



- (51) **International Patent Classification:**
C12N 15/80 (2006.01) C12N 9/22 (2006.01)
- (21) **International Application Number:**
PCT/US2015/065693
- (22) **International Filing Date:**
15 December 2015 (15.12.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
PCT/CN2014/093914
16 December 2014 (16.12.2014) CN
PCT/CN2014/093916
16 December 2014 (16.12.2014) CN
PCT/CN2014/093918
16 December 2014 (16.12.2014) CN
- (71) **Applicant:** DANISCO US INC [US/US]; 925 Page Mill Road, Palo Alto, California 94304 (US).
- (72) **Inventors:** BOWER, Benjamin S.; 925 Page Mill Road, Palo Alto, California 94304 (US). CHAN, Jimmy; 925

Page Mill Road, Palo Alto, California 94304 (US). **GE, Jing**; 925 Page Mill Road, Palo Alto, California 94304 (US). **GU, Xiaogang**; 925 Page Mill Road, Palo Alto, California 94304 (US). **MADRID, Susan Mampusti**; 925 Page Mill Road, Palo Alto, California 94304 (US). **SONG, Danfeng**; 925 Page Mill Road, Palo Alto, California 94304 (US). **SONG, Mingmin**; 925 Page Mill Road, Palo Alto, California 94304 (US). **WARD, Michael**; 925 Page Mill Road, Palo Alto, California 94304 (US).

(74) **Agent:** HONG, Mei; Danisco US, Inc., 925 Page Mill Road, Palo Alto, California 94304 (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, 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.

[Continued on next page]

(54) **Title:** FUNGAL GENOME MODIFICATION SYSTEMS AND METHODS OF USE

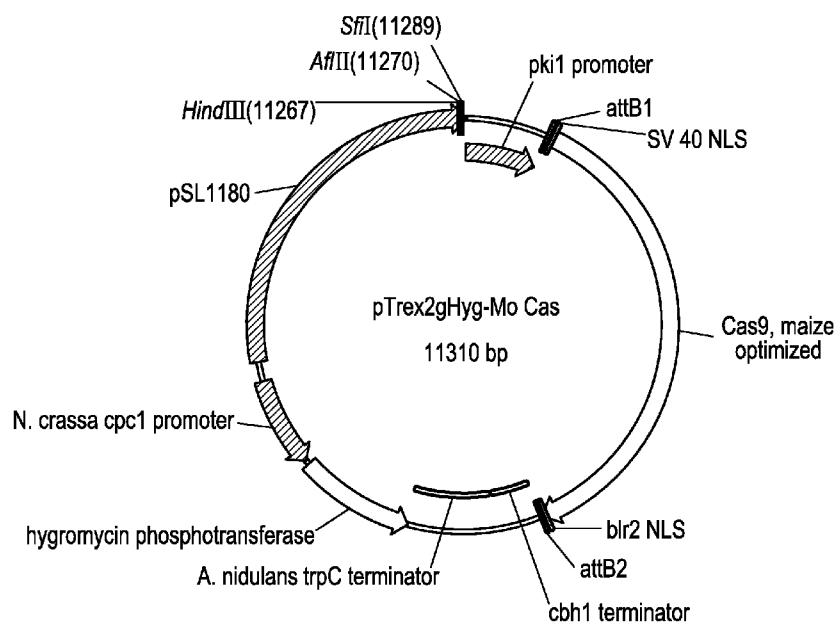


FIG. 2

(57) **Abstract:** Compositions and methods are provided for genome modification at a target site in the genome of a filamentous fungal cell. The methods and compositions are drawn to a guide polynucleotide/Cas endonuclease system for modifying or altering the target site. Aspects in which the filamentous fungal cell being modified has a defective non-homologous end joining pathway are also provided.



(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))
- 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))

FUNGAL GENOME MODIFICATION SYSTEMS AND METHODS OF USE

CROSS-REFERENCE TO RELATED APPLICATIONS

[01] The present application claims priority to PCT Patent Appln. Ser. Nos.
5 PCT/CN2014/093914, PCT/CN2014/093916, and PCT/CN2014/093918, all filed
December 16, 2014, which are hereby incorporated by reference in their entireties.

SEQUENCE LISTING

[02] The sequence listing submitted via EFS, in compliance with 37 C.F.R. §1.52(e),
10 is incorporated herein by reference. The sequence listing text file submitted via EFS
contains the file "40532-WO-PCT-4_2015-832 final_ST25.txt" created on December 3,
2015, which is 165 kilobytes in size.

BACKGROUND

15 [03] Bacteria and archaea have evolved adaptive immune defenses termed clustered
regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas)
systems that can introduce double strand breaks in DNA in a sequence-specific manner.
Cas systems perform their functions through the activity of a ribonucleoprotein complex
that includes short RNA sequences (tracrRNA and crRNA) and an RNA dependent
20 endonuclease (Cas endonuclease) that targets a specific DNA sequence (through
homology to a portion of the crRNA, called the variable targeting domain) and
generates double strand breaks in the target. CRISPR loci were first recognized in *E.*
coli (Ishino et al. (1987) J. Bacteriol. 169:5429-5433; Nakata et al. (1989) J. Bacteriol.
171:3553-3556), with similar interspersed short sequence repeats being subsequently
25 identified in a number of bacterial species, including but not limited to *Haloflex*
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).

30 [04] It is well known that inducing cleavage at a specific target site in genomic DNA
can be used to introduce modifications at or near that site. For example, homologous
recombination for gene targeting has been shown to be enhanced when the targeted

DNA site contains a double-strand break (see, e.g., Rudin et al., Genetics 122:519-534; Smih et al., Nucl. Acids Res. 23:5012-5019). Given the site-specific nature of Cas systems, genome modification/engineering technologies based on these systems have been described, including in mammalian cells (see, e.g., Hsu et al.; Cell vol. 157, p1262–1278, 5 June 2014 entitled “Development and Applications of CRISPR-Cas9 for Genome Engineering”). The power of the Cas-based genome engineering comes from the ability to target virtually any specific location within a complex genome by designing a recombinant crRNA (or equivalently functional polynucleotide) in which the DNA-targeting region (variable targeting domain) of the crRNA is homologous to the desired target site in the genome and combining it with a Cas endonuclease (through any convenient means) into a functional complex in a host cell.

[05] Although Cas-based genome engineering technologies have been applied to a number of different host cell types, the efficient use of such systems in fungal cells has proven to be difficult. Thus, there still remains a need for developing efficient and effective Cas-based genome engineering methods and compositions for modifying/altering a genomic target site in a fungal cell.

BRIEF SUMMARY

[06] Compositions and methods are provided employing a guide RNA /Cas endonuclease system for promoting homologous recombination of a donor DNA with a genomic locus in a fungal cell, e.g., a filamentous fungal cell.

[07] Aspects of the present disclosure are drawn to methods for homologous recombination of a donor DNA with a genomic locus in a fungal cell. In some embodiments, the method includes: a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA comprising a domain with homology to a genomic locus of the fungal cell, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to act at a target site in or near the genomic locus of the fungal cells; and b) identifying at least one fungal cell from the population in which homologous recombination of the donor DNA with the genomic locus has occurred, where the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells.

[08] In one aspect, the present disclosure are drawn to a method for homologous recombination of a donor DNA with a genomic locus in a fungal cell, the method

including: a) introducing into a fungal cell a Cas endonuclease, a guide RNA, and a donor DNA comprising a domain with homology to a genomic locus of the fungal cell, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to act at a target site in or near the genomic locus of the fungal cell; and b) identifying if homologous recombination of the donor DNA with the genomic locus has occurred in the fungal cell, where the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells.

[09] We have found that, in some embodiments, inhibiting or inactivating the non-homologous end joining (NHEJ) mechanism at the target site (i.e., the site of Cas endonuclease activity) in the fungal cells enhances homologous recombination of the donor DNA at the genomic locus. Therefore, aspects of the present invention include performing the homologous recombination methods as described herein under conditions in which the non-homologous end joining (NHEJ) mechanism at the target site in the fungal cells is not activated, non-functional, or reduced.

[010] Rendering non-functional (inactivating) or reducing the NHEJ pathway at the target site in the filamentous fungal cell can be achieved in any convenient manner and can be either a long term (or stable) phenotype of the host cell or a short term (or transient) phenotype of the host cell. For example, long term inactivation of the NHEJ pathway can be achieved by chromosomal genetic alteration of one or more genes involved the NHEJ pathway so that its activity is reduced or eliminated from the host cell (e.g., deletion of a gene in the NHEJ pathway). This results in the obtainment of a progeny cell having the desired genetic alteration (homologous recombination between the donor DNA and the genomic DNA at the desired location) that still has a non-functional/inactivated or reduced NHEJ pathway. Alternatively, blocking the function of or reducing the NHEJ pathway at the target site in the host cell can be done transiently. For example, transient inactivation of the NHEJ pathway can be achieved by introducing into the host cell a transient recombinant DNA construct that expresses an inhibitory RNA or a dominant negative protein whose expression inhibits the expression or the activity of one or more specific components of the NHEJ pathway.

[011] After obtaining a progeny cell having the desired genetic alteration, the transient recombinant DNA construct can be eliminated from the progeny cell, e.g., by removing selection pressure for maintenance of the transient recombinant DNA construct. In this way, the desired progeny cell will have a normally functioning NHEJ pathway.

Examples of NHEJ pathway components that can be rendered non-functional or have a reduction in activity include ku80, ku70, rad50, mre11, xrs2, lig4, xrs, or any desired combination thereof. In one particular embodiment, the fungal cell has an inactivation or reduction in the expression and/or activity of ku80. It is noted here that the term “non-functional” when in reference to a particular component of the NHEJ pathway encompasses cases in which the component is absent from the cell (e.g., by gene deletion) as well as cases in which the component is present but non-functional (e.g., a non-functional mutant protein).

[012] Alternatively, one can employ a Cas endonuclease that has nicking endonuclease activity (i.e., cleaves only one strand of DNA at the target site; also referred to herein as Cas nickases) rather than double-strand break activity. Inducing nicks at the targets site does not activate the NHEJ pathway at the target site as would a double-strand break, but does improve homologous recombination between the genomic locus of interest (one that includes or is near to the target site for the Cas nickase) and the donor DNA. Examples of Cas nickases include Cas endonuclease variants as described below.

[013] Several different types of CRISPR-Cas systems have been described and can be classified as Type I, Type II, and Type III CRISPR-Cas systems (see, e.g., the description in Liu and Fan, CRISPR-Cas system: a powerful tool for genome editing. *Plant Mol Biol* (2014) 85:209-218). In certain aspects, the CRISPR-Cas system is a Type II CRISPR-Cas system employing a Cas9 endonuclease or variant thereof (including, e.g., a Cas nickase). The Cas9 endonuclease may be any convenient Cas9 endonuclease, including but not limited to Cas9 endonucleases, and functional fragments thereof, from the following bacterial species: *Streptococcus* sp. (e.g., *S. pyogenes*, *S. mutans*, and *S. thermophilus*), *Campylobacter* sp. (e.g., *C. jejuni*), *Neisseria* sp. (e.g., *N. meningitides*), *Francisella* sp. (e.g., *F. novicida*), and *Pasteurella* sp. (e.g., *P. multocida*). Numerous other species of Cas9 can be used. For example, functional Cas9 endonucleases or variants thereof containing an amino acid sequence that has at least 70% identity to any one of SEQ ID NOs:45 and 48 to 53 may be employed, e.g., at least 80% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, and including up to 100% identity to any one of SEQ ID NOs:45 and 48 to 53.

[014] In certain embodiments, introducing the Cas endonuclease and/or the guide RNA into the fungal cells includes introducing one or more DNA constructs comprising expressions cassettes for the Cas endonuclease, the guide RNA, or both into the fungal cells. The one or more DNA constructs, once in the fungal cells, express the Cas endonuclease and/or the guide RNA. In certain embodiments, the DNA construct is a circular DNA construct that includes: an expression cassette for the Cas endonuclease, an expression cassette for the guide RNA, and the donor DNA, where the Cas endonuclease can be either a double-strand break Cas endonuclease or a Cas nickase.

[015] In certain embodiments, the introducing step includes directly introducing a Cas endonuclease polypeptide, a guide RNA, or both into the fungal cells. Any combination of direct introduction and using DNA constructs can be employed (e.g., introducing a DNA construct with an expression cassette for a Cas endonuclease into the fungal cell and directly introducing a guide RNA into the cell, either simultaneously or sequentially as desired).

[016] In certain of the methods described herein, the Cas expression cassette in the DNA construct includes a Cas endonuclease encoding gene that is optimized for expression in the fungal cell. For example, a Cas endonuclease encoding gene that is optimized for expression in filamentous fungal cells includes a sequence that has at least 70% sequence identity to SEQ ID NO:44 (encoding Cas9 from *S. pyogenes*; SEQ ID NO:45).

[017] In some instances, the Cas endonuclease is operably linked to one or more nuclear targeting signal (also referred to as a nuclear localization signal/sequence; NLS). SEQ ID NO:7 and SEQ ID NO:8 provide an example of a filamentous fungal cell optimized Cas9 gene with NLS sequences at the N- and C-termini and the encoded amino acid sequence, respectively. Many different NLSs are known in eukaryotes. They include monopartite, bipartite and tripartite types. Any convenient NLS can be used, the monopartite type being somewhat more convenient with examples including the SV40 NLS, a NLS derived from the *T. reesei* blr2 (blue light regulator 2) gene, or a combination of both.

[018] In certain embodiments, the donor DNA comprises a polynucleotide sequence of interest, and wherein homologous recombination at the genomic locus results in the insertion of the polynucleotide sequence of interest in the genomic locus.

[019] In some embodiments of the methods, the introducing step comprises introducing into the fungal cells a DNA construct comprising a sequence encoding a selectable marker or phenotypic marker as described herein. In certain embodiments, the DNA construct comprises both the sequence encoding the selectable marker and the donor DNA. In some embodiments, the DNA construct comprises a sequence encoding the Cas endonuclease, the sequence encoding the selectable marker, and the donor DNA. In some embodiments, the DNA construct comprises a sequence encoding the guide RNA, the sequence encoding the selectable marker, and the donor DNA. In particular embodiments, the DNA construct comprises a sequence encoding the Cas endonuclease, a sequence encoding the guide RNA, a sequence encoding a selectable marker, and the donor DNA. In certain embodiments, the DNA construct is a linear DNA construct. In certain embodiments, the DNA construct is a circular DNA construct.

[020] Fungal cells that find use in the subject methods can be filamentous fungal cell species. In certain embodiments, the fungal cell is a Eumycotina or Pezizomycotina fungal cell. In some embodiments, the fungal cell is selected from *Trichoderma*, *Penicillium*, *Aspergillus*, *Humicola*, *Chrysosporium*, *Fusarium*, *Neurospora*, *Myceliophthora*, *Thermomyces*, *Hypocrea*, and *Emericella*. The filamentous fungi *Trichoderma reesei*, *P. chrysogenum*, *M. thermophila*, *Thermomyces lanuginosus*, *A. oryzae* and *A. niger* are of particular interest. Other fungal cells, including species of yeast, can also be employed.

[021] The target site selected by a user of the disclosed methods can be located within a region of a gene of interest selected from the group consisting of an open reading frame, a promoter, a regulatory sequence, a terminator sequence, a regulatory element sequence, a splice site, a coding sequence, a polyubiquitination site, an intron site, and an intron enhancing motif. Examples of genes of interest include genes encoding acetyl esterases, aminopeptidases, amylases, arabinases, arabinofuranosidases, carboxypeptidases, catalases, cellulases, chitinases, cutinase, deoxyribonucleases, epimerases, esterases, α -galactosidases, β -galactosidases, α -glucanases, glucan lysases, endo- β -glucanases, glucoamylases, glucose oxidases, α -glucosidases, β -glucosidases, glucuronidases, hemicellulases, hexose oxidases, hydrolases, invertases, isomerases, laccases, lipases, lyases, mannosidases, oxidases, oxidoreductases, pectate lyases, pectin acetyl esterases, pectin depolymerases, pectin methyl esterases, pectinolytic enzymes, peroxidases, phenoloxidases, phytases,

polygalacturonases, proteases, rhamno-galacturonases, ribonucleases, transferases, transport proteins, transglutaminases, xylanases, hexose oxidases, and combinations thereof. Target genes encoding regulatory proteins such as transcription factors, repressors, proteins that modifies other proteins such as kinases, proteins involved in post-translational modification (e.g., glycosylation) can be subjected to Cas mediated editing as well as genes involved in cell signaling, morphology, growth rate, and protein secretion. No limitation in this regard is intended.

[022] In certain embodiments, the homologous recombination of the donor DNA with the genomic locus results in a modification of the DNA sequence at or near the target site, wherein the modification is selected from the group consisting of a deletion of one or more nucleotides, an insertion of one or more nucleotides, insertion of an expression cassette encoding a protein of interest, a substitution of one or more nucleotides, and any combination thereof. In some embodiments, the modification is originally present in the donor DNA. In certain embodiments, the protein of interest encoded by the expression cassette is an enzyme. In particular embodiments, the protein of interest is a hemicellulase, a peroxidase, a protease, a cellulase, a xylanase, a lipase, a phospholipase, an esterase, a cutinase, a pectinase, a keratinase, a reductase, an oxidase, a phenol oxidase, a lipoxygenase, a ligninase, a pullulanase, a tannase, a pentosanase, a mannanase, a beta-glucanase, an arabinosidase, a hyaluronidase, a chondroitinase, a laccase, an amylase, a glucoamylase, a variant thereof, a functional fragment thereof, or a hybrid or mixture of two or more thereof. In yet other particular embodiments, the protein of interest is a peptide hormone, a growth factor, a clotting factor, a chemokine, a cytokine, a lymphokine, an antibody, a receptor, an adhesion molecule, a microbial antigen, a variant thereof, a functional fragment thereof, or a hybrid or mixture of two or more thereof.

[023] In some embodiments of the methods, the step of identifying a fungal cell having a genomic modification at or near the site of interest includes culturing the population of cells from step (a) under conditions to select for or screen for the homologous recombination or the modification. Such conditions include antibiotic selection conditions, conditions that select for or screen for auxotrophic cells, and the like. In some embodiments, the identifying step comprises culturing the population of cells from step (a) under conditions to screen for unstable transformants.

[024] The method of any preceding claim, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising a sequence encoding a selectable marker and the donor DNA, and wherein the identifying step comprises culturing the population of cells from step (a) under conditions to screen for unstable transformants that have lost the selectable marker yet retained the donor DNA.

[025] Other aspects of the present disclosure are drawn to recombinant fungal cells produced by the methods described above as well as those for use as parental host cells in performing the methods.

[026] Thus, in certain embodiments, aspects of the present disclosure include recombinant fungal cells that including a first recombinant DNA construct comprising an expression cassette for a Cas endonuclease. In certain embodiments, the NHEJ pathway in the recombinant fungal cell is non-functional (inactivated) or reduced, e.g., where one or more components of the NHEJ pathway are inactivated, nonfunctional, or have reduced activity (e.g., ku80, ku70, rad50, mre11, xrs2, lig4, xrs, or combinations thereof). For example, the fungal cell can have an inactivated/reduced activity form of ku80. In certain other embodiments, the NHEJ pathway in the recombinant fungal cell is functional.

[027] In certain embodiments, the Cas endonuclease expressed from the expression cassette is a Cas9 endonuclease or variant thereof. Alternatively, the Cas endonuclease expressed from the expression cassette is a Cas nickase.

[028] As described above, in some cases the Cas endonuclease is a Cas9 endonuclease (or variant thereof). The Cas9 endonuclease may be any convenient Cas9 endonuclease including but not limited to Cas9 endonucleases, and functional fragments thereof, from the following bacterial species: *Streptococcus* sp. (e.g., *S. pyogenes*, *S. mutans*, and *S. thermophilus*), *Campylobacter* sp. (e.g., *C. jejuni*), *Neisseria* sp. (e.g., *N. meningitides*), *Francisella* sp. (e.g., *F. novicida*), and *Pasteurella* sp. (e.g., *P. multocida*). Numerous other species of Cas9 can be used. In certain of the fungal cells described herein, the first recombinant DNA construct includes a Cas endonuclease gene that is optimized for expression in the fungal cell. For example, a Cas endonuclease encoding gene that is optimized for expression in filamentous fungal cells includes a sequence that has at least 70% sequence identity to SEQ ID NO:44 (encoding Cas9 from *S. pyogenes*; SEQ ID NO:45). In some instances, the Cas endonuclease polypeptide is operably linked to one or more nuclear targeting signal

(also referred to as a nuclear localization signal/sequence; NLS). Any convenient NLS can be used, with examples including the SV40 NLS (SEQ ID NO:46), a NLS derived from the *T. reesei* blr2 (blue light regulator 2) gene (SEQ ID NO:47), or a combination of both. In some embodiments, the recombinant DNA construct comprises a promoter operably linked to a filamentous fungal cell optimized polynucleotide sequence encoding a Cas9 endonuclease or variant thereof.

[029] In certain aspects, the recombinant fungal cell described above further includes a second recombinant DNA construct capable of expressing a guide RNA, optionally through an expression cassette, where the guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to act at a target site in the genome of the recombinant fungal cell, where by “act” is meant that the Cas endonuclease cleaves the DNA as expected (making either double-stranded cut or a nick). In some embodiments, the recombinant DNA construct or the expression cassette for the guide RNA comprises a DNA polymerase III dependent promoter functional in a Eufungiomycete or Pezizomycete, wherein the promoter is operably linked to the DNA encoding the guide RNA. In some embodiments, the promoter is derived from a *Trichoderma* U6 snRNA gene. In certain embodiments, the promoter comprises a nucleotide sequence with at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 40 or 41. In specific embodiments, the promoter comprises the sequence of SEQ ID NO: 40 or 41. In some embodiments, the recombinant DNA construct or the expression cassette for the guide RNA comprises a guide RNA-encoding DNA with an intron sequence from a *Trichoderma* U6 snRNA gene. In some embodiments, the intron sequence derived from *Trichoderma* U6 snRNA gene comprises a nucleotide sequence with at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 42. In specific embodiments, the intron sequence derived from *Trichoderma* U6 snRNA gene comprises the sequence of SEQ ID NO: 42.

[030] In some instances, the recombinant fungal cell further includes a donor DNA that contains a polynucleotide of interest (which is intended by the user of the disclosed method to be inserted into the genome of the fungal cell at or near the target site in the genome via homologous recombination). Thus, in certain embodiments, the fungal cell has the polynucleotide of interest inserted at/near the target site. In some instances, the donor DNA comprises at least one of the following modifications in the domain with homology to the genomic locus of the fungal cell, as compared to the sequence of the

genomic locus: a deletion of one or more nucleotides, an insertion of one or more nucleotides, insertion of an expression cassette encoding a protein of interest, a substitution of one or more nucleotides, and any combination thereof.

[031] As noted above, we have shown that Cas-targeted homologous recombination is enhanced in cells in which the NHEJ pathway at the Cas target site is non-functional, reduced or inhibited. However, it is not required to have a non-functional, reduced or inhibited NHEJ pathway for a successful or even efficient Cas-targeted homologous recombination to occur.

[032] In some embodiments, the recombinant fungal cell comprises a DNA construct comprising a sequence encoding a selectable marker or phenotypic marker as described herein. In certain embodiments, the DNA construct comprises both the sequence encoding the selectable marker and a donor DNA as described herein. In some embodiments, the DNA construct comprises a sequence encoding a Cas endonuclease as described herein, the sequence encoding the selectable marker, and the donor DNA. In some embodiments, the DNA construct comprises a sequence encoding a guide RNA as described herein, the sequence encoding the selectable marker, and the donor DNA. In particular embodiments, the DNA construct comprises a sequence encoding the Cas endonuclease, a sequence encoding the guide RNA, a sequence encoding a selectable marker, and the donor DNA. In certain embodiments, the DNA construct is a linear DNA construct. In certain embodiments, the DNA construct is a circular DNA construct. In certain embodiments, the DNA construct is at least partly integrated into or homologously recombined with the genome of the fungal cell. In particular embodiments, at least part or all of the donor DNA comprised in the DNA construct is integrated into or homologously recombined with the genome of the fungal cell, but the selectable marker-encoding sequence, the Cas endonuclease-encoding sequence, or the guide RNA-encoding sequence is not integrated into or homologously recombined with the genome of the fungal cell.

[033] Fungal cells that find use in the subject methods include filamentous fungal cell species selected from *Trichoderma*, *Penicillium*, *Aspergillus*, *Humicola*, *Chrysosporium*, *Fusarium*, *Neurospora*, *Myceliophthora*, *Thermomyces*, *Hypocrea*, and *Emericella*. The filamentous fungi *Trichoderma reesei*, *P. chrysogenum*, *M. thermophila*, *Thermomyces lanuginosus*, *A. oryzae* and *A. niger* are of particular interest. Other fungal cells, including species of yeast, can also be employed.

[034] The target site selected by a user of the disclosed methods can be located within a region of a gene of interest selected from the group consisting of: an open reading frame, a promoter, a regulatory sequence, a terminator sequence, a regulatory element sequence, a splice site, a coding sequence, a polyubiquitination site, an intron site, and an intron enhancing motif. Examples of genes of interest include genes encoding acetyl esterases, aminopeptidases, amylases, arabinases, arabinofuranosidases, carboxypeptidases, catalases, cellulases, chitinases, chymosin, cutinase, deoxyribonucleases, epimerases, esterases, α -galactosidases, β -galactosidases, α -glucanases, glucan lysases, endo- β -glucanases, glucoamylases, glucose oxidases, α -glucosidases, β -glucosidases, glucuronidases, hemicellulases, hexose oxidases, hydrolases, invertases, isomerases, laccases, lipases, lyases, mannosidases, oxidases, oxidoreductases, pectate lyases, pectin acetyl esterases, pectin depolymerases, pectin methyl esterases, pectinolytic enzymes, peroxidases, phenoloxidases, phytases, polygalacturonases, proteases, rhamno-galacturonases, ribonucleases, thaumatin, transferases, transport proteins, transglutaminases, xylanases, hexose oxidases, and combinations thereof. No limitation in this regard is intended.

[035] Certain aspects of the present invention include recombinant polynucleotides that include a promoter sequence operably linked to a nucleotide sequence encoding a filamentous fungal cell optimized Cas9 endonuclease, where the filamentous fungal cell optimized Cas9 endonuclease is capable of binding to and creating a double strand break in a genomic target sequence in the filamentous fungal genome when complexed with a guide RNA. Examples of the filamentous fungal cell optimized Cas9 endonuclease gene include SEQ ID NO:44 and SEQ ID NO:7 and synonymous variants thereof that have improved expression in a filamentous fungal cell as compared to its parental native Cas9 encoding nucleotide sequence.

[036] Additional recombinant polynucleotide sequences include those having a filamentous fungal cell-derived RNA polymerase III (pol III) driven promoter sequence in operable linkage to a heterologous gene. In some embodiments, the filamentous fungal cell-derived RNA pol III driven promoter sequence comprises a U6 gene promoter, e.g., SEQ ID NO:40, SEQ ID NO:41, or functional variants thereof. In some cases, the recombinant polynucleotide further includes an intron in the heterologous sequence derived from an RNA pol III transcribed gene (e.g., from the U6 gene, e.g., SEQ ID NO:42) and/or a transcriptional terminator from an RNA pol III transcribed gene

(e.g., from the U6 gene, e.g., SEQ ID NO:43). In a particular embodiment, the heterologous sequence encodes a guide RNA.

[037] Additional embodiments of the methods and compositions of the present disclosure are shown herein.

5

BRIEF DESCRIPTION OF THE DRAWINGS

[038] The disclosure can be more fully understood from the following detailed description and the accompanying drawings, which form a part of this application.

[039] **FIG. 1** depicts the nucleotide sequence of a putative *T. reesei* U6 gene (SEQ ID NO:1). Elements of interest are indicated, including the TATA box (underlined), the transcriptional start site (downward arrow), the A –box (underlined), the Intron (forward arrow), the B-box (underlined; within the Intron of the gene), the sequences that are identical to the human U6 gene (in bold italics), and the terminator (underlined).

[040] **FIG. 2** shows a schematic of the pTrex2gHyg-Mo Cas plasmid.

15 **[041]** **FIG. 3** shows a schematic of the p219M plasmid.

[042] **FIG. 4** shows a schematic of the *T. reesei* ad3A gene with PCR primer sites and intronic regions shown.

[043] **FIG. 5** shows a schematic of the *T. reesei* glucoamylase gene (TrGA) with PCR primer and intronic regions shown.

20 **[044]** **FIG. 6** shows a schematic of the pTrex2gHygMoCasgPyr2TS6 plasmid which includes telomere sequences.

[045] **FIG. 7** shows a plasmid map of pET30a-Cas9-D10A nickase.

[046] **FIGS. 8A and 8B** show plasmid maps of pMD18T (T7-Spy-TrGA_sgF1) (**FIG. 8A**) and pMD18T (T7-Spy-TrGA_sgR1) (**FIG. 8B**).

25 **[047]** **FIG. 9** is an agarose gel showing results of a SpyCas9 nuclease assay. Lane 1, DNA ladder; lane 2, control; lane 3 and 4, SpyCas9 assay in the presence of TrGA_sgR1; lane 5 and 6, SpyCas9 assay in the presence of TrGA_sgF1.

[048] **FIG. 10** is an agarose gel showing the results of SpyCas9(D10A) assay. Lane 1, DNA ladder; lane 2, SpyCas9(D10A) with TrGA_sgF1 alone; lane 3, SpyCas9(D10A) with TrGA_sgR1 alone; lane 4, SpyCas9(D10A) with both RNAs.

30

[049] FIG. 11 is a Schematic diagram of TrGA deletion cassette.

[050] FIG. 12 shows 24-well plates with the morphology of transformants growing on Vogel's-starch agar. The transformants with the retarded growth phenotype are indicated by the circles. The morphology of 3 controls are shown in the square at the bottom right. "cellu" = quad-delete strain of *T. reesei* with additional deletions of the endoglucanase-3, endoglucanase-4, endoglucanase-5, endoglucanase-6, mannanase-1; "ΔAA" = "cellu" strain of *T. reesei* with alpha-amylase gene deleted; "ΔGA" = "cellu" strain of *T. reesei* with glucoamylase gene deleted; and "empty" = no cells (empty well). Selected clones #1 to #9 are indicated.

[051] FIGS. 13A and 13B. Confirmation of TrGA deletion in *T. reesei*. (FIG. 13A) Schematic diagram showing the structure of TrGA locus in wild type and TrGA deletion strain, and the binding site of primers Fw1, R3, KOF1 (5'-gaacaatcttcttgcagtgttggtc-3') (SEQ ID NO:79), KOR2 (5'-ggcagactacaagtctactagtactac-3') (SEQ ID NO:58), R1 (5'-gaggaagtcctgctttaggcaggc-3') (SEQ ID NO:80) and F4 (5'-cgacagagcagtcatatggggatacg-3') (SEQ ID NO:81). (FIG. 13B) Agarose gels showing results of PCRs. The PCR products were analyzed using 0.8% agarose gel, running at 140 volts for 30 min.

[052] FIG. 14 shows a plasmid map of pTrexMoCasGATS11-HDR.

DETAILED DESCRIPTION

[053] The present disclosure includes compositions and methods that find use in promoting homologous recombination of a donor DNA with a genomic locus in a fungal cell. The methods employ a functional guide RNA/Cas endonuclease complex which recognizes a desired target site and introduces a double strand break or nick at the site, which thereby promotes and/or enhances homologous recombination at or near the target site. In certain aspects, the non-homologous end joining (NHEJ) mechanism at the target site in the fungal cells is not activated, non-functional, or reduced, which we demonstrate herein improves the efficiency of the desired homologous recombination event.

