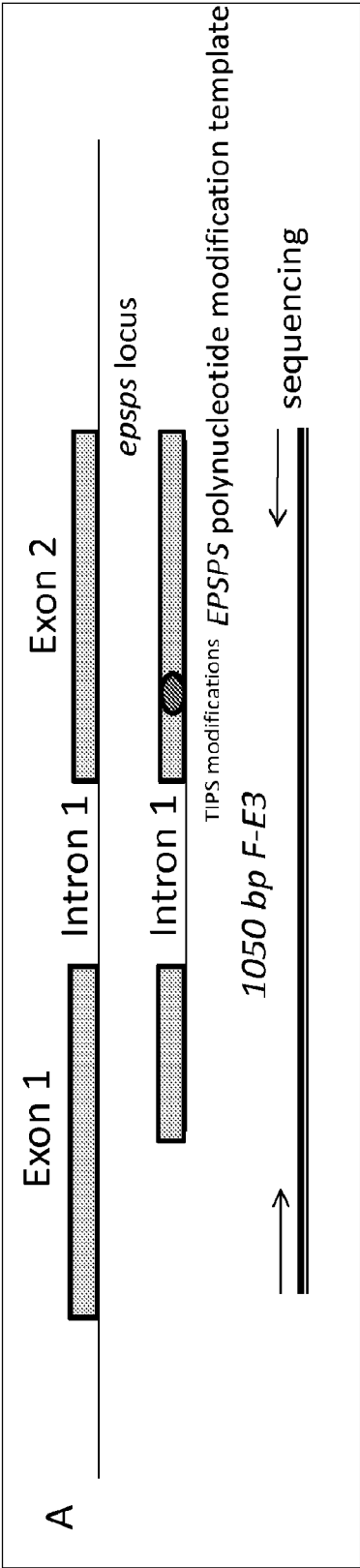


FIGURE 17



B

TIPS edited *EPSPS* nucleotide sequence with the moCas target sequence (underlined) not changed.

nucleotide modifications

↓ ↓ ↓

SEQ ID NO: 209 AATGCTGGAA**TC**GGCAATGCGGG**T**CATTGACAGCAGCTGTTACTGCTGGT

Wild-type *EPSPS* nucleotide sequence with the Cas target sequence not changed.

