



(43) International Publication Date
18 September 2014 (18.09.2014)

(51) International Patent Classification:

C12N 15/82 (2006.01) A01H 5/00 (2006.01)
C12N 15/86 (2006.01)

(21) International Application Number:

PCT/US2014/028445

(22) International Filing Date:

14 March 2014 (14.03.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/790,694 15 March 2013 (15.03.2013) US

(71) Applicant: REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; 1000 Westgate Drive, Suite 160, St. Paul, Minnesota 55114-8658 (US).

(72) Inventors: VOYTAS, Daniel, F.; 2197 Folwell Avenue, Falcon Heights, Minnesota 55108 (US). ATKINS, Paul; 2129 Midlothian Rd., Roseville, Minnesota 55113 (US). BALTES, Nicholas, J.; 369 Old Highway 8 SW Apt. 306, New Brighton, Minnesota 55112 (US).

(74) Agents: KAYTOR, Elizabeth, N. et al.; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (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, 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).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: ENGINEERING PLANT GENOMES USING CRISPR/CAS SYSTEMS

(57) Abstract: Materials and methods for gene targeting using Clustered Regularly Interspersed Short Palindromic Repeats/CRISPR-associated (CRISPR/Cas) systems are provided herein.



WO 2014/144155 A1

ENGINEERING PLANT GENOMES USING CRISPR/Cas SYSTEMS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims benefit of priority from U.S. Provisional Application Serial No. 61/790,694, filed on March 15, 2013.

5

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

This invention was made with government support under GM 834720 awarded by the National Institutes of Health, and DBI0923827 awarded by the National Science Foundation. The government has certain rights in the invention.

10

TECHNICAL FIELD

This document relates to materials and methods for gene targeting in plants, and particularly to methods for gene targeting that include using CRISPR/Cas systems.

15

BACKGROUND

Technologies enabling the precise modification of DNA sequences within living cells can be valuable for both basic and applied research. Precise genome modification – either targeted mutagenesis or gene targeting (GT) – relies on the DNA-repair machinery of the target cell. With respect to targeted mutagenesis, sequence-specific nuclease (SSN)-mediated DNA double-strand breaks (DSBs) are frequently repaired by the error-prone non-homologous end joining (NHEJ) pathway, resulting in mutations at the break site. On the other hand, if a donor molecule is co-delivered with a SSN, the ensuing DSB can stimulate homologous recombination (HR) of sequences near the break site with sequences present on the donor molecule. Consequently, any modified sequence carried by the donor molecule will be stably incorporated into the genome (referred to as GT). Attempts to implement GT in plants often are plagued by extremely low HR frequencies. The majority of the time, donor DNA molecules integrate illegitimately via NHEJ. This process occurs regardless of the size of the homologous “arms;” increasing the length of homology to approximately 22 kb results in no significant enhancement in GT (ThykJaer

25

et al., *Plant Mol Biol*, 35:523-530, 1997). However, introducing a DSB with a SSN can greatly increase the frequency of GT by HR (Shukla et al., *Nature* 459:437-441, 2009; and Townsend et al., *Nature* 459:442-445, 2009).

5

SUMMARY

This document is based in part on the discovery that the Clustered Regularly Interspersed Short Palindromic Repeats/CRISPR-associated (CRISPR/Cas) system can be used for plant genome engineering. The CRISPR/Cas system provides a relatively simple, effective tool for generating modifications in genomic DNA at selected sites.

10 CRISPR/Cas systems can be used to create targeted DSBs or single-strand breaks, and can be used for, without limitation, targeted mutagenesis, gene targeting, gene replacement, targeted deletions, targeted inversions, targeted translocations, targeted insertions, and multiplexed genome modification through multiple DSBs in a single cell directed by co-expression of multiple targeting RNAs. This technology can be used to

15 accelerate the rate of functional genetic studies in plants, and to engineer plants with improved characteristics, including enhanced nutritional quality, increased resistance to disease and stress, and heightened production of commercially valuable compounds.

In one aspect, this document features a method for modifying the genomic material in a plant cell. The method can include (a) introducing into the cell a nucleic acid comprising a crRNA and a tracrRNA, or a chimeric cr/tracrRNA hybrid, where the

20 crRNA and tracrRNA, or the cr/tracrRNA hybrid, is targeted to a sequence that is endogenous to the plant cell; and (b) introducing into the cell a Cas9 endonuclease molecule that induces a double strand break at or near the sequence to which the crRNA and tracrRNA sequence is targeted, or at or near the sequence to which the cr/tracrRNA

25 hybrid is targeted. The introducing steps can include delivering to the plant cell a nucleic acid encoding the Cas9 endonuclease and a nucleic acid encoding the crRNA and tracrRNA or the cr/tracrRNA hybrid, where the delivering is by a DNA virus (e.g., a geminivirus) or an RNA virus (e.g., a tobnavirus). The introducing steps can include delivering to the plant cell a T-DNA containing a nucleic acid sequence encoding the

30 Cas9 endonuclease and a nucleic acid sequence encoding the crRNA and tracrRNA or the

cr/tracrRNA hybrid, where the delivering is via *Agrobacterium* or *Ensifer*. The nucleic acid sequence encoding the Cas9 endonuclease can be operably linked to a promoter that is constitutive (e.g., a cauliflower mosaic virus 35S promoter), cell specific, inducible, or activated by alternative splicing of a suicide exon. The introducing steps can include

5 microprojectile bombardment of nucleic acid encoding Cas9 and the crRNA and tracrRNA or the cr/tracrRNA hybrid. The nucleic acid sequence encoding the Cas9 endonuclease can be operably linked to a promoter that is constitutive, cell specific, inducible, or activated by alternative splicing of a suicide exon. The plant cell can be from a monocotyledonous plant (e.g., wheat, maize, rice, or *Setaria*), or from a

10 dicotyledonous plant (e.g., tomato, soybean, tobacco, potato, cassava, or *Arabidopsis*). The method can further include screening the plant cell after the introducing steps to determine if a double strand break has occurred at or near the sequence targeted by the crRNA and tracrRNA or the cr/tracrRNA hybrid. The method also can include regenerating a plant from the plant cell, and in some embodiments, the method can

15 include cross breeding the plant to obtain a genetically desired plant lineage.

In another aspect, this document features a plant cell containing a nucleic acid encoding a polypeptide having at least 80% sequence identity with SEQ ID NO:12, as well as a plant cell containing a nucleic acid encoding a polypeptide that includes an amino acid sequence having at least 80% sequence identity with amino acids 810 to 872

20 of SEQ ID NO:12.

In another aspect, this document features a virus vector containing a nucleotide sequence that encodes a Cas9 polypeptide. The virus vector can contain a nucleotide sequence encoding a polypeptide with an amino acid sequence having at least 90% identity to SEQ ID NO:12. The virus vector can be from a tobnavirus or a geminivirus.

25

In another aspect, this document features a T-DNA containing a nucleic acid sequence encoding a polypeptide that has an amino acid sequence having at least 80% sequence identity with amino acids 810 to 872 of SEQ ID NO:12. This document also features an *Agrobacterium* strain containing the T-DNA.

25

In yet another aspect, this document features a method for expressing a Cas

30 protein in a plant cell. The method can include providing an *Agrobacterium* or *Ensifer*

vector containing a T-DNA that includes a nucleic acid sequence encoding a polypeptide having an amino acid sequence with at least 80% sequence identity to amino acids 810 to 872 of SEQ ID NO:12, where the polypeptide-encoding sequence is operably linked to a promoter; bringing the *Agrobacterium* or *Ensifer* vector into contact with the plant cell;
5 and expressing the nucleic acid sequence in the plant cell. The promoter can be an inducible promoter (e.g., an estrogen inducible promoter). The method can further include contacting the plant cell with a nucleic acid encoding a guide RNA that associates with the Cas protein. The plant cell can be a protoplast.

Unless otherwise defined, all technical and scientific terms used herein have the
10 same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used to practice the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the
15 present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and from the claims.
20

DESCRIPTION OF DRAWINGS

FIG. 1 is a schematic of a pMDC32 plasmid (a standard T-DNA expression plasmid) containing a Cas9 coding sequence and a cr/tracrRNA hybrid sequence. The nucleotide sequence of the plasmid is set forth in SEQ ID NO:6.

25 FIG. 2 is a schematic of a pFZ19 plasmid (an estrogen-inducible T-DNA expression vector) containing a Cas9 coding sequence and a cr/tracrRNA hybrid sequence. The nucleotide sequence of the plasmid is set forth in SEQ ID NO:7.

FIG. 3 is a schematic of a pNJB121 plasmid (a geminivirus-replicon T-DNA vector) containing a Cas9 coding sequence and a cr/tracrRNA hybrid. The nucleotide
30 sequence of the plasmid is set forth in SEQ ID NO:8.