[054] Before the present compositions and methods are described in greater detail, it is to be understood that the present compositions and methods are not limited to particular embodiments described, and as such may, of course, vary. It is also to be

understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present compositions and methods will be limited only by the appended claims.

[055] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the present compositions and methods. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and are also encompassed within the present compositions and methods, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the present compositions and methods.

[056] Certain ranges are presented herein with numerical values being preceded by the term "about." The term "about" is used herein to provide literal support for the exact number that it precedes, as well as a number that is near to or approximately the number that the term precedes. In determining whether a number is near to or approximately a specifically recited number, the near or approximating unrecited number may be a number which, in the context in which it is presented, provides the substantial equivalent of the specifically recited number. For example, in connection with a numerical value, the term "about" refers to a range of -10% to +10% of the numerical value, unless the term is otherwise specifically defined in context. In another example, the phrase a "pH value of about 6" refers to pH values of from 5.4 to 6.6, unless the pH value is specifically defined otherwise.

[057] The headings provided herein are not limitations of the various aspects or embodiments of the present compositions and methods which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification as a whole.

[058] The present document is organized into a number of sections for ease of reading; however, the reader will appreciate that statements made in one section may apply to other sections. In this manner, the headings used for different sections of the disclosure should not be construed as limiting.

[059] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the present compositions and methods belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present compositions and methods, representative illustrative methods and materials are now described.

[060] All publications and patents cited in this specification are herein incorporated by reference as if each individual publication or patent were specifically and individually indicated to be incorporated by reference and are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present compositions and methods are not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

[061] In accordance with this detailed description, the following abbreviations and definitions apply. Note that the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “an enzyme” includes a plurality of such enzymes, and reference to “the dosage” includes reference to one or more dosages and equivalents thereof known to those skilled in the art, and so forth.

[062] It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as “solely,” “only” and the like in connection with the recitation of claim elements, or use of a “negative” limitation.

[063] As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present compositions and methods described herein. Any recited method can be carried out in the order of events recited or in any other order which is logically possible.

Definitions

[064] As used herein, a polypeptide referred to as a “Cas endonuclease” or having “Cas endonuclease activity” relates to a CRISPR associated (Cas) polypeptide encoded by a Cas gene where the Cas protein is capable of cutting a target DNA sequence when functionally coupled with one or more guide polynucleotides (see, e.g., US Patent 8697359 entitled “CRISPR-Cas systems and methods for altering expression of gene products”). Variants of Cas endonucleases that retain guide polynucleotide directed endonuclease activity are also included in this definition, including Cas variants that have nicking endonuclease activity, i.e., they introduce single strand nick at a double-stranded DNA target site (see definition below). (It is noted that wild-type Cas endonucleases identified to date introduce double-strand breaks at the target site.) A Cas endonuclease is guided by the guide polynucleotide to recognize and cleave a specific target site in double stranded DNA, e.g., at a target site in the genome of a cell. Several different types of CRISPR-Cas systems have been described and can be classified as Type I, Type II, and Type III CRISPR-Cas systems (see, e.g., the description in Liu and Fan, CRISPR-Cas system: a powerful tool for genome editing. *Plant Mol Biol* (2014) 85:209-218). In certain aspects, the CRISPR-Cas system is a Type II CRISPR-Cas system employing a Cas9 endonuclease or variant thereof (including, e.g., a Cas nickase). The Cas9 endonuclease may be any convenient Cas9 endonuclease, including but not limited to Cas9 endonucleases, and functional fragments thereof, from the following bacterial species: *Streptococcus* sp. (e.g., *S. pyogenes*, *S. mutans*, and *S. thermophilus*), *Campylobacter* sp. (e.g., *C. jejuni*), *Neisseria* sp. (e.g., *N. meningitidis*), *Francisella* sp. (e.g., *F. novicida*), and *Pasteurella* sp. (e.g., *P. multocida*). Numerous other species of Cas9 can be used. For example, functional Cas9 endonucleases or variants thereof containing an amino acid sequence that has at least 70% identity to any one of SEQ ID NOs:45 and 48 to 53 may be employed, e.g., at least 80% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, and including up to 100% identity to any one of SEQ ID NOs:45 and 48 to 53. In other embodiments, the Cas endonuclease or variant thereof is a Cpf1 endonuclease of the Type II CRISPR-Cas system. Cpf1 mediates robust DNA interference with features distinct from Cas9. Cpf1 lacks tracrRNA and utilizes a T-rich protospacer-adjacent motif. It cleaves DNA via a staggered DNA double-stranded break. See, e.g., Zetsche et al., *Cell* (2015) 163:759–771.

[065] As used herein, a “Cas nickase” is a Cas endonuclease that, when functionally coupled with one or more guide polynucleotides, is capable of introducing a single-strand nick into a target double stranded DNA sequence. Cas nickases can be generated recombinantly by inactivating one of the two nuclease domains in a parent Cas endonuclease (e.g., by site-directed mutagenesis). One non-limiting example of a Cas nickase is the Cas9 nickase described in Sander and Joung (Nature Biotechnology, 2013, 1-9) in which the RuvC domain is inactivated by a D10A mutation. As mentioned above, the general term “Cas endonuclease” encompasses both double-strand cutting and nicking Cas polypeptides. For example, if a guide RNA is described as being capable of directing a Cas endonuclease to a desired target site, it would do so for both a double-strand cutting Cas endonuclease and a nicking Cas polypeptide (as defined below).

[066] As used herein, the term “guide polynucleotide” relates to a polynucleotide sequence that can form a complex with a Cas endonuclease and enables the Cas endonuclease to recognize and cleave a DNA target site. The guide polynucleotide can be a single molecule or a double molecule. The guide 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 linkage resulting in circularization. A guide polynucleotide that solely comprises ribonucleic acids is also referred to as a “guide RNA”.

[067] 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 nucleotide sequence in a target DNA (also called the “protospacer” or “target site” below) 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 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). 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 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 certain embodiments, the RNA that guides the RNA/Cas9 endonuclease complex is a duplexed RNA comprising a duplex crRNA-tracrRNA.

[068] 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 the guide RNA/Cas endonuclease

complex can direct the Cas endonuclease to a fungal cell genomic target site, enabling the Cas endonuclease to introduce a double strand break into the genomic target site.

[069] 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 in a target cell.

[070] 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. The % complementation between the first nucleotide sequence domain (VT domain) and the target sequence is 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 is 100% complementary. The VT 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 VT domain comprises a contiguous stretch of 12 to 30 nucleotides. The VT domain can be composed of a DNA sequence, a RNA sequence, a modified DNA sequence, a modified RNA sequence, or any combination thereof.

[071] 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.

[072] 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.

5 **[073]** 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 polynucleotide to a subcellular location, a modification or sequence that provides for
 10 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
 15 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
 20 permeability.

[074] As used herein, the term "guide polynucleotide/Cas endonuclease system" (and equivalents) includes a complex of a Cas endonuclease and a guide polynucleotide (single or double) that is capable of introducing a double strand break at a DNA target site. The Cas endonuclease unwinds the DNA duplex in close proximity of the DNA
 25 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 appropriately oriented at the 3' end of the target sequence.

[075] The terms "functional fragment", "fragment that is functionally equivalent", "functionally equivalent fragment", and the like, are used interchangeably and refer to a
 30 portion or subsequence of a parent polypeptide that retains the qualitative enzymatic activity of the parent polypeptide. For example, a functional fragment of a Cas endonuclease retains the ability to create a double-strand break with a guide

polynucleotide. It is noted here that a functional fragment may have altered quantitative enzymatic activity as compared to the parent polypeptide.

[076] The terms “functional variant”, “variant that is functionally equivalent”, “functionally equivalent variant”, and the like are used interchangeably and refer to a variant of a parent polypeptide that retains the qualitative enzymatic activity of the parent polypeptide. For example, a functional variant of a Cas endonuclease retains the ability to create a double-strand break or a nick (depending on the variant in question) with a guide polynucleotide. It is noted here that a functional variant may have altered quantitative enzymatic activity as compared to the parent polypeptide.

[077] Fragments and variants can be obtained via any convenient method, including site-directed mutagenesis and synthetic construction.

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

[079] A “codon-modified gene” or “codon-preferred gene” or “codon-optimized gene” is a gene having its frequency of codon usage designed to mimic the frequency of preferred codon usage of the host cell. The nucleic acid changes made to codon-optimize a gene are “synonymous”, meaning that they do not alter the amino acid sequence of the encoded polypeptide of the parent gene. However, both native and variant genes can be codon-optimized for a particular host cell, and as such no limitation in this regard is intended.

[080] “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.

[081] “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 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 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. As is well-known in the art, promoters can be categorized according to their strength and/or the conditions under which they are active, e.g., constitutive promoters, strong promoters, weak promoters, inducible/repressible promoters, tissue-specific/developmentally regulated promoters, cell-cycle dependent promoters, etc.

[082] “RNA transcript” refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. “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. “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, under certain conditions, 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” refers to antisense RNA, ribozyme RNA, or other RNA that may not be translated into a polypeptide but yet has an effect on cellular processes. The terms “complement” and “reverse complement” are used interchangeably herein with respect to mRNA transcripts, and are meant to define the antisense RNA of the message.

[083] As used herein, “functionally attached” or “operably linked” means that a regulatory region or functional domain of a polypeptide or polynucleotide sequence having a known or desired activity, such as a promoter, enhancer region, terminator, signal sequence, epitope tag, etc., is attached to or linked to a target (e.g., a gene or polypeptide) in such a manner as to allow the regulatory region or functional domain to control the expression, secretion or function of that target according to its known or

desired activity. For example, a promoter is operably linked with a coding sequence when it is capable of regulating the expression of that coding sequence (i.e., the coding sequence is under the transcriptional control of the promoter).

[084] Standard recombinant DNA and molecular cloning techniques used herein are well known in the art.

[085] “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 and is well known in the art.

[086] The term “recombinant,” when used in reference to a biological component or composition (e.g., a cell, nucleic acid, polypeptide/enzyme, vector, etc.) indicates that the biological component or composition is in a state that is not found in nature. In other words, the biological component or composition has been modified by human intervention from its natural state. For example, a recombinant cell encompass a cell that expresses one or more genes that are not found in its native parent (i.e., non-recombinant) cell, a cell that expresses one or more native genes in an amount that is different than its native parent cell, and/or a cell that expresses one or more native genes under different conditions than its native parent cell. Recombinant nucleic acids may differ from a native sequence by one or more nucleotides, be operably linked to heterologous sequences (e.g., a heterologous promoter, a sequence encoding a non-native or variant signal sequence, etc.), be devoid of intronic sequences, and/or be in an isolated form. Recombinant polypeptides/enzymes may differ from a native sequence by one or more amino acids, may be fused with heterologous sequences, may be truncated or have internal deletions of amino acids, may be expressed in a manner not found in a native cell (e.g., from a recombinant cell that over-expresses the polypeptide due to the presence in the cell of an expression vector encoding the polypeptide), and/or be in an isolated form. It is emphasized that in some embodiments, a recombinant polynucleotide or polypeptide/enzyme has a sequence that is identical to its wild-type counterpart but is in a non-native form (e.g., in an isolated or enriched form).

[087] The terms “plasmid”, “vector” and “cassette” refer to an extra chromosomal element that carries a polynucleotide sequence of interest, e.g., a gene of interest to be expressed in a cell (an “expression vector” or “expression cassette”). Such elements are generally in the form of double-stranded DNA and 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. The polynucleotide sequence of interest may be a gene encoding a polypeptide or functional RNA that is to be expressed in the target cell. Expression cassettes/vectors generally contain a gene with operably linked elements that allow for expression of that gene in a host cell.

[088] 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.

[089] “Introduced” in the context of inserting a polynucleotide or polypeptide into a cell (e.g., a recombinant DNA construct/expression construct) refers to any method for performing such a task, and includes any means of “transfection”, “transformation”, “transduction”, physical means, or the like, to achieve introduction of the desired biomolecule.

[090] By “introduced transiently”, “transiently introduced”, “transient introduction”, “transiently express” and the like is meant that a biomolecule is introduced into a host cell (or a population of host cells) in a non-permanent manner. With respect to double stranded DNA, transient introduction includes situations in which the introduced DNA does not integrate into the chromosome of the host cell and thus is not transmitted to all daughter cells during growth as well as situations in which an introduced DNA molecule that may have integrated into the chromosome is removed at a desired time using any convenient method (e.g., employing a cre-lox system, by removing positive selective pressure for an episomal DNA construct, by promoting looping out of all or part of the integrated polynucleotide from the chromosome using a selection media, etc.). No limitation in this regard is intended. In general, introduction of RNA (e.g., a guide RNA, a messenger RNA, ribozyme, etc.) or a polypeptide (e.g., a Cas polypeptide) into host cells is considered transient in that these biomolecules are not replicated and indefinitely passed down to daughter cells during cell growth. With respect to the Cas/guide RNA complex, transient introduction covers situations when either of the components is introduced transiently, as both biomolecules are needed to exert targeted Cas endonuclease activity. Thus, transient introduction of a Cas/guide RNA

complex includes embodiments where either one or both of the Cas endonuclease and the guide RNA are introduced transiently. For example, a host cell having a genome-integrated expression cassette for the Cas endonuclease (and thus not transiently introduced) into which a guide RNA is transiently introduced can be said to have a transiently introduced Cas/guide RNA complex (or system) because the functional complex is present in the host cell in a transient manner. In certain embodiments, the introducing step includes: (i) obtaining a parental fungal cell population that stably expresses the Cas endonuclease, and (ii) transiently introducing the guide RNA into the parental fungal cell population. Conversely, the introducing step can include: (i) obtaining a parental fungal cell population that stably expresses the guide RNA, and (ii) transiently introducing the Cas endonuclease into the parental fungal cell population.

[091] “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.

[092] “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 (the resulting host cell is sometimes referred to herein as a “stable transformant”). 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 integration or stable inheritance (sometimes referred to herein as “unstable transformation”, and the resulting host cell sometimes referred to herein as an “unstable transformant”). Host organisms containing the transformed nucleic acid fragments are referred to as “transgenic” organisms.

[093] “Fungal cell”, “fungi”, “fungal host cell”, and the like, as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth et al., supra) and all mitosporic fungi (Hawksworth et al., supra). In certain embodiments, the fungal host cell is a yeast cell, where by “yeast” is meant ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast

belonging to the Fungi Imperfecti (Blastomycetes). As such, a yeast host cell includes a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell. Species of yeast include, but are not limited to, the following:

Saccharomyces carlsbergensis, *Saccharomyces cerevisiae*, *Saccharomyces*
 5 *diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces*
norbensis, *Saccharomyces oviformis*, *Kluyveromyces lactis*, and *Yarrowia lipolytica* cell.

[094] The term “filamentous fungal cell” includes all filamentous forms of the subdivision Eumycotina or Pezizomycotina. Suitable cells of filamentous fungal genera include, but are not limited to, cells of *Acremonium*, *Aspergillus*, *Chrysosporium*,
 10 *Corynascus*, *Chaetomium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*,
Myceliophthora, *Neurospora*, *Paecilomyces*, *Penicillium*, *Scytalidium*, *Talaromyces*,
Thermoascus, *Thielavia*, *Tolypocladium*, *Hypocrea*, and *Trichoderma*.

[095] Suitable cells of filamentous fungal species include, but are not limited to, cells of *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*,
 15 *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium*
lucknowense, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*,
Fusarium culmorum, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium*
heterosporum, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*,
Fusarium roseum, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium*
 20 *sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium*
trichothecioides, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*,
Hypocrea jecorina, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora*
intermedia, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*,
Penicillium funiculosum *Phanerochaete chrysosporium*, *Talaromyces flavus*, *Thielavia*
 25 *terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*,
Trichoderma reesei, and *Trichoderma viride*.

[096] The terms “target site”, “target sequence”, “genomic target site”, “genomic target sequence” (and equivalents) are used interchangeably herein and refer to a polynucleotide sequence in the genome of a fungal cell at which a Cas endonuclease
 30 cleavage is desired to promote a genome modification, e.g., homologous recombination with a donor DNA. The context in which this term is used, however, can slightly alter its meaning. For example, the target site for a Cas endonuclease is generally very specific and can often be defined to the exact nucleotide sequence/position, whereas in some

cases the target site for a desired genome modification can be defined more broadly than merely the site at which DNA cleavage occurs, e.g., a genomic locus or region where homologous recombination is desired. Thus, in certain cases, the genome modification that occurs via the activity of Cas/guide RNA DNA cleavage is described as occurring “at or near” the target site. The target site can be an endogenous site in the fungal cell genome, or alternatively, the target site can be heterologous to the fungal cell 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. In certain other cases, when the donor DNA comprises a domain with homology to a genomic locus of the fungal cells, the Cas endonuclease and guide RNA introduced to the fungal cells are capable of forming a complex that enables the Cas endonuclease to act at a target site in or near the genomic locus of the fungal cells. In some embodiments, the Cas endonuclease cut site (or target site) on the genomic DNA is in the homologous region between the donor DNA and the genomic locus, where homologous recombination can occur. In other embodiments, the cut site is near the homologous region between the donor DNA and the genomic locus which can be anywhere from 1 bp to about 10 kb away from the homologous region, e.g., 1 bp, 2 bp, 5 bp, 10 bp, 20 bp, 50 bp, 100 bp, 250bp, 500bp, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, or 10 kb away from the site of homologous region.

[097] 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 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 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.

[098] The term “derived from” encompasses the terms “originated from,” “obtained from,” “obtainable from,” “isolated from,” and “created from,” and generally indicates

that one specified material find its origin in another specified material or has features that can be described with reference to the another specified material.

[0099] As used herein, the term “hybridization conditions” refers to the conditions under which hybridization reactions are conducted. These conditions are typically classified by degree of “stringency” of the conditions under which hybridization is measured. The degree of stringency can be based, for example, on the melting temperature (T_m) of the nucleic acid binding complex or probe. For example, “maximum stringency” typically occurs at about $T_m - 5^\circ\text{C}$ (5°C below the T_m of the probe); “high stringency” at about $5-10^\circ\text{C}$ below the T_m ; “intermediate stringency” at about $10-20^\circ\text{C}$ below the T_m of the probe; and “low stringency” at about $20-25^\circ\text{C}$ below the T_m . Alternatively, or in addition, hybridization conditions can be based upon the salt or ionic strength conditions of hybridization, and/or upon one or more stringency washes, e.g.: 6X SSC = very low stringency; 3X SSC = low to medium stringency; 1X SSC = medium stringency; and 0.5X SSC = high stringency. Functionally, maximum stringency conditions may be used to identify nucleic acid sequences having strict identity or near-strict identity with the hybridization probe; while high stringency conditions are used to identify nucleic acid sequences having about 80% or more sequence identity with the probe. For applications requiring high selectivity, it is typically desirable to use relatively stringent conditions to form the hybrids (e.g., relatively low salt and/or high temperature conditions are used).

[0100] As used herein, the term “hybridization” refers to the process by which a strand of nucleic acid joins with a complementary strand through base pairing, as known in the art. More specifically, “hybridization” refers to the process by which one strand of nucleic acid forms a duplex with, i.e., base pairs with, a complementary strand, as occurs during blot hybridization techniques and PCR techniques. A nucleic acid sequence is considered to be “selectively hybridizable” to a reference nucleic acid sequence if the two sequences specifically hybridize to one another under moderate to high stringency hybridization and wash conditions. Hybridization conditions are based on the melting temperature (T_m) of the nucleic acid binding complex or probe. For example, “maximum stringency” typically occurs at about $T_m - 5^\circ\text{C}$ (5° below the T_m of the probe); “high stringency” at about $5-10^\circ\text{C}$ below the T_m ; “intermediate stringency” at about $10-20^\circ\text{C}$ below the T_m of the probe; and “low stringency” at about $20-25^\circ\text{C}$ below the T_m . Functionally, maximum stringency conditions may be used to identify sequences having strict identity or near-strict identity with the hybridization probe; while

intermediate or low stringency hybridization can be used to identify or detect polynucleotide sequence homologs.

[0101] Intermediate and high stringency hybridization conditions are well known in the art. For example, intermediate stringency hybridizations may be carried out with an
5 overnight incubation at 37 °C in a solution comprising 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate and 20 mg/mL denatured sheared salmon sperm DNA, followed by washing the filters in 1x SSC at about 37 - 50 °C. High stringency hybridization conditions may be hybridization at 65 °C and 0.1X SSC (where 1X SSC =
10 0.15 M NaCl, 0.015 M Na citrate, pH 7.0). Alternatively, high stringency hybridization conditions can be carried out at about 42°C in 50% formamide, 5X SSC, 5X Denhardt's solution, 0.5% SDS and 100 µg/mL denatured carrier DNA followed by washing two times in 2X SSC and 0.5% SDS at room temperature and two additional times in 0.1X SSC and 0.5% SDS at 42°C. And very high stringent hybridization conditions may be
15 hybridization at 68 °C and 0.1X SSC. Those of skill in the art know how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

[0102] The phrase "substantially similar" or "substantially identical," in the context of at least two nucleic acids or polypeptides, means that a polynucleotide or polypeptide
20 comprises a sequence that has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99% identical to a parent or reference sequence, or does not include amino acid substitutions, insertions, deletions, or modifications made only to circumvent the present description without adding functionality.

[0103] "Sequence identity" or "identity" in the context of nucleic acid or polypeptide sequences refers to 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.
25

[0104] The term "percentage of sequence identity" refers to the value determined by
30 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%, 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.

[0105] Sequence alignments and percent identity or similarity calculations may be determined using a variety of comparison methods designed to detect homologous 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” will mean any set of values or parameters that originally load with the software when first initialized.

[0106] 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™ 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 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.

[0107] The “Clustal W method of alignment” corresponds to the alignment method 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 identity” by viewing the “sequence distances” table in the same program.

[0108] 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 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 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.

[0109] 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 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%, 96%, 97%, 98% or 99%.

[0110] “Gene” includes a nucleic acid fragment that encodes and is capable to express a functional molecule such as, but not limited to, a specific polypeptide (e.g., an enzyme) or a functional RNA molecule (e.g., a guide RNA, an anti-sense RNA, ribozyme, etc.), and includes regulatory sequences preceding (5’ non-coding sequences) and/or following (3’ non-coding sequences) the coding sequence. “Native

gene” refers to a gene as found in nature with its own regulatory sequences. A recombinant gene refers to a gene that is regulated by a different gene’s regulatory sequences which could be from a different organism or the same organism.

[0111] A “mutated gene” is a gene that has been altered through human intervention.

5 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 comprises an alteration that results from a guide polynucleotide/Cas endonuclease system as disclosed herein. A mutated fungal cell is a fungal cell comprising a mutated gene.

10 **[0112]** As used herein, a “targeted mutation” is 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 DNA of the target sequence as disclosed herein or known in the art.

[0113] The term “donor DNA” or “donor nucleic acid sequence” or “donor polynucleotide” refers to a polynucleotide that contains a polynucleotide sequence of interest that is to be inserted at or near a target site or to replace a region at or near a target site, generally in conjunction with the activity of a Cas/guide polynucleotide complex (where the guide polynucleotide defines the target site, as detailed above). As such, the polynucleotide sequence of interest in the donor DNA may include a novel
 15 region to be inserted at or near the target site and/or a modified polynucleotide sequence when compared to the nucleotide sequence to be replaced/edited at or near the target site. In certain embodiments, the donor DNA construct further comprises a first and a second region of homology that flank the polynucleotide sequence of interest. The first and second regions of homology of the donor DNA share homology
 20 to a first and a second genomic region, respectively, present in or flanking the target site of the fungal cell genome. By “homology” is meant DNA 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 fungal cell genome. A region of homology can be of any length that is sufficient
 25 to promote homologous recombination at the cleaved target site. For example, the region of homology 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 in length such that the region of homology has sufficient homology to undergo homologous recombination with the corresponding genomic region. "Sufficient homology" indicates
5 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
10 comprising localized similarities such as contiguous nucleotides having 100% sequence identity, and percent sequence identity over a portion of the length of the sequences.

[0114] 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, 100-250 bp,
15 150-300 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, 10, 11, 12, 13, 14, 15,
20 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%, 95%, 96%, 97%, 98%, 99% or 100%.

25 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 described by the predicted ability of two
30 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, Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes, (Elsevier, New York).

[0115] As used herein, a “genomic region” is a segment of a chromosome in the genome of a fungal 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.

[0116] As used herein, “homologous recombination” includes the exchange of DNA fragments between two DNA molecules at the sites of homology and is well described in the art.

[0117] 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.

[0118] 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 and antibiotics, such as, chlorimuron ethyl, benomyl, Basta, and hygromycin phosphotransferase (HPT); DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers, dominant heterologous marker-*amdS*); 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.

5 *Methods for Modifying a Fungal Cell Genome*

[0119] Methods are provided employing a guide RNA /Cas endonuclease system for promoting homologous recombination of a donor DNA with a genomic locus in a fungal cell, e.g., a filamentous fungal cell.

[0120] Aspects of the present disclosure include methods for homologous
10 recombination of a DNA sequence with a genomic locus in the genome of a fungal cell by transiently introducing a Cas endonuclease/guide polynucleotide complex into the cell along with a donor DNA that includes a domain with homology to the genomic locus. The Cas endonuclease/ guide polynucleotide complex is capable of acting at a desired target site in the genome of the fungal cell, where by “acting” is meant that the
15 Cas endonuclease, guided by sequences in the guide polynucleotide (as defined above), cleaves either one or both strands of the DNA at the target site.

[0121] Introduction of the Cas endonuclease, guide polynucleotide, and the donor DNA can be done in any convenient manner, including transfection, transduction, transformation, electroporation, particle bombardment, cell fusion techniques, etc.

20 Each of these components can be introduced simultaneously or sequentially as desired by the user. For example, a fungal cell can first be stably transfected with a Cas expression DNA construct followed by introduction of a guide polynucleotide into the stable transfectant (either directly or using a guide polynucleotide expressing DNA construct)with. This set up may even be advantageous as the user can generate a
25 population of stable Cas transfectant fungal cells into which different guide polynucleotides can be introduced independently (in some cases, more than one guide polynucleotide can be introduced into the same cells should this be desired). In some embodiments, a Cas expressing fungal cell is obtained by the user, and thus the user does not need to introduce a recombinant DNA construct capable of expressing a Cas
30 endonuclease into the cell, but rather only need introduce a guide polynucleotide into the Cas expressing cell.

[0122] In certain embodiments, a guide polynucleotide is introduced into the fungal cell by introducing a recombinant DNA construct that includes an expression cassette (or

gene) encoding the guide polynucleotide. In some embodiments, the expression cassette is operably linked to a eukaryotic RNA pol III promoter. These promoters are of particular interest as transcription by RNA pol III does not lead to the addition of a 5' cap structure or polyadenylation that occurs upon transcription by RNA polymerase II from an RNA pol II dependent promoter. In certain embodiments, the RNA pol III promoter is a filamentous fungal cell U6 polymerase III promoter (e.g., SEQ ID NO:40 and functional variants thereof, e.g., SEQ ID NO:41; described in further detail below).

[0123] When a double-strand break is induced in the genomic DNA of a host cell (e.g., by the activity of a Cas endonuclease/guide RNA complex at a target site, the complex having double-strand endonuclease activity), the cell's DNA repair mechanism is activated to repair the break which, due to its error-prone nature, 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).

[0124] Surprisingly, we have found in filamentous fungi that non-homologous insertion of transformed DNA at the double-strand break is highly favored over simple end-joining between the two ends of the chromosomal DNA at a double-strand break. Therefore, in cases where the Cas endonuclease or guide RNA is provided by transformation with an expression cassette containing DNA construct or constructs, those DNA constructs, or fragments thereof, are inserted at the double-strand break at high frequency. This insertion occurs in the absence of homology between DNA sequences on the Cas endonuclease or guide RNA expression constructs and the sequences around the double-strand break. This process is also problematic when homologous recombination between a donor DNA and a genomic locus is desired, as insertion of the entire donor DNA is favored over homologous recombination. We have found that undesirable insertion of transformed DNA occurs even when it is in the form of a vector including telomere sequences that is expected to be maintained autonomously in the fungal cell.

[0125] DNA taken up by transformation may integrate in a stable fashion in the genome or it may be transiently maintained. Transient maintenance can be recognized by an

unstable phenotype. For example, DNA uptake can be recognized by selection for a marker gene present on the transforming DNA. After transformation and selection, the transformants may be grown under non-selective conditions for several generations before transfer back to selective conditions. A stable transformant will be able to grow after transfer back to selective conditions whereas an unstable transformant will be unable to grow after transfer back to selective conditions due to loss of the transforming DNA. As shown in the Examples section below, we have demonstrated that it is possible to transiently express Cas endonuclease and/or guide RNA in fungal cells/unstable transformants.

[0126] In embodiments where unstable transformants are desired, a plasmid with telomere sequences to encourage autonomous replication can be used. Other types of plasmids that are designed for autonomous replication, such as those with autonomous replication sequences, centromere sequences or other sequences, can also be employed. Surprisingly, in *Trichoderma reesei* we have found that one can use plasmids with no known origin of replication, autonomous replication sequence, centromere or telomere sequences. By screening those transformants that show an unstable phenotype with respect to the selectable marker, efficient target site gene modification without vector DNA insertion is obtained (e.g., homologous recombination with a homologous region in a donor DNA).

[0127] Certain embodiments of the present disclosure include integrating a Cas endonuclease expression cassette and first selectable marker in the genome of a fungus, optionally flanked by repeats to allow subsequent removal (loop-out) of the expression cassette and first selectable marker, to produce a Cas endonuclease expressing host cell. These cells can be employed in numerous ways to obtain a genetic modification of interest, including homologous recombination with a donor DNA.

[0128] For example, a Cas endonuclease expressing host cell can be transformed with a DNA construct including a guide RNA expression cassette containing a second selectable marker (and optionally a separate donor DNA). Host cells that are selected for using the second selectable marker will express the guide RNA from this DNA construct, which enables Cas endonuclease activity and targeting to a defined target site of interest in the genome. Screening these host cells for transformants that show an unstable phenotype with respect to the second selectable marker will enable

obtaining host cells with a modified site of interest (e.g., homologous recombination with the donor DNA) without DNA construct insertion.

[0129] As another example, a Cas endonuclease expressing host cell can be induced to uptake an in vitro synthesized guide RNA to enable Cas endonuclease activity and targeting to a defined site in the genome. In some cases, it will be desirable to induce uptake of both guide RNA and a separate DNA construct bearing a selectable marker gene to allow for selection of those cells that have taken up DNA and, at high frequency, are expected to have simultaneously taken up guide RNA. As above, screening those transformants that show an unstable phenotype with respect to the selectable marker for the genetic modification of interest (e.g., homologous recombination with a donor DNA) without vector DNA insertion is obtained.

[0130] As yet another example, a Cas endonuclease expressing host cell can be used to create a “helper strain” that can provide, in trans, the Cas endonuclease to a “target strain”. In brief, a heterokaryon can be created between the helper strain and the target strain, e.g., by fusion of protoplasts from each strain or by anastomosis of hyphae depending on the species of filamentous fungus. Maintenance of the heterokaryon will depend on appropriate nutritional or other marker genes or mutations in each parental strain and growth on suitable selective medium such that the parental strains are unable to grow whereas the heterokaryon, due to complementation, is able to grow.