SEQ ID NO: 210 AATGCTGGAA**CT**GGCAATGCGGG**C**CATTGACAGCAGCTGTTACTGCTGGT

FIGURE 18

Trait A

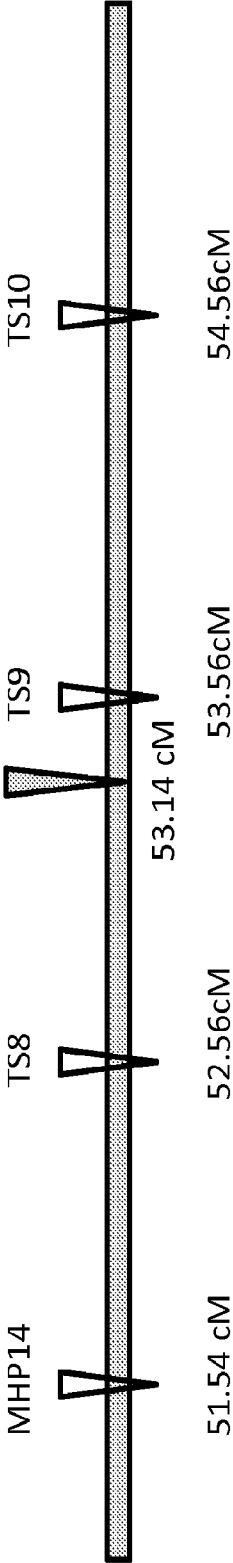


FIGURE 19

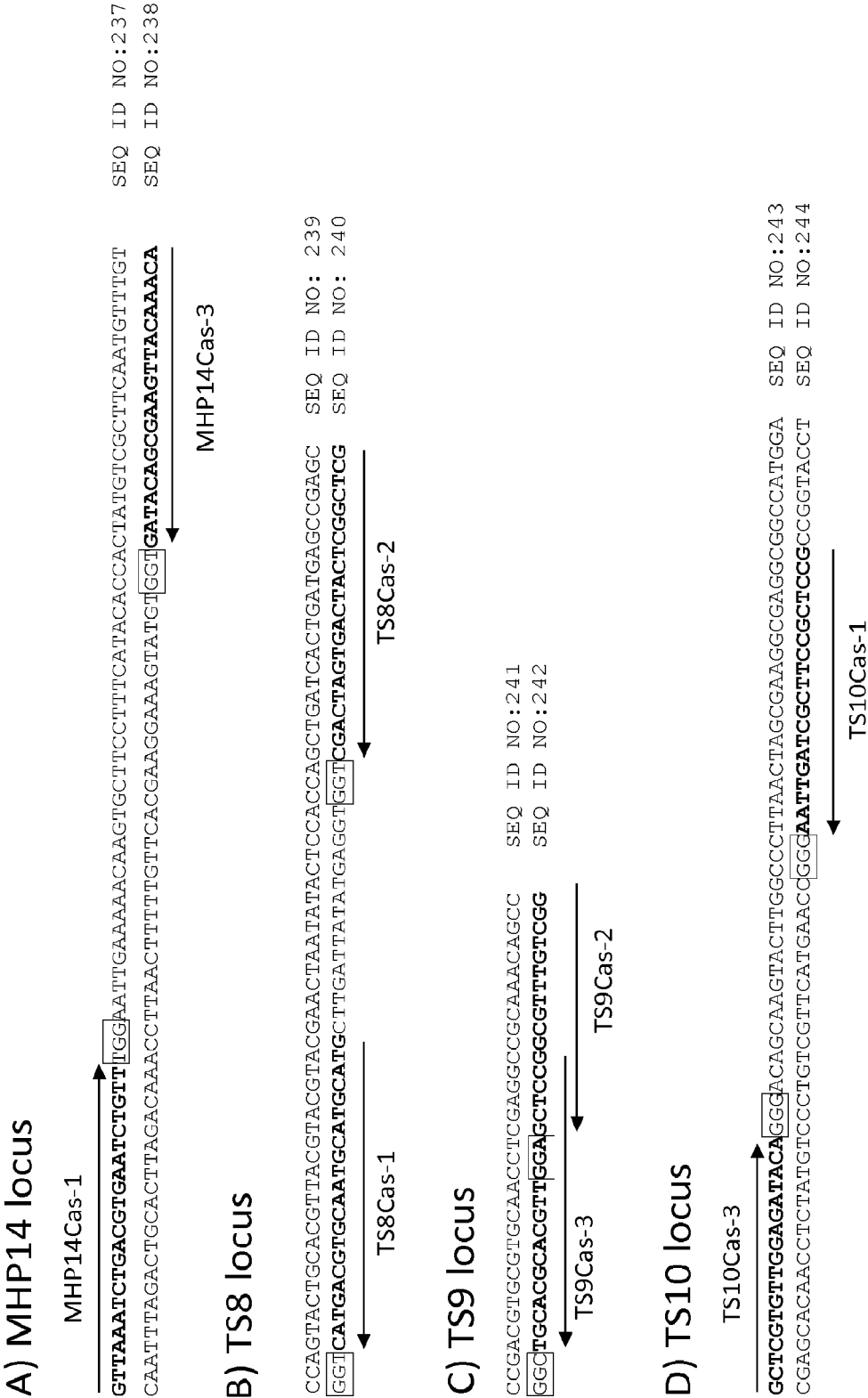


FIGURE 20

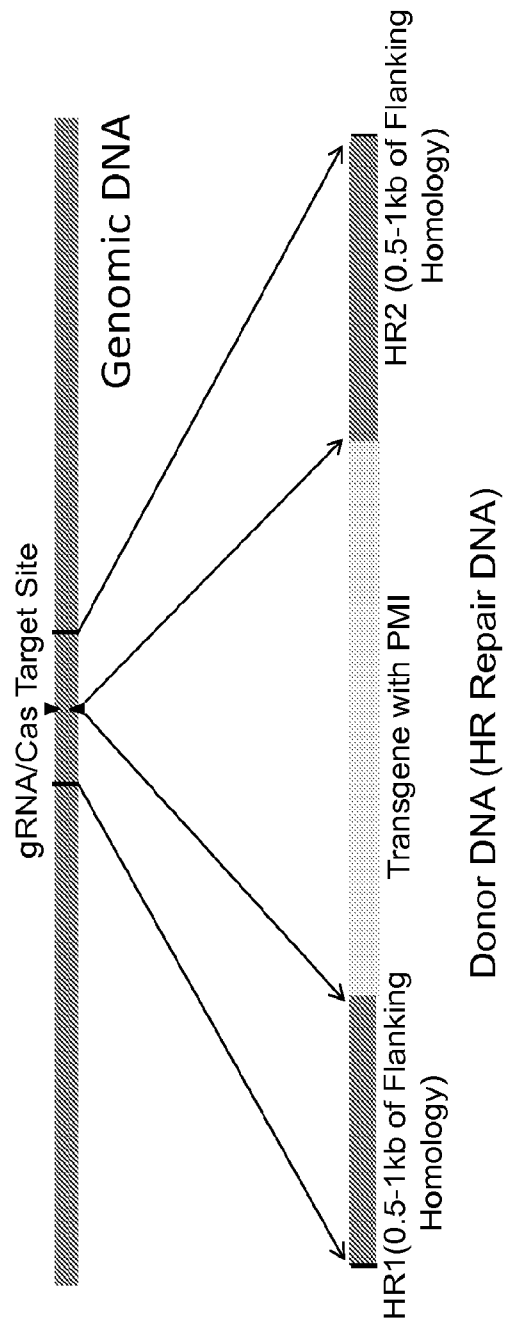


FIGURE 21

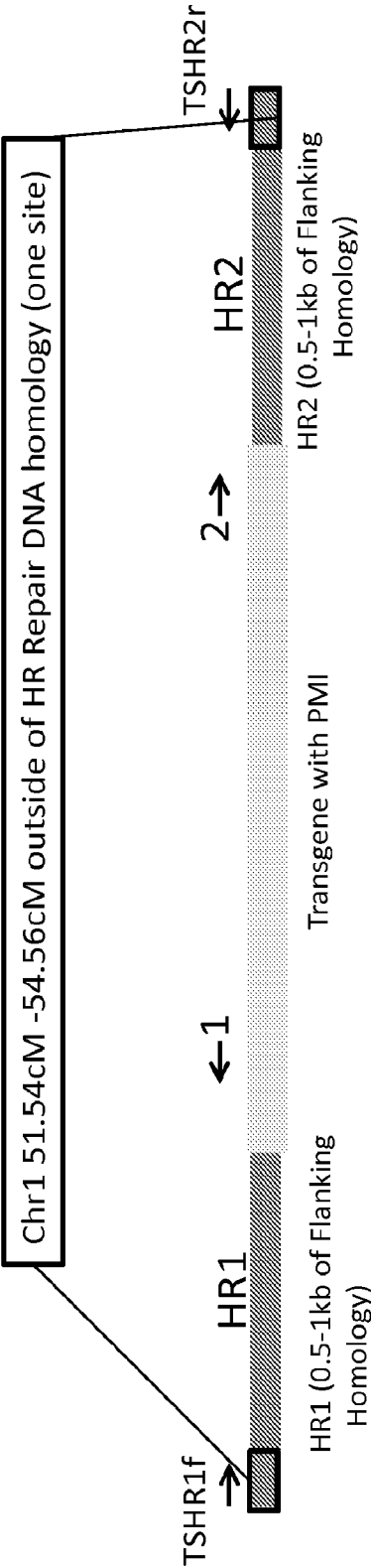


FIGURE 22

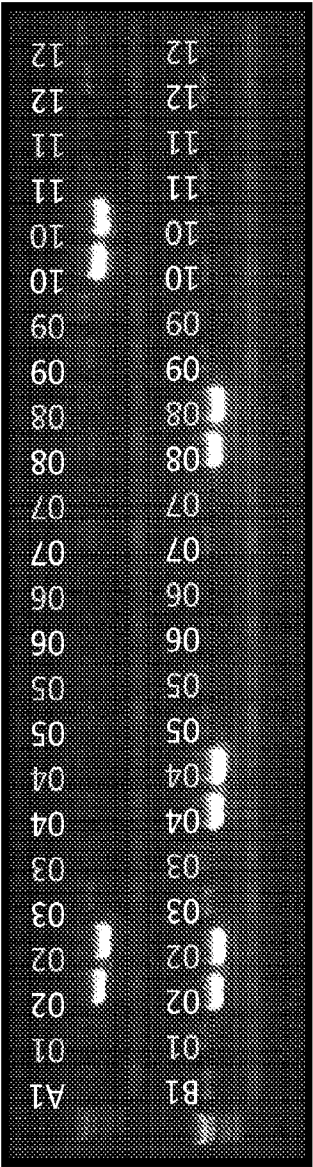
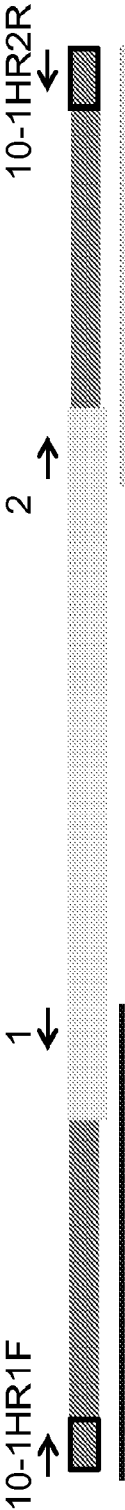
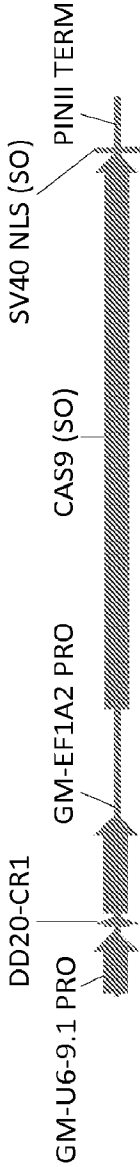