FIGS. 4A-4D provide evidence of CRISPR/Cas function in plant cells in which a Cas9 coding sequence and a cr/tracrRNA hybrid were delivered by *Agrobacterium* or geminivirus replicons. FIG. 4A is an illustration of a T-DNA harboring a plant codon-optimized Cas9 sequence. The cr/tracrRNA hybrid (designated sgRNA) was placed downstream of the *Arabidopsis* AtU6-26 promoter (PU6). The “lollypops” indicate the long intergenic region (LIR) that is important for replication mediated by replicase (Rep). The gray box represents the short intergenic region (SIR) that also is important for replicon function. The unlabeled gray arrow is a 35S promoter that can drive Cas9 expression upon circularization of the replicon. Cas9 expression also can be driven by the LIR, which functions as a promoter. The entire construct depicted is referred to as an LSL T-DNA. FIG. 4B is a picture of an agarose gel containing PCR products, demonstrating circularization of the geminivirus replicon in plant cells. PCR primers (small arrows in FIG. 4A) were used to amplify DNA from cells infected with *Agrobacterium* T-DNA carrying the replicon. Only in the presence of the plasmid encoding the geminivirus replicase (pRep) did circularization and amplification of the replicon occur. FIG. 4C shows detection of Cas9-induced mutations at the *Nicotiana tabacum* *SurA/SurB* loci. Tobacco leaf tissue was syringe infiltrated with two strains of *Agrobacterium* containing pREP and the LSL T-DNA depicted in FIG. 4A; this was done to test for CRISPR/Cas9-mediated mutagenesis using geminivirus replicons. Alternatively, leaf tissue was infiltrated with single strain of *Agrobacterium* containing only the LSL T-DNA; this was done to test for CRISPR/Cas9-mediated mutagenesis by standard *Agrobacterium* T-DNA delivery. Five days post infiltration, genomic DNA was isolated and used as a template in a PCR reaction designed to amplify the Cas9 target site within *SurA/SurB*. The resulting amplicons were digested with AlwI, and bands were separated by gel electrophoresis. FIG. 4D shows sequences (SEQ ID NOS:1-5) that resulted from cleavage-resistant amplicons in the sample transformed with the LSL T-DNA and pREP T-DNA. PAM, protospacer adjacent motif.

FIG. 5 is a schematic of a reporter plasmid encoding a non-functional yellow fluorescent protein (YFP).

FIG. 6 is a graph plotting fluorescence levels as evidence of CRISPR/Cas function in protoplasts using a YFP reporter plasmid. Tobacco protoplasts were prepared and transformed with various constructs to test for targeted cleavage by CRISPR/Cas9, and YFP fluorescence was measured by flow cytometry. Column 1 shows levels of fluorescence observed from cells transformed with the YFP reporter and constructs expressing Cas9 and the cr/tracr RNA expressed from the AtU6-26 promoter. Column 2 shows levels of fluorescence observed from cells transformed with the reporter, Cas9 and the cr/tracr RNA expressed from the At7SL2-2 promoter. Column 3 shows fluorescence observed in cells transformed with the reporter only (negative control); column 4 shows fluorescence in cells transformed with a construct that expresses YFP (positive control).

DETAILED DESCRIPTION

Efficient genome engineering in plants can be enabled by introducing targeted double-strand breaks (DSBs) in a DNA sequence to be modified. The DSBs activate cellular DNA repair pathways, which can be harnessed to achieve desired DNA sequence modifications near the break site. Targeted DSBs can be introduced using sequence-specific nucleases (SSNs), a specialized class of proteins that includes transcription activator-liked (TAL) effector endonucleases, zinc-finger nucleases (ZFNs), and homing endonucleases (HEs). Recognition of a specific DNA sequence is achieved through interaction with specific amino acids encoded by the SSNs. Prior to the development of TAL effector endonucleases, a challenge of engineering SSNs was the unpredictable context dependencies between amino acids that bind to DNA sequence. While TAL effector endonucleases greatly alleviated this difficulty, their large size (on average, each TAL effector endonuclease monomer contains 2.5-3 kb of coding sequence) and repetitive nature may hinder their use in applications where vector size and stability is a concern (Voytas, *Annu Rev Plant Biol*, 64: 327-350, 2013).

This document is based in part on the discovery that the CRISPR/Cas system can be used as a simple, effective tool for plant genome engineering. CRISPR/Cas molecules are components of a prokaryotic adaptive immune system that uses RNA base pairing to direct DNA cleavage. Directing DNA DSBs requires two components: the Cas9 protein,

which functions as an endonuclease, and CRISPR RNA (crRNA) and tracrRNA (tracrRNA) sequences that aid in directing the Cas9/RNA complex to target DNA sequence (Makarova et al., *Nat Rev Microbiol*, 9(6):467-477, 2011). The modification of a single targeting RNA can be sufficient to alter the nucleotide target of a Cas protein. In
5 some cases, crRNA and tracrRNA can be engineered as a single cr/tracrRNA hybrid to direct Cas9 cleavage activity (Jinek et al., *Science*, 337(6096):816-821, 2012). The CRISPR/Cas system can be used in bacteria, yeast, humans, and zebrafish, as described elsewhere (see, e.g., Jiang et al., *Nat Biotechnol*, 31(3):233-239, 2013; Dicarlo et al.,
Nucleic Acids Res, doi:10.1093/nar/gkt135, 2013; Cong et al., *Science*, 339(6121):819-
10 823, 2013; Mali et al., *Science*, 339(6121):823-826, 2013; Cho et al., *Nat Biotechnol*, 31(3):230-232, 2013; and Hwang et al., *Nat Biotechnol*, 31(3):227-229, 2013).

The utility of the CRISPR/Cas system in plants has not previously been demonstrated. The CRISPR/Cas system originates from prokaryotic cells with relatively small genomes, in which Cas9 is stably expressed in cells in the presence of significant
15 RNase III activity. Thus, when the plant cell work described herein was initiated, there was uncertainty as to whether expression of a Cas9 transgene would be possible in plant cells, and whether Cas9 would properly cooperate with RNA-guides and RNase III activity in the plant context. In addition, expression of heterologous proteins in plant cells is generally challenging due to different codon usage. Further, some toxicity from
20 Cas9 expression in plants was expected, as the large size of plant genomes increases the probability that nonspecific cleavage of genomic DNA may induce genotoxicity to the cells. The CRISPR/Cas9 system is reported to operate with specific recognition sequences comprising 10-20 nucleotides, which is less specific than most other rare-cutting endonuclease systems such as TAL effector endonucleases, meganucleases, and
25 zinc finger nucleases.

As described herein, CRISPR/Cas systems can be used to create targeted DSBs or single-strand breaks, and can be used for, without limitation, targeted mutagenesis, gene targeting, gene replacement, targeted deletions, targeted inversions, targeted translocations, targeted insertions, and multiplexed genome modification through
30 multiple DSBs in a single cell directed by co-expression of multiple targeting RNAs. This

technology can be used to accelerate the rate of functional genetic studies in plants, and to engineer plants with improved characteristics, including enhanced nutritional quality, increased resistance to disease and stress, and heightened production of commercially valuable compounds. Proof-of-concept experiments can be performed in plant leaf tissue
5 by targeting DSBs to reporter genes and endogenous loci. The technology then can be adapted for use in protoplasts and whole plants, and in viral-based delivery systems. Finally, multiplex genome engineering can be demonstrated by targeting DSBs to multiple sites within the same genome.

In general, the systems and methods described herein include at least two
10 components: the RNAs (crRNA and tracrRNA, or a single cr/tracrRNA hybrid) complementary (and thus targeted) to a particular sequence in a plant cell (e.g., in a plant genome, or in an extrachromosomal plasmid, such as a reporter), and a Cas9 endonuclease that can cleave the plant DNA at the target sequence. A representative Cas9 coding sequence is shown in nucleotides 9771 to 14045 of SEQ ID NO:6 (also
15 nucleotides 4331 to 8605 of SEQ ID NO:7, and nucleotides 9487 to 13761 of SEQ ID NO:8). In some cases, a system also can include a nucleic acid containing a donor sequence targeted to a plant sequence. The endonuclease can to create targeted DNA double-strand breaks at the desired locus (or loci), and the plant cell can repair the double-strand break using the donor DNA sequence, thereby incorporating the
20 modification stably into the plant genome.

The Cas9 protein includes two distinct active sites – a RuvC-like nuclease domain and a HNH-like nuclease domain, which generate site-specific nicks on opposite DNA strands (Gasiunas et al., *Proc Natl Acad Sci USA* 109(39):E2579-E2586, 2012). The RuvC-like domain is near the amino terminus of the Cas9 protein and is thought to cleave
25 the target DNA noncomplementary to the crRNA, while the HNH-like domain is in the middle of the protein and is thought to cleave the target DNA complementary to the crRNA. A representative Cas9 sequence from *Streptococcus thermophilus* is set forth in SEQ ID NO:11 (*see, also, UniProtKB number Q03JI6*), and a representative Cas9 sequence from *S. pyogenese* is set forth in SEQ ID NO:12 (*see, also, UniProtKB number Q99ZW2*). Thus, the methods described herein can be carried out using a nucleotide
30

sequence encoding a Cas9 polypeptide having the sequence of SEQ ID NO:11 or SEQ ID NO:12. In some embodiments, however, the methods described herein can be carried out using a nucleotide sequence encoding a Cas9 functional variant having at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) sequence identity with SEQ ID NO:11 or SEQ ID NO:12. Further, Cas9 can be split into two portions, with one portion including the HNH domain and the other including the RuvC domain. The HNH domain may have some cleavage activity by itself in association with the RNA-guide, so this document also contemplates the use of Cas9 polypeptides containing an HNH domain with at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) sequence identity with the HNH domain within SEQ ID NO:11 (e.g., amino acids 828 to 879 of SEQ ID NO:11) or SEQ ID NO:12 (e.g., amino acids 810 to 872 of SEQ ID NO:12).