Either at the time of heterokaryon formation or subsequently, a guide RNA is introduced by transfection (and optionally a donor DNA). The guide RNA may be directly introduced or introduced via a DNA construct having a Cas endonuclease expression cassette and a selectable marker gene. Cas endonuclease is expressed from the gene in the helper strain nucleus and is present in the cytoplasm of the heterokaryon. The Cas endonuclease associates with the guide RNA to create an active complex that is targeted to the desired target site(s) in the genome. Subsequently, spores are recovered from the heterokaryon and subjected to selection or screening to recover the target strain with a modification at or near the target site (e.g., homologous recombination with the donor DNA at a genomic locus). In cases in which an expression cassette is used to introduce the guide RNA, heterokaryons are chosen in which the guide RNA expression construct is not stably maintained.

[0131] As noted above, methods of the present disclosure include introducing a DNA construct into the cell (or donor DNA) that has DNA sequence homology with regions of

the chromosomal DNA on each side of the target site of the Cas/guide RNA complex. The intent is for the DNA fragment (e.g., a linear DNA fragment) to integrate by homologous integration/recombination, repairing the cleavage in the DNA at the target site and, in most cases, introducing changes to genome at the desired locus (i.e., at or
5 near the target site of the Cas/guide RNA complex). In many organisms, a double-strand break in the chromosomal DNA stimulates homologous integration of the linear DNA fragment at that site. Surprisingly, in filamentous fungi with a functioning NHEJ pathway we have found that, even when a donor fragment is introduced, insertion of DNA by non-homologous insertion at the double-strand break is highly favored over
10 homologous recombination of the linear DNA fragment.

[0132] With respect to DNA repair in fungal cells, we have found that in the presence of a functioning NHEJ pathway, error-prone repair is highly favored over homologous recombination at a double strand break site. In other words, with respect to DNA repair of a double strand break in filamentous fungal cells, we have found that in the
15 presence of a functioning NHEJ pathway, non-homologous insertion of DNA at the break is highly favored over (1) non-homologous end joining without DNA insertion and (2) homologous recombination at the double strand break site with a donor DNA. Therefore, in certain aspects of the present invention, the functioning of the non-homologous end joining (NHEJ) pathway at the target site in the fungal cell in the
20 population is inhibited, not activated, non-functional, or reduced. This may be achieved in any convenient manner, some of which are described below.

[0133] In some embodiments, the functioning of the non-homologous end joining (NHEJ) pathway at the target site in the fungal cell is inhibited, not activated, non-functional, or reduced by altering the fungal host cell such that one or more
25 components of the NHEJ pathway are non-functional or have reduced activity (e.g., deleted from the genome or mutated to be non-functional or less active). This alteration of the fungal cell can be achieved by any convenient means, including gene deletion, gene mutation, expression of a dominant-interfering recombinant protein, gene replacement, gene expression inhibition, e.g., using antisense RNA/RNAi methodology,
30 and the like. In certain aspects, the one or more components of the NHEJ pathway are selected from the group consisting of ku80, ku70, rad50, mre11, xrs2, lig4, and xrs. As but one example, a fungal cell that finds use in aspects of the present invention includes a genetic modification that inhibits the expression and/or activity of ku80.

[0134] In additional embodiments, the functioning of the non-homologous end joining (NHEJ) pathway at the target site in the fungal cell is inhibited, not activated, non-functional, or reduced by using a Cas endonuclease that has DNA nicking activity, i.e., it cleaves only one strand of the DNA at the target site (also called Cas nickases).

5 Unlike double-strand breaks in the DNA, nicks do not activate the NHEJ pathway but are sufficient to promote homologous recombination at or near the site of cleavage with a donor DNA having one or more region of homology. Numerous Cas nickases, which are recombinant variants of wild-type Cas endonucleases, have been described in the art (see, e.g., definition above) and may be used in the disclosed methods.

10 **[0135]** In some instances, the donor DNA includes a first region and a second region that are homologous to corresponding first and second regions in the genome of the fungal cell, wherein the regions of homology generally include or surround the target site at which the genomic DNA is cleaved by the Cas endonuclease. These regions of homology promote homologous recombination with their corresponding genomic
15 regions of homology resulting in exchange of DNA between the donor DNA and the genome. As such, the provided methods result in the integration of the polynucleotide of interest of the donor DNA at or near the cleavage site in the target site in the fungal cell genome, thereby altering the original target site, thereby producing an altered genomic target site.

20 **[0136]** 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 fungal cell genome can be at least 50%, 55%, 60%,
25 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or even 100% sequence identity, such that the sequences undergo homologous recombination.

[0137] 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
30 significant sequence homology to the genomic sequence 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 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.

- 5 **[0138]** As with the Cas endonuclease and the guide polynucleotide expression constructs, the donor DNA may be introduced by any convenient means (as discussed elsewhere herein).

- [0139]** In certain embodiments, the Cas endonuclease is a Cas9 endonuclease (see, e.g., WO 2013141680 entitled "RNA-directed DNA Cleavage by the Cas9-crRNA Complex"). Examples of Cas9 endonucleases include those from *Streptococcus* sp. (e.g., *S. pyogenes*, *S. mutans*, and *S. thermophilus*), *Campylobacter* sp. (e.g., *C. jejuni*), *Neisseria* sp. (e.g., *N. meningitidis*), *Francisella* sp. (e.g., *F. novicida*), and *Pasteurella* sp. (e.g., *P. multocida*) (see, e.g., Cas9 endonucleases described in Fonfara et al., *Nucleic Acids Res.*, 2013, pages 1-14: incorporated herein by reference).
- 10
- 15 In some embodiments, the Cas endonuclease is encoded by an optimized Cas9 endonuclease gene, e.g., optimized for expression in a fungal cell (e.g., Cas9 encoding genes containing SEQ ID NO:44, e.g., SEQ ID NO:7, as described below).

- [0140]** In certain instances, the Cas endonuclease gene is operably linked to one or more polynucleotides encoding nuclear localization signals such that the Cas endonuclease/guide polynucleotide complex that is expressed in the cell is efficiently transported to the nucleus. Any convenient nuclear localization signal may be used, e.g., a polynucleotide encoding an SV40 nuclear localization signal present upstream of and in-frame with the Cas codon region and a polynucleotide encoding a nuclear localization signal derived from the *T. reesei* *blr2* (blue light regulator 2) gene present downstream and in frame with the Cas codon region. Other nuclear localization signals can be employed.
- 20
- 25

- [0141]** In certain embodiments of the disclosure, the guide polynucleotide is a guide RNA that includes a crRNA region (or crRNA fragment) and a tracrRNA region (or tracrRNA fragment) of the type II CRISPR/Cas system that can form a complex with a type II Cas endonuclease. As indicated above, the guide RNA/Cas endonuclease complex can direct the Cas endonuclease to a fungal cell genomic target site, enabling the Cas endonuclease to introduce a double strand break into the genomic target site. In some cases, the RNA that guides the RNA/ Cas9 endonuclease complex is a duplex
- 30

that includes a crRNA and a separate tracrRNA. In other instances, the guide RNA is a single RNA molecule that includes both a crRNA region and a tracrRNA region (sometimes referred to herein as a fused guide RNA). One advantage of using a fused guide RNA versus a duplexed crRNA-tracrRNA is that only one expression cassette
5 needs to be made to express the fused guide RNA.

[0142] Host cells employed in the methods disclosed herein may be any fungal host cells are from the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the
10 Oomycota (as cited in Hawksworth et al., supra) and all mitosporic fungi (Hawksworth et al., supra). In certain embodiments, the fungal host cells are yeast cells, e.g., *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell. Species of yeast include, but are not limited to, the following: *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces*
15 *diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, *Kluyveromyces lactis*, and *Yarrowia lipolytica* cell. In additional embodiments, the fungal cells are filamentous fungal cells including but not limited to species of *Trichoderma*, *Penicillium*, *Aspergillus*, *Humicola*, *Chrysosporium*, *Fusarium*, *Neurospora*, *Myceliophthora*, *Hypocrea*, and *Emericella*.
20 For example, the filamentous fungi *T. reesei* and *A. niger* find use in aspects of the disclosed methods.

[0143] Virtually any site in a fungal cell genome may be targeted using the disclosed methods, so long as the target site includes the required protospacer adjacent motif (PAM). In the case of the *S. pyogenes* Cas9, the PAM has the sequence NGG (5' to 3';
25 where N is A, G, C or T), and thus does not impose significant restrictions on the selection of a target site in the genome. Other known Cas9 endonucleases have different PAM sites (see, e.g., Cas9 endonuclease PAM sites described in Fonfara et al., *Nucleic Acids Res.*, 2013, pages 1-14: incorporated herein by reference).

[0144] The length of the target site can vary, and includes, for example, target sites that
30 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 cleavage site can be within the target sequence or the

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.

[0145] In some cases, active variant target sequences in the genome of the fungal cell can also be used, meaning that the target site is not 100% identical to the relevant sequence in the guide polynucleotide (within the crRNA sequence of the guide polynucleotide). 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 variant target sequences retain biological activity and hence are capable of being recognized and cleaved by a 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.

[0146] Target sites of interest include those located within a region of a gene of interest. Non-limiting examples of regions within a gene of interest include an open reading frame, a promoter, a transcriptional regulatory element, a translational regulatory element, a transcriptional terminator sequence, an mRNA splice site, a protein coding sequence, an intron site, and an intron enhancing motif.

[0147] In certain embodiments, modification of the genome of the fungal cell results in a phenotypic effect that can be detected and, in many instances, is a desired outcome of the user. Non-limiting examples include acquisition of a selectable cell growth phenotype (e.g., resistance to or sensitivity to an antibiotic, gain or loss of an auxotrophic characteristic, increased or decreased rate of growth, etc.), expression of a detectable marker (e.g., fluorescent marker, cell-surface molecule, chromogenic enzyme, etc.), and the secretion of an enzyme the activity of which can be detected in culture supernatant.

[0148] When modification of the genome of the fungal cell results in a phenotypic effect, a donor DNA is often employed that includes a polynucleotide of interest that is (or encodes) a phenotypic marker. Any convenient phenotypic marker can be used, including any selectable or screenable marker that allows one to identify, or select for or against a fungal cell that contains it, often under particular culture conditions. Thus, in

some aspects of the present invention, the identification of fungal cells having a desired genome modification includes culturing the fungal population of cells that have received the Cas endonuclease and guide polynucleotide (and optionally a donor DNA) under conditions to select for or screen for cells having the modification at the target site. Any type of selection system may be employed, including assessing for the gain or loss of an enzymatic activity in the fungal cell (also referred to as a selectable marker), e.g., the acquisition of antibiotic resistance or gain/loss of an auxotrophic marker.

[0149] In some instances, the genomic modification in the fungal cells is detected directly using any convenient method, including sequencing, PCR, Southern blot, restriction enzyme analysis, and the like, including combinations of such methods.

[0150] In some embodiments, specific genes are targeted for modification using the disclosed methods, including genes encoding enzymes, e.g., acetyl esterases, aminopeptidases, amylases, arabinases, arabinofuranosidases, carboxypeptidases, catalases, cellulases, chitinases, cutinase, deoxyribonucleases, epimerases, esterases, α -galactosidases, β -galactosidases, α -glucanases, glucan lysases, endo- β -glucanases, glucoamylases, glucose oxidases, α -glucosidases, β -glucosidases, glucuronidases, hemicellulases, hexose oxidases, hydrolases, invertases, isomerases, laccases, lipases, lyases, mannosidases, oxidases, oxidoreductases, pectate lyases, pectin acetyl esterases, pectin depolymerases, pectin methyl esterases, pectinolytic enzymes, peroxidases, phenoloxidases, phytases, polygalacturonases, proteases, rhamno-galacturonases, ribonucleases, transferases, transport proteins, transglutaminases, xylanases, hexose oxidases, and combinations thereof.

[0151] There are numerous variations for implementing the methods described herein. For example, instead of having the Cas expression cassette present as an exogenous sequence in the fungal host cell, this cassette can be integrated into the genome of the fungal host cell. Generating this parental cell line would allow a user to simply introduce a desired guide RNA (e.g., as a guide RNA expression vector) which would then target the genomic site of interest as detailed elsewhere herein. In some of these embodiments, the integrated Cas gene can be designed to include polynucleotide repeats flanking it for subsequent loop-out /removal from the genome if needed.

Compositions of Fungal Cells

[0152] Aspects of the present invention include a transgenic fungal cell that finds use in carrying out the methods described above as well as the resulting fungal cell having a modified genome. Thus, embodiments of the present invention include a recombinant fungal cell produced by any aspect of the methods described herein as well as any parental fungal cell employed to produce them.

[0153] Certain embodiments of the present invention are drawn to a recombinant fungal cell comprising a Cas endonuclease, where the Cas endonuclease is expressed from a recombinant DNA construct in the cell (a first recombinant DNA construct). In some embodiments, the recombinant fungal cell has a non-functional or reduced-activity NHEJ pathway. The embodiments of the Cas endonuclease and the polynucleotide encoding it described in detail above find use in the fungal cell compositions in this section (a few of which are set forth below). This fungal cell finds use as a parent for generating a fungal cell with a desired genome modification of interest, where generating the genome modification includes introducing a guide polynucleotide (e.g., via an expression cassette) into the cell thus allowing the formation of a Cas/guide polynucleotide complex that drives the genetic modification (as described above). In certain aspects, one or more components of the NHEJ pathway are non-functional or have reduced activity in the recombinant fungal cells, e.g., ku80, ku70, rad50, mre11, xrs2, lig4, xrs, and combinations thereof. In one particular embodiment, the recombinant fungal cell has a genetic modification that inhibits the expression and/or activity of ku80. Any convenient genetic modification to achieve disruption of the NHEJ pathway component(s) may be employed, including but not limited to gene deletion, gene mutation, expression of a dominant-interfering recombinant protein, gene replacement, gene expression inhibition, e.g., using antisense RNA/RNAi methodology, and the like.

[0154] In certain aspects, the recombinant fungal cell further includes a second recombinant DNA construct capable of expressing a guide RNA, where 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 target site in the genome of the recombinant fungal cell. The embodiments of the guide polynucleotide and the polynucleotide encoding it described in detail above find use in the fungal cell compositions in this section (a few of which are set forth below). Expression of the guide RNA may be driven by a eukaryotic RNA pol III promoter, wherein certain embodiments the RNA pol III promoter is a filamentous fungal cell U6 gene promoter

(e.g., SEQ ID NO:40 or SEQ ID NO:41 as described in further detail below) and functional variants thereof. The action of this complex can result in the modification of the genomic DNA sequence at the target site of the fungal cell (as described above), thus generating a fungal cell having a modification at (or near) the target site. The modification can include a deletion of one or more nucleotides, an insertion of one or more nucleotides, a substitution of one or more nucleotides, or any combination thereof. In some embodiments, the fungal cell further includes a donor DNA that has a polynucleotide of interest. The polynucleotide of interest in the donor DNA can be present in the genome at (or near) the target site (inserted into the genome, e.g., by homologous recombination), a process that is driven by the action of the Cas/guide polynucleotide complex at the target site.

[0155] In certain embodiment, the Cas endonuclease encoded by the recombinant polynucleotide is a Cas9 endonuclease. Any Cas9 endonuclease may be encoded, including but not limited to Cas9 endonucleases, and functional fragments thereof, from the following bacterial species: *Streptococcus* sp. (e.g., *S. pyogenes*, *S. mutans*, and *S. thermophilus*), *Campylobacter* sp. (e.g., *C. jejuni*), *Neisseria* sp. (e.g., *N. meningitides*), *Francisella* sp. (e.g., *F. novicida*), and *Pasteurella* sp. (e.g., *P. multocida*) (see, e.g., Cas9 endonucleases described in Fonfara et al., *Nucleic Acids Res.*, 2013, pages 1-14: incorporated herein by reference). In some embodiments, the polynucleotide encoding the Cas endonuclease gene is one that has been optimized for expression in a filamentous fungal host cell, e.g., the polynucleotide shown in SEQ ID NO:44 (which is a filamentous fungal cell codon optimized version of the *S. pyogenes* Cas9 endonuclease) or SEQ ID NO:7 (which contains SEQ ID NO:44 and also includes N- and C-terminal NLS sequences). Additional codon-optimized Ca9 genes may be employed, including synonymous variants of SEQ ID NO:44 or SEQ ID NO:7. As described above, the Cas endonuclease can be operably linked to one or more nuclear localization signal, which function to enhance the cytoplasmic to nuclear transit of the Cas endonuclease to its site of action, i.e., in the nucleus of the cell. Any convenient nuclear localization signal may be employed, including the SV40 nuclear localization signal, a nuclear localization signal derived from the *T. reesei* blr2 (blue light regulator 2) gene, or a combination of both.

[0156] Any of a wide variety of filamentous fungal host cells find use in the present invention, including fungal host cells from the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth et al., In, Ainsworth and

Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth et al., supra) and all mitosporic fungi (Hawksworth et al., supra). In certain embodiments, the fungal host cells are yeast cells, e.g., *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*,

5 *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell. Species of yeast include, but are not limited to, the following: *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, *Kluyveromyces lactis*, and *Yarrowia lipolytica* cell. In additional embodiments, the fungal cells are filamentous
10 fungal cells including but not limited to species of *Trichoderma*, *Penicillium*, *Aspergillus*, *Humicola*, *Chrysosporium*, *Fusarium*, *Neurospora*, *Myceliophthora*, *Thermomyces*, *Hypocrea*, and *Emericella*. For example, the filamentous fungi *Trichoderma reesei*, *P. chrysogenum*, *M. thermophila*, *Thermomyces lanuginosus*, *A. oryzae* and *A. niger* find use in aspects of the disclosure.

15 **[0157]** As detailed above, the present invention is drawn generally to methods and compositions useful in modifying a target site of interest in the genome of a filamentous fungal cell. The particular target site of interest is determined by the user of such methods and composition and include sites that are located within a region of a gene of interest, including: a promoter, a regulatory sequence, a terminator sequence, a
20 regulatory element sequence, a splice site, a coding sequence, a polyubiquitination site, an intron site, and an intron enhancing motif. In addition, any genes of interest can be selected by a user, including but not limited to: acetyl esterases, aminopeptidases, amylases, arabinases, arabinofuranosidases, carboxypeptidases, catalases, cellulases, chitinases, cutinase, deoxyribonucleases, epimerases, esterases, α -galactosidases, β -
25 galactosidases, α -glucanases, glucan lysases, endo- β -glucanases, glucoamylases, glucose oxidases, α -glucosidases, β -glucosidases, glucuronidases, hemicellulases, hexose oxidases, hydrolases, invertases, isomerases, laccases, lipases, lyases, mannosidases, oxidases, oxidoreductases, pectate lyases, pectin acetyl esterases, pectin depolymerases, pectin methyl esterases, pectinolytic enzymes, peroxidases,
30 phenoloxidases, phytases, polygalacturonases, proteases, rhamno-galacturonases, ribonucleases, transferases, transport proteins, transglutaminases, xylanases, hexose oxidases, and combinations thereof.

Recombinant Polynucleotides

[0158] Aspects of the present invention are drawn to recombinant polynucleotides that find use in the methods and compositions described herein.

[0159] Embodiments of the disclosure include a recombinant polynucleotide DNA construct having a promoter sequence operably linked to a fungal cell optimized nucleotide sequence encoding a Cas endonuclease. Embodiments of the disclosure include a recombinant polynucleotide DNA construct having a promoter sequence operably linked to a fungal cell optimized nucleotide sequence encoding a Cas endonuclease or a bacterial cell optimized nucleotide sequence encoding a Cas endonuclease. As described above, the Cas endonuclease encoded in the fungal cell optimized nucleotide sequence as well as the bacterial cell optimized nucleotide sequence are capable of acting at a target site in when complexed with a guide RNA. Any Cas endonuclease may be encoded by the optimized nucleotide sequences, including but not limited to a Cas9 endonuclease, e.g., Cas9 from *S. pyogenes*. In certain embodiments, the fungal cell optimized nucleotide sequence is optimized for expression in a filamentous fungal cell. For example, a filamentous fungal cell optimized sequence can encode a Cas9 endonuclease and contain the nucleotide sequence shown in SEQ ID NO:44 (100% identity) or encodes a Cas9 endonuclease and contains a nucleotide sequence that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to SEQ ID NO:44. In certain embodiments, the bacterial cell optimized nucleotide sequence is optimized for expression in an *E. coli* cell. For example, an *E. coli* cell optimized sequence can encode a Cas9 endonuclease and contain the nucleotide sequence shown in SEQ ID NO:65 (100% identity) or encodes a Cas9 endonuclease and contains a nucleotide sequence that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to SEQ ID NO:65.

[0160] Embodiments of the disclosure further relate to a *Trichoderma* sp. RNA polymerase III driven promoters. Transcription of a gene by RNA pol III from an RNA pol III directed promoter does not lead to the addition of a 5' cap structure or polyadenylation that occurs upon transcription by RNA polymerase II from an RNA pol II dependent promoter. As described in the Examples below, we have identified an RNA pol III driven promoter sequence in *T. reesei* that is associated with the U6 gene as well as the transcription terminator sequence. The full promoter sequence is set forth in SEQ ID NO:40 and the terminator sequence is set forth in SEQ ID NO:43. In addition, a shorter version of the U6 gene RNA pol III driven promoter was identified and is set

forth in SEQ ID NO:41. Thus, aspects of the invention include a promoter that function as RNA pol III driven promoter and having a nucleotide sequence that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or up to 100% identical to SEQ ID NO:40 or SEQ ID NO:41. This RNA pol III directed promoter sequence finds use in expressing any heterologous sequence of interest. Thus, aspects of the disclosure include recombinant polynucleotide sequences having a *T. reesei* derived RNA pol III driven promoter sequence operably linked to a heterologous sequence of interest. In certain embodiments, the heterologous sequence is one that encodes a guide polynucleotide, e.g., a tracrRNA, a crRNA, or a single guide RNA. Guide RNA encoding polynucleotides targeted to specific genomic sites of interest are described in detail in the Examples and include SEQ ID NOs:2 to 6. In certain embodiments, the recombinant polynucleotide further includes a transcriptional terminator sequence is operably linked to the heterologous sequence of interest (which is operably linked to the RNA pol III promoter), e.g., at a site downstream of the heterologous sequence of interest (where “downstream” refers to the direction of transcription, as is common in the art). The terminator sequence includes, in some embodiments, the polynucleotide sequence shown in SEQ ID NO:43 or functional variants thereof. Thus, in certain embodiments, the recombinant polynucleotide includes an RNA pol III promoter operably linked to a heterologous sequence of interest (e.g., a guide RNA) operably linked to a terminator.

[0161] Non-limiting examples or embodiments of compositions and methods disclosed herein are as follows:

1. A method for homologous recombination of a donor DNA with a genomic locus in a filamentous fungal cell, the method comprising:
 - a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA comprising a domain with homology to a genomic locus of the fungal cell, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to act at a target site in or near the genomic locus of the fungal cells; and
 - b) identifying at least one fungal cell from the population in which homologous recombination of the donor DNA with the genomic locus has occurred,wherein the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells.

2. The method of embodiment 1, wherein the non-homologous end joining (NHEJ) mechanism at the target site in the fungal cells is not activated, non-functional, or reduced.

5

3. The method of embodiment 2, wherein the non-homologous end joining (NHEJ) pathway in the fungal cells comprises one or more non-functional or reduced-activity components.

10

4. The method of embodiment 3, wherein the one or more non-functional or reduced-activity components are selected from the group consisting of ku80, ku70, rad50, mre11, xrs2, lig4, xrs, and combinations thereof.

15

5. The method of embodiment 4, wherein the one or more non-functional or reduced-activity components is ku80.

6. The method of any preceding embodiment, wherein the Cas endonuclease is a Cas nickase.

20

7. The method of any one of embodiments 1-5, wherein the Cas endonuclease is a Cas9 endonuclease or variant thereof.

25

8. The method of embodiment 7, wherein the Cas9 endonuclease or variant thereof comprises a full length Cas9 or a functional fragment thereof from a species selected from the group consisting of: *Streptococcus* sp., *S. pyogenes*, *S. mutans*, *S. thermophilus*, *Campylobacter* sp., *C. jejuni*, *Neisseria* sp., *N. meningitides*, *Francisella* sp., *F. novicida*, and *Pasteurella* sp., *P. multocida*.

30

9. The method of embodiment 8, wherein the Cas9 endonuclease or variant thereof comprises an amino acid sequence that has at least 70% identity to any one of SEQ ID NOs:45 and 48 to 53.

10. The method of any preceding embodiment, wherein the donor DNA comprises a polynucleotide sequence of interest, and wherein homologous

recombination at the genomic locus results in the insertion of the polynucleotide sequence of interest in the genomic locus.

11. The method of any preceding embodiment, wherein the introducing step comprises introducing a DNA construct comprising an expression cassette for the Cas endonuclease into the fungal cells.

12. The method of any preceding embodiment, wherein the introducing step comprises introducing a DNA construct comprising an expression cassette for the guide RNA into the fungal cells.

13. The method of any preceding embodiment or any one of embodiments 48-54, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising a sequence encoding a selectable marker.

14. The method of embodiment 13, wherein the DNA construct comprises both the sequence encoding the selectable marker and the donor DNA.

15. The method of any preceding embodiment, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising: a sequence encoding the Cas endonuclease, a sequence encoding the guide RNA, a sequence encoding a selectable marker, and the donor DNA.

16. The method of any one of embodiments 11 to 15, wherein the DNA construct is a linear DNA construct.

17. The method of any one of embodiments 11 to 15, wherein the DNA construct is a circular DNA construct.

18. The method of any one of embodiments 11 and 15-17, wherein the expression cassette for the Cas endonuclease or the sequence encoding the Cas endonuclease comprises a Cas coding sequence that is optimized for expression in the filamentous fungal cell.

19. The method of embodiment 18, wherein the Cas coding sequence is a Cas9 coding sequence comprising a polynucleotide sequence that is at least 70% identical to SEQ ID NO:44.

5 20. The method of any one of embodiments 1 to 10, 12 to 14, and 16-17, wherein the introducing step comprises directly introducing the Cas endonuclease into the fungal cells.

10 21. The method of any one of embodiments 1 to 11, 13 to 14, and 16 to 20, wherein the introducing step comprises directly introducing the guide RNA into the fungal cells.

15 22. The method of any preceding embodiment, wherein the Cas endonuclease is operably linked to a nuclear localization signal.

23. The method of any preceding embodiment, wherein the fungal cell is a Eumycotina or Pezizomycotina fungal cell.

20 24. The method of any preceding embodiment, wherein the fungal cell is selected from the group consisting of Trichoderma, Penicillium, Aspergillus, Humicola, Chrysosporium, Fusarium, Myceliophthora, Neurospora, Hypocrea, and Emericella.

25 25. The method of any preceding embodiment, wherein the target site is located within a region of a gene of interest selected from the group consisting of an open reading frame, a promoter, a regulatory sequence, a terminator sequence, a regulatory element sequence, a splice site, a coding sequence, a polyubiquitination site, an intron site, and an intron enhancing motif.

30 26. The method of any preceding embodiment, wherein the homologous recombination results in a modification of the DNA sequence at or near the target site, wherein the modification is selected from the group consisting of a deletion of one or more nucleotides, an insertion of one or more nucleotides,

insertion of an expression cassette encoding a protein of interest, a substitution of one or more nucleotides, and any combination thereof.

27. The method of any preceding embodiment, wherein the identifying step comprises culturing the population of cells from step (a) under conditions to select for or screen for the homologous recombination or the modification.

28. The method of any preceding embodiment, wherein the identifying step comprises culturing the population of cells from step (a) under conditions to screen for unstable transformants.

29. The method of any preceding embodiment, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising a sequence encoding a selectable marker and the donor DNA, and wherein the identifying step comprises culturing the population of cells from step (a) under conditions to screen for unstable transformants that have lost the selectable marker yet retained the donor DNA.

30. A recombinant filamentous fungal cell produced by the method of any preceding embodiment.

31. A recombinant filamentous fungal cell comprising a first recombinant DNA construct comprising an expression cassette for a Cas endonuclease.

32. The recombinant fungal cell of embodiment 30 or 31, wherein the recombinant filamentous fungal cell comprises one or more non-functional or reduced activity component in the NHEJ pathway.

33. The recombinant fungal cell of embodiment 32, wherein the one or more components of the NHEJ pathway are selected from the group consisting of ku80, ku70, rad50, mre11, xrs2, lig4, and xrs.

34. The recombinant fungal cell of embodiment 33, wherein the fungal cell comprises a genetic modification that inhibits the expression and/or activity of ku80.

5 35. The recombinant fungal cell of embodiment 31, wherein the Cas endonuclease is a Cas nickase.

10 36. The recombinant fungal cell of any one of embodiments 30 to 35, further comprising a second recombinant DNA construct comprising an expression cassette for a guide RNA, wherein the guide RNA and Cas endonuclease are capable of forming a complex that enables the Cas endonuclease to act at a target site in the genome of the recombinant filamentous fungal cell.

15 37. The recombinant fungal cell of any one of embodiments 30 to 36 further comprising a donor DNA comprising a polynucleotide of interest.

38. The recombinant fungal cell of any one of embodiments 30 to 37, wherein the Cas endonuclease is a Cas9 endonuclease or variant thereof.

20 39. The recombinant fungal cell of embodiment 38, wherein the Cas9 endonuclease or variant thereof comprises an amino acid sequence that has at least 70% identity to any one of SEQ ID NOs:45 and 48 to 53.

25 40. The recombinant fungal cell of any one of embodiments 31 to 39, wherein the expression cassette for the Cas endonuclease comprises a polynucleotide sequence that is at least 70% identical to SEQ ID NO:44.

30 41. The recombinant fungal cell of any one of embodiments 31 to 40, wherein the expression cassette for the Cas endonuclease comprises a Cas endonuclease gene that is optimized for expression in the filamentous fungal cell.

42. The recombinant fungal cell of any one of embodiments 31 to 41, wherein the Cas endonuclease is operably linked to a nuclear localization signal.

43. The recombinant fungal cell of embodiment 42, wherein the nuclear localization signal is selected from the group consisting of SV40 nuclear localization signal (SEQ ID NO:46), a nuclear targeting signal derived from the *T. reesei* blr2 (blue light regulator 2) gene (SEQ ID NO:47), and a combination of both.

44. The recombinant fungal cell of any one of embodiments 30 to 43, wherein the fungal cell is a filamentous fungal cell selected from the group consisting of: *Trichoderma*, *Penicillium*, *Aspergillus*, *Humicola*, *Chrysosporium*, *Fusarium*, *Myceliophthora*, *Neurospora*, *Hypocrea*, and *Emericella*.

45. A recombinant DNA construct comprising a promoter operably linked to a filamentous fungal cell optimized polynucleotide sequence encoding a Cas9 endonuclease or variant thereof.

46. The recombinant DNA construct of embodiment 45, wherein the filamentous fungal cell optimized polynucleotide sequence is at least 70% identical to SEQ ID NO:44.

47. The method of embodiment 12, wherein the expression cassette for the guide RNA comprises a DNA polymerase III dependent promoter functional in a *Euscomycete* or *Pezizomycete*, and wherein the promoter is operably linked to the DNA encoding the guide RNA.

48. The method of embodiment 47, wherein the promoter is derived from a *Trichoderma* U6 snRNA gene.

49. The method of embodiment 48, wherein the promoter comprises a nucleotide sequence with at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 40 or 41.

50. The method of embodiment 49, wherein the promoter comprises the sequence of SEQ ID NO: 40 or 41.

51. The method of any one of embodiments 12 and 47-50, wherein the expression cassette for the guide RNA comprises a guide RNA-encoding DNA with an intron sequence from a *Trichoderma* U6 snRNA gene.