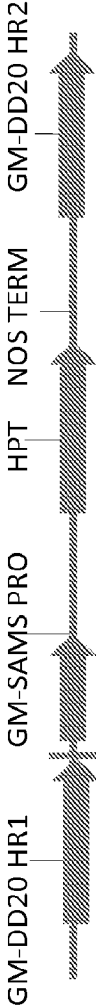
FIGURE 23

A. Linked gRNA and Cas9 gene expression cassettes



U6-9.1:DD20CR1+EF1A2:CAS9

B. Repair DNA cassette with homologous regions.



DD20HR1-SAMS:HPT-DD20HR2

FIGURE 24

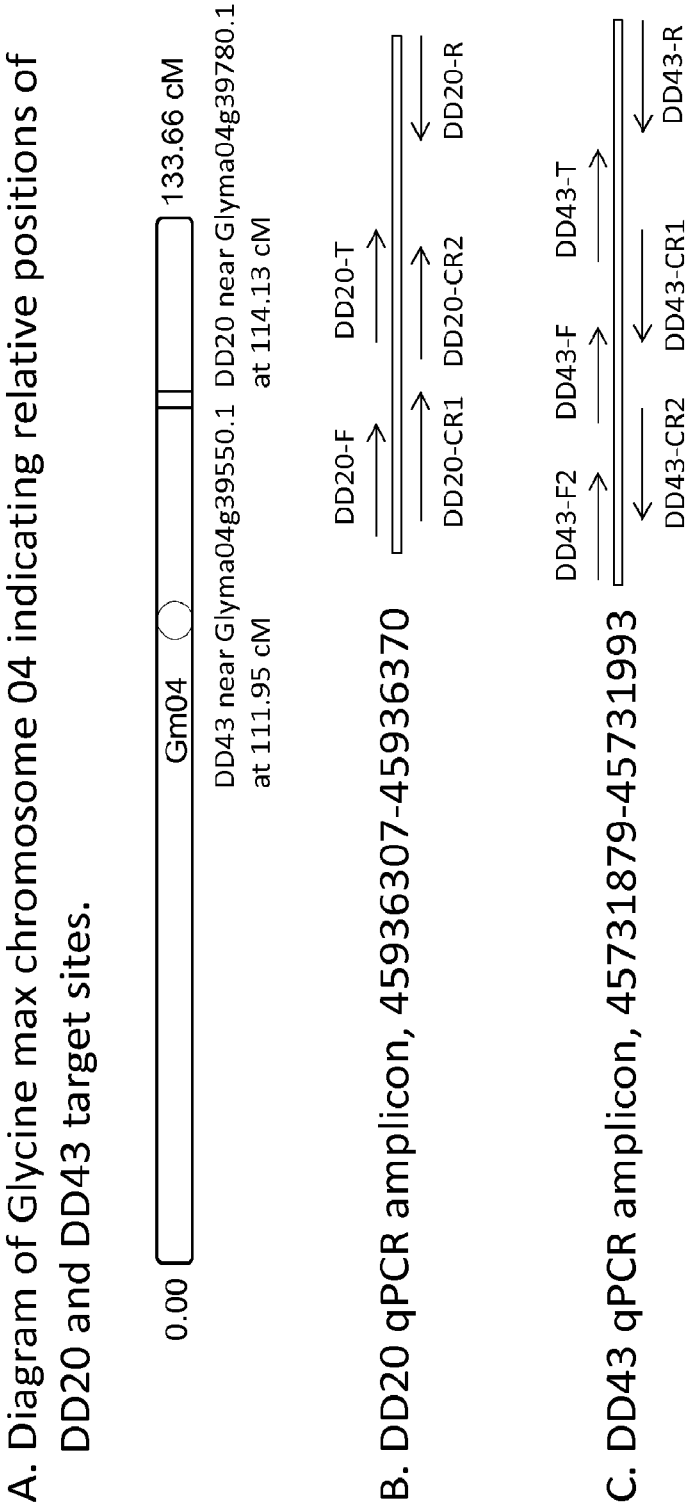
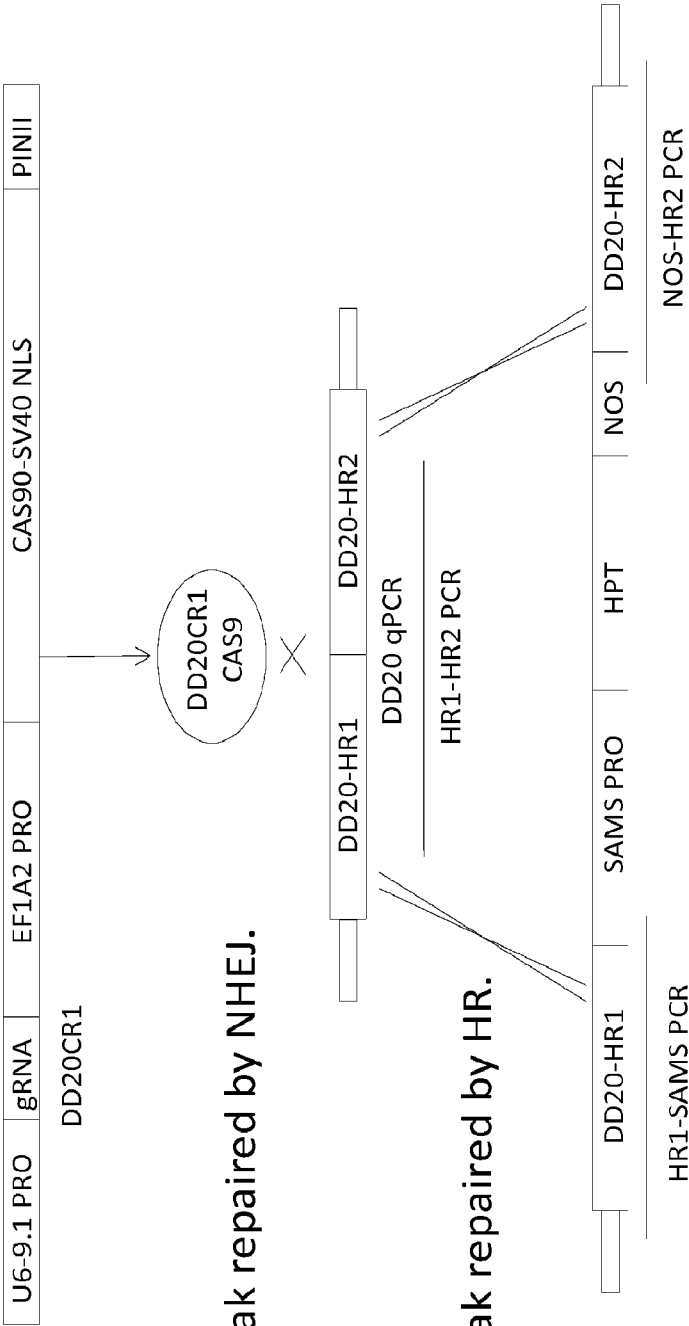


FIGURE 25

A. Transiently expressed gRNA and Cas9 assembled in vivo capable to cleave genomic DNA at target DD20.



B. DNA break repaired by NHEJ.

C. DNA break repaired by HR.

DD20CR1 target site

SUBSTITUTE SHEET (RULE 26)