The percent sequence identity between a particular nucleic acid or amino acid sequence and a sequence referenced by a particular sequence identification number is determined as follows. First, a nucleic acid or amino acid sequence is compared to the sequence set forth in a particular sequence identification number using the BLAST 2 Sequences (B12seq) program from the stand-alone version of BLASTZ containing BLASTN version 2.0.14 and BLASTP version 2.0.14. This stand-alone version of BLASTZ can be obtained online at fr.com/blast or at ncbi.nlm.nih.gov. Instructions explaining how to use the B12seq program can be found in the readme file accompanying BLASTZ. B12seq performs a comparison between two sequences using either the BLASTN or BLASTP algorithm. BLASTN is used to compare nucleic acid sequences, while BLASTP is used to compare amino acid sequences. To compare two nucleic acid sequences, the options are set as follows: -i is set to a file containing the first nucleic acid sequence to be compared (e.g., C:\seq1.txt); -j is set to a file containing the second nucleic acid sequence to be compared (e.g., C:\seq2.txt); -p is set to blastn; -o is set to any desired file name (e.g., C:\output.txt); -q is set to -1; -r is set to 2; and all other options are left at their default setting. For example, the following command can be used to generate an output file containing a comparison between two sequences: C:\B12seq -i c:\seq1.txt -j c:\seq2.txt -p blastn -o c:\output.txt -q -1 -r 2. To compare two amino acid sequences, the options of B12seq are set as follows: -i is set to a file containing the first

amino acid sequence to be compared (e.g., C:\seq1.txt); -j is set to a file containing the second amino acid sequence to be compared (e.g., C:\seq2.txt); -p is set to blastp; -o is set to any desired file name (e.g., C:\output.txt); and all other options are left at their default setting. For example, the following command can be used to generate an output file
5 containing a comparison between two amino acid sequences: C:\BI2seq -i c:\seq1.txt -j c:\seq2.txt -p blastp -o c:\output.txt. If the two compared sequences share homology, then the designated output file will present those regions of homology as aligned sequences. If the two compared sequences do not share homology, then the designated output file will not present aligned sequences.

10 Once aligned, the number of matches is determined by counting the number of positions where an identical nucleotide or amino acid residue is presented in both sequences. The percent sequence identity is determined by dividing the number of matches either by the length of the sequence set forth in the identified sequence (e.g., SEQ ID NO:11), or by an articulated length (e.g., 100 consecutive nucleotides or amino
15 acid residues from a sequence set forth in an identified sequence), followed by multiplying the resulting value by 100. For example, an amino acid sequence that has 1300 matches when aligned with the sequence set forth in SEQ ID NO:11 is 93.7 percent identical to the sequence set forth in SEQ ID NO:11 (i.e., $1300 \div 1388 \times 100 = 93.7$). It is noted that the percent sequence identity value is rounded to the nearest tenth. For
20 example, 75.11, 75.12, 75.13, and 75.14 is rounded down to 75.1, while 75.15, 75.16, 75.17, 75.18, and 75.19 is rounded up to 75.2. It also is noted that the length value will always be an integer.

As used herein, the term “functional variant” is intended to refer to a catalytically active mutant of a protein or a protein domain. Such a mutant can have the same level of
25 activity, or a higher or lower level of activity as compared to the parent protein or protein domain.

The construct(s) containing the crRNA, tracrRNA, cr/tracrRNA hybrid, endonuclease coding sequence, and, where applicable, donor sequence, can be delivered to a plant, plant part, or plant cell using, for example, biolistic bombardment.

30 Alternatively, the system components can be delivered using *Agrobacterium*-mediated

transformation. In some embodiments, the system components can be delivered in a viral vector (e.g., a vector from a DNA virus such as, without limitation, geminivirus (e.g., cabbage leaf curl virus, bean yellow dwarf virus, wheat dwarf virus, tomato leaf curl virus, maize streak virus, tobacco leaf curl virus, or tomato golden mosaic virus) or nanovirus (e.g., Faba bean necrotic yellow virus), or a vector from an RNA virus such as, without limitation, tobnavirus (e.g., tobacco rattle virus, tobacco mosaic virus), potexvirus (e.g., potato virus X), or hordeivirus (e.g., barley stripe mosaic virus).

After a plant, plant part, or plant cell is infected or transfected with an endonuclease encoding sequence and a crRNA and a tracrRNA, or a cr/tracrRNA hybrid (and, in some cases, a donor sequence), any suitable method can be used to determine whether GT or targeted mutagenesis has occurred at the target site. In some embodiments, a phenotypic change can indicate that a donor sequence has been incorporated into the target site. PCR-based methods also can be used to ascertain whether a genomic target site contains targeted mutations or donor sequence, and/or whether precise recombination has occurred at the 5' and 3' ends of the donor. One method to detect targeted mutations, referred to herein as "PCR digest," is described by Zhang et al. (*Proc Natl Acad Sci USA* 107:12028-12033, 2010). Methods to detect precise recombination include southern blotting using a probe with homology to the donor sequence.

In some embodiments, the methods provided herein can include introducing into a plant, plant part, or plant cell a nucleic acid that includes a crRNA and a tracrRNA, or a chimeric cr/tracrRNA hybrid, where the crRNA and tracrRNA, or the cr/tracrRNA hybrid, is targeted to a nucleotide sequence that is endogenous to the plant cell, and also introducing into the plant, plant part, or plant cell a Cas9 endonuclease molecule (e.g., a Cas9 polypeptide or a portion thereof, such as a portion of a Cas9 polypeptide that includes the HNH domain, or a nucleic acid encoding a Cas9 polypeptide or a portion thereof), where the Cas9 endonuclease molecule induces a double strand break at or near the sequence to which the crRNA and tracrRNA sequences (or the cr/tracrRNA hybrid) are targeted.

The plants, plant parts, and plant cells used in the methods provided herein can be from any species of plant. In some embodiments, for example, the methods provided herein can utilize monocotyledonous plants, portions thereof, or cells therefrom. Exemplary monocotyledonous plants include, without limitation, wheat, maize, rice, orchids, onion, aloe, true lilies, grasses (e.g., *Setaria*), woody shrubs and trees (e.g., palms and bamboo), and food plants such as pineapple and sugar cane. Exemplary dicotyledonous plants include, without limitation, tomato, cassava, soybean, tobacco, potato, *Arabidopsis*, rose, pansy, sunflower, grape, strawberry, squash, bean, pea, and peanut.

In some embodiments, the methods described herein can include screening the plant, plant part, or plant cell to determine if a DSB has occurred at or near the sequence targeted by the crRNA and tracrRNA or the cr/tracrRNA hybrid. For example, the PCR-digest assay described by Zhang et al. (*supra*) can be used to determine whether a DSB has occurred. Other useful methods include, without limitation, the T7 assay, the Surveyor assay, and southern blotting (if a restriction enzyme binding sequence is present at or near the predicted cleavage site).

In addition, in some embodiments in which a plant part or plant cell is used, the methods provided herein can include regenerating a plant from the plant part or plant cell. The methods also can include breeding the plant (e.g., the plant into which the nucleic acids were introduced, or the plant obtained after regeneration of the plant part or plant cell used as a starting material) to obtain a genetically desired plant lineage. Methods for regenerating and breeding plants are well established in the art.

Also provided herein are plants, plant parts, and plant cells containing a nucleic acid that encodes a Cas9 polypeptide with an amino acid sequence that is at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) identical to the amino acid sequence set forth in SEQ ID NO:11 or SEQ ID NO:12, or a nucleic acid that encodes a Cas9 polypeptide containing an amino acid sequence that is at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) identical to amino acids 828 to 879 of SEQ ID NO:11, or amino acids 810 to 872 of SEQ ID NO:12.

This document also provides virus vectors that contain nucleotide sequences encoding Cas9 polypeptides. For example, a virus vector can include a nucleotide sequence encoding a polypeptide having an amino acid sequence that is at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) identical to the amino acid sequence set forth in SEQ ID NO:11 or SEQ ID NO:12. In some embodiments, a virus vector can have a nucleotide sequence encoding a Cas9 polypeptide that includes an amino acid sequence with at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) sequence identity to amino acids 828 to 879 of SEQ ID NO:11, or amino acids 810 to 872 of SEQ ID NO:12. The vector can be from any suitable type of virus, such as a tobnavirus or a geminivirus, for example.

Also provide herein are T-DNA molecules that contain a nucleic acid sequence encoding a Cas9 polypeptide having an amino acid sequence that is at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) identical to the amino acid sequence set forth in SEQ ID NO:11 or SEQ ID NO:12. In some embodiments, a T-DNA can include a nucleotide sequence encoding a Cas9 polypeptide that includes an amino acid sequence with at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) sequence identity to amino acids 828 to 879 of SEQ ID NO:11, or amino acids 810 to 872 of SEQ ID NO:12.

This document also provides *Agrobacterium* strains comprising a T-DNA as described herein.