52. The method of embodiment 51, wherein the intron sequence derived from *Trichoderma* U6 snRNA gene comprises a nucleotide sequence with at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 42.

53. The method of embodiment 52, wherein the intron sequence derived from *Trichoderma* U6 snRNA gene comprises the sequence of SEQ ID NO: 42.

EXAMPLES

[0162] 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.

Section A: Introduction of Cas/guide RNA by Expression Vectors

*Example 1: Identification of *T. reesei* U6 snRNA gene*

[0163] An RNA polymerase III directed promoter is desired for production of guide RNA in *T. reesei* without the addition of a 5' cap structure or polyadenylation that would result from the use of a RNA polymerase II dependent promoter. However, no RNA polymerase III dependent promoter that is functional in *T. reesei* has been described. Known RNA polymerase III dependent promoters from other species were considered to be tested for their ability to function in *T. reesei* including the 5' upstream regions from the *Saccharomyces cerevisiae* snr52 gene, the human U6 snRNA gene, or the corn U6 snRNA gene.

[0164] More desirable was to identify a native *T. reesei* sequence that would function as an RNA polymerase III dependent promoter. The DNA sequence encoding the human U6 small nuclear RNA (snRNA; GenBank accession number M14486) was used to

search the *T. reesei* v2 genome sequence (www.jgi.doe.gov) using the BLAST algorithm. A short region of *T. reesei* DNA sequence was identified with similarity to the human sequence. Examination of the surrounding DNA sequence and comparison with the U6 genes of yeasts, particularly *Schizosaccharomyces pombe* (Marck et al., 2006, Nucleic Acids Research 34:1816-1835), allowed a number of features of the *T. reesei* U6 gene to be putatively identified (SEQ ID NO:1, shown below). The start of the transcribed sequence and the terminator were identified as were an upstream TATA box. An intron apparently interrupts the transcribed region and possible A-box and B-box promoter elements can be recognized within the transcribed region, the latter within the intron. (see FIG. 1).

```

AAAAAAGACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTAACTTCTGCA
GTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTATTATTTTTAT
TTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTATTATTATAATAT
ATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAATAATTTATAG
TAATAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATGAAATGGTATT
ATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTGGCTATAAGTC
TGCTGCAAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTGATGGTAGTCT
ATCGCCTTCGGGCATTTGGTCAATTTATAACGATACAGGTTTCGTTTCGGCTTTTCC
TCGGAACCCCCAGAGGTCATCAGTTCGAATCGCTAACAGGTCAACAGAGAAGATT
AGCATGGCCCCTGCACTAAGGATGACACGCTCACTCAAAGAGAAGCTAAACATTTT
TTTTCTCTTCCAAGTCGTGATGGTTATCTTTTTGCTTAGAGAATCTATTCTTGTGGA
CGATTAGTATTGGTAAATCCCTGCTGCACATTGCGGCGGATGGTCTCAACGGCAT
AATACCCCATTCGTGATGCAGCGGTGATCTTCAATATGTAGTGTAATACGTTGCAT
ACACCACCAGGTTTCGGTGCCTCCTGTATGTACAGTACTGTAGTTCGACTCCTCCG
CGCAGGTGGAACGATTCCCTAGTGGGCAGGTATTTTGGCGGGGTCAAGAA (SEQ
ID NO:1)

```

Example 2: sgRNA sequences to target *T. reesei* genes

[0165] It has been shown that a single guide RNA (sgRNA) molecule can interact with the *Streptococcus pyogenes* Cas9 protein to target this endonuclease in vivo to a specific locus in a eukaryote genome. The sgRNA is a hybrid molecule designed as a

fusion between the tracrRNA and crRNA observed naturally to be components of the *Streptococcus pyogenes* type II CRISPR-Cas system (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). The first 20 nucleotides of the sgRNA are complementary to the target site in the genome. An additional sequence, protospacer adjacent motif (PAM) is also required to be present at the target site in the genome adjacent to the sgRNA-complementary region. In the case of the *S. pyogenes* Cas9 the PAM has the sequence NGG (where N is A, G, C or T).

[0166] The sequence of sgRNA used in these experiments is shown below where the 20 nucleotides designed to be complementary to the target site are shown as N residues (SEQ ID NO:2) (N = A, G, C, or U).

NNNNNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

[0167] sgRNAs were designed to target different loci in the *T. reesei* genome. The sequence of an sgRNA (called gAd3A TS1) to target the *T. reesei* *ad3A* gene (Phosphoribosylamidoimidazole-succinocarboxamide synthase) at a site designated as target site 1 (TS1) is shown below (SEQ ID NO:3). The 20 nucleotide region that is complementary to the *T. reesei* genome sequence is shown in lower case.

guccucgagcaaaaggugccGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

[0168] The sequence of an sgRNA (called gTrGA TS2) to target the *T. reesei* *gla1* (glucoamylase) gene at a site designated as target site 2 (TS2) is shown below (SEQ ID NO:4). The 20 nucleotide region that is complementary to the *T. reesei* genome sequence is shown in lower case.

guucagugcaauaggcgucuGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

[0169] The sequence of an sgRNA (called gTrGA TS11) to target the *T. reesei* *gla1* (glucoamylase) gene at a site designated as target site 11 (TS11) is shown below (SEQ ID NO:5). The 20 nucleotide region that is complementary to the *T. reesei* genome sequence is shown in lower case.

gccaauggcgacggcagcacGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGU
CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

[0170] The sequence of an sgRNA (called gPyr2 TS6) to target the *T. reesei* *pyr2*

5 (orotate phosphoribosyltransferase) gene at a site designated as target site 6 (TS6) is shown below (SEQ ID NO:6). The 20 nucleotide region that is complementary to the *T. reesei* genome sequence is shown in lower case.

gcacagcgggaugcccuuguGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGU
CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

10

Example 3: Cas9 DNA and protein sequences for expression in *T. reesei*

[0171] A codon optimized *Streptococcus pyogenes* Cas9-encoding gene, including NLS sequences, was designed, synthesized and tested for expression in *T. reesei* (SEQ ID NO:7). The encoded protein (SEQ ID NO:8) has an N- terminal SV40 nuclear

15 localization signal (NLS; SEQ ID NO:46) and a C-terminal NLS derived from the *T. reesei* *blr2* (blue light regulator 2) gene (SEQ ID NO:47; both are underlined in SEQ ID NO:8 below).

SEQ ID NO:7

atggcaccgaagaagaagcgcaaggtgatggacaagaagtacagcatcggcctcgacatcggcaccaactcgggtgg
20 gctggggccgctcatcacggacgaatataaggtcccgtcgaagaagttcaaggtcctcggaatacagaccgccacagca
tcaagaaaaacttgatcggcgccctcctgttcgatagcggcgagaccgcggaggcgaccaggctcaagaggaccgcc
aggagacggtacactaggcgcaagaacaggatctgctacctgcaggagatcttcagcaacgagatggcgaaggtgg
acgactccttctccaccgcctggaggaatcattcctgttgaggaggacaagaagcatgagcggcacccaatcttcgg
caacatcgctcgacgaggtggcctaccacgagaagtacccgacaatctaccacctccggaagaaactggtggacagca
25 cagacaaggcggacctccggctcatctaccttgccctcgcgcatatgatcaagttccgcggccacttcctcatcgagggc
gacctgaacccggacaactccgacgtggacaagctgttcacccagctcgtgcagacgtacaatcaactgttcgaggaga
accccataaacgctagcggcgtggacgccaaggccatcctcgcggccaggctctcgaaatcaagaaggctggagaac
cttatcgcgagttgccaggcgaaaagaagaacggcctcttcggcaaccttattgcgctcagcctcggcctgacgcca
acttcaaataaaacttcgacctcgcgaggacgccaagctccagctctcaaaggacacctacgacgacgacctcgaca
30 acctcctggcccagataggagaccagtacgcggacctcttcctcgccgccaagaacctctccgacgctatcctgctcagc
gacatccttcgggtcaacaccgaaattaccaaggcaccgctgtccgccagcatgattaaacgctacgacgagcaccatc
aggacctcacgctgctcaaggcactcgtccgccagcagctccccgagaagtacaaggagatcttcttcgaccaatcaaa
aaacggctacgcgggatataatcgacggcggtgccagccaggaagagttctacaagttcatcaaaccaatcctggagaa
gatggacggcaccgaggagttgctggtcaagctcaacaggaggacctcctcaggaagcagaggaccttcgacaac

ggctccatcccgcatcagatccacctggggaactgcatgccatcctgcggcgccaggaggacttctacccgttctgaa
ggataaccgggagaagatcgagaagatcttgacgttccgcatcccatactacgtgggcccgtggctcgcggaactcc
cggttcgctggatgacccggaagtcggaggagaccatcacaccctggaactttgaggaggtggtcgataagggcgct
agcgctcagagcttcatcgagcgcatgaccaacttcgataaaaacctgcccgaatgaaaaagtcctccccaagcactcgc
5 tgctctacgagtacttaccgtgtacaacgagctcaccaaggtaaaatagtcaccgagggcatgcggaagccggcggt
cctgagcggcgagcagaagaaggcgatagtgacccctcttcaagaccaacaggaaggtagccgtgaagcaattaa
aagaggactacttcaagaaaatagagtgcctcgactccgtggagatctcgggcgtggaggatcgggtcaacgcctcactc
ggcagctatcacgacctcctcaagatcattaaagacaaggacttctcgacaacgaggagaacgaggacatcctcgag
gacatcgtcctcacccctgacctgttcgaggaccgcgaatgatcgaggagaggctgaagacctacgcgcacctgttcg
10 acgacaaggctcatgaaacagctcaagaggcgccgctacactgggtgggaaggctgtcccgcaagctcattaatggca
tcagggacaagcagagcggcaagaccatcctggacttctcaagtcggacgggttcgccaaccgcaacttcatgcagc
tcattcacgacgactcgctcacgttcaaggaagacatccagaaggcacaggtagcgggcagggtagctccctccacg
aacacatcgccaacctggccggtcgccggccattaaaaaggcatcctgcagacgggtcaaggtagctcgacgagctc
gtgaaggtagtgccggcacaagcccgaataatcgatcatagagatggccagggagaaccagaccacccaaaaa
15 gggcagaagaactcgcgcgagcggatgaaacggatcgaggaggcgattaaagagctcgggtcccagatcctgaag
gagcaccctgtggaaaatacccagctccagaatgaaaagctctactctactacctgcagaacggccgcgacatgtac
gtggaccaggagctggacattaatcggtatcggactacgacgtcgaccacatcgtagccgagtcgttctcaaggacg
atagcatcgacaacaaggtagctcaccgggtcgataaaaatcggggcaagagcgacaacgtgcccagcgaggaggt
cgtgaagaagatgaaaaactactggcgccagctcctcaacgcgaactgatcaccagcgcaagttcgacaacctga
20 cgaaggcggaacgcgggtggcttgagcgaactcgataaggcggttcataaaaaggcagctggtcgagacgcgcca
gatcacgaagcatgtcgcccagatcctggacagccgcatgaataactaagtagatgaaaacgacaagctgatccggg
aggtagaaggtagcacgtgaagtccaagctcgtgtcggaacttcgcaaggacttccagttctacaaggtagcgagatc
aacaactaccaccagcccacgacgcctacctgaatgcgggtggtcgggaccgccctgatcaagaagtacccgaagct
ggagtcggagttcgtgtacggcgactacaaggcttacgacgtgcgcaaaatgatcgccaagtccgagcaggagatcgg
25 caaggccacggcaaaaatacttcttactcgaacatcatgaacttctcaagaccgagatcacctcggaacggcgag
atccgcaagcgcccgtcatcgaaaccaacggcgagacgggagatcgtctgggataagggccgggatttcgagac
ggtagcgaaggtagcttccatgccgcaagtcaatatcgtgaaaaagacggagggtccagacgggagggttcagcaagg
agtcacatctccgaagcgcaactccgacaagctcatcgcgaggaagaaggattgggacccgaaaaaatatggcggc
ttcgacagcccagccgtcgcatacagcgtcctcgtcgtggcgaagggtggagaagggaagtcataagaagctcaagtcc
30 gtgaaggagctgctcgggatcacgattatggagcgggtcctccttcgagaagaacccgatcgacttctagaggccaagg
gatataaggaggtagaagaaggacctgattattaaactgccgaagtactcgctcttcgagctggaaaacggccgcaaga
ggatgctcgctccgcaggcgagttgcagaagggaacgagctcgccctcccgagcaatactcaatttctgtacctc
gctagccactatgaaaagctcaagggcagcccggaggacaacgagcagaagcagctctcgtggagcagcacaag
cattacctggacgagatcatcgagcagatcagcgagttctgaagcgggtgatcctcgccgacgcgaacctggacaag

gtgctgtcgccatataacaagcaccgacgacaaaccaatacgcgagcaggccgaaaatatcatccacctcttcaccctca
 ccaacctcgccgctccggcagccttcaagtacttcgacaccacgattgaccggaagcggtacacgagcacgaaggag
 gtgctcgatgcgacgctgatccaccagagcatcacagggctctatgaaacacgcatcgacctgagccagctggcgga
 gacaagaagaagaagctcaagctctag

5

SEQ ID NO:8

MAPKKKRKYMDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEE
 DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLI
 10 EGD LNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLP
 GEKKNGLFGNLI ALSGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADL
 FLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
 SIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
 15 SEETITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
 KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
 FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKV
 MKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFK
 EDIQAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDLVKVMGRHKPENIVIEMA
 20 RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
 MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
 KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
 MNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TAL
 IKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI
 25 RKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRN
 SDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS SF
 EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVLADANLDKVLSA
 YNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITG
 30 LYETRIDLSQLGGDKKKKLKL

Example 4: Construction of Cas9 expression vectors

[0172] The synthetic DNA sequence encoding Cas9 shown above was inserted into pENTR/D-TOPO so that it would be between flanking *attL1* and *attL2* sites to enable

transfer by Invitrogen™ Gateway® cloning technology (Thermo Fisher Scientific Inc., Grand Island, NY) into suitable expression vectors. A Gateway compatible expression vector, pTrex2gHyg, was available that comprises the following features; the promoter region from the *T. reesei pki1* (pyruvate kinase) gene and terminator region from the *T. reesei cbh1* (cellobiohydrolase I) gene separated by Gateway cloning sites, a bacterial hygromycin phosphotransferase gene functionally linked to the *Neurospora crassa cpc1* (cross pathway control 1) promoter region and the *Aspergillus nidulans trpC* (trifunctional protein with glutamine amido transferase, indoleglycerolphosphate synthase and phosphoribosylanthranilate isomerase activity) terminator region, and bacterial vector sequences for selection and maintenance in *E. coli*. The cas9 gene was cloned into pTrex2gHyg using the Gateway cloning procedure to give pTrex2gHyg MoCas (see FIG. 2).

Example 5: Construction of sgRNA expression vectors

[0173] Synthetic DNA sequences were obtained that encode the gAd3A TS1 sgRNA flanked by different putative RNA polymerase III dependent promoters and terminators. Each of these synthetic DNA sequences also had restriction enzyme recognition sites (EcoRI and BamHI) at either end.

[0174] The following sequence encodes the gAd3A TS1 sgRNA (underlined) with the *Saccharomyces cerevisiae* snr52 promoter and *S. cerevisiae* sup4 terminator (denoted gAd3A TS1-1; SEQ ID NO:9):

gaattcggatccTCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGA
AACTTGATGTTTTCTTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACA
ACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTCAC
ACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAAGTTG
GTGCGCATGTTTCGGCGTTTCGAACTTCTCCGCAGTGAAAGATAAATGATCgtcctcg
agcaaaagggtgccGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATC
AACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTTTGTTTTTTATGTCTgaattcg
gatcc

[0175] The following sequence encodes the gAd3A TS1 sgRNA (underlined) with the *T. reesei* U6 promoter and terminator (denoted gAd3A TS1-2; SEQ ID NO:10):

gaattcggatccAAAAACACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTA
 ACTTCTGCAGTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTA
 TTATTTTTATTTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTTATTT
 ATTATAATATATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAA
 5 TAATTTATAGTAATAAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATG
 AAATGGTATTATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTG
 GCTATAAGTCTGCTGCAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTG
 ATGGTAGTCTATCgtcctcgagcaaaaggtgccGTTTTAGAGCTAGAAATAGCAAGTTAAAA
TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTT
 10 TTCTCTTgaattcggatcc

[0176] The following sequence encodes the gAd3A TS1 sgRNA (underlined) with the *T. reesei* U6 promoter, terminator and an intron (in italics) (denoted gAd3A TS1-3; SEQ ID NO:11):

gaattcggatccAAAAACACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTA
 ACTTCTGCAGTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTA
 TTATTTTTATTTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTTATTT
 ATTATAATATATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAA
 TAATTTATAGTAATAAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATG
 20 AAATGGTATTATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTG
 GCTATAAGTCTGCTGCAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTG
 ATGGTAGTCTATCgtcctcgagcaaaaggtgccGTTTTAGAGCTAGAGTTCGTTTCGGCTTT
TCCTCGGAACCCCCAGAGGTCATCAGTTCGAATCGCTAACAGAAATAGCAAGTTAAA
ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTT
 25 TTTCTCTTgaattcggatcc

[0177] Plasmid p219M (FIG. 3) is an *E. coli* vector containing the *T. reesei* *pyr4* (orotidine monophosphate decarboxylase) gene including its native promoter and terminator. This vector was digested with EcoRI and BamHI and the ends were
 30 dephosphorylated. Each of the above synthetic DNA molecules was digested with EcoRI and BamHI and ligated with the cut p219M to create a series of vectors containing an sgRNA expression cassette and the *pyr4* gene. Each vector was designated by the name of the sgRNA that it encoded (for example, p219M gAd3A

TS1-1 incorporates the gAd3A expression cassette with the *S. cerevisiae* snr52 promoter and sup4 terminator).

[0178] Guide RNA expression cassettes with a shorter *T. reesei* U6 promoter region were obtained as synthetic DNA. An example is provided here that includes the sequence for an sgRNA targeting the *T. reesei gla1* gene at TS11 (SEQ ID NO:12; intron sequence is underlined).

AATTCCTAAAGAAACAGCATGAAATGGTATTATGTAAGAGCTATAGTCTAAAGGCA
CTCTGCTGGATAAAAATAGTGGCTATAAGTCTGCTGCAAACTACCCCCAACCTCG
TAGGTATATAAGTACTGTTTGATGGTAGTCTATCgccaatggcgacggcagcacGTTTTAGA
10 GCTAGAGTTCGTTTCGGCTTTTCCTCGGAACCCCCAGAGGTCATCAGTTCGAATC
GCTAACAGAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG
CACCGAGTCGGTGGTGCTTTTTTTTCTCTT

[0179] The above gRNA expression cassette was amplified by PCR using primers gRNA fwd *af*III (5'- cgtcagcttaagAATTCCTAAAGAAACAGCATGAAATGG; SEQ ID NO:13) and gRNA rev *sfi*I (5'-

cgtcagggccacgtgggccAAGAGAAAAAAGCACCCCGACTCGG; SEQ ID NO:14).

These primers add an *af*III to the 5' end and an *sfi*I site to the 3' end of the guide RNA expression cassette. The PCR product was purified using a Qiagen PCR Purification Kit according to the manufacturer's directions. The PCR product was then digested with *Sfi*I and *Af*III and cleaned again on a Qiagen PCR Purification Kit. Plasmid pTrex2g/Hyg MoCas was digested with *Sfi*I and *Af*III and dephosphorylated using the Roche Rapid alkaline phosphatase kit (Roche Diagnostics Corp., IN). The digested plasmid and PCR product were finally ligated using the Roche Rapid DNA ligase kit to create pTrex2g/Hyg MoCas gTrGA TS11B. Other sgRNA expression cassettes were inserted into pTrex2g/Hyg MoCas in a similar manner.

Example 6: Cas9-mediated gene inactivation in Trichoderma reesei

[0180] A series of experiments are described below in which a *Trichoderma reesei*

strain was either co-transformed with two separate expression vectors, one for production of Cas9 and one for production of gRNA, or was transformed with a single vector for expression of both Cas9 and gRNA. These experiments demonstrate that the 5' upstream region from the *T. reesei* U6 gene promoted gRNA transcription only when

the U6 intron is also present within the gRNA transcribed region. The experiments also demonstrate that targeted gene inactivation can occur with high efficiency in *T. reesei* transformants.

5 Inactivation of the *ad3A* gene

[0181] A strain of *Trichoderma reesei* derived from the publicly available strain RL-P37 in which the genes (*cbh1*, *cbh2*, *egl1*, and *egl2*) encoding the four major secreted cellulases were deleted was used. This strain also lacked a functional *pyr4* gene. Biolistic transformation (as described in US20060003408A1) was used to co-transform
 10 with a mixture of equal amounts of pTrex2gHyg MoCas (FIG. 2) and either p219M gAd3A TS1-1, p219M gAd3A TS1-2, or p219M gAd3A TS1-3. Transformants were selected on agar plates with Vogel's minimal medium containing 2% glucose, 100 mg/L hygromycin B and 200 mg/L adenine. After selection on the first plates transformant colonies were picked to fresh plates of the same selective medium. During growth on
 15 the second plate it was possible to distinguish between stable and unstable hygromycin-resistant transformants. Stable transformants grew more rapidly, the colonies had a smooth outline and the mycelium was more dense. Unstable transformants grew slower, had less dense mycelium and colonies had a ragged irregular outline. After growth on the second plate transformants were transferred to
 20 Vogel's medium with glucose, without hygromycin and with 14mg/L adenine to screen for those which exhibited a red/brown color indicating that they were adenine auxotrophs. Five stable and 23 unstable transformants were obtained with p219M gAd3A TS1-1 and all were adenine prototrophs. Eleven stable and 38 unstable transformants were obtained with p219M gAd3A TS1-2 and all 11 stable and 29 of the
 25 unstable transformants were adenine prototrophs. Nineteen stable and 2 unstable transformants were obtained with p219M gAd3A TS1-3 and all were adenine auxotrophs. Clearly, adenine auxotrophs were only obtained with gAd3A TS1-3 that utilizes the *T. reesei* U6 promoter, intron and terminator to control transcription of sgAd3A TS1. Adenine auxotrophy indicates targeted Cas9 cleavage at the native *T.*
 30 *reesei ad3A* locus. It can be concluded that Cas9-mediated gene inactivation is efficient because all transformants with gAd3A TS1-3 that were tested were adenine auxotrophs.

[0182] In order to determine the mutations at the *ad3A* locus in co-transformants with pTrex2gHyg MoCas and p219M gAd3A TS1-3, genomic DNA was extracted from 10

stable adenine auxotrophic transformants. This DNA was used as template for PCR using several different primer pairs designed to generate products that spanned the Cas9 target site or were upstream or downstream of the target site. PfuUltra II Fusion HS DNA polymerase (Agilent Technologies, Santa Clara, CA) was used for the PCR according to the manufacturer's directions. In each case, the extension time was that suggested by the manufacturer for the expected size of the PCR product as described below. The sizes of the PCR products were evaluated by agarose gel electrophoresis.

[0183] A PCR product of the expected size (872 bp) was obtained in all transformants using Ad3 5' fwd + Ad3 5' rev primers (5'- tgaacacagccaccgacatcagc [SEQ ID NO:15] and 5'- gctggtgaggggttgctattg [SEQ ID NO:16] respectively) that amplify a region on the 5' side of the TS1 target site.

[0184] A PCR product of the expected size (1214 bp) was obtained in all transformants using Ad3 5' fwd + Ad3a 5005 rev primers (5'- tgaacacagccaccgacatcagc [SEQ ID NO:15] and 5'- gattgcttggaggaggacat [SEQ ID NO:17] respectively) that amplify a region on the 5' side of the TS1 target site.

[0185] A PCR product of the expected size (904 bp) was obtained in all transformants using Ad3 3' fwd + Ad3 3' rev primers (5'- cgaggccactgatgaagttgttc [SEQ ID NO:18] and 5'- cagtttccaaggctgccaacgc [SEQ ID NO:19] respectively) that amplify a region on the 3' side of the TS1 target site.

[0186] A PCR product of the expected size (757 bp) was obtained in all transformants using Ad3a 5003 fwd + Ad3mid rev primers (5'- ctgatcttgccacctggaaatc [SEQ ID NO:20] and 5'- ctctctatcatttgccaccctcc [SEQ ID NO:21] respectively) that amplify a region on the 3' side of the TS1 target site.

[0187] The above PCR results demonstrated that the genomic DNA preparations were of a quality sufficient to obtain PCR products from either upstream or downstream of the Cas9 target site.

[0188] No PCR product could be obtained for any transformants using Adfrag fwd + Adfrag rev primers (5'- ctccattcaccctcaattctcc [SEQ ID NO:22] and 5'- gttcccttggcgggtgcttgatc [SEQ ID NO:23] respectively) spanning the TS1 target site in ad3A. The expected size for this PCR product presuming no large size change caused by Cas9 activity was approximately 764 bp.

[0189] No PCR product could be obtained for any transformants using Adfrag fwd + Ad3 3' rev primers (5'- ctccattcacccctcaattctcc [SEQ ID NO:22] and 5'- cagttttccaaggctgccaacgc [SEQ ID NO:19] respectively) spanning the TS1 target site in ad3A. The expected size for this PCR product presuming no large size change caused by Cas9 activity was approximately 2504 bp.

[0190] No PCR product could be obtained for any transformants using Ad3a 2k fwd + Ad3a 2k rev primers (5'- caatagcacaaccctcaccagc [SEQ ID NO:24] and 5'- gaacaacttcacgtggcctcg [SEQ ID NO:25] respectively) spanning the TS1 target site in ad3A. The expected size for this PCR product presuming no large size change caused by Cas9 activity was approximately 1813 bp.

[0191] Five of the transformants also gave no PCR product using Adfrag fwd + Ad3 mid rev primers (5'- ctccattcacccctcaattctcc [SEQ ID NO:22] and 5'- ctctctatcatttgccaccctcc [SEQ ID NO:21] respectively) spanning the TS1 target site. The expected size for this PCR product presuming no large size change caused by Cas9 activity was approximately 1438 bp.

[0192] Based on published data, Cas9-mediated inactivation of genes typically involves error-prone repair of a double-strand break in the DNA at the target site. The end result is small deletions or insertions (indels) at the target site. The above results from PCR analysis were surprising in that it was not possible to obtain a PCR product of the expected size that spanned the target site suggesting that inactivation of ad3A was not due to small insertions or deletions (indels) at the target site. Instead, these data are consistent with the possibilities that inactivation of *ad3A* was caused by a chromosomal rearrangement or large insertion at the target site.

Inactivation of the glucoamylase (GA) gene

[0193] A strain of *Trichoderma reesei* derived from the publicly available strain RL-P37 in which the genes (*cbh1*, *cbh2*, *egl1*, and *egl2*) encoding the four major secreted cellulases were deleted was used. This strain also lacked a functional *pyr4* gene. This strain was co-transformed using the biolistic method with a mixture of equal amounts of pTrex2gHyg MoCas and p219M gTrGA TS2. Transformants were selected on agar plates with Vogel's minimal medium containing 1% glucose, 100 ug/ml hygromycin B and 2 mg/ml uridine. After selection on the first plates transformant colonies were picked to fresh plates of the same selective medium. During growth on the second plate

it was possible to distinguish between stable and unstable hygromycin-resistant transformants. Seventeen stable and 4 unstable transformants were obtained. These transformants were transferred to Vogel's agar plates without glucose and with 1% insoluble starch to screen for presence or absence of secreted glucoamylase. Colonies
5 able to secrete glucoamylase grow well and sporulate. Colonies unable to secrete glucoamylase grow with very sparse mycelium and are clearly distinguishable. Fourteen of the 17 stable transformants were unable to secrete glucoamylase and all 4 of the unstable transformants did not secrete glucoamylase.

[0194] In order to determine the mutations at the gla1 (glucoamylase) locus in co-
10 transformants with pTrex2gHyg MoCas and p219M gTrGA TS2 genomic DNA was extracted from 5 stable glucoamylase non-producing transformants. This DNA was used as template for PCR using different primer pairs designed to generate products that spanned the Cas9 target site or were upstream or downstream of the target site. PfuUltra II Fusion HS DNA polymerase (Agilent Technologies) was used for the PCR
15 according to the manufacturer's directions. In each case, the extension time was that suggested by the manufacturer for the expected size of the PCR product as described below. The sizes of the PCR products were evaluated by agarose gel electrophoresis.

[0195] No PCR product could be obtained for any transformants using glaA + glaB
primers (5'- ccgttagtgaagatccttgccg [SEQ ID NO:26] and 5'- gtcgaggattgtgcttcatacctc
20 [SEQ ID NO:27] respectively) spanning the TS2 target site in gla1. The expected size for this PCR product presuming no large size change caused by Cas9 activity was approximately 1371 bp.

[0196] A band of the expected size (364 bp) was obtained in all transformants using
glaA + glaJ primers (5'- ccgttagtgaagatccttgccg [SEQ ID NO:26] and 5'-
25 tgccgactttgtccagtgttcg [SEQ ID NO:30] respectively) that amplify a region on the 5' side of the TS2 target site.

[0197] A band of the expected size (520 bp) was obtained in 4 of the transformants
using glaK + glaB primers (5'- ttacatgtggacgcgagatagcg [SEQ ID NO:31] and 5'-
gtcgaggattgtgcttcatacctc [SEQ ID NO:27] respectively) that amplify a region on the 3'
30 side of the TS2 target site. One of the transformants gave no PCR product with this primer pair.

[0198] A separate experiment intended to demonstrate inactivation of the gla1 gene by
targeted Cas9 action was performed using a strain of *T. reesei* derived from RL-P37

and having an inactive *pyr4* gene. Protoplasts of this strain were transformed with pTrex2gHyg MoCas gTrGA TS11 using a polyethylene glycol-mediated procedure (as described below). Transformants were selected on agar plates of Vogel's minimal medium with 2% glucose, 2 mg/ml uridine, 1.1M sorbitol and 100 ug/ml hygromycin B.

5 After selection on the first plates transformant colonies were picked to fresh plates of the same selective medium without sorbitol. During growth on the second plate it was possible to distinguish between stable and unstable hygromycin-resistant transformants. Transformants were transferred to Vogel's agar plates without glucose and with 1% insoluble starch to screen for presence or absence of secreted
10 glucoamylase. Five stable transformants, designated B#1, B#2, B#4, B#5 and B#6, which did not secrete glucoamylase were selected for further analysis. Genomic DNA was extracted from each of these transformants.

[0199] PCR was performed using genomic DNA as template and primers *gla1repF* and *gla1repR* (5'-gtgtgtctaataatgcctccaccac [SEQ ID NO:32] and 5'-gatcgtgctagcgcgtgctgttg [SEQ ID NO:23] respectively) that generate a product of 983 bp from the wild-type *gla1*
15 locus spanning the TS11 target site. The PCR conditions included gradually reducing the primer annealing temperature with each PCR cycle and a long extension time to determine if there had been a large insertion at the target site. The specific PCR conditions were as follows.

20 Step 1: 94C for 1 minute

Step 2: 94C for 25 seconds

Step 3: 63C for 30 seconds (temperature reduced by 0.2C per cycle)

Step 4: 70C for 8 minutes

Steps 2-4 repeated 24 more times

25 Step 5: Hold at 4C

[0200] A clear PCR product of greater than 12 kb was obtained from two of the transformants (B#1 and B#6) suggesting an increase of greater than 11 kb in the DNA region spanning the target site. The other three transformants gave only non-specific PCR products that appeared as low intensity bands on agarose gel electrophoresis.