FIGURE 26-B

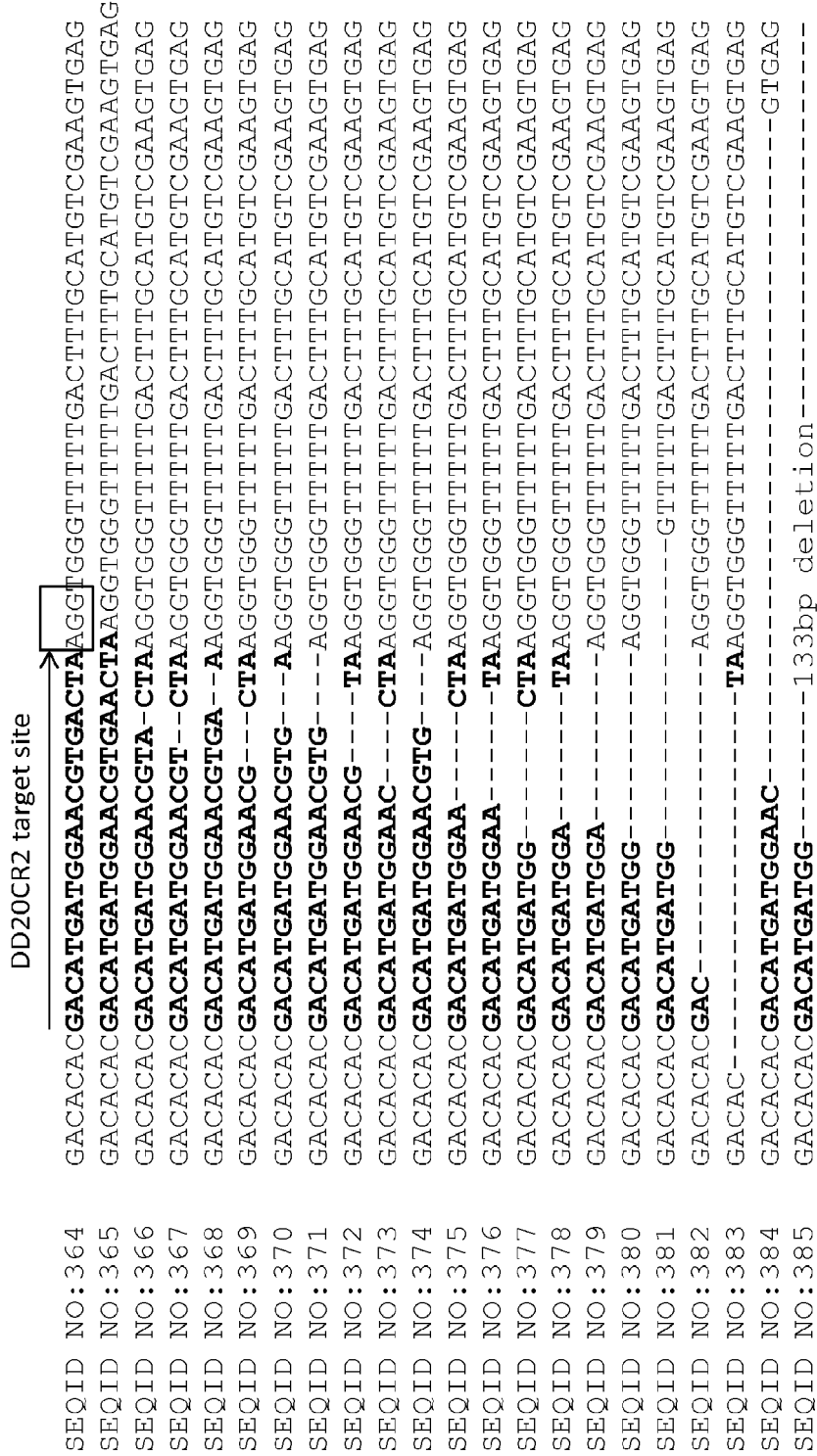


FIGURE 26-C

DD43CR1 target site

SEQID NO: 386 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**CGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 387 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 388 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**-GTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 389 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**-GTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 390 AGCCTTACAACCTCACAA**GTCCCTTGACTTG**-CGTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 391 AGCCTTACAACCTCACAA**GTCCCTTGACTTG**-GTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 392 AGCCTTACAACCTCACAA**GTCCCTTGACTTG**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 393 AGCCTTACAACCTCACAA**GTCCCTTGACTT**-GTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 394 AGCCTTACAACCTCACAA**GTCCCTTGACTT**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 395 AGCCTTACAACCTCACAA**GTCCCTTGACT**-GTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 396 AGCCTTACAACCTCACAA**GTCCCTTGACT**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 397 AGCCTTACAACCTCACAA**GTCCCTTGTA**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 398 AGCCTTACAACCTCACAA**GTCCCTTG**-GTACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 399 AGCCTTACAACCTCACAA**GTCCCTTG**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 400 AGCCTTACAACCTCACAA**GTCCCTT**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 401 AGCCTTACAACCTCACAA**GTCCCT**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 402 AGCCTTACAACCTCACAA**GTCC**-TACCGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 403 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**-----AGAAAAAGAGG
 SEQID NO: 404 AGCCTTACAACCTCACAA**GTCC**-----TAAATTAA^AGGTTATTCTAGAAAAAGAGG
 Insertion starts at ^ with the insert size indicated.
 SEQID NO: 405 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^167GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 406 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^38--GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 407 AGCCTTACAACCTCACAA**GTCCCTTGTA**^130-----GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 408 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^171GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 409 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^220GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 410 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^190GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 411 AGCCTTACAACCTCACAA**GTCCCTT**^110-----GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 412 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^125GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 413 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^154GTACGGAGGGTATTCTAGAAAAAGAGG
 SEQID NO: 414 AGCCTTACAACCTCACAA**GTCCCTTGACTTGTA**^177-----GAGGGTATTCTAGAAAAAGAGG

FIGURE 27 A

Expected Site of Cleavage

PAM

LIGCas-1	Count	SEQ ID NO:
CTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGAGGATATATATACCTCACACGTACCGGTACGCGTATATATAC	55	
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGTCCGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	2116	415
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	1156	416
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	473	417
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	161	418
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	133	419
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	82	420
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	77	421
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	55	422
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	39	423
TCCTCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCCCCGGGACGAGGATATATATACCTCACACGTACGCGTACGCGTATATATAC	39	424

FIGURE 27-B

Expected Site of Cleavage

PAM

LIGCas-2	Count	SEQ ID NO:
CTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	55	55
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	1048	425
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	743	426
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	543	427
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	220	428
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	193	429
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	159	430
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	137	431
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	94	432
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	93	433
GAAGCTGTAACGATTTACGCACCTGCTGGGAATTGTACCGTACGTGCGCGGGAGGATATATATACCTCACACGTACCGGTACGGGTATATATAC	60	434

Expected Site of Cleavage

PAM

LIGCas-3		SEQ ID NO:
CGCAAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTACGGTGAGGTTAATATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	Count	76
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTACGGTGTGAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	1208	435
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTACGGTGTGAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	453	436
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTACGGTGTGAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	338	437
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTAC--TGTCAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	145	438
AAGGCGCAAATGAGTAGCAGCGCACGTATATATA-----TGTCAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	94	439
AAGGCGCAAATGAGTAGCAGCGCACGTATATATA-----TGTCAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	53	440
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACG--TGTCAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	42	441
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTAC-----TAATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	37	442
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTACG--TGTCAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	20	443
AAGGCGCAAATGAGTAGCAGCGCACGTATATAACGGCTACGCGTACG--TGTCAGGTTATATATCTCCGCCGGGCACGTACGGTACAATTCCAG	19	444