In addition, this document provides methods for expressing a Cas protein in a plant, a plant part, or a plant cell. Such methods can include, for example, (a) providing an *Agrobacterium* or *Ensifer* vector containing a T-DNA that includes a nucleic acid sequence encoding a Cas9 polypeptide having an amino acid sequence with at least 80% (e.g., at least 85%, at least 90%, at least 95%, or at least 98%) sequence identity to SEQ ID NO:11 or SEQ ID NO:12, where the Cas9-encoding sequence is operably linked to a promoter, (b) bringing the *Agrobacterium* or *Ensifer* vector into contact with a plant, plant part, or plant cell, and (c) expressing the nucleic acid sequence in the plant, plant part, or plant cell. The promoter can be, for example, a constitutive promoter (e.g., a CaMV 35S promoter), an inducible promoter (e.g., an estradiol-induced XVE promoter;

Zuo et al., *Plant J* 24:265-273, 2000), a cell specific promoter, or a promoter that is activated by alternative splicing of a suicide exon. In some embodiments, such methods also can include contacting the plant, plant part, or plant cell with a nucleic acid encoding a guide RNA that associates with the Cas protein, and expressing the guide RNA.

5 The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

Example 1 – Plasmids for expressing CRISPR/Cas components

10 To demonstrate functionality of the CRISPR/Cas systems for genome editing in plants, plasmids were constructed to encode Cas9, crRNA and tracrRNA, the cr/tracrRNA hybrid, and RNA polymerase III promoters (e.g., AtU6-26 or At7SL-2) from which to express the crRNA, tracrRNA, or cr/tracrRNA hybrid. Plant codon-optimized Cas9 coding sequence was synthesized and cloned into a MultiSite Gateway
15 entry plasmid. Additionally, crRNA and tracrRNA, or cr/tracrRNA hybrid, driven by the RNA polymerase III (PolIII) promoters AtU6-26 and At7SL2-2, were synthesized and cloned into a second MultiSite Gateway entry plasmid. To enable efficient reconstruction of the crRNA sequences (serving to redirect CRISPR/Cas-mediated DSBs), inverted type-IIS restriction enzyme sites (e.g., *BsaI* and *Esp3I*) were inserted within the crRNA
20 nucleotide sequence. By digesting with the appropriate type-IIS restriction enzyme, target sequences can be efficiently cloned into the crRNA sequence using oligonucleotides. Entry plasmids for both Cas9 and the expression of the crRNA and tracrRNA or the cr/tracrRNA hybrid, from a RNA polymerase III promoter (AtU6-26 or At7SL2-2), were recombined into pMDC32 (a standard T-DNA expression plasmid with a 2x35S
25 promoter; FIG. 1 and SEQ ID NO:6), pFZ19 (an estrogen-inducible T-DNA expression vector; FIG. 2 and SEQ ID NO:7; Zuo et al., *Plant J.* 24(2):265-273, 2000), and pNJB121 (a geminivirus-replicon T-DNA vector; FIG. 3 and SEQ ID NO:8).

Example 2 – CRISPR/Cas activity in somatic plant tissue

To demonstrate the capacity for CRISPR/Cas systems to function as SSNs, the geminivirus-replicon T-DNA vector, pNJB121, was modified to encode both Cas9 and cr/tracrRNA hybrid sequences (FIG. 4A). Targeting RNA sequences (encoded by
5 nucleotide sequence within the crRNA; responsible for directing Cas9 cleavage) were designed to be homologous to sequences within the endogenous *SuRA* and *SuRB* genes. The sequence of the targeting portion of the crRNA that matched the *SuR* loci was 5'-GUGGGAGGAUCGGUUCUAUA (SEQ ID NO:9; the 5' G does not match the *SuR* loci, but is needed for transcription by RNA polymerase III). Although pNJB121 is a
10 geminivirus-replicon, in the absence of replicase (Rep), no amplification occurs. Therefore, pNJB121 in the absence of Rep is a standard T-DNA vector and no replicons are formed. The modified pNJB121 plasmid delivered to *Nicotiana tabacum* leaf tissue by syringe infiltration with *Agrobacterium tumefaciens*. Five days after infiltration, *SuRA/SuRB* sequences were assessed for Cas9-mediated mutations using PCR-digest
15 (FIG. 4C). The presence of mutations at the corresponding target sequences indicated functionality of CRISPR/Cas systems in plant leaf cells.

Example 3 – CRISPR/Cas activity in protoplasts

To further demonstrate the activity of CRISPR/Cas systems in plants, targeted
20 mutagenesis of DNA sequence within *Arabidopsis thaliana* and *Nicotiana tabacum* protoplasts is assessed. Targeting crRNA sequences are redesigned to be homologous to sequences present within the endogenous *ADH1* or *TT4* genes (*Arabidopsis*), or the integrated *gus:nptII* reporter gene or *SuRA/SuRB* (*Nicotiana*). Protoplasts are isolated from *Arabidopsis* and *Nicotiana* leaf tissue and transfected with plasmids encoding Cas9
25 and the *ADH1*- or *TT4*-targeting crRNAs, or Cas9 and the *gus:nptII*- or *SuRA/SuRB*-targeting crRNA, respectively. Genomic DNA is extracted 5-7 days post transfection and assessed for mutations at the corresponding target sequences. Detecting mutations within the *ADH1*, *TT4*, *gus:nptII* or *SuRA/SuRB* genes indicates the functionality of CRISPR/Cas systems to target endogenous genes in plant protoplasts.

In initial studies, the CRISPR/Cas system was assessed for the ability to cleave an extrachromosomal reporter plasmid, using methods similar to those described by Zhang et al. (*Plant Physiol* 161:20-27, 2013). The reporter plasmid encodes a non-functional yellow fluorescent protein (YFP; FIG. 5 and SEQ ID NO:10). YFP expression is
5 disrupted by a direct repeat of internal coding sequence that flanks a target sequence for the Cas9/crRNA complex. The generation of targeted DSBs at the Cas9/crRNA target sequence results in recombination of the direct repeat sequences, thereby restoring YFP gene function. A sequence from the tobacco *SuRA/SuRB* loci was cloned into the YFP reporter between the direct repeats. A cr/tracrRNA hybrid construct that targets this site
10 was then generated. The sequence of the portion of the crRNA that targets the *SuR* loci was 5'- GUGGGAGGAUCGGUUCUAUA (SEQ ID NO:9; again, the 5' G does not match the *SuR* loci, but it is needed for transcription by RNA polymerase III). *Nicotiana tabacum* protoplasts were transformed with plasmids encoding Cas9, a cr/tracrRNA hybrid, and the YFP reporter, and restoration of YFP expression as a result of
15 CRISPR/Cas nuclease activity was monitored by flow cytometry. Using a positive control plasmid that encodes YFP, 94.7% of the cells were transformed and expressed YFP (FIG. 6, column 4). Cells transformed with the reporter alone gave activity levels barely above background (FIG. 6, column 3). When cells were transformed with constructs expressing Cas9 and a cr/tracr RNA, significant activity was observed,
20 indicating the Cas9/crRNA complex cleaved the target. For the cr/tracrRNA expressed from the AtU6-26 promoter, 18.8% of the cells fluoresced (FIG. 6, column 1). When the cr/tracr RNA was expressed from the At7SL2-2 promoter, 20.7% of the cells were YFP positive (FIG. 6, column 2). Detection of YFP-expressing cells indicated the functionality of CRISPR/Cas systems in plant protoplasts.

Example 4 – Multiplex genome engineering in protoplasts using CRISPR/Cas systems

The ability of CRISPR/Cas systems to create multiple DSBs at different DNA sequences is assessed using plant protoplasts. To direct Cas9 nuclease activity to *TT4*, *ADH1*, and the extrachromosomal YFP reporter plasmid (within the same *Arabidopsis*
30 protoplast), crRNA and tracrRNA or cr/tracrRNA hybrid plasmid is modified to express

multiple crRNA targeting sequences. These sequences are designed to be homologous to sequences present within *TT4*, *ADHI* and the YFP reporter plasmid. Following transfection with Cas9, crRNA, tracrRNA, or the cr/tracrRNA hybrid, and YFP reporter plasmids into *Arabidopsis* protoplasts, YFP-expressing cells are quantified and isolated, and genomic DNA is extracted. Observing mutations within the *ADHI* and *TT4* genes in YFP-expressing cells suggests that CRISPR/Cas can facilitate multiplex genome engineering in *Arabidopsis* cells.

To demonstrate multiplex genome engineering in *Nicotiana* protoplasts, plasmids containing multiple crRNA are modified to encode sequences that are homologous to the integrated *gus:nptII* reporter gene, *SuRA/SuRB*, and the YFP reporter plasmid. Similar to the methods described in *Arabidopsis* protoplasts, *Nicotiana* protoplasts are transfected with Cas9, crRNA, tracrRNA, or the cr/tracrRNA hybrid, and YFP reporter plasmids. YFP-expressing cells are quantified and isolated, and genomic DNA is extracted. Observing mutations within the integrated *gus:nptII* reporter gene and *SuRA/SuRB* in YFP-expressing cells suggests that CRISPR/Cas can facilitate multiplex genome engineering in tobacco cells.