30 Sequence analysis of the >12 kb PCR product from B#6 demonstrated that DNA derived from plasmid pTrex2gHyg MoCas gTrGA TS11 was inserted at the TS11 target site.

[0201] PCR was performed using genomic DNA samples B#2, B#4, and B#5 and primer pair 1553R and 1555F (5'- CCGTGATGGAGCCCGTCTTCT [SEQ ID NO:34] and 5'- CGCGGTGAGTTCAGGCTTTTTC [SEQ ID NO:35] respectively). Primer 1553R binds to the *gla1* gene on the 3' side of target site 11. Primer 1555F binds near the start codon of the hygromycin phosphotransferase (*hygB*) gene on the plasmid pTrex2gHyg MoCas gTrGA TS11. The same PCR conditions were used as above. PCR products of 4.5 kb and 6.5 were obtained for transformants B#4 and B#5 respectively. PCR products should only be obtained if the plasmid with the *hygB* gene had inserted into the *gla1* gene. Presumably, the inserted plasmid DNA in transformants B#4, and B#5 was so large that it was not possible to obtain a PCR product using primers *gla1*repF and *gla1*repR.

[0202] Taken together, the PCR data demonstrated that stable hygromycin-resistant transformants with glucoamylase inactivation have arisen through insertion of large segments of the Cas9 and guide RNA expression vector at the target site in the *gla1* gene.

Inactivation of the *pyr2* gene

[0203] Transformants of *T. reesei* strains QM6a or RL-P37 were generated by PEG-mediated transformation of protoplasts with derivatives of plasmid pTrex2gHyg MoCas that included guide RNA expression cassettes targeting different positions within the *T. reesei pyr2* gene. Inactivation of this gene confers uridine auxotrophy and resistance to 5-fluoroorotic acid (FOA). Transformants were initially selected on medium containing hygromycin B. Upon transfer to fresh agar plates containing hygromycin B they were scored as stable or unstable. Transformants were then transferred to agar plates of Vogel's minimal medium with 2 mg/ml uridine and 1.2 mg/ml FOA. The ability to grow in the presence of FOA is indicative of uridine auxotrophy due to Cas9-mediated inactivation of the *pyr2* gene.

[0204] Genomic DNA was extracted from some of the FOA resistant hygromycin stable and unstable transformants for PCR analysis. The primers used for this analysis were *pyr2*F (5'-gtataagagcaggaggaggaggag [SEQ ID NO:36]) and *pyr2*R (5'-gaacgcctcaatcagtcagtcg [SEQ ID NO:37]) designed to amplify a region of the *pyr2* locus spanning the target sites and approximately 0.8kb in length.

[0205] Among the QM6a transformants shown to be FOA resistant 18 stable and 5 unstable hygromycin resistant transformants were tested using the PCR protocol with an extension time sufficient to amplify the region of the pyr2 locus presuming the size to be similar to that in a wild-type strain. None of the stable transformants gave a PCR product with this short extension time whereas 2 of the unstable transformants did give a PCR product. DNA sequence analysis of these two PCR products showed that one had a single nucleotide deletion and the other had a 111 nt deletion at the expected target site.

[0206] Among the RL-P37 transformants shown to be FOA resistant 4 stable and 2 unstable hygromycin resistant transformants were tested using the PCR protocol with a short extension time. None of the stable transformants gave a PCR product with this short extension time whereas both of the unstable transformants did give a PCR product. DNA sequence analysis of these two PCR products showed that one had a single nucleotide deletion and the other had an insertion of 134 nt at the expected target site. This insertion consisted of two small fragments of the pTrex2gHyg vector.

[0207] A different 6 stable hygromycin resistant RL-P37 transformants were analyzed using the PCR protocol described earlier designed to enable amplification of the region of the pyr2 locus presuming a large DNA fragment was inserted at the target site in the pyr2 locus. All 6 transformants gave a large PCR product (between approximately 5 kb and >12 kb depending on the transformant) with this long extension time protocol. DNA sequence analysis of 5 of these PCR products showed that pTrex2gHyg vector DNA, or fragments thereof, was integrated in all cases.

[0208] Taken together, these data show that repair of a double strand break caused by Cas9 predominantly involves integration of large vector fragments in stable transformants. This can be a very efficient method of gene inactivation. This also demonstrates that a DNA fragment or vector bearing a functional gene and having no sequence homology with the target site can integrate in a site-specific manner at the target site following Cas9 cleavage and double strand break formation. In contrast, small deletions or insertions (indels) are associated with inactivation of a gene by Cas9 in unstable transformants. This is the method of choice for gene inactivation if vector integration is undesirable.

Example 7: Expression of cas9 and sgRNA using expression vector with telomeres

[0209] A version of the Cas9 and guide RNA expression vector pTrex2gHyg MoCAS gPyr2 TS6 was constructed that contained *Trichoderma reesei* telomere sequences (shown in FIG. 6). The DNA sequence shown below (SEQ ID NO:38) was inserted into the vector. The underlined regions contain the repeated telomere sequences, each reading in towards center of this fragment. The central portion is a bacterial kanamycin resistance gene with promoter and terminator that enables selection in *E. coli* to ensure maintenance of the telomere repeats. In *Trichoderma*, a vector with telomeres is expected to linearize with the telomere sequences at each end and should be maintained autonomously at low copy number although occasional integration into the chromosomal DNA can also occur.

tcaggaaatagctttaagtagcttattaagttataaaattatatatattttaataataactatatttcttaataaataaggtattttaag
ctttatatataaataataataaaaataatattatatagctttttattaataaataaaaataagctaaaaatataaaaaaaatag
ctttaaaatacttattttaattagaattttatatattttaataataaagatcttttactttttataagcttcctacctaataatttta
cttttttactattttactatatcttaataaaggctttaaaaaatataaaaaaaatcttcttatattataagctataaggattatat
atatatttttttaattttaagtaagttataaagctagaattaaagtttaatttttaaggctttatttaaaaaaaggcagtaata
gcttataaaagaaatttcttttttatactaaaagtactttttttaataagggttaggggttaggggttactcacaccgaccatcc
caaccacatcttaggggttaggggttaggggttaggggttaggggttaggggttaggggttaaacaaagccacgtt
gtgtctcaaaatctctgatgttacattgcacaagataaaaaatatcatcatgaacaataaaaactgtctgcttacataaacag
taatacaaggggtgttatgagccatattcaacgggaacgtctgtctcgaggccgcgattaaattccaacatggatgtctga
tttatatgggtataaatgggctcgcgataatgtcgggcaatcagggtcgcacaatctatcgattgtatgggaagcccgatgcg
ccagagtgtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggctgacg
gaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcattgttactcaccactgcgatccccgggaaa
acagcattccaggtattagaagaatacctgattcaggtgaaaatattgttgatgcgctggcagtggtcctgcgccggttgca
ttcgattcctgtttgtaattgtcctttaacagcgatcgcgtatttcgtctcgtcaggcgcaatcacgaatgaataacggtttgg
tgatgcgagtgattttgatgacgagcgtaatggctggcctgttgacaagctctggaaagaaatgcataagcttttgccattct
caccggattcagtcgtcactcatggtgatttctcacttgataacctattttgacgaggggaaattaataggttgattgatgttg
gacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtagttttctccttcattacagaa
acggccttttcaaaaatatggtattgataatcctgatataaataattgcagtttcatttgatgctcgatgagttttctaatacaga
attggttaattgggtgtaacactggcagagcattacgctgacttgacgggacggcggcgtttgtgaataaatcgaacttttgct
gagttgaaggatcagatcacgcattctccgacaacgcagaccgttccgtggcaaagcaaaagtcaaaatcaccaact
gggtccacctacaacaaagctctcatcaaccgtggctccctcactttctggctggatgatggggcgattcaggcctggatga
gtcagcaaacaccttcttcacgaggcagacctaagcggtttaaacctaaccctaaccctaaccctaaccctaaccctaacc
ctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaacc
ctaacctaatggggtcgatctgaaccgaggat
gaggggtctatagactaatctacaggccgtacatggtgtgattgcagatgcgacgggcaagggttacagtgtccagaagg

aggagagcgccataggtattgtaatagaccagctttacataataatcgccctgttgctactgactgatgaccttctccctaac
cagtttcctaattaccactgcagtgaggataaccctaactcgctctgggggtattattatactgattagcaggtggccttatatagt
gctgaagtactataagagtttctgcgggaggaggtggaaggactataaaactggacacagttagggatagagtgatgaca
agacctgaatgttatcctccggtgtggtatagcgaattggctgaccttcagatggtaatggtttaggcagggttttgcagag
5 ggggacgagaacgcgttctgcgatttaacggctgctgccccaagctttacggttctctaattggcgccgc (SEQ ID
 NO:38)

[0210] This vector was inserted into *T. reesei* strain RL-P37 by PEG-mediated transformation of protoplasts. Transformants were selected for hygromycin resistance and transferred to fresh agar plates with hygromycin. The majority of transformants showed an unstable hygromycin resistance phenotype. Individual transformed colonies were transferred to minimal medium agar plates containing 2 mg/ml uridine and 1.2 mg/ml 5-fluoroorotic acid to select for those that were able to grow and thus had a Pyr-minus phenotype. Eight out of 142 (6%) of the unstable transformants were Pyr-minus. Analysis by PCR of the *pyr2* locus and sequencing of three of these transformants showed that two had small deletions at the target site (1 bp and 27 bp respectively) and one had a 1 bp deletion combined with an insertion of 68 bp derived from the bacterial vector portion of pTrex2gHyg MoCAS gPyr2 TS6. The other 5 transformants did not give a PCR product despite using PCR conditions designed to amplify large DNA fragments [PCR conditions: Step 1: 94 °C for 1 minute; Step 2: 94 °C for 25 seconds; Step 3: 63C for 30 seconds (temperature reduced by 0.2C per cycle); Step 4: 70 °C for 8 minutes; Steps 2-4 repeated 24 more times; Step 5: Hold at 4 °C. Polymerase: PfuUltra II Fusion HS DNA polymerase (Agilent Technologies)].

[0211] These results demonstrate that expression of Cas9 and guide RNA from an autonomously replicating vector enables Cas9 targeting to a specific locus (*pyr2* in this case). The resulting gene inactivation can occur without insertion of vector DNA at the target site.

Example 8: Gene editing by homologous integration

[0212] *Trichoderma reesei* strain T4(1)7 was used for the following experiments. This is a strain derived from RL-P37 by screening for increased cellulase productivity and having a single point mutation that inactivates the *pyr2* gene making the strain a uridine auxotroph.

[0213] A synthetic DNA fragment called Gla1rep having the sequence shown below (SEQ ID NO:39) was designed and custom-made.

gtgtgtctaatagcctccaccacaggaaccaaaccggccttgacctctgggaagaagtcaatgggagctcattcttactgtt
 gccaaaccagcaccgaggtatgaagcaaactcctcgacattcgctgctactgcacatgagcattgttactgaccagctctac
 5 agcacttgtcgagggcgccactcttgcctgactcctggccagtcgggaagcgcttattcatctgttgcctccaggtttgtg
 ctttctccaacgattctgggtgtcgtctggtggatacgtcgactccaacagtatgtctttcactgtttatatgagattggccaata
 ctgatagctcgctctagtcacaccaacgagggcaggactggcaaggatgtcaactccgtcctgacttccatccacacc
 ttgatcccaaccttggctgtgacgcaggcacctccagccatgcagtgacaaagcgctctccaacctcaaggttgtgtcg
 actccttccgctccatctacggcgtgaacaagggcattcctg**CG**gtgctgccgtcgccattggccggtatgcagaggatgt
 10 gtactacaacggcaaccttggatcttgcctacatttgcctgctgccgagcagctgtacgatgccatctacgtctggaagaag
 acgggtcccatcacggtgaccgccacctccctggccttctccaggagcttgttctggcgtgacggccgggacctactcc
 agcagctcttcgacctttaccaacatcatcaacgccgtctcgacatacgccgatggcttctcagcgaggctgccaagtac
 gtccccgccgacgggtcgctggccgagcagtttgaccgcaacagcggcactccgctgtctgcgcttcacctgacgtggtc
 gtacgcctcgttcttgacagccacggcccgtcgggctggcatcgtgccccctcgtgggccaacagcagcgctagcacg
 15 atc

[0214] The Gla1rep sequence is 982 bp of the gla1 locus from within the ORF. It spans the TS11 target site (underlined). A single “C” nucleotide within the “CCG” PAM sequence of the wild type Gla1 gene, right upstream of the TS11 target site, has been
 20 deleted to create a frame shift in the Gla1 coding sequence and to destroy the PAM adjacent to TS11, thereby preventing cleavage by Cas9. The remaining two nucleotides of the PAM are shown in upper case bold font.

[0215] The Gla1rep fragment was amplified by PCR for use in transformation using the primers gla1rep F and gla1rep R (5'-gtgtgtctaatagcctccaccac [SEQ ID NO:32] and 5'-
 25 gatcgtgctagcgctgctgttg [SEQ ID NO:33] respectively).

[0216] Protoplasts of *T. reesei* strain T4(1)7 were co-transformed by the PEG-mediated method with pTrex2gHyg MoCas gTrGA TS11B (2 ug) plus Gla1rep (8 ug). Transformants were selected on agar plates with Vogel's minimal medium containing 50 ug/ml hygromycin B, 2 mg/ml uridine and 1.1M sorbitol. Plasmid pTrex2gHyg MoCas
 30 gTrGA TS11B is the same as pTrex2gHyg MoCas gTrGA TS11 except that the expression cassette for TS11 guide RNA is in the opposite orientation relative to the rest of the plasmid.

[0217] Transformants were picked to fresh agar plates of Vogel's minimal medium with uridine and hygromycin and it was possible to distinguish between stable and unstable hygromycin resistant phenotypes. Transformants were transferred to agar plates of Vogel's minimal medium with uridine and 1% insoluble starch as sole carbon source in order to score for glucoamylase positive or negative phenotypes. Approximately 83% of the stable hygromycin resistant transformants were negative for glucoamylase production whereas 15% of the unstable hygromycin resistant transformants were negative for glucoamylase production. Seven unstable transformants with glucoamylase-minus phenotype were transferred to non-selective agar medium (Vogel's + uridine) and allowed to grow for 1 week. When subsequently picked to plates of Vogel's + uridine + hygromycin they were all hygromycin-sensitive demonstrating loss of the hygromycin resistance gene associated with pTrex2gHyg MoCas gTrGA TS11B.

[0218] Genomic DNA was isolated from 5 unstable hygromycin-sensitive and glucoamylase-negative transformants obtained with pTrex2gHyg MoCas gTrGA TS11B plus Gla1rep (transformants #31, 107, 114, 118 and 120) and used as template in PCR (program as described above) using primers glaA and glaD (5'- ccgtagtgtaagatccttgccg [SEQ ID NO:26] and 5'- gagagacgcaggatgactcaaag [SEQ ID NO:28] respectively) designed to amplify approx. 3.2 kb spanning TS11 or glaK [SEQ ID NO:31] (see above) and glaH 5'- tgccgtgggtcattggcatattc [SEQ ID NO:29]. The PCR products were sequenced using gla1rep F and gla1rep R (5'-gtgtgtctaatgcctccaccac [SEQ ID NO:32] and 5'- gatcgtgctagcgctgctgttg [SEQ ID NO:33] respectively) as primers to determine the alterations at the target site TS11. One of the transformants showed PCR and sequencing results consistent with homologous recombination of Gla1rep at the gla1 locus that introduced the single bp deletion at the PAM associated with TS11 and inactivated the gla1 gene. Two of the transformants had small indels at the TS11 target site whereas the other two showed insertion of fragments of Gla1rep into the Cas9 cleavage site rather than homologous integration across this site.

[0219] The above experiment was repeated in which protoplasts were co-transformed with pTrex2gHyg MoCas gTrGA TS11A (identical to pTrex2gHyg MoCas gTrGA TS11B except that the guide RNA expression cassette was in the opposite orientation within the vector) plus a linear DNA fragment designed to integrate by homologous recombination at the gla1 locus. However, instead of using the 982 bp Gla1rep DNA fragment as donor for homologous recombination at the target site TS11 in the

glucoamylase gene a longer, approximately 2 kb, fragment called Gla1repL was used. The central portion of Gla1repL was the same sequence as Gla1rep but the 5' and 3' ends of the fragment were extended to include more of the upstream and downstream portions of the gla1 locus. *Trichoderma reesei* strain RL-P37 was used in this experiment instead of strain T4(1)7 used above. As a control, protoplasts were co-transformed with Gla1repL and pTrex2gHyg MoCas to determine the frequency with which Gla1repL integrates at the gla1 locus in the absence of active Cas9. Following transformation and phenotypic screening transformants could be assigned to the following categories.

Transforming DNA	No. Hyg ^R transformants	No. glucoamylase-minus transformants (% among respective Hyg ^R transformants)
pTrex2gHyg MoCAS + Gla1repL	26 unstable 19 stable	0 2 (10%)
pTrex2gHyg MoCAS gTrGA TS11A + Gla1repL	52 stable 16 unstable	46 (88%) 14 (87%)

[0220] Genomic DNA was isolated from 5 stable and 5 unstable hygromycin-sensitive and glucoamylase-negative transformants obtained with pTrex2gHyg MoCas gTrGA TS11A plus Gla1repL (stable transformants #51, 52, 60, 61 and 67; unstable transformants 338, 41, 65, 66 and 68) and used as template in PCR using primers glaA and glaD (see above). The PCR product was expected to be 3.2 kb if no insertion or large deletion had occurred at the target site TS11 in the gla1 gene.

[0221] Three of the 5 stable transformants (#52, 60 and 61) gave a PCR product of approximately 3.2 kb whereas the other 2 gave larger products indicative of an insertion of DNA at TS11. The 3 PCR products of approximately 3.2 kb were sequenced using glarepF as a primer. For two transformants (#52 and 61) the sequencing results were consistent with integration of Gla1repL by homologous recombination at the gla1 locus and the other had a mixed signal that could not easily be interpreted.

[0222] Only one of the 5 unstable transformants (#66) gave a PCR product of approximately 3.2 kb whereas the other 4 gave larger products indicative of an insertion

of DNA at TS11. The one PCR product of approximately 3.2 kb was sequenced using glarepF as a primer and the results were consistent with integration of Gla1repL by homologous recombination at the gla1 locus.

[0223] Taken together, these results show that homologous integration of a linear DNA fragment can be stimulated by Cas9 cleavage at a targeted locus. However, small indels or large insertions of DNA by non-homologous end joining (NHEJ) are also common occurrences. Use of a larger homologous linear DNA fragment helps to improve the frequency of homologous integration at the target site versus other events. It is possible to obtain homologous integration at the target site in unstable hygromycin resistant transformants from which the pTrex2g MoCas-based vector can subsequently be removed by allowing growth on medium without hygromycin.

Example 9: Gene editing in a NHEJ-deficient strain of *T. reesei*

[0224] A strain (MAD6) derived from a “quad-delete strain” of *Trichoderma reesei* (derived from RL-P37 and having the cellobiohydrolase 1, cellobiohydrolase 2, endoglucanase 1, and endoglucanase 2 genes deleted (Δ cbh1, Δ cbh2, Δ egl1, and Δ egl2 strain; see WO 92/06184 and WO 05/001036)) and having deletions in the native endoglucanase-3 and betaglucosidase-1 genes was used for experiments designed to determine the role of non-homologous end joining (NHEJ) DNA insertion at Cas9 target site. The MAD6 strain was also deleted for a native gene, orthologous to human ku80, essential for the major NHEJ pathway for DNA recombination (see US20130149742 A1 “Filamentous fungal host strains and DNA constructs, and methods of use thereof” for a description of how the MAD6 strain was made). The strain was co-transformed with pTrex2gHyg MoCAS gTrGA TS11B plus donor Gla1rep fragment described above. Integration of this fragment by homologous recombination at the *gla1* locus would inactivate the *gla1* gene and remove the TS11 target site by deleting one bp from the PAM sequence. Transformants were obtained from protoplasts by the PEG-mediated method. Selection for transformants was on Vogel’s minimal medium containing 1.1M sorbitol and 100 ug/mL hygromycin B. Out of 91 transformants transferred to fresh agar plates of minimal medium with hygromycin only 4 had a stable hygromycin resistant phenotype (confirmed by their ability to grow when re-plated on medium with hygromycin following a period of growth under non-selective conditions). All transformants were transferred to Vogel’s minimal medium with 1% insoluble starch as sole carbon source, and 17 (18%) were shown to be glucoamylase-negative, including

the 4 stable transformants. PCR and DNA sequence analysis showed that 12 of the 13 unstable transformants and one of the stable transformants had the single bp deletion at the TS11 PAM expected if donor *Gla1rep* had integrated by homologous recombination at the *gla1* locus. The other unstable transformant had the wild-type *gla1* sequence even though it had a glucoamylase-negative phenotype. The other three stable transformants did not give a clear PCR product of the size expected for the *gla1* locus and vector or donor *Gla1rep* insertion may have occurred in these. All of the unstable glucoamylase-negative transformants were grown on medium without hygromycin and transferred back onto medium with hygromycin. None were able to grow indicating that they had lost the pTrex2gHyg MoCAS gTrGA TS11B vector.

[0225] These results clearly show that vector or donor DNA fragment insertion at the Cas9 target site is minimized in a strain deficient for NHEJ. As a result, a high frequency of very specific gene editing (deletion of a single bp) is possible through homologous recombination of a donor DNA fragment in unstable transformants with transient expression of Cas9 and guide RNA.

Section B: Direct introduction of Cas Nickase/guide RNA

Example 10: Heterologous expression of CRISPR SpyCas9-D10A nickase in *E.coli*

[0226] *E. coli* codon-optimized Streptococcus pyogenes Cas9-D10A (SpyCas9-D10A) nickase gene was synthesized and inserted into the expression vector pET30a at NcoI and HindIII sites by Generay (Shanghai, China), resulting in the plasmid pET30a-SpyCas9-D10A nickase (FIG. 7). As indicated in the plasmid map in FIG. 7, the full coding sequence of the expression cassette contains, in 5' to 3' orientation, a sequence encoding an N-terminal His6 tag / thrombin / S•Tag™ / enterokinase region (SEQ ID NO:68, including a start codon for methionine), a sequence encoding an SV40 nuclear localization signal (SEQ ID NO:69), a sequence encoding the SpyCas9-D10A nickase (SEQ ID NO:65), and a sequence encoding the BLR2 nuclear localization signal (SEQ ID NO:70) all in operable linkage. This entire coding sequence is shown in SEQ ID NO: 54. The amino acid sequence of the N-terminal His6 tag / thrombin / S•Tag™ / enterokinase region encoded by SEQ ID NO:68 is shown in SEQ ID NO:67 (including the methionine at position 1), the amino acid sequence of the SV40 nuclear localization signal encoded by SEQ ID NO:69 is shown in SEQ ID NO:46, the amino acid sequence of the SpyCas9-D10A nickase encoded by SEQ ID NO:65 is shown in SEQ ID NO:66,

and the amino acid sequence of the BLR2 nuclear localization signal encoded by SEQ ID NO:70 is shown in SEQ ID NO:47. The amino acid sequence encoded by SEQ ID NO: 54 is shown in SEQ ID NO:55.

[0227] The pET30a-SpyCas9-D10A nickase plasmid was transformed into Rosetta2

(De3)plysS E. coli strain (Novagen®, EMD Biosciences, Inc., Merck KGaA, Darmstadt, Germany) and the transformation products were spread on Luria Agar plates supplemented with 34ppm Chloramphenicol and 50ppm Kanamycin. Colonies were picked and subjected to fermentation in 25ml of Invitrogen MagicMedia™ (Thermo Fisher Scientific Inc.) in a 250ml shake flask for 24 hours at 30°C at 300 rpm.

[0228] The amino acid sequence of the wild-type Cas9 protein from *Streptococcus pyogenes*, from which the SpyCas9-D10A sequence is derived, is set forth as SEQ ID NO:45.

Example 11: Purification of SpyCas9-D10A

[0229] For purification of SpyCas9(D10A), a combination of affinity, hydrophobic interaction and size exclusion chromatographic steps were applied. Two liters of crude broth were obtained and centrifuged. Cells were pelleted and resuspended in 400ml lysis buffer (20mM HEPES, pH7.5, 500mM NaCl, 0.1% Triton X-100, 1mM DTT and 1 mM TCEP, protease inhibitor cocktail purchased from Roche) and lysed via ultrasonicator (35% power, 20 min, 2s on/3s off) (SCIENT2-II D, Ningbo Scientz Biotechnology Co., LTD., Zhejiang, China). The lysate was cleared by centrifugation at 20,000g for 40 min.

[0230] The clarified lysate was incubated with Ni-NTA resin (GE Healthcare) overnight at 4°C, 30 rpm in a Rolling Incubator (Kylin-Bell Lab Instruments Co., Ltd., Haimen, China). After centrifugation, the resin was transferred to a XK26/20 column (GE Healthcare) and connected to AKTA Explorer system (GE Healthcare). After washing extensively with equilibration buffer (20 mM HEPES, pH 7.5, 300 mM NaCl, 0.1% Triton X-100) and wash buffer (25 mM imidazole in equilibration buffer), the target protein was eluted with 50, 250 and 500 mM imidazole in equilibration buffer. The desired protein was found with relatively high purity in 50 and 250 mM imidazole eluates, which were pooled and further processed separately.

[0231] To the active fraction collected from the affinity step, ammonium sulfate was added to 0.6 M and loaded onto a 20 ml phenyl-Sepharose HP column (GE

Healthcare). The column was eluted with a gradient of 0.6 M to 0.0 M ammonium sulfate in HEPES buffer at pH 7.5. The purity of each fraction was evaluated by SDS-PAGE gel which revealed that the protein of interest mainly present in the flow-through fraction.

- 5 **[0232]** Finally, the protein was further purified by size exclusion chromatography on a Superdex 200 16/60 column (GE Healthcare) in 20 mM HEPES pH7.5, 150 mM KCl and 10% glycerol. The protein-containing pure fractions were pooled and concentrated using an Amicon 30 KDa membrane filter (Millipore). The two batches of purified protein (from 50 mM imidazole and 250 mM imidazole elution, respectively) were stored in 20
10 mM HEPES buffer with 150 mM KCl and 40% glycerol at pH7.5 at -20°C until use.

Example 12: Nickase in vitro assay

Preparation of substrate DNA fragment for in vitro nickase cleavage assays

- [0233]** Genomic DNA was extracted from a *Trichoderma reesei* strain derived from RL-
15 P37 and having the cellobiohydrolase 1, cellobiohydrolase 2, endoglucanase 1, and endoglucanase 2 genes deleted (Δ cbh1, Δ cbh2, Δ egl1, and Δ egl2 strain; also called “quad-delete strain”; see WO 92/06184 and WO 05/001036)) using a ZF Fungal/Bacterial DNA miniprep kit from Zymo Research Corporation (Irvine, CA). With 1 ng of extracted genomic DNA as template a DNA fragment containing the
20 *Trichoderma reesei* glucoamylase (TrGA) gene (Gene ID: 18483895) and its partial 5'-UTR (SEQ ID NO:56) was amplified by PCR using KOD-Plus PCR kit (Toyobo Co., LTD, Japan) and 0.4 μ M of each forward and reverse primers: 5'-gactgtctccaccatgtaattttc-3' (SEQ ID NO:57) and 5'-ggcagactacaagtctactagtactac-3' (SEQ ID NO:58). PCR products were purified and concentrated with a DNA Clean &
25 ConcentratorTM-5 kit from Zymo Research Corporation, and its DNA concentration was determined by NanoDropTM Spectrophotometer (Thermo Fisher Scientific Inc.).