Figure 28

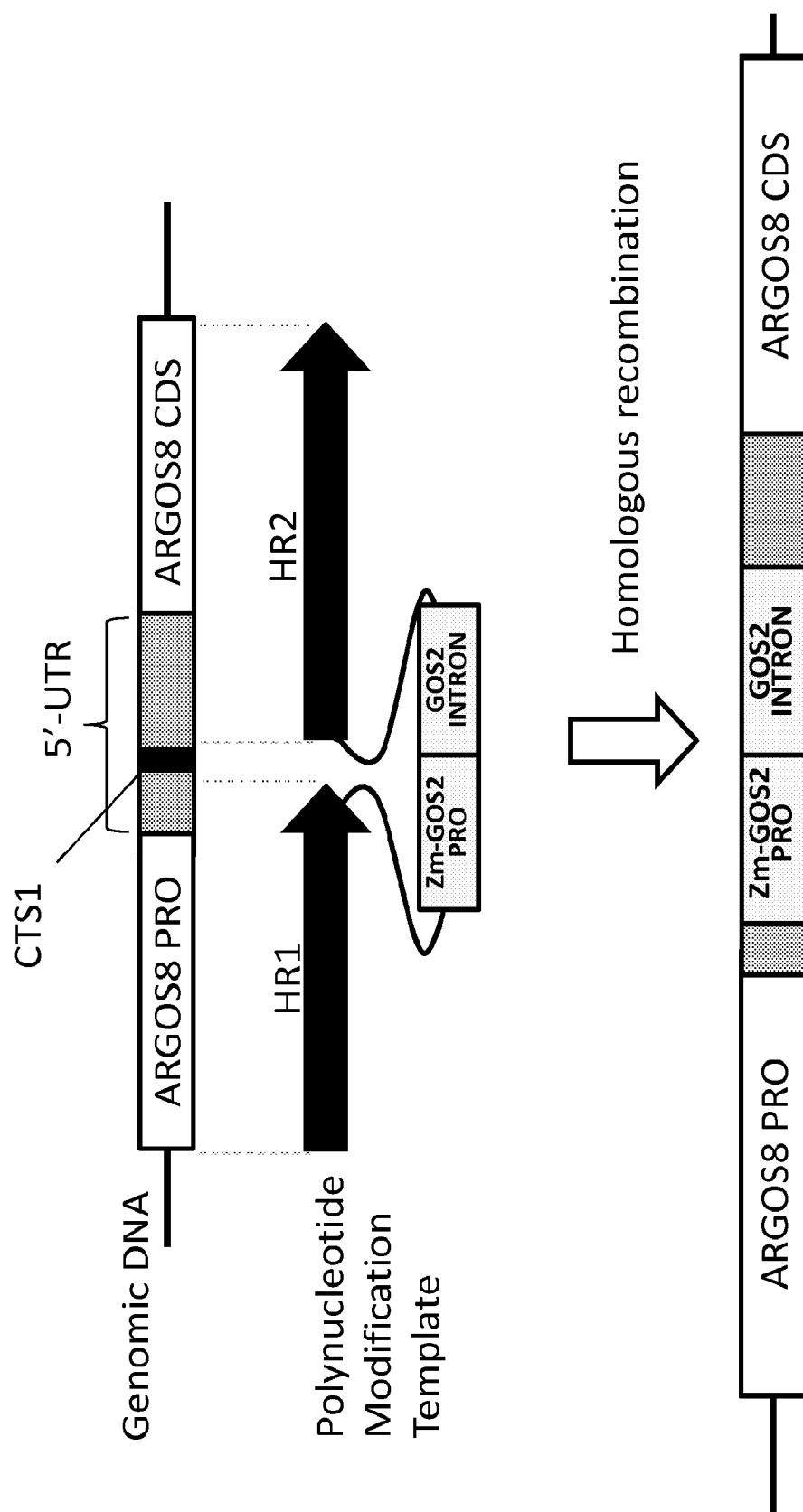


Figure 29A

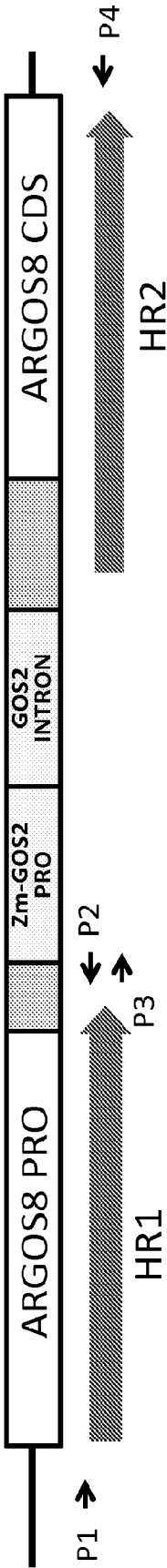


Figure 29C

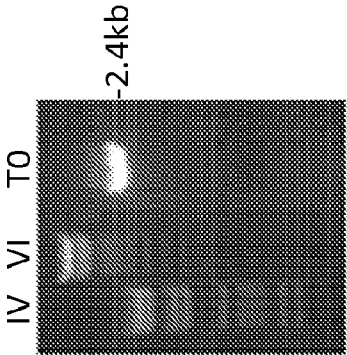


Figure 29B

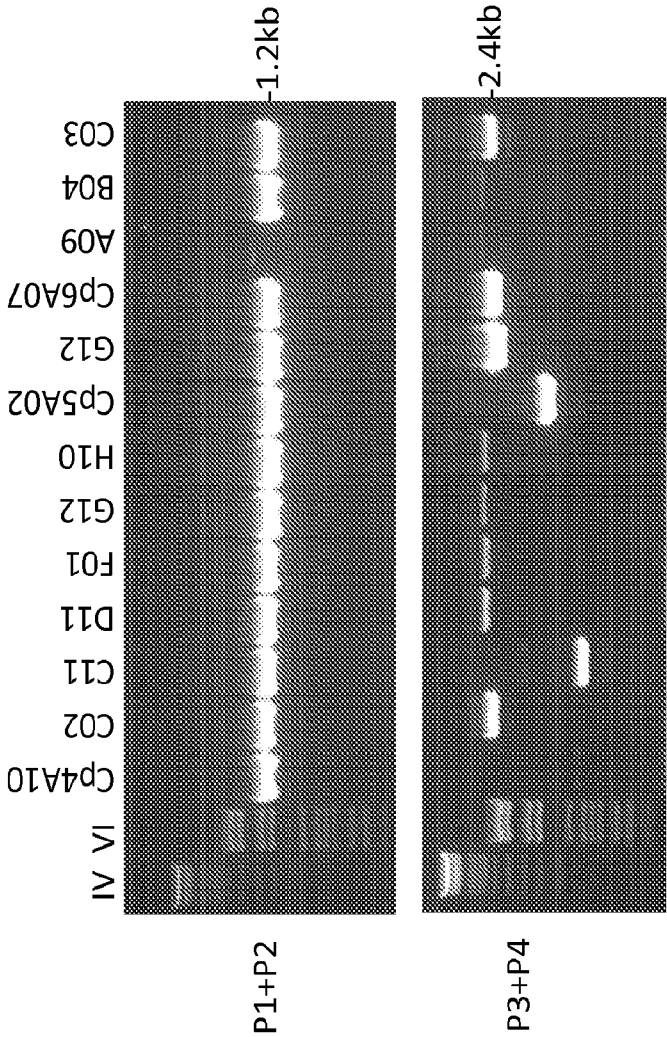


Figure 30

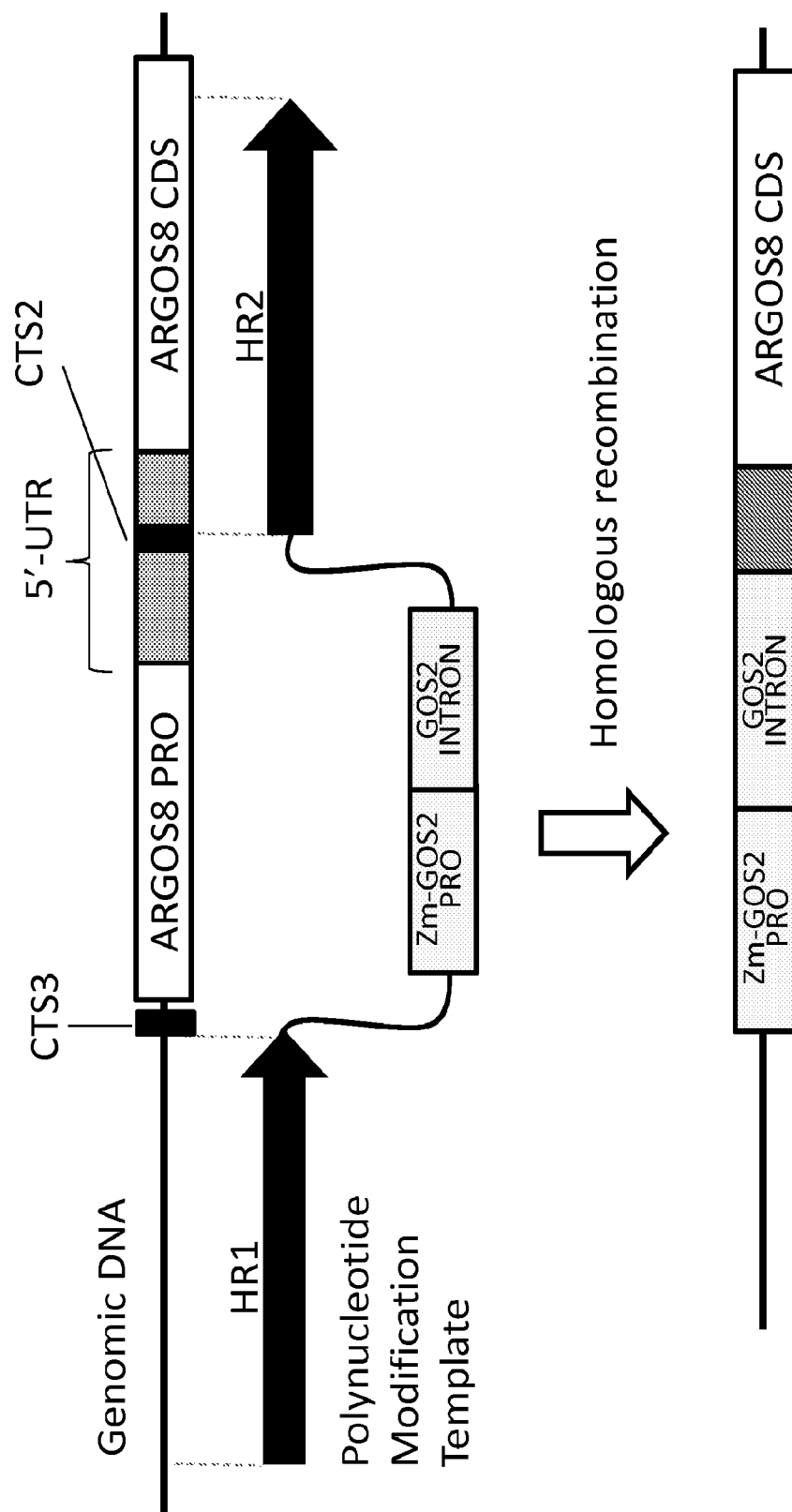


Figure 31A

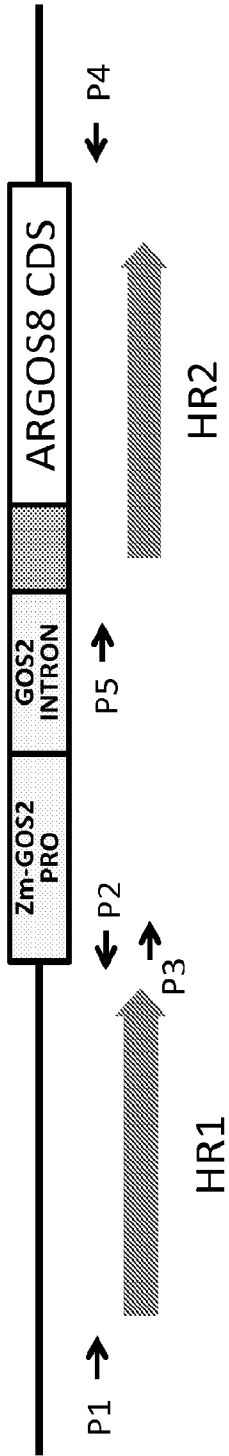


Figure 31B

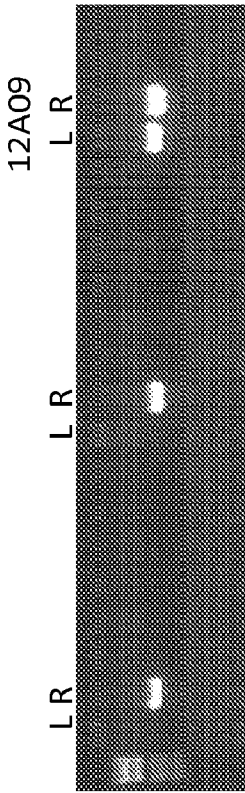