Example 5 – CRISPR/Cas activity in planta

To demonstrate CRISPR/Cas activity in *planta*, pFZ19 T-DNA is modified to encode both Cas9 and the crRNA and tracrRNA, or the cr/tracrRNA hybrid sequences. Target DNA sequences are present within the endogenous *ADHI* or *TT4* genes. The resulting T-DNA is integrated into the *Arabidopsis thaliana* genome by floral dip using *Agrobacterium*. Cas9 expression is induced in primary transgenic plants by direct exposure to estrogen. Genomic DNA from somatic leaf tissue is extracted and assessed for mutations at the corresponding genomic locus by PCR-digest. Observing mutations within the *ADHI* or *TT4* genes demonstrates CRISPR/Cas activity in *planta*. Alternatively, CRISPR/Cas activity can be assessed by screening T2 seeds (produced from induced T1 patents) for heterozygous or homozygous mutations at the corresponding genomic locus. Furthermore, the capacity for CRISPR/Cas to carry out multiplex genome engineering is assessed by modifying plasmids containing multiple

crRNAs with homologous sequences to both *ADHI* and *TT4*. The resulting T-DNA plasmid is integrated into the *Arabidopsis* genome, Cas9 expression is induced in primary transgenic plants, and CRISPR/Cas activity is assessed by evaluating the *ADHI* and *TT4* genes in both T1 and T2 plants. Observing mutations in both the *ADHI* and *TT4* genes suggests CRISPR/Cas can facilitate multiplex genome engineering in *Arabidopsis* plants.

Example 6 – Viral delivery of CRISPR/Cas components

Plant viruses can be effective vectors for delivery of heterologous nucleic acid sequence, such as for RNAi reagents or for expressing heterologous proteins. Useful plant viruses include both RNA viruses (e.g., tobacco mosaic virus, tobacco rattle virus, potato virus X, and barley stripe mosaic virus) and DNA viruses (e.g., cabbage leaf curl virus, bean yellow dwarf virus, wheat dwarf virus, tomato leaf curl virus, maize streak virus, tobacco leaf curl virus, tomato golden mosaic virus, and Faba bean necrotic yellow virus; Rybicki et al., *Curr Top Microbiol Immunol*, 2011; and Gleba et al., *Curr Opin Biotechnol* 2007, 134-141). Such plant viruses can be modified for the delivery of CRISPR/Cas9 components. Proof-of-concept experiments were performed in *Nicotiana tabacum* leaf cells using DNA viruses (geminivirus replicons; Baltes et al., *Plant Cell* 26:151-163, 2014). To this end, crRNA sequences were modified to contain homology to the endogenous *SuRA/SuRB* loci. The resulting plasmids were cloned into pNJB121 (a T-DNA destination vector with *cis*-acting elements required for geminivirus replication (LSL T-DNA)) along with Cas9 (FIG. 4A). Co-delivery of LSL T-DNA along with T-DNA encoding replicase protein (Rep; REP T-DNA) by *Agrobacterium* resulted in the replicational release of geminiviral replicons (FIG. 4B). The T-DNA was delivered to tobacco leaf tissue by syringe infiltration with *Agrobacterium*. Five to seven days after infiltration, *SuRA/SuRB* sequences were assessed for Cas9-mediated mutations using PCR-digest (FIG. 4C). Digestion-resistant PCR amplicons were cloned and sequenced. The presence of mutations at the corresponding target sequences indicates that plant viruses are effective vectors for delivery of CRISPR/Cas components (FIG. 4D).

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended
5 claims. Other aspects, advantages, and modifications are within the scope of the following claims.

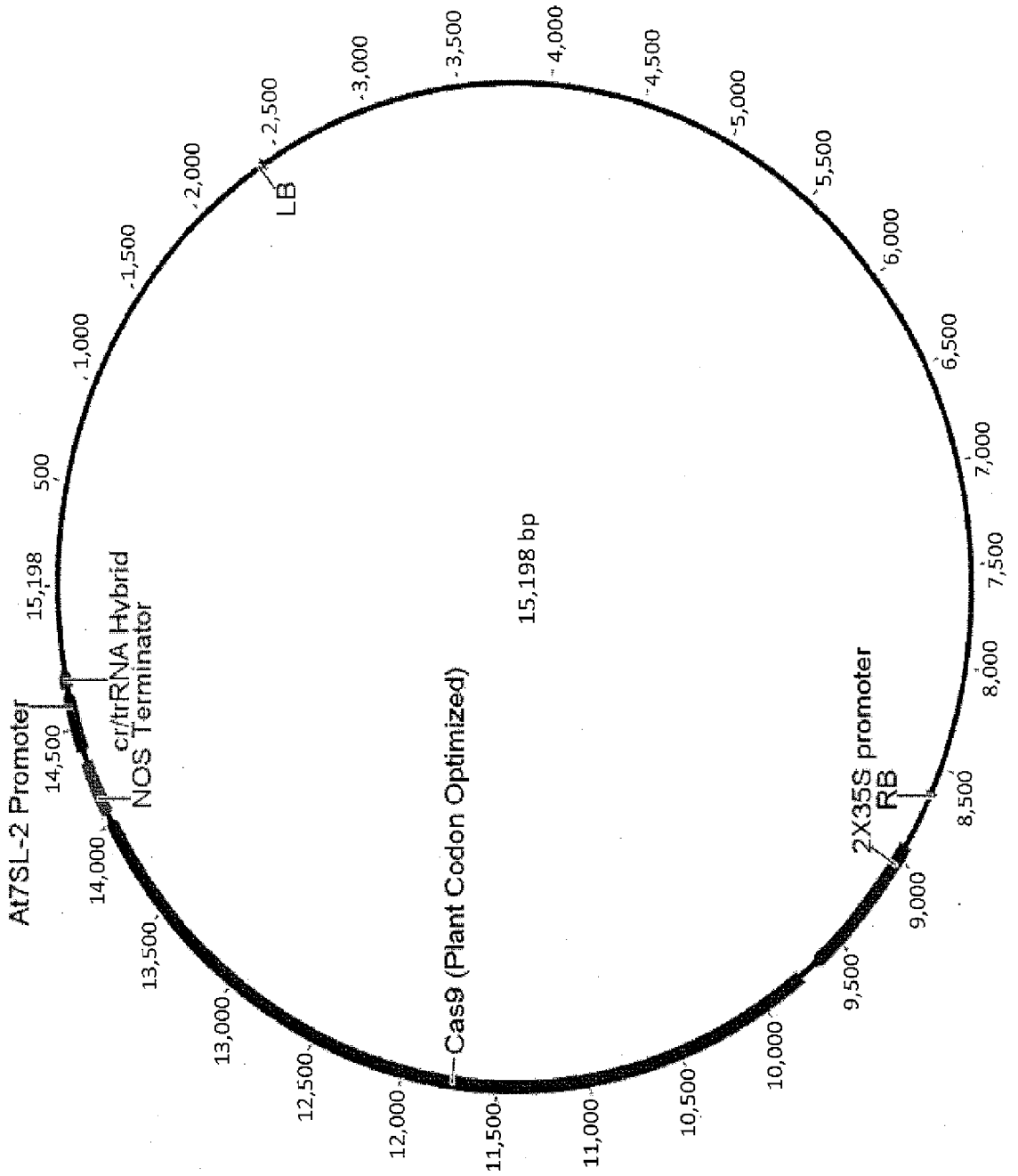
WHAT IS CLAIMED IS:

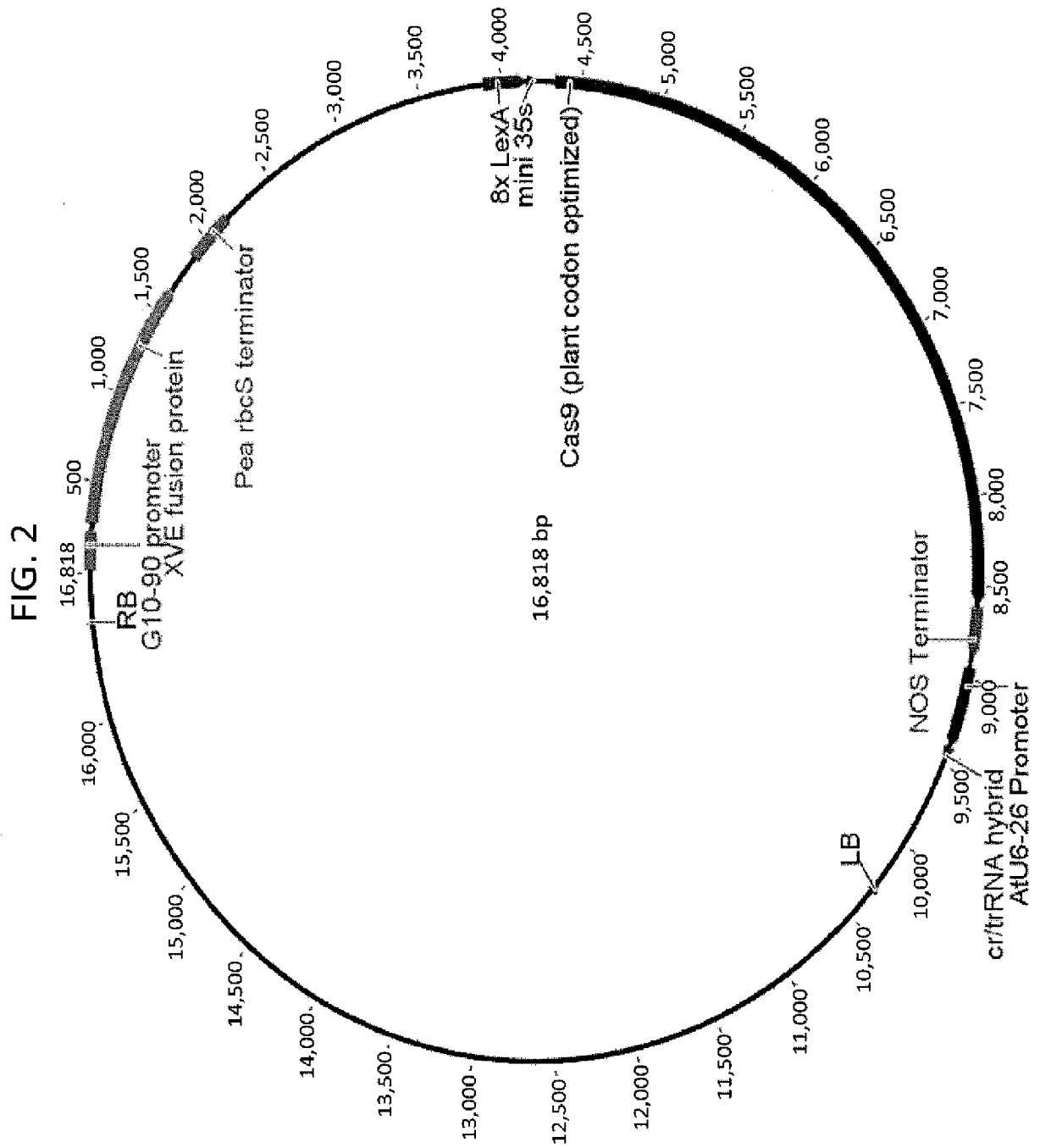
1. A method for modifying the genomic material in a plant cell, comprising:
 - (a) introducing into the cell a nucleic acid comprising a crRNA and a tracrRNA, or a chimeric cr/tracrRNA hybrid, wherein the crRNA and tracrRNA, or the cr/tracrRNA hybrid, is targeted to a sequence that is endogenous to the plant cell; and
 - (b) introducing into the cell a Cas9 endonuclease molecule that induces a double strand break at or near the sequence to which the crRNA and tracrRNA sequence is targeted, or at or near the sequence to which the cr/tracrRNA hybrid is targeted.
2. The method of claim 1, wherein the introducing steps comprise delivering to the plant cell a nucleic acid encoding the Cas9 endonuclease and a nucleic acid encoding the crRNA and tracrRNA or the cr/tracrRNA hybrid, and wherein the delivering is by a DNA or RNA virus.
3. The method of claim 2, wherein the DNA virus is a geminivirus.
4. The method of claim 2, wherein the RNA virus is a tobnavirus.
5. The method of claim 1, wherein the introducing steps comprise delivering to the plant cell a T-DNA containing a nucleic acid sequence encoding the Cas9 endonuclease and a nucleic acid sequence encoding the crRNA and tracrRNA or the cr/tracrRNA hybrid, and wherein the delivering is via *Agrobacterium* or *Ensifer*.
6. The method of claim 5, wherein the nucleic acid sequence encoding the Cas9 endonuclease is operably linked to a promoter that is constitutive, cell specific, inducible, or activated by alternative splicing of a suicide exon.
7. The method of claim 1, wherein the introducing steps comprise microprojectile bombardment of nucleic acid encoding Cas9 and the crRNA and tracrRNA or the cr/tracrRNA hybrid.
8. The method of claim 1, wherein the plant cell is from a monocotyledonous plant.

9. The method of claim 8, wherein the monocotyledonous plant is wheat, maize, rice, or *Setaria*.
10. The method of claim 1, wherein the plant cell is from a dicotyledonous plant.
11. The method of claim 10, wherein the dicotyledonous plant is tomato, soybean, tobacco, potato, cassava, or *Arabidopsis*.
12. The method of claim 1, further comprising screening the plant cell after the introducing steps to determine if a double strand break has occurred at or near the sequence targeted by the crRNA and tracrRNA or the cr/tracrRNA hybrid.
13. The method of claim 1, further comprising regenerating a plant from the plant cell.
14. The method of claim 13, further comprising cross breeding the plant to obtain a genetically desired plant lineage.
15. A plant cell comprising a nucleic acid encoding a polypeptide having at least 80% sequence identity with SEQ ID NO:12.
16. A plant cell comprising a nucleic acid encoding a polypeptide that comprises an amino acid sequence having at least 80% sequence identity with amino acids 810 to 872 of SEQ ID NO:12.
17. A virus vector comprising a nucleotide sequence that encodes a Cas9 polypeptide.
18. The virus vector of claim 17, wherein the vector comprises a nucleotide sequence encoding a polypeptide with an amino acid sequence having at least 90% identity to SEQ ID NO:12.
19. The virus vector of claim 17, wherein the virus is a tobnavirus or a geminivirus.

20. A T-DNA comprising a nucleic acid sequence encoding a polypeptide that comprises an amino acid sequence having at least 80% sequence identity with amino acids 810 to 872 of SEQ ID NO:12.
21. An *Agrobacterium* strain comprising the T-DNA of claim 20.
22. A method for expressing a Cas protein in a plant cell, comprising:
 - providing an *Agrobacterium* or *Ensifer* vector containing a T-DNA that comprises a nucleic acid sequence encoding a polypeptide having an amino acid sequence with at least 80% sequence identity to amino acids 810 to 872 of SEQ ID NO:12, wherein the polypeptide-encoding sequence is operably linked to a promoter;
 - bringing the *Agrobacterium* or *Ensifer* vector into contact with the plant cell; and
 - expressing the nucleic acid sequence in the plant cell.
23. The method of claim 22, wherein the promoter is an inducible promoter.
24. The method of claim 23, wherein the inducible promoter is an estrogen inducible promoter.
25. The method of claim 22, further comprising contacting the plant cell with a nucleic acid encoding a guide RNA that associates with the Cas protein.
26. The method of claim 22, wherein the plant cell is a protoplast.