- [0234]** SEQ ID NO:56 (below) shows the nucleotide sequences of the substrate DNA fragment. The UTR sequences are shown in lowercase while the TrGA gene is shown in uppercase. Two selected VT domains, TrGA_sgF1 and TrGA_sgR1, are shown in
30 bold and underlined, respectively (note that these sequences overlap).

gactgtctccaccatgtaattttccctgcgactccatataacgccgatcgtaaatcttcttcttcttctccttctcaacaa
acaacggatctgtgctttgcggtcccctgcgttcacgcgtcagggtcgactgctctgcagctcgataactccatggagccat
caacttgctatggtgtcaatcatcctatcgacaggtccaagaacaagccggcctccggctgcctcattcgctgtcgcaaga

cggcttgagtgtgtggctggaggattcgggggccccatattccaaccctttttccaaggccgctcggccggtgaggttgag
 gaaaaccatgggttgccctacatattatcgatgctggtgtttggtagtagcaatgtttgcggtggcagtttgagccgagccctcgt
 cttgggcttctgacccaggcaacgccatctgactagctgcgccgaaggaaggatgattcattgtacgacgccagtcattg
 gaatcttcaagtaaaagccccgacgaaccgacatgtcagatatcagaattctcctggctgggtggggttggttgagactgc
 5 ttacggagtcgatgcctcgtgactgtcatggccgcgtccagccctcctgggactctgtccgatattatgacacgagtaaagcc
 tgcattgatgtcagtttgctgcgtctcatgtcgagaacaacacacctgggtgctacataggcaatactacctcgtagcttcaaa
 gttgactgttttgcttgatgtctttgatcatgccatccatcccttgcttgacgtgcattgtggatctctacgtccagacggggag
 aaagcttgtctgtgataaagtacgatgatgcattgatgcctgtggctacggccctttatccccatcgtcatgcattctatatta
 atccaggagactctcctcctggcatgggtgagtacaagtgcaggagacatgtagaagcagagccacgcaacgtcttga
 10 catctgtacctattttgggccaaaaatcgagacccaccagctcgtcctaccttacatgtgaagatcttagcccacaatcctac
 tgtttactagtattactgcacagctgtcatcacgagtcctcgggtgcttgtaaaccagctcagctcctgagcacatgcagt
 aacgccgactcggcgtcatttcgccacacccaatttgacctgagggatgctggaagctgctgagcagatcccgttaccg
 attcatggcactactacatccatcagcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcc
 cggcattgtcgatgcagagatgcagatttcagcgggagattctagggtagggcgactactactactaataccacctagtca
 15 gtatgtatctagcaccggaggctaggcgggttagtgagcgggaacctgggtcattccatcgcaaccaggatcccgccacttcgt
 tgcgcttctgccccacggggcgggagttggcagaggcagaatgcggagcagccccctgtctgccctggccggggcct
 gttgaagcaagcagacgagagcagagcgggtgagaagcgggtggtgacgcttgacggtagcgaagacgagcagagaat
 cccgttaagccgaggctgggctccccccccgcgtcatcatcatgccatcctgctcttcagccactcgtctcctgcctcgt
 cgctccccctcccccgattagctgcgcattgtctcctgacagcgtgactaatgacgcgttgccagccatttcgctga
 20 cgcattcccgcatctgagctagctcgtcacgctggcaatcttggcccaggcagagcagcaagacggcgggcatgattg
 ggccgtgccctggcgggcatcagctggccatccgctgccacccgagaccgcatcaccgacttgcggatctctccgagc
 agcaggaggctgatcctggccggcgagacgattgaaaagggtgccggggccggagcaggacagcggcgagagc
 gagcagagagagaggaaaagaagaaggctgactgtcttattttcagccagccccggctcaacagaagcagaggagaa
 ggcaacgacgtcaacgacgacgacgacgacgaagacgggtgaagtccgttagttgaagatccttgccgtcaca
 25 caccatctcgtggatattgctttccctgccgttgcttgccacctgttccctcttcttcccccttcttctcattccgagcgt
 actgggtcctactccgcagccttcgggttgcccttctcttctgacacattgcaccgcccgtcgcggcacttgggccccggag
 aattcgcccttctgcagcattttggccctcagttcccatggggacggtccacacttctcttggccctgcagacctttgt
 cgtcgggtccgagtcggaagaagctcagcttgagcgttgagtagcatctacgcgcgaatcactggacaaagtccgcaa
 gacgaagccgtcgtgcctgctgctgctgtgtactgcgacaggcgctccgactgggggcatcggcataataaaaagat
 30 gcccgccttcgcatggacctggccatgagccactcggcatcggctctctctcaacgcttctctcacacatcctcttcat
 tccgcccattcATGCACGTCTGTGCGACTGCGGTGCTGCTCGGCTCCGTTGCCGTTCAA
 AAGGTCCT**GGAAGACCAGGATCAAGCGGTCTGTCCGACGTCACCAAGAGGTCT**
 GTTGACGACTTCATCAGCACCGAGACGCCTATTGCACTGAACAATCTTCTTTGCAA
 TGTTGGTCCTGATGGATGCCGTGCATTGGGCACATCAGCTGGTGCGGTGATTGCA

TCTCCCAGCACAATTGACCCGGACTGTAAGTTGGCCTTGATGAACCATATCATATA
TCGCCGAGAAGTGGACCGCGTGCTGAGACTGAGACAGACTATTACATGTGGACG
CGAGATAGCGCTCTTGTCTTCAAGAACCTCATCGACCGCTTCACCGAAACGTACG
ATGCGGGCCTGCAGCGCCGCATCGAGCAGTACATTACTGCCCAGGTCACTCTCCA
5 GGGCCTCTCTAACCCCTCGGGCTCCCTCGCGGACGGCTCTGGTCTCGGCGAGCC
CAAGTTTGAGTTGACCCTGAAGCCTTTCACCGGCAACTGGGGTCGACCGCAGCG
GGATGGCCCAGCTCTGCGAGCCATTGCCTTGATTGGATACTCAAAGTGGCTCATC
AACAACAACCTATCAGTCGACTGTGTCCAACGTCATCTGGCCTATTGTGCGCAACGA
CCTCAACTATGTTGCCAGTACTGGTCAGTGCTTGCTTGCTCTTGAATTACGTCTT
10 TGCTTGTGTGTCTAATGCCTCCACCACAGGAACCAAACCGGCTTTGACCTCTGGG
AAGAAGTCAATGGGAGCTCATTCTTTACTGTTGCCAACCAGCACCGAGGTATGAA
GCAAATCCTCGACATTCGCTGCTACTGCACATGAGCATTGTTACTGACCAGCTCTA
CAGCACTTGTGAGGGCGCCACTCTTGCTGCCACTCTTGGCCAGTCGGGAAGCG
CTTATTCATCTGTTGCTCCCCAGGTTTTGTGCTTTCTCCAACGATTCTGGGTGTG
15 TCTGGTGGATACGTCGACTCCAACAGTATGTCTTTTCACTGTTTATATGAGATTGG
CCAATACTGATAGCTCGCCTCTAGTCAACACCAACGAGGGCAGGACTGGCAAGGA
TGTCAACTCCGTCCTGACTTCCATCCACACCTTCGATCCCAACCTTGGCTGTGACG
CAGGCACCTTCCAGCCATGCAGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTGT
CGACTCCTTCCGCTCCATCTACGGCGTGAACAAGGGCATTCTGCCGGTGCTGCC
20 GTCGCCATTGGCCGGTATGCAGAGGATGTGTACTACAACGGCAACCCTTGGTATC
TTGCTACATTTGCTGCTGCCGAGCAGCTGTACGATGCCATCTACGTCTGGAAGAA
GACGGGCTCCATCACGGTGACCGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGTT
CCTGGCGTGACGGCCGGGACCTACTCCAGCAGCTCTTCGACCTTTACCAACATCA
TCAACGCCGTCTCGACATACGCCGATGGCTTCCTCAGCGAGGCTGCCAAGTACGT
25 CCCC GCCGACGGTTCGCTGGCCGAGCAGTTTGACCGCAACAGCGGCACTCCGCT
GTCTGCGCTTCACCTGACGTGGTTCGTACGCCTCGTTCTTGACAGCCACGGCCCGT
CGGGCTGGCATCGTGCCCCCCTCGTGGGCCAACAGCAGCGCTAGCACGATCCCC
TCGACGTGCTCCGGCGCGTCCGTGGTTCGATCCTACTCGCGTCCCACCGCCACG
TCATTCCCTCCGTGCGCAGACGCCCAAGCCTGGCGTGCCTTCCGGTACTCCCTACA
30 CGCCCCTGCCCTGCGCGACCCCAACCTCCGTGGCCGTACCTTCCACGAGCTCG
TGTCGACACAGTTTGGCCAGACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTGG
GCAACTGGAGCACGAGCGCCGCGGTGGCTCTGGACGCCGTCAACTATGCCGATA
ACCACCCCCTGTGGATTGGGACGGTCAACCTCGAGGCTGGAGACGTGCTGGAGT
ACAAGTACATCAATGTGGGCCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCAA

CCACACTTACACGGTTCCTGCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGAC
ACctggcagtcgtaatgaatcggcaaggggtagtactagtagactgtagtctgcc (SEQ ID NO:56)

***In vitro* transcription**

5 **[0235]** Two VT domains in the TrGA gene, TrGA_sgF1 and TrGA_sgR1 and their specific PAMs, were identified for downstream activity and transformation experiments. Oligonucleotides containing the T7 promoter and single-guide RNA sequences, with either TrGA_sgF1 (SEQ ID NO:59) or TrGA_sgR1 (SEQ ID NO:60) were synthesized and inserted into the pMD18T vector by Generay, resulted in pMD18T(T7-Spy-
10 TrGA_sgF1) (FIG. 8A) or pMD18T (T7-Spy-TrGA_sgR1) (FIG. 8B), respectively. DNA fragments for the in vitro transcription were amplified from either pMD18T (T7-Spy-TrGA_sgF1) or pMD18T (T7-Spy-TrGA_sgR1) by PCR with 0.4 μ M of each forward and reverse primers: 5'-tgatgacggtgaaaacctc-3' (SEQ ID NO:71) and 5'-aaaagcaccgactcgg-3' (SEQ ID NO:72). PCR products were purified and concentrated with the DNA Clean
15 & ConcentratorTM-5 kit from Zymo Research Corporation, and its DNA concentration was determined by the NanoDropTM Spectrophotometer .

[0236] With the above specific PCR product as template, RNA for VT domain TrGA_sgF1 or TrGA_sgR1 was generated by in vitro transcription using MEGAshortscriptTM T7 transcription kit from Invitrogen, Thermo Fisher Scientific Inc.
20 according to the manufacturer's instructions. Transcribed RNAs were purified using MEGAcleanTM Transcription Clean-Up kit from Invitrogen, Thermo Fisher Scientific Inc.. The RNA concentration was measured by NanoDropTM.

[0237] SpyCas9 in vitro assays were performed to confirm the function of the synthesized single-guide RNAs. To initiate the assay, 1 μ g of purified SpyCas9, 200 ng
25 of substrate DNA fragment and 200 ng of single-guided RNA (or water as control) were mixed together in 15 μ l reaction buffer containing 50 mM HEPES pH 7.3, 150 mM KCl, 0.5 mM DTT and 10 mM MgCl₂. Assays were carried out at 37° C for 20 min, followed by the addition of 2 μ g of Proteinase K (Sigma, Cat No. P6556). The reaction was continued at 40° C for 20 min and terminated by an additional incubation at 80° C for 20
30 min. And the reaction results were analyzed using 0.8% agarose gel, running at 140 volts for 30 min, and the result is shown in FIG. 9. In FIG. 9, Lane 1 is a DNA ladder (molecular weights are shown on the left), Lane 2 is a control SpyCas9 reaction (no guide RNA), Lanes 3 & 4 show SpyCas9 in the presence of TrGA_sgR1, and Lanes 5 & 6 show SpyCas9 in the presence of TrGA_sgF1. The size of intact DNA substrate (SEQ

ID NO:56) is 4.9 Kb (Lane 2), while the sizes of cut products in the presence of either sgRNA are 2.8Kb and 2.1Kb. As shown in FIG. 9, in the presence of specific single-guide RNA, SpyCas9 can successfully cut substrate DNA fragment into the desired sizes, confirming the correct function of the synthesized RNAs.

- 5 **[0238]** SEQ ID NO:59 (below) shows the oligonucleotide sequences for the transcription of the T7 promoter, CER domain, and the VT domain TrGA_sgF1. The VT domain is shown in upper case, while the T7 promoter and CER domain region are shown in bold and lower case, respectively.

10 **taatacgactcactata**GGGAAGACCAGGATCAAGgttttagagctagaaatagcaagttaaaataaggct
agtcggttatcaacttgaaaaagtggcaccgagtcggtgc

- 15 **[0239]** SEQ ID NO:60 (below) shows the oligonucleotide sequences for the transcription of the T7 promoter, CER domain, and the VT domain TrGA_sgR1. The VT domain is shown in upper case, while the T7 promoter and CER domain region are shown in bold and lower case, respectively.

taatacgactcactataGGACAGACCGCTTGATCCgttttagagctagaaatagcaagttaaaataaggcta
gtccggttatcaacttgaaaaagtggcaccgagtcggtgc

***In vitro* nickase cleavage assays with purified SpyCas9(D10A)**

- 20 **[0240]** The *in vitro* nickase cleavage assay is a two-step reaction. For the first step, 1 µg of purified SpyCas9(D10A), 200 ng of substrate DNA fragment and 200 ng of single-guided RNA (or water as control) were mixed together in 15 µl reaction buffer containing 50 mM HEPES pH 7.3, 150 mM KCl, 0.5 mM DTT and 10 mM MgCl₂. The reaction was performed as described in SpyCas9 nuclease assay. Following the
- 25 termination of the first-step reaction, 1 µg of SpyCas9(D10A) and 200 ng of specific single-guide RNA were added. After repeating the first-step reaction, the reaction results were subsequently analyzed using 0.8% agarose gel, running at 140 volts for 30 min, and the result is shown in FIG. 10. In FIG. 10, Lane 1 is the DNA ladder (molecular weights shown on the left), Lane 2 shows a reaction of SpyCas9(D10A) with
- 30 substrate DNA and TrGA_sgF1 alone, Lane 3 shows a reaction of SpyCas9(D10A) with substrate DNA and TrGA_sgR1 alone, and Lane 4 shows a reaction of SpyCas9(D10A) with substrate DNA and both sgRNAs. The size of intact DNA substrate (SEQ ID NO:56) would be 4.9 Kb, while the sizes of cut products with both sgRNAs would be

2.8Kb and 2.1Kb. As shown in FIG. 10, the substrate DNA fragment is only cut in the presence of both RNAs (Lane 4), suggesting that SpyCas9(D10A) is an active nickase.

Example 13: In vivo nickase uptake experiment

5 Protoplast preparation

[0241] For protoplast preparation, 5×10^8 spores of a quad-delete strain of *T. reesei* (described above) with additional deletions of the endoglucanase-3, endoglucanase-4, endoglucanase-5, endoglucanase-6, mannanase-1, and alpha-amylase genes, but with normal NHEJ mechanism (grown on a PDA plate for 5 days at 30° C) were inoculated
 10 into 50 ml germination medium (recipe described in US Patent No. 8,679,815) in a 250 ml shake flask with 4 baffles and incubated at 27 °C for 17 hours at 170 rpm. The mycelia were recovered by transferring the liquid volume into 50 ml conical tubes and spinning at 3000 rpm for 10 minutes. The supernatant was decanted and the mycelial pellets were washed twice using 1.2 M MgSO_4 -10 mM Na-phosphate buffer and
 15 resuspended in 15 ml lysing enzyme buffer (dissolve Lysing Enzymes from *Trichoderma harzianum* (Sigma catalog #L1412)) using in 1.2 M MgSO_4 – 10 mM Na-phosphate buffer (pH 5.8), 50 mg/ml). The cell suspension was transferred into a 250 ml shake flask with 4 baffles and shaken at room temperature for at least 2 hours at 200 rpm. The protoplasts were harvested by filtration through Miracloth (Calbiochem
 20 Art. No. 475855) folded in a glass funnel into a Greiner tube. 0.6 M Sorbitol - 0.1 M Tris-HCl buffer was added carefully on top of the filtered protoplasts. The protoplasts were collected by centrifugation for 15 minutes at 4000 rpm. The middle phase containing the protoplasts was transferred into a new tube and added at least an equal volume of 1.2 M Sorbitol - 10 mM Tris-HCl buffer. The protoplasts were collected by centrifugation for
 25 5 minutes at 4000 rpm, and washed two times using 1.2 M sorbitol, 10 mM Tris-HCl buffer. The pellet was resuspended into at least 1ml 1.2 M Sorbitol - 10 mM Tris-HCl pH 7.5 - 10 mM CaCl_2 buffer and the number of protoplasts counted under a microscope. The protoplast suspension was diluted using 4 parts of 1.2 M Sorbitol – 10 mM Tris-HCl – 10 mM CaCl_2 and 1 part of 25% PEG6000 – 50 mM CaCl_2 – 10mM Tris-HCl until
 30 5×10^8 per ml for the future transformation.

Preparation of deletion cassette

[0242] A TrGA deletion cassette was constructed and schematically depicted in FIG. 11. It contained a *pyr2* expression cassette including *pyr2* promotor, *pyr2* CDS and

pyr2 terminator, a 500 bp repeat sequence for subsequent loop out, flanked by 5' and 3' TrGA-homologous regions .

[0243] TrGA knockout transformants can be screened on Vogel's agar plates without glucose and with 1% insoluble starch (Vogel's-starch medium). TrGA knockout

5 transformants grow poorly on this medium compared to strains with an intact TrGA gene. The nucleotide sequence of the TrGA knockout cassette is 4248 base pairs in length: bases 1-1000 correspond to the TrGA 5' homologous region; bases 1001-2730 correspond to the *pyr2* expression cassette; bases 2739-3248 correspond to the 500 bp repeat; and bases 3249-4248 correspond to the TrGA 3' homologous region. The
10 nucleotide sequence of the TrGA knockout cassette is provided as SEQ ID NO:61 (shown below):

```

ccctgcctcgctgcctcccctcccctccccgattagctgcgcatgttctcctgacagcgtgactaatgacgcgttgccagccc
attcgctgacgcatcccgcatctgagtctagctcgtcacgctggcaatcttgcccaggcagagcagcaagacggcg
ggcatgattggccgtgccctggcgggcatcagctggccatccgctgccacccgagaccgcatcaccgactgtcgcat
15 ctctccgagcagcaggaggctgatcctggccggcgagacgattgaaaagggtgccggggccggagcaggacagcg
gcgagagcgagcgagagagaggaaaagaagaaggctgactgtcttatttcagccagccccggctcaacagaagca
gaggagaaggcgaacgacgtcaacgacgacgacgacgacgaagacgggtgaagtccgttagttgaagatccttg
ccgtcacaacaccatctcgtggatattgcttcccctgccgttgcggtgccacctgttccctcttcttcccccttcttctcatt
ccgagcgctactggttctactccgcagccttcggttgcccttcttcttgtagaccattgcaccgcccgtcgcggcacttg
20 gccccggagaattcgccctttcgcagcattttggccctcagttcccatggggacgggtccacacttctcttggccctgc
agacctttgtcgctcgggtccgagtcggaagaagctcagtttgagcgcttgagtagcatctacgcgcgaatcactggacaa
agtcggcaagacgaagccgtcgtcgctgctgctgttactgcgacaggcgctccgactgggggcatcggcata
ataaaaagatgccgccttcgcatggacctggccatgagccactcggcatcggtctctcttcaacgcttctctcacac
atcctccttattccgcccattatggtttaaacctcgagtttataagtgacaacatgcttcaaagcgctcatggctggcaca
25 agcctggaaagaaccaacacaaagcactgcagcaaatacagctgaattcgtcaccaattaagtgaacatcaacctga
aggcagagtatgaggccagaagcacatctggatcgagatcatggattgcccctctgttggaagtgaagaatctagaaa
gatggcggggatgagataagagcgatgggggggcacatcatctccaagacaaacaacctttgcagagtcaggcaat
tttcgtataagagcaggaggaggaggtccagtcatttcacagcggtaaaatacactctagacaatttcaagatgagttct
gccttgggtgacttatagccatcatcatacctagacagaagcttggggatactaagaccaacgtacaagctcgactgta
30 cgcttgacttccatgtgaaaactcgatacggcgcgctctaaattttatagctcaaccactccaatccaacctctgcatccct
ctcactcgtcctgatctactgttcaaatcagagaataaggacactatccaaatccaacagaatggctaccacctcccagct
gctgcctacaagcaggacttctcaaataccgcatcgacggcggtcctcaagttggcagcttcgagctcaagtcca
agcggatatccccctacttcttcaacgcggggaattccacacggcgcgctcgccggcgccatcgctccgcctttgca
aagaccatcatcgaggccaggagaaggccggcctagagttcgacatcgttctcgcccgccgctacaaggggcatccc

```

gctgtgctccgccatcaccatcaagctcggcgagctggcgccccagAACCTGGACCGCGTCTCCTACTCGTTGACCGCA
 aggaggccaaggaccacggcgaggcggaacatcgtcggcgcttcgctcaagggaagagggctctgattgtcga
 cgacgtcatcaccgccggcaccgccaagagggacgccattgagaagatcaccaaggaggcgccatcgtcgccgg
 catcgtcgtggccctggaccgcatggagaagctccccgctcggatggcgacgactccaagcctggaccgagtgccatt
 5 ggcgagctgaggaaggagtacggcatccccatctttgccatctcactctggatgacattatcgatggcatgaagggttt
 gctaccctgaggatatcaagaacacggaggattaccgtgccaaagtacaaggcgactgactgattgaggcggtcaatgt
 cagaaggagagaaagactgaaaagggtggaaagaaggagcaaattgtgttattattatttctatctcgaatcttctaga
 tcttgcgtaaataaacaagcgtaactagctagcctccgtacaactgctgaattgatacccgatggagggcagttattttat
 tttgttttcaagattttccattcgccgttgaactcgtctcacatcgcggtattgcccggttgccatgtgttctcactaccccaa
 10 gtccctcacgggtgtctcactttcttctcctttatcctccctattttttcaagtcagcgacagagcagtcatatggggatacgt
 gcaactgggactcacaacaggccatcttatggcctaatagccggcggttgatccactagtcattgtttaaacagcacat
 gcagtaacgccgactcggcgctatttcgccacaccaatttgacctgagggatgctggaagctgctgagcagatcccg
 taccgattcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaa
 agtcccggcattgtcgatgcagagatgcagatttcagcggcgattctagggtagggcgactactactactaataccacct
 15 agtcagtatgtatctagcaccggaggctaggcggttagtgacgggaacctgggtcattccatcgcaaccaggatcccg
 acttcgttgcgcttctgccccacggggcgagggtggcagaggcagaatgcggagcagcccctgtctgcctggccg
 gggcctgttgaagcaagcagacgagagcagagcggttgagaagcggtggtgacgcttgacggtacgaagacgagc
 gagaatcccgtaagccgaggctgggctaattaattaatgaatcggcaagggtagtactagtagactttagtctgccgg
 attattgattggagttggtcagtagaatgaaccacgggaatatcgggtcaccgggacatttgggatatagcgtttcgagaag
 20 ctgctggttgacgacattggagaaggatgccctttacgacttataccgctatgccgggtatattaatttagccgttatgaa
 ctcaaagagacgatgataatgatgacgagtaattgttcgtttcaatttcgaaagctgactccacgaagaatatgccaatg
 acccacggcatgaagcctgaactgggctgtgtaacactttaattgcctgacggcggaacaaacaaaggcggcagca
 atgttgagaccgtgtgataaaccaagggtcccgagggagagagagagagagagagagagagagagagagagtaggtgaaaga
 atgagtccgcctttgagtcactcgtctctctctccccctctctcactctctglatcccatcaacctcttccctgttctctctat
 25 cgcattcatgcgtttgcatcttccatttcattttcccttgagccccatctatgcaaactcatcatccggcgccctcgatggaat
 ccttgaccttgatgagaatcgccgtcatccaaggctccagcctgctcgtcgggtcgaactggaacagcagctcgctaaac
 tcatcctggctgtggttgcgacggcggtgcacaggtcctcgagcagcttgacttgattgagaggagaactcggggctctt
 tggcggtaggactcgacggcgcgggcggtgccgaccatgtcgcccggtggcgagggtggcagatgccggccttgaagca
 gtaggtcgagaggctccactcatggtgccgttgcgcatcatggtgttgatgatgcggctcgatcgtctcgatggcgccgtagt
 30 agtcgccgtcagggcgggcgaggctggcggtactcgtccagagctt

Transformation

[0244] To knock out TrGA gene in *T. reesei* using a Spycas9 nickase premix and deletion cassette, 25µg Spycas9 nickase protein in storage buffer (20 mM HEPES pH

7.5, 150 mM KCl, and 40% glycerol) was mixed with 20µg sgRNA (TrGA_sgR1) dissolved in nuclease-free water in 3 µl NEB buffer 3 (New England Biolabs) gently to obtain a 30 µl premix, and incubated for 30 min at room temperature. 30 µl premix was added to 200 µL protoplasts (1×10^8) with 20µg (20µl) deletion cassette and kept on ice for 30 min. After incubation on ice for 30 min, protoplasts were added to cooled molten sorbitol/Vogel agar (1.1 M sorbitol in minimal Vogel's agar with 2% glucose) to be used as the top layer of the minimal Vogel's agar plate (Davis et al., (1970) Methods in Enzymology 17A, pgs. 79-143 and Davis, Rowland, NEUROSPORA, CONTRIBUTIONS OF A MODEL ORGANISM, Oxford University Press, (2000)). The plates were incubated at 30° C for a week. The detailed steps are described in US Patent No. 8,679,815 (incorporated herein by reference).

[0245] Ninety transformants were selected and inoculated into four 24-well plates with 1 ml fresh Vogel agar per well alongside 3 controls: the quad-delete strain with additional deletions of the endoglucanase-3, endoglucanase-4, endoglucanase-5, endoglucanase-6, mannanase-1 (*T. reesei* strain described above; "cellu" in FIG. 12), the "cellu" strain with an additional deletion of the alpha-amylase gene (AA deletion; Δ AA in FIG. 12), and the "cellu" strain with an additional glucoamylase gene deletion (GA deletion; Δ GA in FIG. 12). After one week, the transformants were transferred into another four 24-well plates with 1 ml fresh Vogel's-starch agar per well. After one week, the morphology of these 90 strains was observed. 47 transformants have the retarded growth as compared with those of the "cellu" strain and the "cellu" strain with the AA deletion (Δ AA), but looked similar to that of the "cellu" strain with the GA deletion (Δ GA) in Vogel's-starch plate (FIG. 12). On Vogel's-starch agar, the sole carbon source was starch in comparison to the regular Vogel's agar with glucose as carbon source. Glucoamylase consecutively hydrolyzes α -1,4 glycosidic bonds from the non-reducing ends of starch, resulting in the production of glucose, which can be utilized by the fungal strain. When glucoamylase is deleted, the retarded growth can be observed due to the limited availability of glucose.

Strain verification

[0246] Forty-seven out of 90 strains showed retarded growth in Vogel's-starch agar. Nine of them (see FIG. 12, from #1 to #9) were selected randomly to perform PCR screening using Fw1 (5'-cactactacatccatacgcagcaaacaatgg-3' (SEQ ID NO:62)) and R3

(5'-ggtaagaagcacatgccagagttcg -3'(SEQ ID NO:63)) as primers. FIG. 13A shows the genomic loci of wild type and TrGA deletion strains as well as the primer annealing sites (the arrows indicate the direction of polymerization when used in PCR reactions).

[0247] The 9 strains were pre-screened by PCR with primer pair FW1 and R3. The PCR conditions for amplifying the PCR product were as follows: Step 1: 95 °C for 5 min. Step 2: 95 °C for 30 sec. Step 3: 60 °C for 30 sec. Step 4: 68 °C for 3 min. Steps 2, 3 and 4 were repeated for an additional 29 cycles. Step 5: 68 °C for 10 min. In the ideal condition, two PCR fragments (1.9-kb and 5.2-kb) would be expected from the PCR pre-screening. In this result, the 5.2-kb product could not be observed clearly.

However, the 1.9-kb fragment as compared with the 5.1-kb product from spores of celluligher (C1) and the PCR product from spores of celluligher plus TrGA deletion cassette (C2) could confirm for certain that the TrGA deletion cassette was integrated at the TrGA locus by homologous recombination (FIG.. 13B,. top gel).

[0248] Two (#1 and #3) out of the 9 stains were selected to do the further PCR confirmation with primer pair FW1/R1, F4/R3, and KOF1/KOR2, respectively. The TrGA deletion cassette integrated at the TrGA locus by homologous recombination was confirmed by PCR with primer pair FW1/R1 and F4/R3 further, when 2.0-kb and 2.2-kb fragments were obtained from the spores of #1 and #3 as compared with the result from C2 (spores of celluligher plus TrGA deletion cassette) (FIG. 13B, bottom left gel).

PCR with another pair of primers, KOF1 and KOR2, also confirmed the result, when no PCR product was got from the spores of #1 and #3 while a product of 2.1 kb was amplified from celluligher (C1) (FIG. 13B, bottom right gel, first 4 lanes, including molecular weight marker).

[0249] The PCR products from the spores of #1 and #3 for the whole region of TrGA locus by using primers TrGAF2 (5' gactgtctccaccatgtaattttc 3'(SEQ ID NO:64)) and R3 (SEQ ID NO:63) were obtained and their DNA sequence determined. The sequencing result showed that TrGA gene has been replaced by *pyr2* expression cassette (see PCR result in FIG. 13B, bottom right gel, last 2 lanes; data not shown for sequencing result).

[0250] The results above demonstrate that the SpyCas9 nickase and sgRNA were able to promote homologous recombination in filamentous fungus, allowing for homologous recombination-based gene deletion on the filamentous fungus *T. reesei*. Specifically,

47 out of 90 transformants (~52%) showed the retarded growth phenotype on Vogel's-starch plates, indicating that the TrGA gene has been disrupted. The intended homologous recombination event was verified by PCR product sequencing, confirming that the deletion cassette was successfully incorporated into the TrGA locus in the host cells by homologous recombination and as a result, the TrGA gene was replaced by pyr2 expression cassette. Our data demonstrate that directly introducing a functional SpyCas9 nickase complex into a target fungal cell in addition to the donor DNA (i.e., the DNA that is to be homologously recombined at the genomic locus of interest), the homologous recombination ratio in fungi can be significantly increased.

Section C: Introduction of Cas/guide RNA and donor DNA on a Single Expression Vector

Example 14: Deletion of a single base from the gla1 gene by use of cas9 expression plasmid containing a directed homologous donor fragment for gene editing

[0251] In this example, a gene disruption was directed by incorporating a homologous *gla1* donor fragment containing a single base deletion into the plasmid pTrex2gHyg MoCas gTrGA TS11B (described in example 8). The plasmid pTrexMoCasGATS11-HDR (FIG. 14) was created by inserting the approximately 1kb synthetic DNA fragment Gla1rep (SEQ ID NO:39, labeled as "Tr-gla 1kb homologous fragment" in FIG. 14) into the unique EcoRV restriction site of the plasmid pTrex2gHygMoCasgTrGATS11B. The EcoRV site resides in a polylinker sequence between the pSL1180 sequence and the *N. crassa* *cpc1* promoter of the hygromycin resistance marker (*hph*). The primers 1556F (5'-ATGCGCAAATTTAAAGCGCTGATgtgtgtctaatgcctccaccac [SEQ ID NO:73]) and 1557R (5'-ATATGGATCTGCGCGCGATCGATgatcgtgctagcgctgctgttg [SEQ ID NO:74]) were used to amplify Gla1rep and addend homologous tails to overlap both sides of the EcoRV site in pTrex2gHygMoCasgTrGATS11B to facilitate Gibson Assembly of the fragment with the EcoRV-cut pTrex2gHygMoCasgTrGATS11B plasmid (using NEB Gibson Assembly Master Mix, New England Biolabs, Beverly, MA).

[0252] The pTrexMoCasGATS11-HDR plasmid was used to transform *T. reesei* strain RL-P37, a strain with normal NHEJ mechanism, using protoplasts and PEG-mediated DNA uptake. This strain was auxotrophic for *pyr2*, incidental for this experiment, but requiring uridine (at 2 g/L) to be included in all growth media.

[0253] Transformants were initially grown on selective media agar plates of Vogel's minimal medium + 1.2M sorbitol + 75 ppm hygromycin. After growth, 80 transformants of both stable and unstable phenotypes were transferred to second agar plates of Vogel's-starch medium. After growth on second plates, the presence or absence of a starch clearing zone was used as a reporter to determine if the *gla1* gene was disrupted. Transformants were transferred from the second plates onto non-selective third plates of Vogel's minimal medium. After a few days of growth, transformants were transferred from the third plates of Vogel's minimal medium to selective fourth plates of Vogel's minimal medium + 75 ppm hygromycinB. On these fourth plates transformants could be accessed for growth to determine whether transformants maintained the ability to grow on hygromycin-containing media after growth on the non-selective media of the second and third plates.

[0254] It was found that only 11 of 80 transformants produced a clearing zone on Vogel's-starch medium, indicating *gla1* disruption, and did not grow on the fourth hygromycin selective plates of Vogel's minimal medium + 75 ppm hygromycin. These transformants presumably had alterations of the *gla1* gene and did not incorporate the hygromycin marker from the plasmid into their genome.

[0255] Genomic DNA was isolated from these 11 *gla1*- transformants, from mycelia of the third Vogel's minimal medium plates. Primers glaK (SEQ ID NO:31) and glaH (SEQ ID NO:29) were used to PCR amplify the *gla1* gene from genomic DNA to determine status of the TS11 locus. PCR products were isolated and sequenced using primers 1538F 5'-CCACCACAGGAACCAAACC (SEQ ID NO:75), 1539R 5'-CTGCGACGGAGGGAATGACG (SEQ ID NO:76), 1540F 5'-GGGCAGGACTGGCAAGGATGT (SEQ ID NO:77) and 1541R 5'-GCCGTCACGCCAGGAACAAG (SEQ ID NO:78).

[0256] The sequencing result revealed that 8 of these 11 *gla1*- transformants contained the single base deletion of the Gla1rep sequence. Two *gla1*- transformants contained *gla1* deletions at the TS11 locus of 9 and 100 bases. One *gla1*- transformant contained an insertion of 470 bases at TS11.

[0257] Incorporation of the donor Gla1rep fragment into the plasmid pTrex2gHyg MoCas TrGA TS11 generated a single plasmid which directed gene editing. A high frequency of the hygromycin-unstable transformants contained the single base deletion in the *gla1* gene that was engineered into the donor Gla1rep fragment. Unstable transformants were generated on the first selective agar medium, Vogel's minimal

medium + 75 ppm hygromycinB, which subsequently lost the transforming cas9 plasmid. Such strains are more advantageous because they do not have the inconvenience of NHEJ incorporation of the gene editing plasmid into the genome.

5 **[0258]** In another experiment for which data is not shown here, a substitutional mutation of gla1 was successfully introduced into T. reesei strain P37 using the same homologous integration method as described herein, with a gla1 homologous DNA fragment containing substitutional nucleotides instead of deletions inserted in pTrex2gHyg MoCas gTrGA TS11B.