Figure 31C

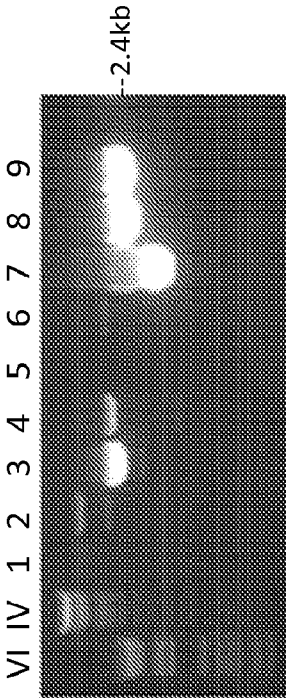


Figure 31D

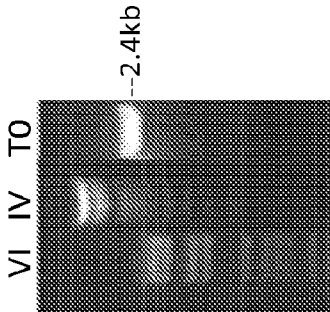


Figure 32A

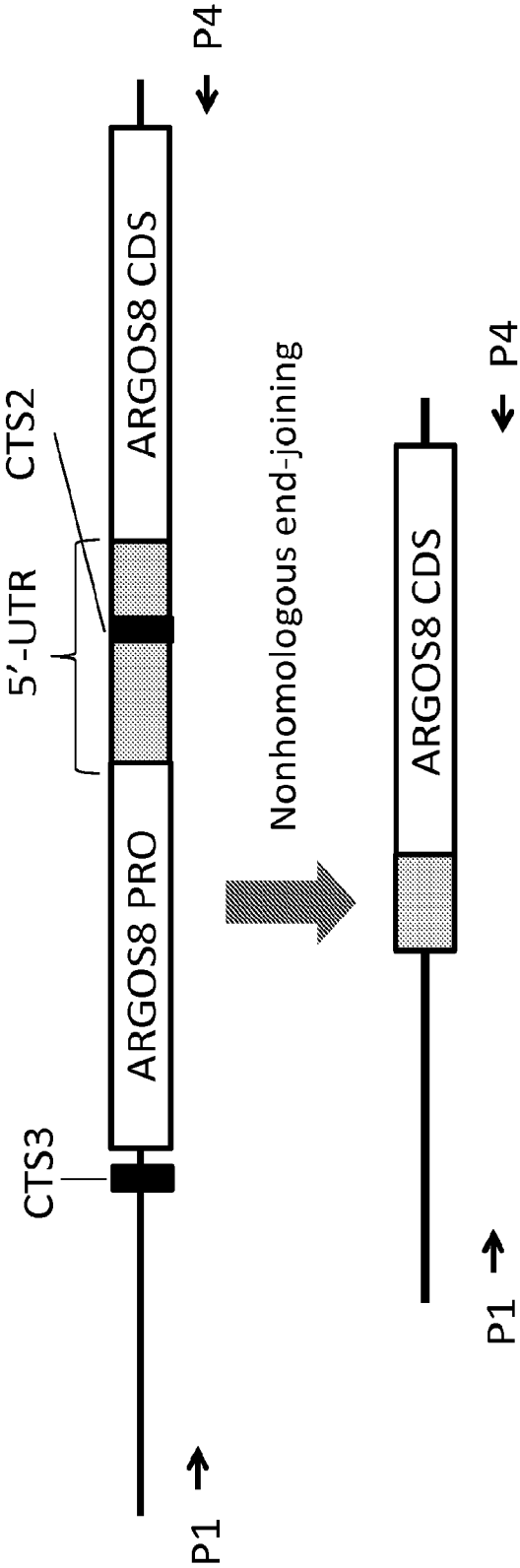


Figure 32B

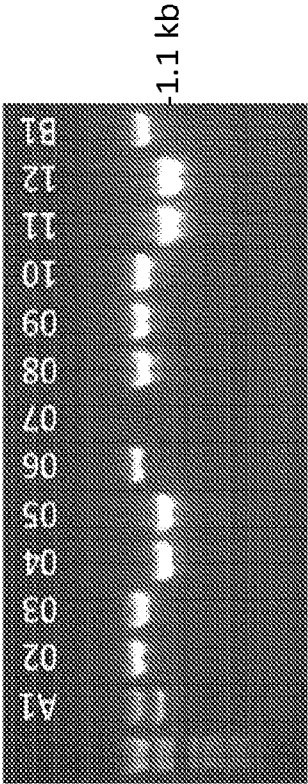


Figure 33

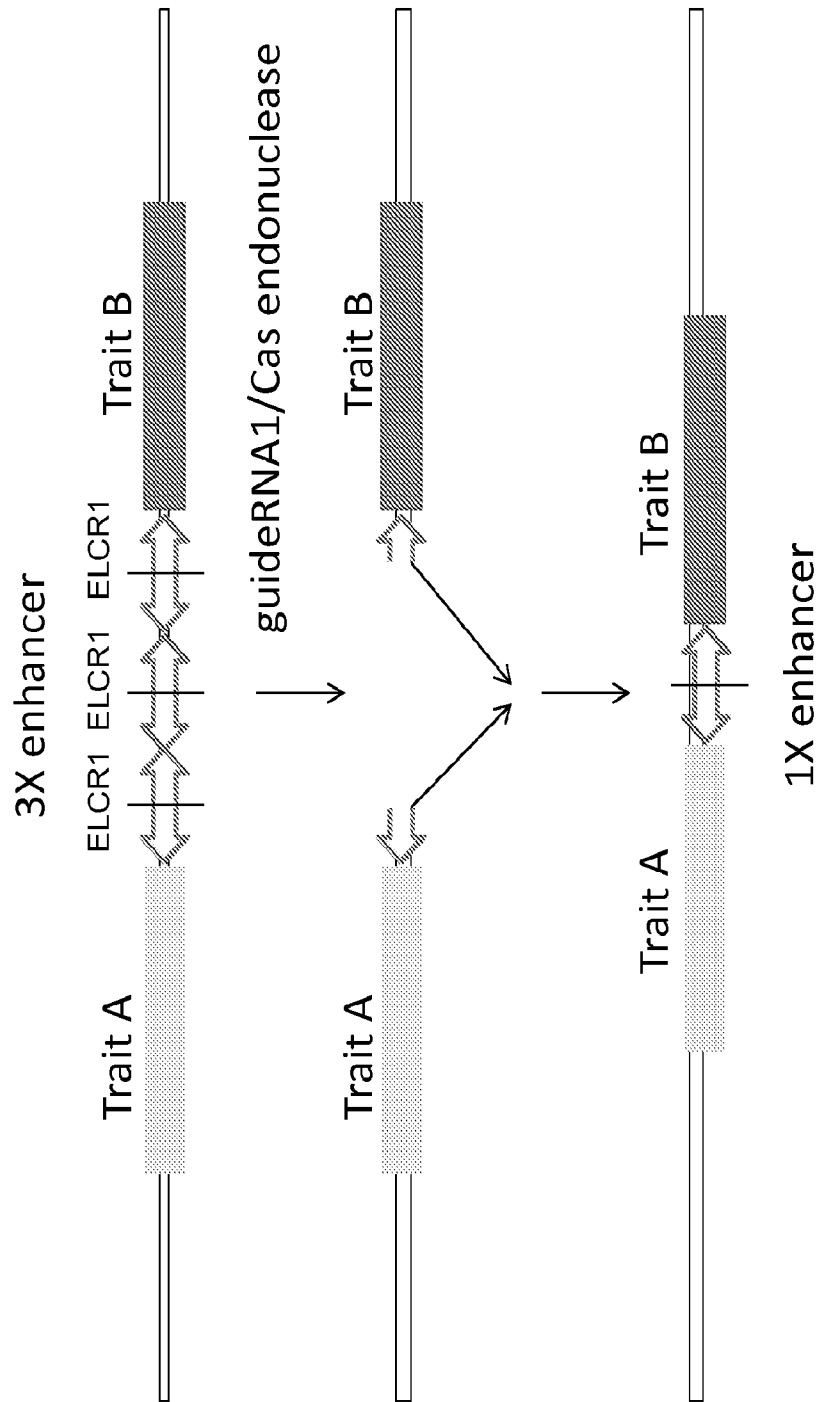


Figure 34 A

VEDAKEEV	Maize
GKESKEEI	Petunia
GKSEEEI	Tomato
EKDAKEEV	Sorghum
VEDSKEEV	Rice
GKDGKEEI	Amarathus