FIG. 1





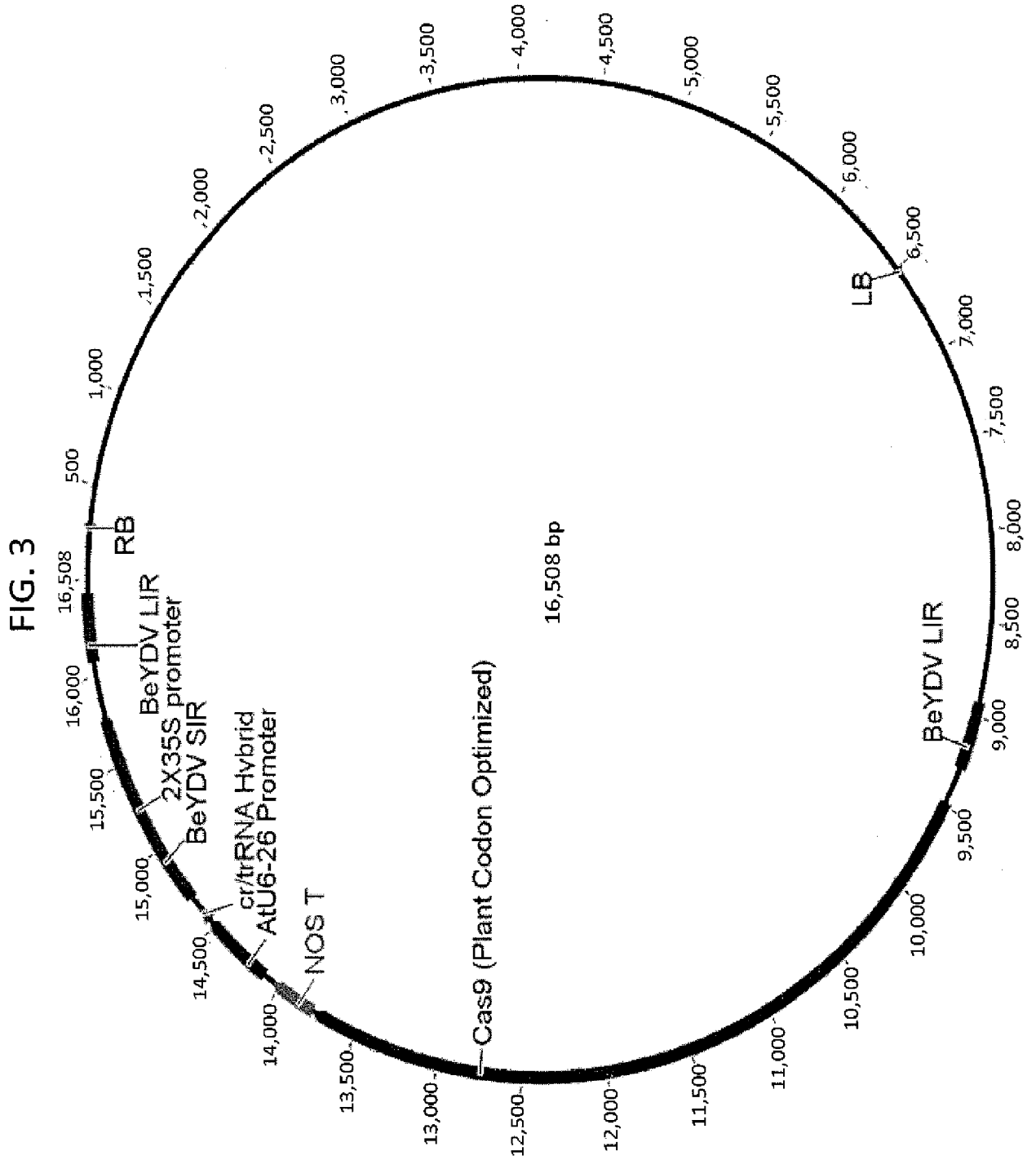


FIG. 4A



FIG. 4B

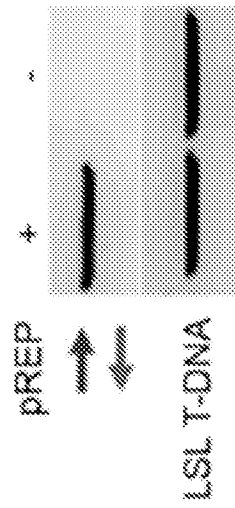


FIG. 4C

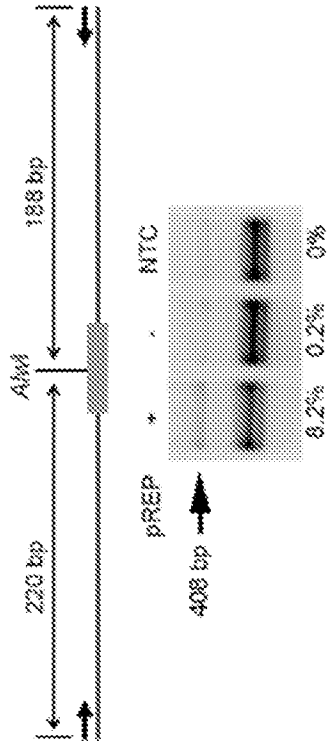
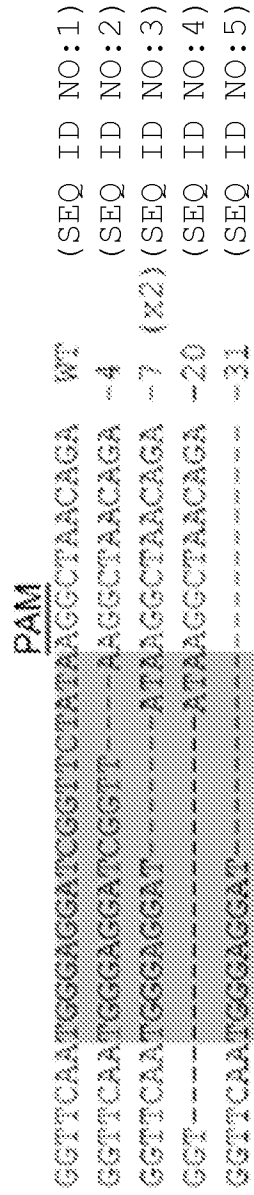


FIG. 4D



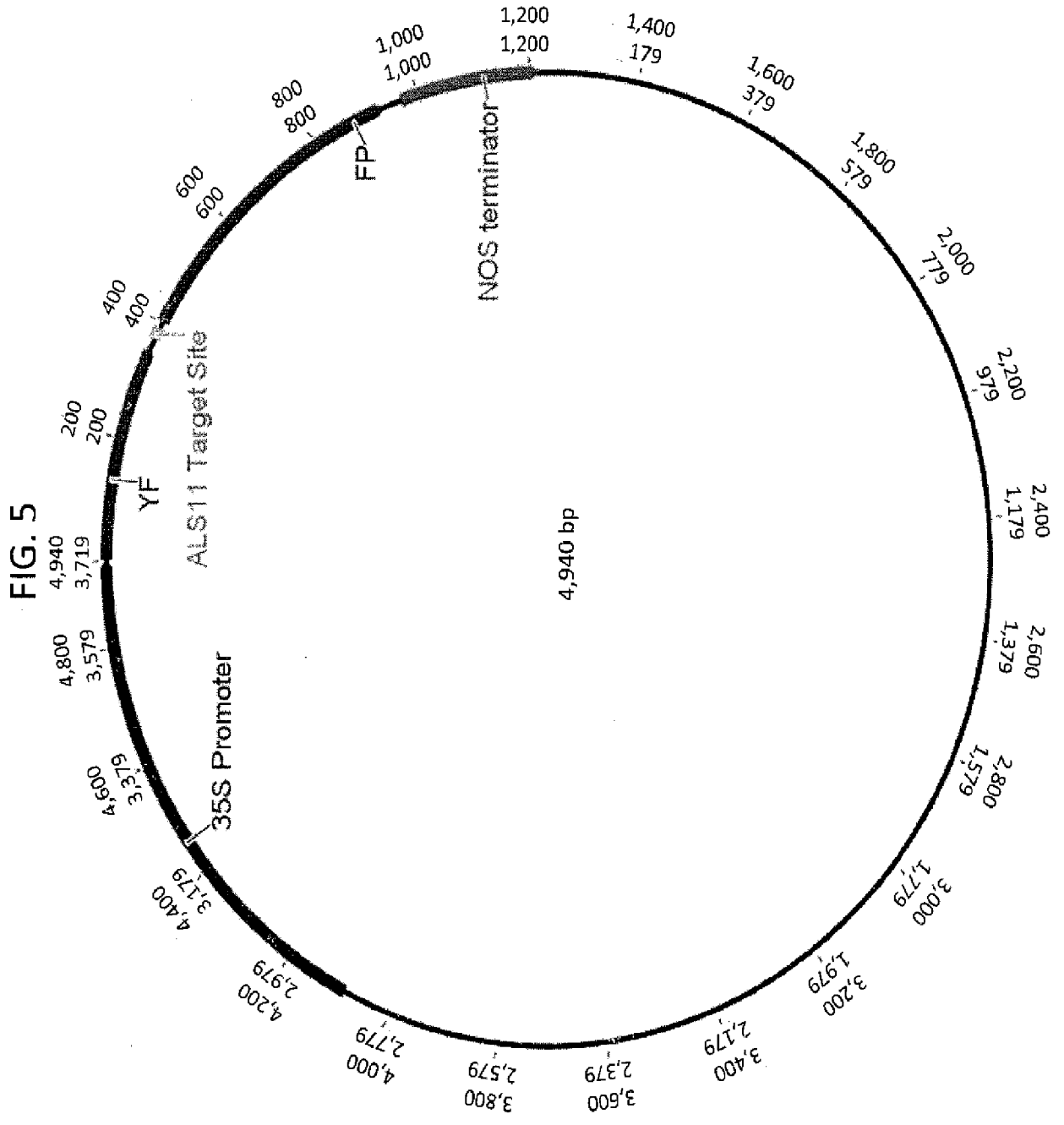
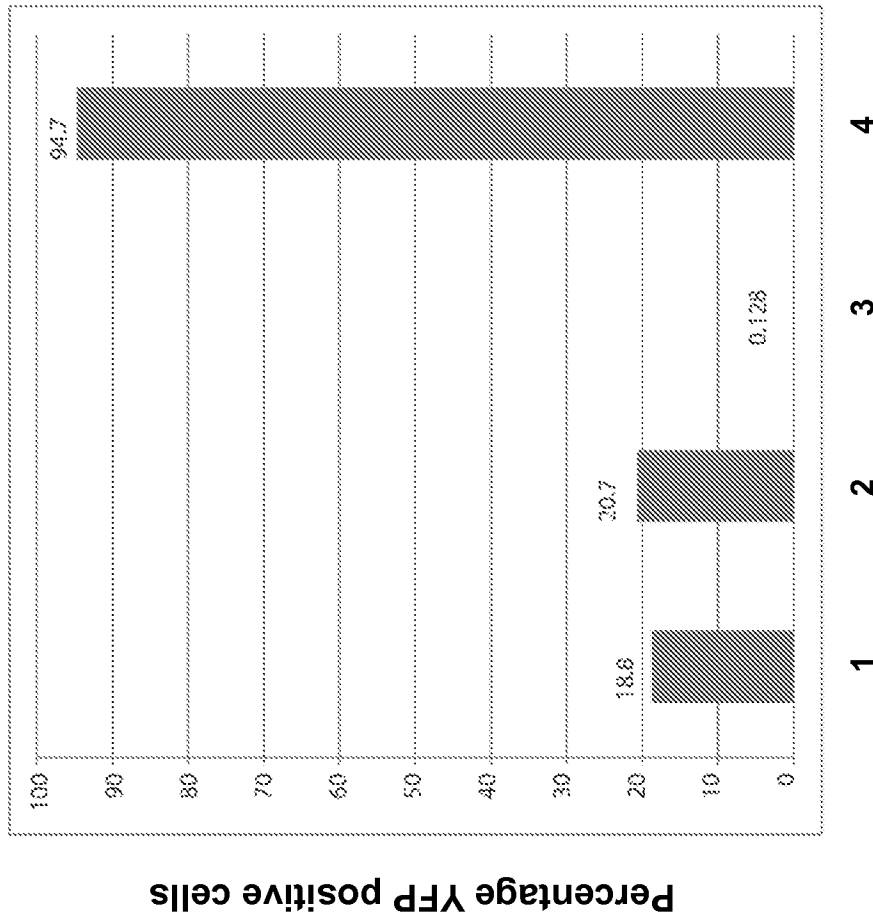