10

[0259] These results demonstrate the utility of creating a single cas9 editing plasmid incorporating a homologous fragment along with cas9 and target site guide RNA. A single plasmid vector is created that directs both the specific cas9 targeted cleavage and gene editing homologous recombination at the targeted locus. By screening
15 unstable hygromycin resistance transformants, the incidence of NHEJ insertion of plasmid DNA into the target locus is minimized. Additionally, by screening unstable hygromycin resistant transformants in this example, transformants may be found that have incorporated the homologous donor fragment into the target locus by homologous recombination. Yet, such transformants have conveniently lost the plasmid that
20 directed the targeted recombination.

[0260] Although the foregoing compositions and methods have been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings herein that
25 certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

[0261] Accordingly, the preceding merely illustrates the principles of the present compositions and methods. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown
30 herein, embody the principles of the present compositions and methods and are included within its spirit and scope. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the present compositions and methods and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to

such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the present compositions and methods as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include

5 both currently known equivalents and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. The scope of the present compositions and methods, therefore, is not intended to be limited to the exemplary embodiments shown and described herein.

Sequences:

SEQ ID NO:1

Putative *T. reesei* U6 gene

AAAAAACACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTAACTTCTGCA
 5 GTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTATTATTTTTAT
 TTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTTATTATTATAATAT
 ATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAATAATTTATAG
 TAATAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATGAAATGGTATT
 ATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTGGCTATAAGTC
 10 TGCTGCAAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTGATGGTAGTCT
 ATCGCCTTCGGGCATTTGGTCAATTTATAACGATACAGGTTTCGTTTCGGCTTTTCC
 TCGGAACCCCCAGAGGTCATCAGTTCGAATCGCTAACAGGTCAACAGAGAAGATT
 AGCATGGCCCCTGCACTAAGGATGACACGCTCACTCAAAGAGAAGCTAAACATTTT
 TTTTCTCTTCCAAGTCGTGATGGTTATCTTTTTGCTTAGAGAATCTATTCTTGTGGA
 15 CGATTAGTATTGGTAAATCCCTGCTGCACATTGCGGCGGATGGTCTCAACGGCAT
 AATACCCCATTCGTGATGCAGCGGTGATCTTCAATATGTAGTGTAATACGTTGCAT
 ACACCACCAGGTTTCGGTGCCTCCTGTATGTACAGTACTGTAGTTCGACTCCTCCG
 CGCAGGTGGAAACGATTCCCTAGTGGGCAGGTATTTTGGCGGGGTCAAGAA

20 SEQ ID NO:2

sequence of sgRNA (N is sequence complementary to target site)

NNNNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGG
 CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

25 SEQ ID NO:3

sgRNA: gAd3A TS1

guccucgagcaaaaggugccGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGU
 CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

30 SEQ ID NO:4

sgRNA: gTrGA TS2

guucagugcaauaggcgucGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGU
 CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

SEQ ID NO:5

sgRNA: gTrGA TS11

gccaauggcgacggcagcacGUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGU
CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

5

SEQ ID NO:6

sgRNA: gPyr2 TS6

gcacagcgggaugcccuuguGUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGU
CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC

10

SEQ ID NO:7

Codon optimized *Streptococcus pyogenes* Cas9-encoding gene; with N- and C-terminal
NLS sequences

atggcaccgaagaagaagcgcaaggtgatggacaagaagtacagcatcggcctcgacatcggcaccaactcgggtgg
15 gctgggcccgtcatcacggacgaatataaggtcccgtcgaagaagttcaaggtcctcggcaatacagaccgccacagca
tcaagaaaaacttgatcggcgccctcctgttcgatagcggcgagaccgcggaggcgaccagggtcaagaggaccgcc
aggagacgggtacactaggcgcaagaacaggatctgctacctgcaggagatcttcagcaacgagatggcgaaggtgg
acgactccttctccaccgcctggaggaatcattcctgggtggaggaggacaagaagcatgagcggcacccaatcttcgg
caacatcgtcgacgaggtggcctaccacgagaagtacccgacaatctaccacctccggaagaaactgggtggacagca
20 cagacaaggcggacctccgggtcatctaccttgccctcgcgcatatgatcaagttccgcggccacttcctcatcgagggc
gacctgaacccggacaactccgacgtggacaagctgttcacccagctcgtgcagacgtacaatcaactgttcgaggaga
accccataaacgctagcggcgtggacgccaaggccatcctctcggccagggtctcgaaatcaagaaggctggagaac
cttatcgcgagttgccaggcgaaaagaagaacggcctcttcggcaaccttattgcgctcagcctcggcctgacgccga
acttcaaatacaacttcgacctcgcgaggagcgccaagctccagctctcaaaggacacctacgacgacgacctcgaca
25 acctcctggcccagataggagaccagtacgcggacctcttctcgcgccagaacctctccgacgctatcctgctcagc
gacatccttcgggtcaacaccgaaattaccaaggcaccgctgtccgccagcatgattaaacgctacgacgagcaccatc
aggacctcacgctgctcaaggcactcgtccgccagcagctccccgagaagtacaaggagatcttctcgaccaatcaaa
aaacgggtacgcgggatatatcgacggcgggtccagccaggaagagttctacaagttcatcaaaccaatcctggagaa
gatggacggcaccgaggaggtgctgggtcaagctcaacaggaggacctcctcaggaagcagaggaccttcgacaac
30 ggctccatcccgcacatcagatccacctgggccaactgcatgccatcctcgggcgccaggaggacctctaccggttctgaa
ggataaccgggagaagatcgagaagatcttgacgttcgcacatccatactacgtgggcccgtgggtcgcggaactcc
cggttcgctggatgacccggaagtcggaggagaccatcacaccctggaactttgaggaggtggtcgataagggcgct
agcgctcagagcttcatcgagcgcatgaccaacttcgataaaaacctgcccaatgaaaaagtcctccccaagcactcgc
tgctctacgagtacttaccgtgtacaacgagctcaccaagggtcaaatacgtcaccgagggcatgcggaagccggcggt

cctgagcggcgagcagaagaaggcgatagtgacctcctcttcaagaccaacaggaaggtgaccgtgaagcaattaa
aagaggactacttcaagaaaatagagtgtctcgactccgtggagatctcgggcgtggaggatcggttcaacgcctcactc
ggcacgtatcacgacctcctcaagatcattaaagacaaggacttctcgacaacgaggagaacgaggacatcctcgag
gacatcgtcctcaccctgacctgttcgaggaccgcgaaatgatcgaggagaggctgaagacctacgcgcacctgttcg
5 acgacaaggctcatgaaacagctcaagaggcgccgtacactgggtgggaaggctgtcccgcaagctcattaatggca
tcagggacaagcagagcggcaagaccatcctggacttctcaagtccgacgggttcgccaaccgcaacttcatgcagc
tcattcacgacgactcgctcacgttcaaggaagacatccagaaggcacagggtgagcgggcagggtgactccctccacg
aacacatcgccaacctggccggctcgccggccattaaaaagggcatcctgcagacgggtcaaggctcgtcgacgagctc
gtgaagggtgatgggcccgcacaagcccgaataatcgatagagatggccagggaaccagaccacccaaaaa
10 gggcagaagaactcgcgcgagcggatgaaacggatcgaggaggcgattaaagagctcgggtcccagatcctgaag
gagcacccttggaataaccagctccagaatgaaaagctctaccttactacctgcagaacggccgcgacatgtac
gtggaccaggagctggacattaatcggtatcgactacgacgtcgaccacatcgtgccgcagtcgttctcaaggacg
atagcatcgacaacaagggtgctcaccggctcgataaaaaatcggggcaagagcgacaacgtgccagcgaggaggt
cgtgaagaagatgaaaaactactggcgccagctcctcaacgcgaaactgatcaccagcgcaagttcgacaacctga
15 cgaaggcggaacgcggtggcttgagcgaactcgataaggcgggcttcataaaaaggcagctggtcgagacgcgcca
gatcacgaagcatgtcgcccagatcctggacagccgcatgaataactaagtacgatgaaacgacaagctgatccggg
agggtgaagggtgatcacgctgaagtcgaagctcgtgtcggacttccgcaaggacttccagttctacaagggtccgcgagatc
aacaactaccaccacgcccacgacgcctacctgaatgcggtggtcgggaccgccctgatcaagaagtaccggaagct
ggagtcggagttcgtgtacggcgactacaagggtctacgacgtgcgcaaaatgatcgccaagtcgagcaggagatcgg
20 caaggccacggcaaaatacttcttactcgaacatcatgaacttctcaagaccgagatcaccctcgcaacggcgag
atccgcaagcgcccgtcatcgaaaccaacggcgagacgggcgagatcgtctgggataaggcgccgggatttcgcgac
gggtccgcaagggtgcttccatgccgcaagtcaatatcgtgaaaaagacggagggtccagacgggcgggttcagcaagg
agtccatctctccgaagcgcaactccgacaagctcatcgcgaggaagaaggattgggacccgaaaaaatatggcggc
ttcgacagcccagccgtcgcatcacgctcctcgtcgtggcgaagggtggagaagggaagtcataagaagctcaagtc
25 gtgaaggagctgctcgggatcacgattatggagcgggtcctcctcgagaagaacccgatcgacttcttagaggccaagg
gatataaggagggtcaagaaggacctgattattaaactgccgaagtactcgctcttcgagctggaaaacggccgcaaga
ggatgctcgccctccgaggcgaggtgcagaagggaacgagctcgccctcccgagcaataacgtcaatttctgtacctc
gctagccactatgaaaagctcaagggcagcccggaggacaacgagcagaagcagctcttcgtggagcagcacaag
cattacctggacgagatcatcgagcagatcagcgagttctgaagcgggtgatcctcgccgacgcgaacctggacaag
30 gtgctgtcggcatataacaagcaccgcgacaaaccaataacgcgagcaggccgaaaatatcatccaccttccacctca
ccaacctcggcgctccggcagccttcaagtacttcgacaccacgattgaccggaagcgggtacacgagcacgaaggag
gtgctcgatgcgacgctgatccaccagagcatcacagggtctatgaaacacgcatcgacctgagccagctgggcgga
gacaagaagaagaagctcaagctctag

SEQ ID NO:8

Streptococcus pyogenes Cas9 encoded by SEQ ID NO:7; with N- and C-terminal NLS sequences

MAPKKKRKVMDDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 5 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEE
 DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLI
 EGDLPDNDSDVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLP
 GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADL
 FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 10 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
 SIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
 SEETITPWNFEEVVDKGASASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
 KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
 FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKV
 15 MKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFK
 EDIQAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
 RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
 MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
 KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
 20 MNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA
 IKKYPKLESEFVYGDYKVYDVRKMIKSESEQEIGKATAKYFFYSNIMNFFKTEITLANGEI
 RKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRN
 SDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERS
 EKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 25 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVILADANLDKVLSA
 YNKHARDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQ
 SITGLYETRIDLSQLGGDKKKKLKL

SEQ ID NO:9

30 Synthetic DNA: gAd3A TS1-1 (gAd3A TS1 sgRNA (SEQ ID NO:3) with *Saccharomyces cerevisiae* snr52 promoter and *S. cerevisiae* sup4 terminator)
 gaattcggatccTCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGA
 AACTTGATGTTTTCTTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACA
 ACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTTCAC

ACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAAGTTG
 GTGCGCATGTTTCGGCGTTTCGAACTTCTCCGCAGTGAAAGATAAATGATCggtcctcg
 agcaaaagggtgccGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATC
 AACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTTTGTTTTTTATGTCTgaattcg
 5 gatcc

SEQ ID NO:10

Synthetic DNA: gAd3A TS1-2 (gAd3A TS1 sgRNA (SEQ ID NO:3) with T. reesei U6 promoter and terminator)

10 gaattcggatccAAAAACACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTA
 ACTTCTGCAGTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTA
 TTATTTTTATTTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTTATTT
 ATTATAATATATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAA
 TAATTTATAGTAATAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATG
 15 AAATGGTATTATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTG
 GCTATAAGTCTGCTGCAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTG
 ATGGTAGTCTATCgtcctcgagcaaaagggtgccGTTTTAGAGCTAGAAATAGCAAGTTAAA
 TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTT
 TTCTCTTgaattcggatcc

20

SEQ ID NO:11

Synthetic DNA: gAd3A TS1-3 (gAd3A TS1 sgRNA (SEQ ID NO:3) with T. reesei U6 promoter, terminator and intron)

gaattcggatccAAAAACACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTA
 25 ACTTCTGCAGTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTA
 TTATTTTTATTTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTTATTT
 ATTATAATATATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAA
 TAATTTATAGTAATAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATG
 AAATGGTATTATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTG
 30 GCTATAAGTCTGCTGCAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTG
 ATGGTAGTCTATCgtcctcgagcaaaagggtgccGTTTTAGAGCTAGAGTTCGTTTCGGCTTT
 TCCTCGGAACCCCCAGAGGTCATCAGTTCGAATCGCTAACAGAATAGCAAGTTAAA
 ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTT
 TTTCTCTTgaattcggatcc

SEQ ID NO:12

Guide RNA expression cassettes with a shorter T. reesei U6 promoter region were obtained as synthetic DNA. An example is provided here that includes the sequence for an sgRNA targeting the T. reesei gla1 gene at TS11.

5 AATTCCTAAAGAAACAGCATGAAATGGTATTATGTAAGAGCTATAGTCTAAAGGCA
CTCTGCTGGATAAAAATAGTGGCTATAAGTCTGCTGCAAACTACCCCCAACCTCG
TAGGTATATAAGTACTGTTTGATGGTAGTCTATCgccaatggcgacggcagcacGTTTTAGA
GCTAGAGTTTCGTTTCGGCTTTTCCTCGGAACCCCCAGAGGTCATCAGTTCGAATC
10 GCTAACAGAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG
CACCGAGTCGGTGGTGCTTTTTTTTTCTCTT

SEQ ID NO:13

Primer: gRNA fwd aflII

15 cgtcagcttaagAATTCCTAAAGAAACAGCATGAAATGG

SEQ ID NO:14

Primer: gRNA rev sfil

cgtcagggccacgtgggccAAGAGAAAAAAAAGCACCCGACTCGG
20

SEQ ID NO:15

Primer: Ad3 5' fwd

tgaacacagccaccgacatcagc

25 SEQ ID NO:16

Primer: Ad3 5' rev

gctggtgagggttgtgctattg

SEQ ID NO:17

30 Primer: Ad3a 5005 rev

gattgcttgggaggaggacat

SEQ ID NO:18

Primer: Ad3 3' fwd

cgaggccactgatgaagttgttc

5

SEQ ID NO:19

Primer: Ad3 3' rev

Cagttttccaaggctgccaacgc

10 SEQ ID NO:20

Primer: Ad3a 5003 fwd

ctgatcttgccaccctggaaatc

SEQ ID NO:21

15 Ad3mid rev

ctctctatcattgccaccctcc

SEQ ID NO:22

Primer: Adfrag fwd

20 ctccattcaccctcaattctcc

SEQ ID NO:23

Primer: Adfrag rev

gttccttggcgggtgcttgatc

25

SEQ ID NO:24

Primer: Ad3a 2k fwd

caatagcacaaccctcaccagc

30 SEQ ID NO:25

Ad3a 2k rev

gaacaactcatcagtggcctcg

SEQ ID NO:26

Primer: glaA

ccgtagttgaagatccttgccg

SEQ ID NO:27

5 Primer: glaB

gtcgaggatttgcttcatacctc

SEQ ID NO:28

Primer: glaD

10 gagagacgcaggatgactcaaag

SEQ ID NO:29

Primer: glaH

tgccgtgggtcattggcatattc

15

SEQ ID NO:30

Primer: glaJ

tgccgactttgtccagtgattcg

20 SEQ ID NO:31

Primer: glaK

ttacatgtggacgcgagatagcg

SEQ ID NO:32

25 Primer: gla1repF

gtgtgtctaatgcctccaccac

SEQ ID NO:33

Primer: gla1repR

30 gatcgtgctagcgctgctgttg

SEQ ID NO:34

Primer: 1553R

CCGTGATGGAGCCCGTCTTCT

SEQ ID NO:35

Primer: 1555F

CGCGGTGAGTTCAGGCTTTTTC

5

SEQ ID NO:36

Primer: pyr2F

gtataagagcaggaggaggag

10 SEQ ID NO:37

Primer: pyr2R

gaacgcctcaatcagtcagtcg

SEQ ID NO: 38

15 Bacterial kanamycin resistance gene (with promoter and terminator) between
Trichoderma reesei telomere sequences

tcaggaaatagctttaagtagcttattaagattataatataatataactatattctttaataaaggattttaag
ctttatatataaataataataaaaataatattatatagcttttattaataaaaaatagctaaaaatataaaaaaataag
ctttaaataacttattttaattagaattttatatattttaataataaagatctttactttttataagcttcctacctaataaatttta
20 cttttttactattttactatatcttaataaaggctttaaaaaatataaaaaaatcttcttatattataagctataaggattatat
atatatttttttaattttaaagtaagattataaagctagaattaaagtttaatttttaaggctttattaaaaaaggcagtaata
gcttataaaagaaatttcttttttataactaaaagactttttttaataagggttagggtaggggttactcacaccgaccatcc
caaccacatcttagggtaggggttaggggttaggggttaggggttaggggttaggggttaaacaaagccacggt
gtgtctcaaaatctctgatgttacattgcacaagataaaaaatataatcatcatgaacaataaaactgtctgttacataaacag
25 taatacaaggggtgttatgagccatattcaacgggaacgtctgtctcgaggccgcgattaaattccaacatggatgctga
tttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgattgtatgggaagcccgatgcg
ccagagttgttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggctgacg
gaatttatgcctcttcgaccatcaagcattttatccgtactcctgatgatgcattgttactcaccactgcgatccccgggaaa
acagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtggtcctgcgccggttgca
30 ttogattcctgtttgaattgtcctttaacacgcgatcgcgtatttcgtctcgtcagggcgcaatcacgaatgaataacgggttggt
tgatgcgagtgattttgatgacgagcgtaattggctggcctgttgaaacagctctggaagaaatgcataagcttttgccattct
caccggattcagtcgtcactcatggtgatttctcacttgataacctattttgacgaggggaaattaataggttgattgatgttg
gacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtagttttctccttcattacagaa
acggccttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttcatttgatgctcgatgagttttctaatacaga

attggttaattggttgaactggcagagcattacgctgacttgacgggacggcggccttgggtgaataaatcgaacttttgc
 gagttgaaggatcagatcacgcatttcccgacaacgcagaccgttccgtggcaaagcaaaagttcaaaatcaccaact
 ggtccacctacaacaaagctctcatcaaccgtggctccctcactttctggctggatgatggggcgattcaggcctggtatga
 gtcagcaacaccttcttcacgaggcagacctcagcgggttaaacctaaccctaaccctaaccctaaccctaacc
 5 ctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaaccctaacc
 gaggggttctatagactaatctacaggccgtacatgggtgattgcagatgcgacgggcaagggtacagtgccagaagg
 aggagagcggcataggtattgtaatagaccagcttacataataatcgctgttgctactgactgatgaccttctccctaac
 cagtttctaattaccactgcagtgaggataaccctaactcgctctgggggtattattatactgattagcagggtggcttatatagt
 gctgaagtactataagagtttctcgggaggaggtggaaggactataaactggacacagttagggatagagtgatgaca
 10 agacctgaatgttatctccggtgtggtatagcgaattggctgaccttcagatggtaatgggttaggcagggttttgcagag
 ggggacgagaacgcgttctgcgatttaacggctgctgccgccaagctttacggttctctaattgggcggccgc

SEQ ID NO:39

A synthetic DNA: Glalrep

gtgtgtctaatagcctccaccacaggaaccaaaccggccttgacctctgggaagaagtcaatgggagctcattctttactgtt
 gccaacaggcaccgaggtatgaagcaaactctcgacattcgctgctactgcacatgagcattgttactgaccagctctac
 agcattgtcgagggcgccactcttctgctgccactcttgccagtcgggaagcgcttattcatctgttctccccagggtttgtg
 ctttctccaacgattctgggtgtcgtctgggtggatacgtcgactccaacagtatgtcttttactgtttatatgagattggccaata
 ctgatagctcgctctagtcaacaccaacgagggcaggactggcaaggatgtcaactccgtcctgacttccatccacacc
 20 ttgatcccaaccttggctgtgacgcaggcacctccagccatgcagtgacaaagcgctctccaacctcaagggtgtgtcg
 actccttcgctccatctacggcgtgaacaagggcattctctgCGgtgctgccgtcgccattggccggtatgcagaggatgt
 gtactacaacggcaaccttggatatcttctacatttctgctgctgccgagcagctgtacgatgccatctacgtctggaagaag
 acgggtccatcacggtgaccgccacctccctggccttctccaggagctgttcttgccgtgacggccgggacactacc
 agcagctcttcgacctttaccaacatcatcaacgccgtctcgacatacgccgatggcttctcagcgaggctgccaaagtac
 25 gtccccgccgacggttcgctggccgagcagtttgaccgcaacagcggcactccgctgtctgcgcttcacctgacgtggtc
 gtacgcctcgttcttgacagccacggcccgtcgggctggcatcgtgccccctcgtgggccaacagcagcgctagcacg
 atc

SEQ ID NO:40

30 Full U6 gene promoter sequence (not including transcription start site)

AAAAAACAAGTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTAACTTCTGCA
 GTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTATTATTTTTAT
 TTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTATTTATTATAATAT
 ATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAATAATTTATAG

TAATAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATGAAATGGTATT
 ATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTGGCTATAAGTC
 TGCTGCAAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTGATGGTAGTCT
 ATC

5

SEQ ID NO:41

Truncated/shorter U6 gene promoter sequence (not including transcription start site)

AATTCCTAAAGAAACAGCATGAAATGGTATTATGTAAGAGCTATAGTCTAAAGGCA
 CTCTGCTGGATAAAAATAGTGGCTATAAGTCTGCTGCAAAACTACCCCCAACCTCG
 TAGGTATATAAGTACTGTTTGATGGTAGTCTATC

10

SEQ ID NO:42

U6 gene intron

GTTTCGTTTCGGCTTTTCCTCGGAACCCCCAGAGGTCATCAGTTCGAATCGCTAA
 CAG

15

SEQ ID NO:43

U6 gene transcriptional terminator sequence

TTTTTTTTCTCTT

20

SEQ ID NO:44

Filamentous fungal cell Codon optimized Streptococcus pyogenes Cas9-encoding
 gene; no NLS

atggacaagaagtacagcatcggcctcgacatcggcaccaactcggtagggctgggccgcatcacggacgaataaa
 ggtcccgtcgaagaagttcaaggtcctcggaatacagaccgccacagcatcaagaaaaacttgatcggcgccctcct
 gttagatagcggcgagaccgcggaggcgaccagggtcaagaggaccgccaggagacggtagactaggcgcaaga
 acaggatctgctacctgcaggagatcttcagcaacgagatggcgaagggtggacgactccttctccaccgcctggagga
 atcattcctggtggaggaggacaagaagcatgagcggcacccaatcttcggcaacatcgtagcagaggtggcctacca
 cgagaagtagccgacaatctaccacctccggaagaaactggtggacagcacagacaaggcggacctccggctcatct
 acctgcccctcgcgcatatgatcaagttccgcgccacttctcatcgaggcgacctgaacccggacaactccgacgtg
 gacaagctgttcatccagctcgtgcagacgtacaataactgttcgaggagaacccataaacgctagcggcggtggacg
 ccaaggccatcctctcgccagggtctcgaaatcaagaaggctggagaaccttatcgcgaggtgccaggcgaaaaga
 agaacggcctcttcggcaaccttattgcgctcagcctcggcctgacgccgaactcaaataaaacttcgacctcgcgga
 gacgccaagctccagctctcaaaggacacctacgacgacacctcgacaacctcctggcccagataggagaccagta

30

cgcgacacctcttctcgccgccaagaacctctcgacgctatcctgctcagcgacatccttcgggtcaacaccgaaattac
caaggcaccgctgtccgccagcatgattaaacgctacgacgagcaccatcaggacctcacgctgctcaaggcactcgt
ccgccagcagctccccgagaagtacaaggagatcttctcgaccaatcaaaaaacggctacgcgggatataatcgacgg
cgggtgccagccaggaagagttctacaagttcatcaaaccaatcctggagaagatggacggcaccgaggagttgctggt
5 caagctcaacagggaggacctctcaggaagcagaggaccttcgacaacgggtccatcccgcatcagatccacctgg
gcgaactgcatgccatcctgcggcgccaggaggacttctacccgttctgaaggataaccgggagaagatcgagaag
atcttgacgttccgcatcccatactacgtgggcccgtggctcgcggcaactcccggttcgcttgatgacccggaagtcg
gaggagaccatcacacctggaactttgaggaggtggctgataagggcgctagcgctcagagcttcatcgagcgcatg
accaacttcgataaaaaacctgcccataaaaaagtcctcccaagcactcgctgctctacgagtacttcacctgttaciaa
10 cgagctcaccaagggtcaaatacgtcacccgagggcatgcggaagccggcggttctgagcggcgagcagaagaaggc
gatagtggacctccttctcaagaccaacaggaaggtagccgtgaagcaattaaaagaggactacttcaagaaaataga
gtgcttcgactccgtggagatctcgggctggaggatcggttcaacgcctcactcggcacgtatcacgacctcctcaagat
cattaaagacaaggacttctcgacaacgaggagaacgaggacatcctcgaggacatcgctcctaccctgacctgttc
gaggaccgcgaaatgatcgaggagaggctgaagacctacgcgcacctgttcgacgacaaggatgaaacagctca
15 agaggcgccgtacactggttggggaaggctgtcccgaagctcattaatggcatcagggacaagcagagcggcaag
accatcctggacttctcaagtccgacgggttcgccaaccgcaacttcagtcagctcattcacgacgactcgctcacgttc
aaggaagacatccagaaggcacaggtgagcgggcagggtgactccctccacgaacacatcgccaacctggccgggt
cgccggccattaaaaaggcatcctgcagacgggtcaaggctcgtcgcagagctcgtgaaggtagggccggcacaag
cccgaaaatatcgtcatagagatggccaggggagaaccagaccacccaaaaagggcagaagaactcgcgcgagcg
20 gatgaaacggatcgaggagggcattaaagagctcgggtcccagatcctgaaggagcaccctggaataatccag
ctccagaatgaaaagctctaccttactacctgcagaacggccgcgacatgtacgtggaccaggagctggacattaatc
ggctatcggactacgacgtcgaccacatcgtgccgcagtcgttctcaaggacgatagcatcgacaacaagggtgtcac
ccggtcggataaaaaatcggggcaagagcgacaacgtgccagcgaggaggtcgtgaagaagatgaaaaactactg
gcgccagctcctcaacgcgaaactgatcaccagcgcaagttcgacaacctgacgaaggcggaacgcgggtggctga
25 gcgaactcgataaggcgggctcataaaaaggcagctggtcgagacgcgccagatcacgaagcatgtcgccagatc
ctggacagccgcatgaataactaagtacgatgaaaacgacaagctgatccgggaggtgaaggtagtcacgtgaagtcc
aagctcgtgtcggacttccgcaaggacttccagttctacaagggtccgcgagatcaacaactaccaccacgcccacgacg
cctacctgaatgcggtggtcgggaccgcccctgatcaagaagtacccgaagctggagtcggagttcgtgtacggcgacta
caaggcttacgacgtgcgcaaaatgatcgccaagtcgagcaggagatcggcaaggccacggcaaaaatacttctcta
30 ctcgaaatcatgaacttctcaagaccgagatcacctcgcgaacggcgagatccgcaagcgcccgtcatcgaaac
caacggcgagacgggagagatcgtctgggataagggccgggatttcgcgacgggtccgcaagggtgtctccatgccgca
agtcaatatcgtgaaaaagacggaggtccagacgggcccgttcagcaaggagtcctcctcccgaagcgcaactccg
acaagctcatcgcgaggaagaaggattgggacccgaaaaaataatggcggttcgacagcccagccgtcgcatacag
cgtcctcgtcgtggcgaagggtggagaagggcaagtcaaagaagctcaagtccgtgaaggagctgctcgggatcacgat

tatggagcggctcctccttcgagaagaacccgatcgacttcctagaggccaagggatataaggagggtcaagaaggacct
 gattattaaactgccgaagtactcgctcttcgagctggaaaacggccgcaagaggatgctcgctcccgaggcgagttgc
 agaaggggcaacgagctcgccctcccgagcaaaatacgtcaatttcctgtacctcgctagccactatgaaaagctcaaggg
 cagccccggaggacaacgagcagaagcagctcttcgtggagcagcacaagcattacctggacgagatcatcgagcag
 5 atcagcgagttctcgaagcgggtgatcctcgccgacgcgaacctggacaaggtgctgctggcatataacaagcaccgc
 gacaaaccaatacgcgagcaggccgaaaatatcatccacctcttcacctcaccaacctcggcgctccggcagcctca
 agtacttcgacaccacgattgaccggaagcgggtacacgagcacgaaggaggtgctcgatgcgacgctgatccaccag
 agcatcacagggctctatgaaacacgcatcgacctgagccagctgggcgagac

10 SEQ ID NO:45

Streptococcus pyogenes Cas9 encoded by SEQ ID NO:44; no NLS

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETA
 EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPI
 FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDN
 15 SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLF
 GNLIALLSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS
 DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLEKYKEIFFDQSKN
 GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
 GELHAILRRQEDFYPFLKDNREKIEKILTRIPYYYVGPLARGNSRFAWMTRKSEETITP
 20 WNFEVVDKGAASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKR
 RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
 VSGQGDSLHEHIANLAGSPAIIKGIQTVKVDELVKVMGRHKPENIVIMARENQTTQ
 25 KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELD
 INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLL
 NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
 DKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPKLES
 EFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
 30 GETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKK
 DWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLE
 AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
 EKLKGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVLSAYNKHDKPI

REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDL SQ
LGGD

SEQ ID NO:46

5 SV40 NLS
PKKKRKV

SEQ NO:47

T. reesei blr2 (blue light regulator 2) gene NLS

10 KKKKLKL

SEQ ID NO:48

Streptococcus thermophilus LMD-9 Cas9

MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITA
15 EGRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSLVPDDKRDSKYPI
FGNLVEEKAYHDEFPTIYHLRKYLADSTKKADRLVYLALAHMIKYRGHFLIEGEFNSK
NNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIF
SEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSVFLKAKKLY
DAILLSGFLTVDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNVFKDDTKN
20 GYAGYIDGKTNQEDFYVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHL
QEMRAILDKQAKFYFPFLAKNKERIEKILTFRIPIYYVGPLARGNSDFAWSIRKRNEKITPW
NFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRD
YQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLL
NIINDKEFLDDSSNEAIIIEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWG
25 KLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNI
KEVVKSLPGSPAIIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQQ
RLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDIDRL
SNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKRKTFWYQLLKSCLIS
QRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRVR
30 TVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVVASALLKKYPKLEPEFVYG
DYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKE
SDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAK
EYLDPPKYGGYAGISNSFTVLVKG TIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLE
KGYKDIELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAK

RISNTINENHRKYVENHKKEFEELFYIILEFNENYVGAKKNGKLLNSAFQSWQNHSIDE
LCSSFIGPTGSERKGLFELTSRGSAADEFGLGVKIPRYRDYTPSSLLKDATLIHQSVTGL
YETRIDLAKLGEG