Figure 34 B

K GCTAAAGAGGAAAGTGCAGCTCTTCTTGGGGAATGCTGGAACTGGCAATGGGGCCATTGACAGCAGCTGTTACTGGCTGCTGG

T **P** **moCas9 target sequence**

Figure 34 C

R GCTAGAGAGGAAGTGCAGCTCTTCTTGGGGAATGCTGGAAATCGGAATGGGGTCATTGACAGCAGCTGTTACTGGCTGCTGG

I **S** **moCas9 target sequence**

Figure 35 A

CATATCTG

Figure 35 B

CATCTC...ACGATCAGAT...GCACCGCATGTCCGATGCCTA

Figure 35 C

CAT**A**TCTGCACGATCAGAT**AT**GCACCGCATGTCCGAT**ATCTG**

Figure 36 A

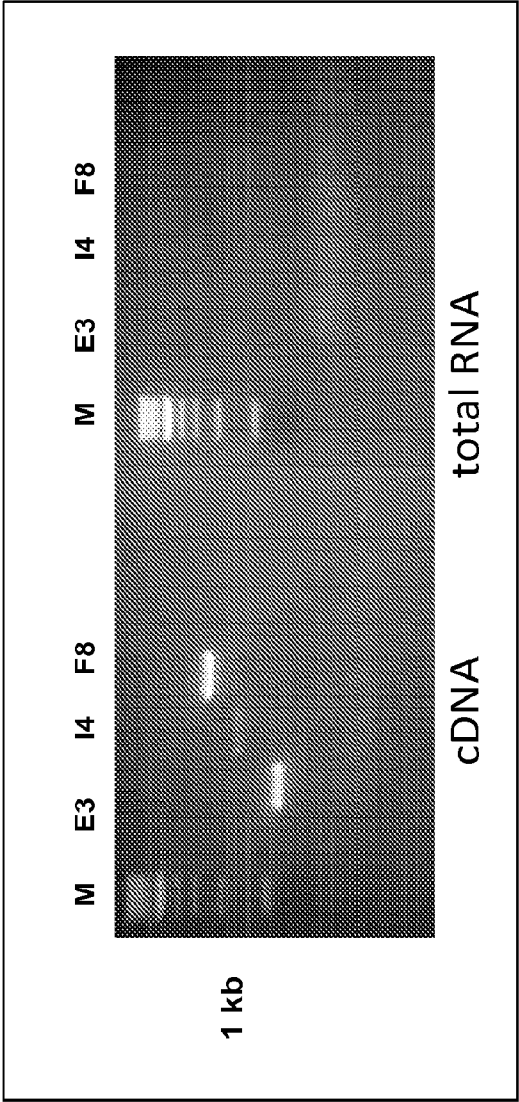