FIG. 6



INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/028445

A. CLASSIFICATION OF SUBJECT MATTER
 INV. C12N15/82 C12N15/86
 ADD. A01H5/00

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

B. FIELDS SEARCHED

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

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

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

EPO-Internal, WPI Data, Sequence Search, BIOSIS, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| X | <p>WO 02/34771 A2 (CHIRON SPA [IT]; INST GENOMIC RESEARCH [US]; TELFORD JOHN [IT]; MASIGN) 2 May 2002 (2002-05-02) page 10, line 17 - page 13, line 27 page 13, line 28 - page 15, line 8; sequence 4212 & DATABASE Geneseq [Online]</p> <p>2 July 2002 (2002-07-02), "Streptococcus polypeptide SEQ ID NO 4212.", retrieved from EBI accession no. GSP:ABP27518 Database accession no. ABP27518 sequence</p> <p style="text-align: center;">----- -/--</p> | 15-26 |



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

"E" earlier application or patent but published on or after the international filing date

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

7 August 2014

Date of mailing of the international search report

14/08/2014

Name and mailing address of the ISA/

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

Authorized officer

Burkhardt, Peter

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2014/028445

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|---|-----------------------|
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| Y | <p>WO 2011/146121 A1 (SANGAMO BIOSCIENCES INC [US]; GREGORY PHILIP D [US]; MILLER JEFFREY C) 24 November 2011 (2011-11-24) paragraph [0184] - paragraph [0188] paragraph [0221] - paragraph [0226] paragraph [0232] paragraph [0364] - paragraph [0373]; claims</p> <p style="text-align: center;">-----</p> | 1-13 |
| Y | <p>SEUNG WOO CHO ET AL: "Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease", NATURE BIOTECHNOLOGY, NATURE PUBLISHING GROUP, NEW YORK, NY, US, vol. 31, no. 3, 1 March 2013 (2013-03-01), pages 230-232, XP002699850, ISSN: 1087-0156, DOI: 10.1038/NBT.2507 [retrieved on 2013-01-29] the whole document</p> <p style="text-align: center;">-----</p> | 1-13 |
| Y | <p>L. CONG ET AL: "Multiplex Genome Engineering Using CRISPR/Cas Systems", SCIENCE, vol. 339, no. 6121, 15 February 2013 (2013-02-15), pages 819-823, XP055067741, ISSN: 0036-8075, DOI: 10.1126/science.1231143 the whole document -& L. CONG ET AL: "Supplementary Material to : Multiplex Genome Engineering Using CRISPR/Cas Systems", SCIENCE, vol. 339, no. 6121, 3 January 2013 (2013-01-03), pages 819-823, XP055067744, ISSN: 0036-8075, DOI: 10.1126/science.1231143 the whole document</p> <p style="text-align: center;">-----</p> | 1-13 |
| Y | <p>P. MALI ET AL: "RNA-Guided Human Genome Engineering via Cas9", SCIENCE, vol. 339, no. 6121, 3 January 2013 (2013-01-03), pages 823-826, XP055111247, ISSN: 0036-8075, DOI: 10.1126/science.1232033 the whole document</p> <p style="text-align: center;">-----</p> <p style="text-align: center;">-/--</p> | 1-13 |

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/028445

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|--|-----------------------|
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| Y | <p>WOONG Y HWANG ET AL: "Efficient genome editing in zebrafish using a CRISPR-Cas system", NATURE BIOTECHNOLOGY, vol. 31, no. 3, 29 January 2013 (2013-01-29), pages 227-229, XP055086625, ISSN: 1087-0156, DOI: 10.1038/nbt.2501 the whole document -& WOONG Y HWANG ET AL: "Supplementary Material to: Efficient genome editing in zebrafish using a CRISPR-Cas system", NATURE BIOTECHNOLOGY, NATURE PUBLISHING GROUP, NEW YORK, NY, US, vol. 31, no. 3, 29 January 2013 (2013-01-29), pages 1-21, XP002718602, ISSN: 1087-0156, DOI: 10.1038/NBT.2501 Retrieved from the Internet: URL:http://www.nature.com/nbt/journal/v31/ n3/full/nbt.2501.html> [retrieved on 2013-01-29]</p> <p style="text-align: center;">-----</p> | 1-13 |
| X,P | <p>WO 2013/176772 A1 (UNIV CALIFORNIA [US]; UNIV VIENNA [AT]; DOUDNA JENNIFER A [US]; JINEK) 28 November 2013 (2013-11-28) paragraph [0371] - paragraph [0381]; sequence 2 & DATABASE Geneseq [Online]</p> <p>16 January 2014 (2014-01-16), "Streptococcus pyogenes CRISPR polypeptide SEQ: 2.", retrieved from EBI accession no. GSP:BAZ49539 Database accession no. BAZ49539 sequence</p> <p style="text-align: center;">-----</p> | 1-13, 15-26 |
| X,P | <p>LI JIAN-FENG ET AL: "Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9", NATURE BIOTECHNOLOGY, vol. 31, no. 8, 23 June 2013 (2013-06-23), pages 688-691, XP055133339, ISSN: 1087-0156, DOI: 10.1038/nbt.2654 the whole document</p> <p style="text-align: center;">-----</p> <p style="text-align: center;">-/--</p> | 1-13, 15-26 |

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/028445

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| X,P | VLADIMIR NEKRASOV ET AL: "Targeted mutagenesis in the model plant <i>Nicotiana benthamiana</i> using Cas9 RNA-guided endonuclease", NATURE BIOTECHNOLOGY, vol. 31, no. 8, 1 August 2013 (2013-08-01) , pages 691-693, XP055129106, ISSN: 1087-0156, DOI: 10.1038/nbt.2654 the whole document ----- | 1-13, 15-26 |

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/028445

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: 14
because they relate to subject matter not required to be searched by this Authority, namely:
Rule 39.1(ii) PCT - Essentially biological process for the production of plants
2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2014/028445

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|--|------------------|-------------------------|------------------|
| WO 0234771 | A2 | 02-05-2002 | |
| | | AT 439372 T | 15-08-2009 |
| | | AU 1412702 A | 06-05-2002 |
| | | CA 2425303 A1 | 02-05-2002 |
| | | DK 1328543 T3 | 23-11-2009 |
| | | EP 1328543 A2 | 23-07-2003 |
| | | EP 1671981 A2 | 21-06-2006 |
| | | EP 2189473 A2 | 26-05-2010 |
| | | EP 2277894 A1 | 26-01-2011 |
| | | EP 2277895 A1 | 26-01-2011 |
| | | EP 2277896 A1 | 26-01-2011 |
| | | EP 2284181 A1 | 16-02-2011 |
| | | EP 2284182 A1 | 16-02-2011 |
| | | EP 2284183 A1 | 16-02-2011 |
| | | ES 2330083 T3 | 04-12-2009 |
| | | HK 1153754 A1 | 11-04-2014 |
| | | IL 154853 A | 16-06-2010 |
| | | JP 4254931 B1 | 15-04-2009 |
| | | JP 4874323 B2 | 15-02-2012 |
| | | JP 2009118848 A | 04-06-2009 |
| | | JP 2010508808 A | 25-03-2010 |
| | | JP 2011239783 A | 01-12-2011 |
| | | JP 2013165714 A | 29-08-2013 |
| | | MX PA03003690 A | 05-05-2004 |
| | | NZ 524888 A | 29-09-2006 |
| | | NZ 540544 A | 31-08-2007 |
| | | NZ 560966 A | 25-06-2010 |
| | | NZ 583028 A | 28-10-2011 |
| | | NZ 593617 A | 27-04-2012 |
| | | NZ 594877 A | 27-07-2012 |
| | | SG 165981 A1 | 29-11-2010 |
| | | US 2006210579 A1 | 21-09-2006 |
| | | US 2006210580 A1 | 21-09-2006 |
| | | US 2006210581 A1 | 21-09-2006 |
| | | US 2006210582 A1 | 21-09-2006 |
| | | US 2006258849 A1 | 16-11-2006 |
| | | US 2010105865 A1 | 29-04-2010 |
| | | US 2013129740 A1 | 23-05-2013 |
| | | WO 0234771 A2 | 02-05-2002 |
| ----- | | | |
| WO 2011146121 | A1 | 24-11-2011 | |
| | | AU 2011256838 A1 | 06-12-2012 |
| | | CA 2798988 A1 | 24-11-2011 |
| | | CN 103025344 A | 03-04-2013 |
| | | EP 2571512 A1 | 27-03-2013 |
| | | JP 2013529083 A | 18-07-2013 |
| | | KR 20130111219 A | 10-10-2013 |
| | | US 2011301073 A1 | 08-12-2011 |
| | | US 2014134723 A1 | 15-05-2014 |
| | | US 2014134740 A1 | 15-05-2014 |
| | | US 2014134741 A1 | 15-05-2014 |
| | | WO 2011146121 A1 | 24-11-2011 |
| ----- | | | |
| WO 2013176772 | A1 | 28-11-2013 | |
| | | US 2014068797 A1 | 06-03-2014 |
| | | WO 2013176772 A1 | 28-11-2013 |
| ----- | | | |