Figure 36 B

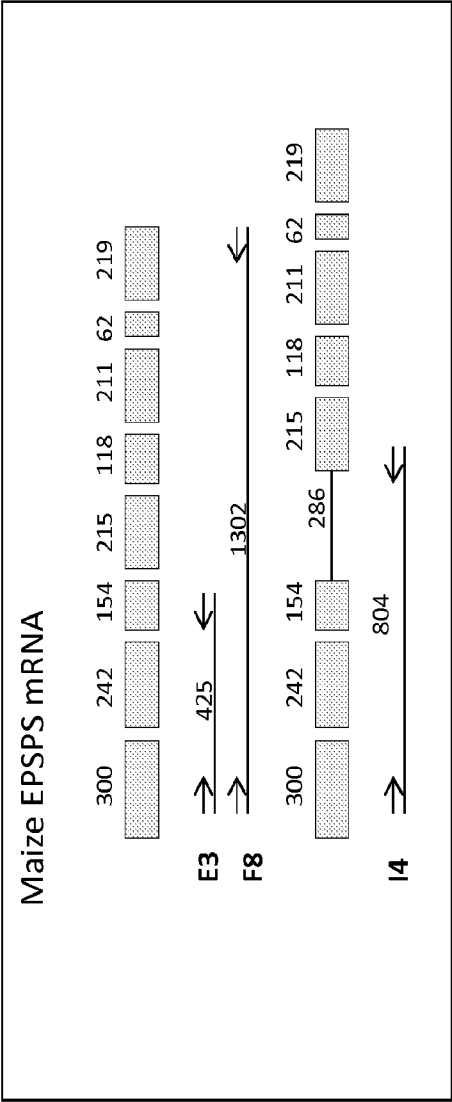


Figure 37

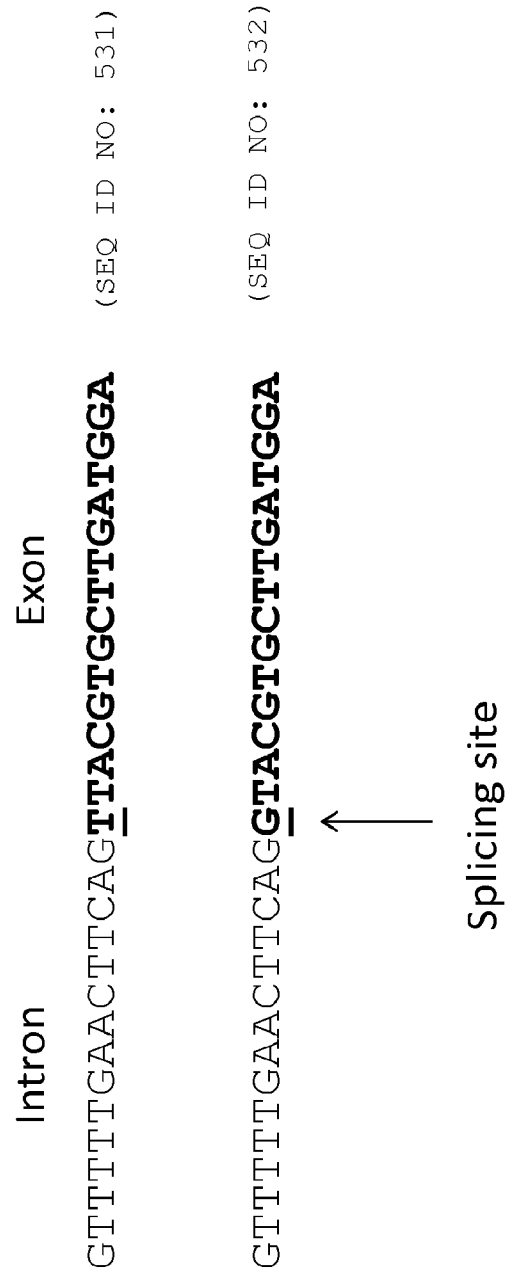


Figure 38

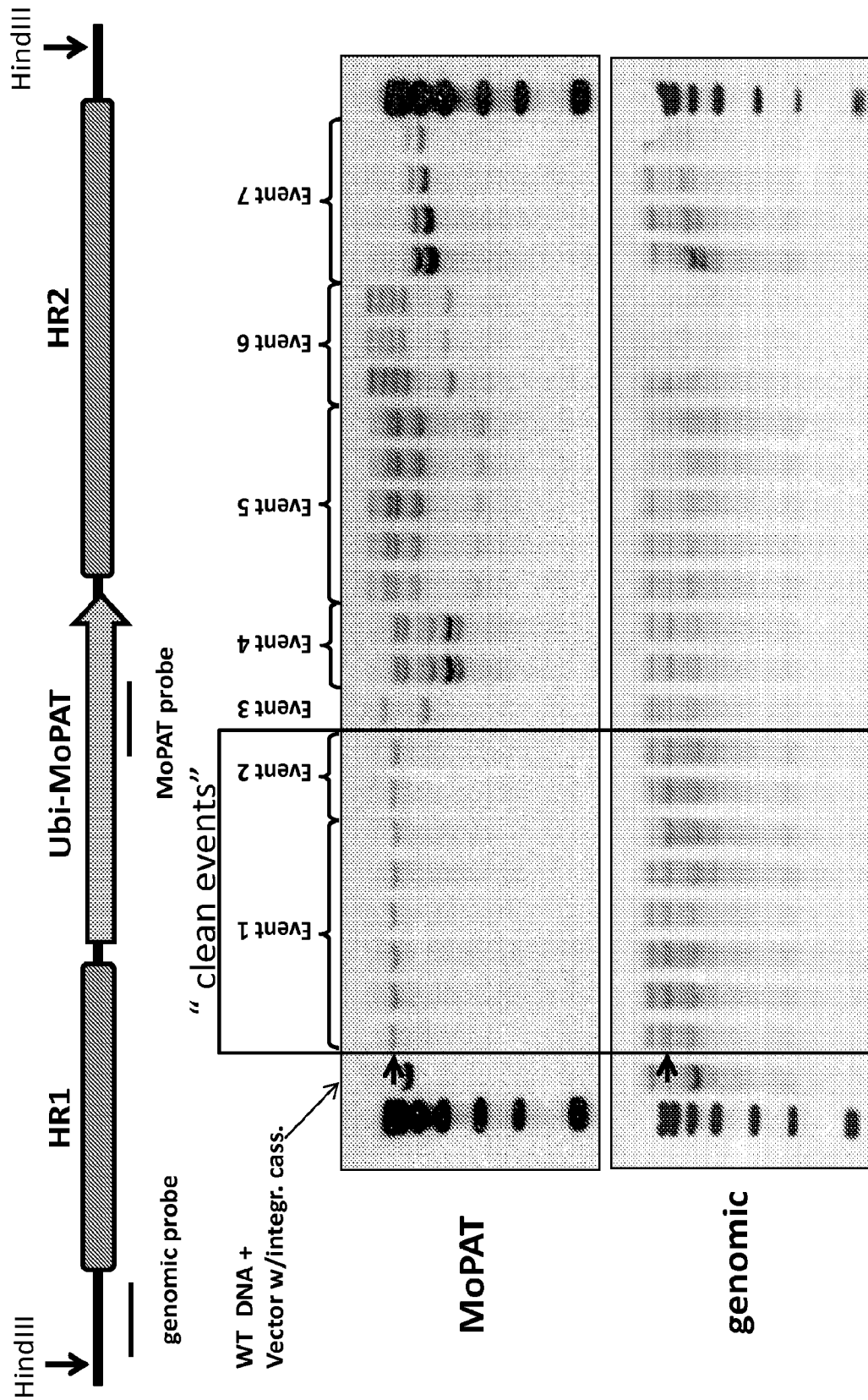
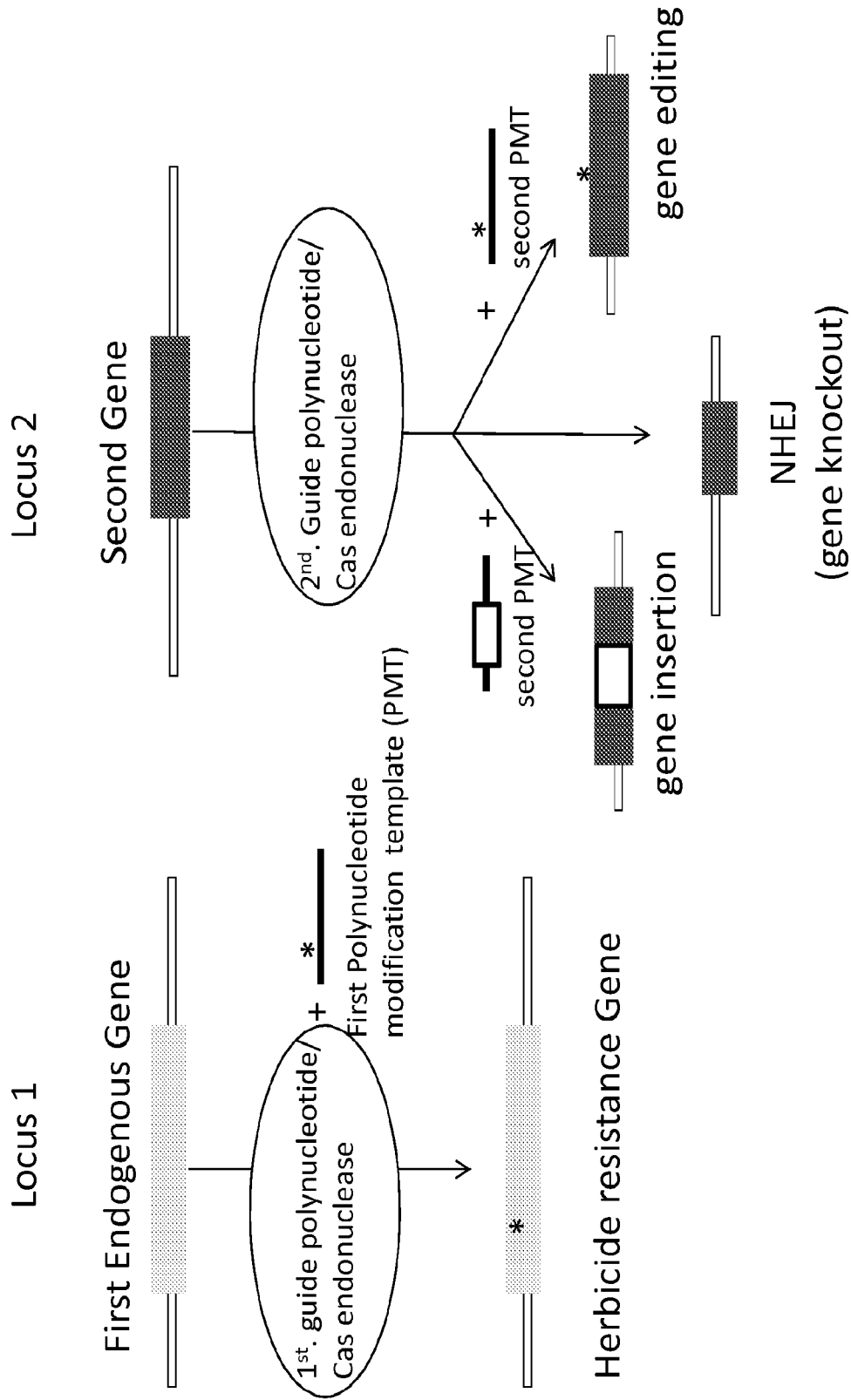


Figure 39



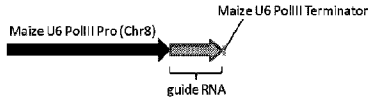
A

Maize codon optimized *Streptococcus pyogenes* **Cas9** (N & C Terminal NLSs)



B

Maize U6 PolIII snRNA promoter driving **guide RNA**



C

Cas9 and **guide RNA** expression cassettes linked together