5 SEQ ID NO:49

Streptococcus mutans UA159 Cas9

MKKPYSIGLDIGTNSVGWAVVTDDYKVPAAKKMKVLGNTDKSHIEKNLLGALLFDSGNT
AEDRRLKRTARRRYTRRRNRILYLQEIFSEEMGKVDDSFHRLSDSFLVTEDKRGERH
PIFGNLEEEVKYHENFPTIYHLRQYLADNPEKVDLRLVYLALAHIIKFRGHFLIEGKFDTR
10 NNDVQRLFQEFLAVYDNTFENSSLQEQNVQVEEILTDKISKSAKKDRVLKLPNEKSN
GRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEEELVLLAQIGDNYAELFLSAK
KLYDSILLSGILTVTDVGTKAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSD
VSKDGYAGYIDGKTNQEAIFYKYLKGLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPH
QIHLQEMRAIIRRQAEFYPPFLADNQDRIEKLITFRIPYYVGPLARGKSDFAWLSRKSAD
15 KITPWNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPHSLLYEKFTVYNELTKVKYK
TEQGKTAFDDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKV
FNASYGTYHDLCKILDKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLNYSDDLTKEQ
VKKLERRHYTGWGRLSAELIHGIRNKESRKTILDYLIDDGNSNRNFMQLINDDALSFK
EIAKAQVIGETDNLNQVVSIDIAGSPAIKKGILQSLKIVDELVKIMGHQPENIVVEMAREN
20 QFTNQGRRNSQQRLKGLTDSIKEFGSQILKEHPVENSQQLQNDRLFLYYLQNGRDMYT
GEELDIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKS
YWSKLLSAKLITQRKFDNLTKAERGGLTDDDKAGFIKQRLVETRQITKHVARILDERFNTET
DENKKIRQVKIVTLKSNLVSNFRKEFELYKVVREINDYHHAHDAYLNAVIGKALLGVYP
QLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKDEHIS
25 NIKKVLSYPQVNIVKKVEEQTGGFSKESILPKGNSDKLIPRKTKKFYWDTKKYGGFDSP
IVAYSILVIADIEKGKSKKLKTVKALVGVTIMEKMTFERDPVAFLEKGYRNVQEENIIKL
PKYSLFKLENGRKRLLASARELQKGNEIVLPNHLGTLTYHAKNIHKVDEPKHLDYVDKH
KDEFKELLDVVSNSFSKKYTLAEGNLEKIKELYAQNNGEDLKELASSFINLLTFTAIGAPA
TFKFFDKNIDRKRYTSTTEILNATLIHQSIITGLYETRIDLNKLGGD

30

SEQ ID NO:50

Campylobacter jejuni Cas9

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLAR
RKARLNHLKHLIANEFKLNIEDYQSFDESLAKAYKGSLSIPYELRFRALNELLSKQDFA

RVILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENS
 KEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFEEVLSVAFYKRALK
 DFSHLVGNCSSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLNKNTGILYTKDDLNALLNE
 VLKNGTLTYKQTKKLLGLSDDYEFKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAK
 5 DITLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKALKLVTPLMLEGKKYDEAC
 NELNLKVAINEDKKDFLPAFNETYYKDEVTNPVVLRAIKEYRKVLNALLKKYGVHKINI
 ELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEKLGLKINSKNILKLRLFKEQKEFC
 AYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMKNVLVFTKQNQEKLNQTPFEAFGN
 DSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQKNFKDRNLNDTRYIARLVNLYTKDYL
 10 DFLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSALRHTWGFSKDRNNHLHHAIDAVI
 IAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQKVLDKIDEI
 FVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVKALELGKIRKVNGKIVKNGDMF
 RVDIFKHKKTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCSLYK
 DSLILIQTKDMQEPEFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKS
 15 IGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK

SEQ ID NO:51

Neisseria meningitides Cas9

MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAEVPKTGDSLAM
 20 ARRLARSVRRLTRRRRAHLLRTRRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAAL
 DRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVAGNAHALQTGDFRT
 PAELALNKFESGHIRNQSRSDYSHTFSRKDLQAEILLLFEKQKEFGNPHVSGGLKEGI
 ETLLMTQRPALSGDAVQKMLGHCTFEPAPKAANKNTYTAERFIWLTCLNNLRILEQGS
 ERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEM
 25 KAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEAL
 LKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEI
 RNPVVLRAISQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDR
 EKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRNEKGYVEIDHA
 LPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFP
 30 RSKKQRILLQKFDEDFKERNLNDTRYVNRFLCQFVADRMRLTGKGGKRVFASNGQI
 TNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTID
 KETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTLEKLRTLLAEKLSSR
 PEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSARKLDEGVSVLRVPLTQLKLKDLEK
 MVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKT

GVWVRNHNGIADNATMVRVDVFEKGDKEYLVPIYSWQVAKGILPDRAVVQGKDEED
WQLIDDSFNFKFSLHPNDLVEVITKKARMTFGYFASCHRG TGNNIRIHDLDHKIGKNGIL
EGIGVKTALS FQKYQIDELGKEIRPCRLKKRPPVR

5 SEQ ID NO:52

Francisella tularensis subsp. novicida Cas9

MNFKILPIAIDLGVKNTGVFSAFYQKGTSLERLDNKNKGKVELSKDSYTLMMNNRTARR
HQRGRIDRKQLVKRLFKLIWTEQLNLEWDKDTQQAISFLFNRRGFSFITDGYSPEYLN
VPEQVKAILMDIFDDYNGEDDLSYKLATEQESKISEIYNKLMQKILEFKLMKLCTDIKD
10 DKVSTKTLKEITSYEFELLADYLANYESLKTQKFSYTDKQGNLKELSYYHHDKYNIQE
FLKRHATINDRILDTLLTDDLDIWNFNFEKFDKNEEKLQNQEDKDHIAHLHHFVFA
VNKIKSEMASGGRHRSQYFQEITNVLDENNHQEGYLKNFCENLHNKKYSNLSVKNLV
NLIGNLSNLELKPLRKYFNDKIHAKADHWDEQKFTETECHWILGEWRVGVKDQDKKD
GAKYSYKDL CNELKQKVTAGLVDFLELDPCRTIPPYLDNNNRKPPKCSLILNPKFL
15 DNQYPNWQQYLQELKKLQSIQNYLDSFETDLKVLKSSKDQPYFVEYKSSNQQIASGQ
RDYKDL DARILQFIFDRVKASDELLLENIYFQAKKLKQKASSELEKLESSKKLDEVIANS
QLSQILKSQHTNGIFEQGTFLHLVCKYYKQRQRARDSRLYIMPEYRYDKKLHKYNNTG
RFDDDNQLLTYCNHKPRQKRYQLLNDLAGVLQVSPNFLKDKIGSDDDLFISKWLVEHI
RGFKKACEDSLKIQKDNRGLLNHNKINIARNTKGKCEKEIFNLICKIEGSEDKKGNKYKHGL
20 AYELGVLLFGEPNEASKPEFDRKIKKFNSIYSFAQIQQIAFAERKGNANTCAVCSADNA
HRMQQIKITEPVEDNKDKIILSAKAQRLPAIPTRIVDGAVKKMATILAKNIVDDNWQNIKQ
VLSAKHQLHIPITESNAFEFEPALADVKGKSLKDRRKALERISPENIFKDKNNRIKEFA
KGISAYSGANLTDGDFDGAKEELDHIIPRSHKKYGTLNDEANLICVTRGDNKNKGNRIF
CLRDLADNYKLKQFETDDLEIEKKIADTIWDANKKDFKFGNYRSFINLTPQEQKAFRH
25 ALFLADENPIKQAVIRAINNRNRTFVNGTQRYFAEVLANNIYLRAKKENLNTDKISFDYF
GIPTIGNGRGIAEIRQLYEKVDSDIQAYAKGDKPQASYSHLIDAMLAF CIAADEHRNDGS
IGLEIDKNYSLYPLDKNTGEVFTKDIFSQIKITDNEFSKDLVRKKAIEGFNTHRQMTRD
GIYAENYLPILHKELENEVRKGYTWKNSEEIKIFKGKKYDIQQLNNLVYCLKFVDKPISIDI
QISTLEELRNILTTNNIAATAEYYYINLKTQKLHEYYIENYNTALGYKKYSKEMEFLRSLA
30 YRSERVKIKSIDDVKQVLDKDSNFIIGKITLPFKKEWQRLYREWQNTTIKDDYEFLKSFF
NVKSITKLHKKVRKDFSLPISTNEGKFLVKRKTWDNNFIYQILNDSDSRADGTKPFIPAF
DISKNEIVEAIDSFTSKNIFWLPKNIELQKVDNKNIFAIDTSKWFEVETPSDLRDIGIATIQ
YKIDNNSRPKVRVKLDYVIDDDSKINYFMNHSLLKSRYPDKVLEILKQSTIIEFESSGFNK
TIKEMLMGMLAGIYNETSNN

SEQ ID NO:53

Pasteurella multocida Cas9

MQTTNLSYILGLDLGASVGWAVVEINENEDPIGLIDVGVRIFERAIEVPKTGESLALSRR
 5 LARSTRRLIRRRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEW
 GAVLLHLIKHRGYLSKRKNESQTNNKELGALLSGVAQNHQLLQSDDYRTPAELALKKF
 AKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQFGNPHCKEHIQQYMTTELLMWQ
 KPALSGEAILKMLGKCTHEKNEFKAAKHTYSAERFVWLTKLNNLRILEDGAERALNEEE
 RQLLINHPYEKSKLTYAQVRKLLGLSEQAIFKHLRYSKENAESATFMELKAWHAIKAL
 10 ENQGLKDTWQDLAKKPDLLDEIGTAFSLYKTDEDIQQYLTKVPNSVINALLVSLNFDK
 FIELSLKSLRKILPLMEQGKRYDQACREIYGHYGEANQKTSQLLPAIPAQEIRNPVVL
 TLSQARKVINAIIRQYGSPARVHIETGRELGKSFKERREIQKQEDNRTKRESAVQKFK
 ELFSDFSSEPKSKDILKFRLYEQQHKGKCLYSGKEINIHRLNEKGYVEIDHALPFSRTWD
 DSFNKVLVLASENQNKGNQTPYEWLQGKINSERWKNFVALVLGSQCSAAKKQRLLT
 15 QVIDDNKFIDRNLDTRYIARFLSNYIQENLLLVGKNKKNVFTPNGQITALLRSRWGLIK
 ARENNNRHHALDAIVVACATPSMQQKITRIFIRFKEVHPYKIENRYEMVDQESGEIISPH
 FPEPWAYFRQEVNIRVFDNHPDVLKEMPLDRPQANHQQFVQPLFVSRAPTRKMSGQ
 GHMETIKSAKRLAEGISVLRIPLTQLKPNLLENMVNKEREPALYAGLKARLAEFNQDPA
 KAFATPFYKQGGQVKAIRVEQVQKSGVLVRENNGVADNASIVRTDVFIKNNKFFLVPI
 20 YTWQVAKGILPNKAIVAHKNEDEWEEMDEGAKFKFSLFPNDLVELKTKKEYFFGYIG
 LDRATGNISLKEHDGEISKGKDG VYRVGVKLALSFEKYQVDELGKNRQICRPQQRQ
 PVR

SEQ ID NO: 54

25 The nucleotide sequence of the synthesized Cas9-D10A Nickase gene in plasmid
 pET30a-Cas9-D10A nickase. the oligonucleotide encoding the N-terminal His6 tag is in
 underlined and italicized, the oligonucleotide encoding the SV40 nuclear localization
 signal is shown in italic and grey highlight and the oligonucleotide encoding the BLR
 nuclear localization signal is shown in underlined and grey highlight.

30 atg*caccatcatcatcatcat*tcttctggtctggtgccacgcggttctggtatgaaagaaaccgctgctgctaaattcgaac
 gccagcacatggacagcccagatctgggtaccgacgacgacgacaaggccatggcccaaaaaagaaacgcaag
 gttatggataaaaaatacagcattggtctggccatcggaaccaacagcgttgggtgggcagtaataacagatgaataca
 aagtgccgtcaaaaaaatttaaggttctggggaatacagatcgccacagcataaaaaagaatctgattggggcattgctg
 ttgattcgggtgagacagctgaggccacgcgtctgaaacgtacagcaagaagacgttacacacgtcgtaaaaatcgtat

ttgctacttacaggaaatTTTTCTAACGAAATGGCCAAGGTAGATGATAGTTTCTCCATCGTCTCGAAGAATCTTTCTGGTTGA
ggaagataaaaaaacgaacgtcacctatctttggcaatatcgtaggatgaagtggcctatcatgaaaaataccctacga
ttatcatcttcgcaagaagttggtgatagtagcgacaaagcggatctgcgttaatatctatctgcgttagcgccatgatca
aatctcgtaggtcatttctaattgaaggtagatctgaatcctgataactctgatgtggacaaattgtttatacaattagtgcaaacct
5 ataatcagctgttcgaggaaaaccccatatgcctctggagttgatgccaaagcgatttaagcgcgagactttctaagtc
ccggcgtctggagaatctgatcgccagttaccaggggaaaagaaaaatggctgtttggtaatctgattgccctcagctctg
gggcttaccctgaacttcaaatccaattttgacctggctgaggacgcaaagctgcagctgagcaaagatacttatgatgat
gacctcgacaatctgctcgccagattggtagcaatatgcggatctgtttctggcagcgaagaatcttcggatgctatctt
gctgtcggatattctgcgtgttaataaccgaaatcaccaaagcgcctctgtctgcaagtagatcaagagatagcagcagca
10 ccaccaggacctgactcttctaaggcactggtacgccaacagcttccggagaaatacaaagaaatattcttcgaccagt
ccaagaatggttacgcgggtacatcgatgggtgcatcacaggaagagttctataaattattaaaccaatccttgaga
aaatggatggcacggaagagttactgttaaaacttaaccgcaagacttgcttagaaagcaacgtacattcgacaacgg
ctccatcccacaccagattcatttaggtgaacttcacgccatcttgcgcagacaagaagatttctatcccttctaaaagaca
atcgggagaaaaatcgagaagatcctgacgttccgcattccctattatgtcggccccctggcacgtggtaattctcggttgcct
15 ggatgacgcgcaaaagttaggaaaccatcaccccttgaaactttgaagaagtcgtggataaaggtagcgcgcagt
ctttatagaaagaatgacgaacttcgataaaaacttgcccaacgaaaaagtcctgccaagcactctctttatatgagta
ctttactgtgtacaacgaactgactaaagtgaataacgttacggaaggtagcgcgcaaacctgcctttcttagtggcgagcag
aaaaaagcaattgtcgatcttctttaaacgaatcgcaaggtaactgtaaaacagctgaaggaagattattcaaaaag
atcgaatgctttgattctgtcgagatctcgggtgtcgaagatcgtttcaacgcttcccttagggacctatcatgatttgcgaagat
20 aataaaaagacaaagactttctcgacaatgaagaaaatgaagatattctggaggatattgtttgaccttgaccttattcgaa
gatagagagatgatcgaggagcgttaaaaacctatgccacctgtttgatgacaaagtcataagcaattaaagcgcc
gcagatatacggggtggggccgcttgagccgcaagttgattaacggtattagagacaagcagagcggaaaaaactatcc
tggatttctcaaatctgacggatttgcaaccgcaattttatgcagcttatacatgatgattcgttaccttcaagaggatatt
tcagaaggctcaggtgtctgggcaagggtgattcactccacgaacatatagcaaatttggccggctctcctgcgattaagaa
25 ggggatcctgcaaacagttaaagttgtggatgaactgtaaaagtaatgggccgccacaagccggagaatatcgtgata
gaaatggcgcgcgagaatcaaacgacacaaaaagggtcaaaagaactcaagagagagaatgaagcgcattgagga
ggggataaaggaacttgatctcaattctgaaagaacatccagttgaaaacactcagctgcaaaatgaaaaattgtac
ctgtactacctgcagaatggaagagacatgtacgtggatcaggaattggatatcaatagactctcggactatgacgtagat
cacattgtccctcagagcttctcaaggatgattctatagataataaagtacttacgagatcggacaaaaatcgcggtaaat
30 cggataacgtcccatcggagggaagtcgttaaaaagatgaaaaactattggcgtcaactgctgaacgccaaagctgatcac
acagcgtaaagtttgataatctgactaaagccgaacgcgggtgggtcttagtgaactcgataaagcaggatttataaacggc
agttagtagaaacgcgccaattacgaaacacgtggctcagatcctcgattctagaatgaatacaaagtacgatgaaaa
cgataaactgatccgtgaagtaaaagtcattacctaataaactgtgtccgatttccgcaaagattttagttttacaa
gggtccgggaaatcaataactatcccatgcacatgatgcataattaaatgcgggttaggcacggcccttataagaaatac

cctaactcgaagtgagttgtttatggggattataaagtgatgacgttcgcaaatgatcgcgaaatcagaacaggaa
atcggtaaggctaccgctaataacttttttattccaacattatgaatttttaagaccgaaataactctcgcgaaatggtgaaat
ccgtaaacggcctcttatagaaaccaatggtgaaacgggagaaaatcgtttgggataaaaggctgtgactttgccaccgttcg
taaagtcctctcaatgccgcaagttaacattgtcaagaagacggaagttcaaacagggggattctccaaagaatctatcct
5 gccgaagcgtaacagtgataaacttattgccagaaaaaagattgggatccaaaaaatacggaggccttgattccct
accgtcgcgtatagtgtgctggtggtgctaaagtcgagaaagggaaaagcaagaaattgaaatcagttaaagaactgc
tgggtattacaattatggaaagatcgctccttgagaaaaatccgatcgacttttagaggccaaggggtataaggaagtga
aaaaagatctcatcatcaaattaccgaagtatagtcttttgagctggaaaacggcagaaaaagaatgctggcctccgcg
ggcgagttacagaaggggaaatgagctggcgctgccttccaaatatgttaattttctgtaccttgccagtcattatgagaaact
10 gaagggcagccccgaagataacgaacagaaacaattatcgtggaacagcataagcactatttagatgaaattataga
gcaaattagtgaattttctaagcgcggtatcctcgcggatgctaatttagacaaagtactgtcagcttataataaacatcggg
ataagccgattagagaacaggccgaaaatatcattcatttgtttacctaaccaaccttgagcaccagctgccttcaaata
ttcgataccacaattgatcgtaaacgggtatacaagtacaaaagaagtcttggacgcaacctcattcatcaatctattactg
gattatatgagacacgcattgatctttcacagctgggcggagacagaagaaaaaactgaaactg

15

SEQ ID NO:55

The amino acid sequence of the Cas9-D10A nickase protein expressed from plasmid pET30a-Cas9-D10A nickase(encoded by SEQ ID NO:54). the N-terminal His6 tag is shown in underline and italics, the mutated amino acid is in bold underline (D10A), the SV40 nuclear localization signal is shown in italic and grey and the BLR nuclear localization signal is shown in underlined and grey.

[illegible]

geirkrplietngetgeivwdkgrdfatvrkvlsmppqvnivkktevtggfskesilpknsdkliarkkdwdpkkyggfdspt
 vaysvlvakvekgkskklksvkellgitimerssfeknpidfleakgykevkkdliiklpkyslfelengrkrmlasagelqk
 gnelaipskyvnflylashyeklkgspedneqkqlfveqhkhyldieiieqisefskrviladanldkvlsaynkhrdkpireq
 aeniihlftlnlgapaafkyfdttidrkrytstkevldatlihqsitglyetridlsqlggdkkkkkl

5

SEQ ID NO: 56

The nucleotide sequences of the substrate DNA fragment. The UTR sequences were shown in lowercase while the TrGA gene was shown in uppercase. Two selected VT domains, TrGA_sgF1 and TrGA_sgR1, were shown in bold and underlined, respectively.

10

gactgtctccaccatgtaattttccctgcgactccatataacgccgcatcgtgaaattttcttctttccttcttcaacaa
 acaacggatctgtgcttgcgggtcccctgcgttcacgcgtcagggcgactgctctgcagctcgataactccatggagccat
 caacttgctatggtgtcaatcatcctatcgacaggtccaagaacaagccggcctccggctgcctcattcgctgtcgcaaga
 cggcttgagtgtgtggtggaggattcgggggccccatattccaaccctttttccaaggccgtcggccggtgaggttgag
 15 gaaaaccatgggtgcctacatattatcgatgctggtgtttggtagtagcaatgtttgcgggtggcagttgagccgagcctcgt
 cttgggcttctgacccaggcaacgccatctgactagctgcgccgaaggaaggatgattcattgtacgacgccagtcattg
 gaatcttcaagtaaaagcccgacgaaccgacatgtcagatatcagaattctcctggctggtggggtggttgagactgc
 ttacggagtcgatgcctcgtgactgtcatggccgcgtccagcctcctgggactctgtccgatattatgacacgagtaaagcc
 tgcattgatgtcagtttgcgtctcatgtcgagaacaacacacctgggtgctacataggcaatactacctcgtagcttcaaa
 20 gttgactgttttgcttgatgtctttgatcatgccatccatcccttgccttgacgtgcatgtggatctctacgtccagacggggag
 aaagcttgcctgtgataaagtacgatgatgcattgatgcctgtggctacggccctttatccccatcgtcatgcattctatatta
 atccaggagactctcctcctggcatgggtgagtacaagtacgaggacatgtagaagcagagccacgcaacgtcttga
 catctgtacctattttgggccaaaaatcgagacccaccagctcgtcctaccttacatgtgaagatcttagccacaatcctac
 tgttttactagtattactgcacagctgtcatcacgagtcctcgggtgcttgtaaaccagctcagctcctgagcacatgcagt
 25 aacgccgactcggcgtcatttcgccacacccaatttgacctgagggatgctggaagctgctgagcagatcccgttaccg
 attcatggcactactacatccatcgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcc
 cggcattgtcgatgcagagatgcagatttcagcgggagcatttagggtagggcgactactactactaataccacctagtc
 gtatgtatctagcaccggaggctaggcggtagtgagcgggaacctgggtcattccatcgcaaccaggatcccgcacttcgt
 tgcgcttctgccccacggggcgggagttggcagaggcagaatgcggagcagcccctgtctgccctggccggggcct
 30 gttgaagcaagcagacgagagcagagcgggttgagaagcgggtggttgacgcttgacggtacgaagacgagcagagaat
 cccgttaagccgaggctgggctccccccccgtcatcatcatgccatcctgctctccagcccactcgtctccctgcctcgt
 cgctccccctcccccgattagctgcgcatttctcctgacagcgtgactaatgacgcgttgccagcccattcgctga
 cgcatcccggcatctgagctagctcgtcacgctggcaatcttgcccaggcagagcagcaagacggcgggcatgattg
 ggccgtgccctggcgggcatcagctggccatccgctgccacccgagaccgcatcaccgacttgcggatctctccgagc

agcaggaggctgatcctggccggcgagacgattgaaaagggctgccggggcccgagcaggacagcggcgagagc
 gagcgagagagaggaaaagaagaaggtcgactgtctattttcagccagccccggctcaacagaagcagaggagaa
 ggcgaacgacgtcaacgacgacgacgacgacgaagacgggtgaagtcggttagttgaagatccttgccgtcaca
 caccatctcgtggatattgctttccctgccgttgccgttgccacctgttccctcttcttcccccttcttctcattccgagcgct
 5 actggttctactccgcagccttcgggttgcccttcttcttgctgaccattgcaccgcccgtcgcggcacttgggccccggag
 aattcggcccttctcgagcattttggccctcagttcccatggggacgggtccacacttctctcttggccctgcagacctttgt
 cgtcgggtccgagtcggaagaagctcagtccttgagcgcttgagtagcatctacgcgcgaatcactggacaaagtcggcaa
 gacgaagccgtcgtcgcctgctgctgctgtgtactgcgacaggcgctccgactgggggcatcggcataataaaaagat
 gcccgccttcgcatggacctggccatgagccactcggcatcggctctctctctcaacgcttctctcacacatctcttcat
 10 tccgcccata**ATGCACGTCCTGTCGACTGCGGTGCTGCTCGGCTCCGTTGCCGTTCAA**
AAGGTCCTGGGAAGACCAGGATCAAGCGGTCTGTCCGACGTCACCAAGAGGTCT
 GTTGACGACTTCATCAGCACCGAGACGCCTATTGCACTGAACAATCTTCTTTGCAA
 TGTTGGTCCTGATGGATGCCGTGCATTCGGCACATCAGCTGGTGCGGTGATTGCA
 TCTCCCAGCACAAATTGACCCGGACTGTAAGTTGGCCTTGATGAACCATATCATATA
 15 TCGCCGAGAAGTGGACCGCGTGCTGAGACTGAGACAGACTATTACATGTGGACG
 CGAGATAGCGCTCTTGTCTTCAAGAACCTCATCGACCGCTTCACCGAAACGTACG
 ATGCGGGCCTGCAGCGCCGCATCGAGCAGTACATTACTGCCCAGGTCACTCTCCA
 GGGCCTCTCTAACCCTCGGGCTCCCTCGCGGACGGCTCTGGTCTCGGCGAGCC
 CAAGTTTGAGTTGACCCTGAAGCCTTTCACCGGCAACTGGGGTGCACCGCAGCG
 20 GGATGGCCCAGCTCTGCGAGCCATTGCCTTGATTGGATACTCAAAGTGGCTCATC
 AACAACAATATCAGTCGACTGTGTCCAACGTCATCTGGCCTATTGTGCGCAACGA
 CCTCAACTATGTTGCCCAGTACTGGTCAGTGCTTGCTTGTCTTGAATTACGTCTT
 TGCTTGTGTGTCTAATGCCTCCACCACAGGAACCAACCGGCTTTGACCTCTGGG
 AAGAAGTCAATGGGAGCTCATTCTTTACTGTTGCCAACCAGCACCGAGGTATGAA
 25 GCAAATCCTCGACATTCGCTGCTACTGCACATGAGCATTGTTACTGACCAGCTCTA
 CAGCACTTGTCGAGGGCGCCACTCTTGCTGCCACTCTTGCCAGTCGGGAAGCG
 CTTATTCATCTGTTGCTCCCCAGGTTTTGTGCTTTCTCCAACGATTCTGGGTGTGCG
 TCTGGTGGATACGTCGACTCCAACAGTATGTCTTTTCACTGTTTATATGAGATTGG
 CCAATACTGATAGCTCGCCTCTAGTCAACACCAACGAGGGCAGGACTGGCAAGGA
 30 TGTCAACTCCGTCCTGACTTCCATCCACACCTTCGATCCCAACCTTGGCTGTGACG
 CAGGCACCTTCCAGCCATGCAGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTGT
 CGACTCCTTCCGCTCCATCTACGGCGTGAACAAGGGCATTCTGCGCGGTGCTGCC
 GTCGCCATTGGCCGGTATGCAGAGGATGTGTACTACAACGGCAACCCTTGGTATC
 TTGCTACATTTGCTGCTGCCGAGCAGCTGTACGATGCCATCTACGTCTGGAAGAA

GACGGGCTCCATCACGGTGACCGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGTT
 CCTGGCGTGACGGCCGGGACCTACTCCAGCAGCTCTTCGACCTTTACCAACATCA
 TCAACGCCGTCTCGACATACGCCGATGGCTTCCTCAGCGAGGCTGCCAAGTACGT
 CCCC GCCGACGGTTCGCTGGCCGAGCAGTTTGACCGCAACAGCGGCACTCCGCT
 5 GTCTGCGCTTCACCTGACGTGGTTCGTACGCCTCGTTCTTGACAGCCACGGCCCGT
 CGGGCTGGCATCGTGCCCCCCTCGTGGGCCAACAGCAGCGCTAGCACGATCCCC
 TCGACGTGCTCCGGCGCGTCCGTGGTCCGATCCTACTCGCGTCCCACCGCCACG
 TCATTCCCTCCGTGCGCAGACGCCCAAGCCTGGCGTGCCTTCCGGTACTCCCTACA
 CGCCCCTGCCCTGCGCGACCCCAACCTCCGTGGCCGTACCTTCCACGAGCTCG
 10 TGTCGACACAGTTTGGCCAGACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTGG
 GCAACTGGAGCACGAGCGCCGCGGTGGCTCTGGACGCCGTCAACTATGCCGATA
 ACCACCCCCTGTGGATTGGGACGGTCAACCTCGAGGCTGGAGACGTCGTGGAGT
 ACAAGTACATCAATGTGGGCCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCAA
 CCACACTTACACGGTTCCTGCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGAC
 15 ACctggcagtcgtaatgaatcggcaagggtagtactagtagactttagtctgcc

SEQ ID NO:57

forward primer for SEQ ID NO:56:

5'- gactgtctccaccatgtaattttc-3'

20

SEQ ID NO:58

TrGA primer KOR2 ; reverse primer for SEQ ID NO:56:

5'-ggcagactacaagtctactagtagtactac-3'

25 SEQ ID NO: 59

The oligonucleotide sequences for the transcription of the T7 promoter, CER domain, and the VT domain TrGA_sgF1. The VT domain is shown in uppercase, while the T7 promoter and CER domain region are shown in bold and lowercase, respectively.

taatacgaactcactataGGGAAGACCAGGATCAAGggttttagagctagaaatagcaaggtaaaataaggct
 30 agtccgttatcaacttgaaaaagtggcaccgagtcggtgc

SEQ ID NO: 60

The oligonucleotide sequences for the transcription of the T7 promoter, CER domain, and the VT domain TrGA_sgR1. The VT domain is shown in uppercase, while the T7 promoter and CER domain region are shown in bold and lowercase, respectively.

taatac**gactcactata**GGACAGACCGCTTGATCCgtttagagctagaaatagcaagttaaataaggcta
gtccggtatcaacttgaaaaagtggcaccgagtcggtgc

SEQ ID NO:61

The nucleotide sequence of the TrGA knockout cassette:

ccctgcctcgtcgccctccccctcccccgattagctgcgcgatgttctcctgacagcgtgactaatgacgcggtgccagccc
attgcctgacgcatcccggcatctgagctagctcgtcacgctggcaatcttggcccaggcagagcagcaagacggcg
ggcatgattgggccgtgccctggcgggcatcagctggccatccgctgccacccgagaccgcatcaccgacttgtcggat
ctctccgagcagcaggaggctgatcctggccggcgagacgattgaaaagggctgccggggcccgagcaggacagcg
gcgagagcgagcgagagagaggaaaagaagaaggctgactgtcttatttcagccagccccgggtcaacagaagca
gaggagaaggcgaacgacgtcaacgacgacgacgacgacgaagacgggtgaagtccgttagttgaagatccttg
ccgtcacaacaccatctcgtggatattgtcttccccctgccgttgcggtgccacctgttccctcttctcttcccccttcttctcatt
ccgagcgctactggttctactccgcagccttcggttgtgcctttctcttctgacattgcaccgcccgtcgcggcacttgg
gccccggagaattcgcccttctgcgacatttggccctcagttcccatggggacggtccacacttctcttggccctgc
agacctttgtcgtcgggtccgagtcggaagaagctcagctctgagcgcttgagtagcatctacgcgcgaatcactggacaa
agtgcggaagacgaagccgtcgtcgctgctgctgctgttactgcgacaggcgctccgactgggggcatcggcata
ataaaaagatgcccgcccttcgcatggacctggccatgagccactcggcatcgggtctctctctcaacgcttctctcacac
atctcttctcattccgccccatcatggtttaaacctcgagttataagtacaacatgctctcaaagcgctcatggctggcaca
agcctggaaagaaccaacacaaagcatactgcagcaaatacagctgaattcgtcaccaattaagtgaacatcaacctga
aggcagagtatgaggccagaagcacatctggatcgcagatcatggattgcccctcttgttgaagatgagaatctagaaa
gatggcgggggtatgagataagagcgatgggggggcacatcatctccaagacaaacaacctttgcagagtcaggcaat
tttctgtataagagcaggaggaggaggtccagtcatttcatcagcggtaaaatcactctagacaatctcaagatgagttct
gccttgggtgacttatagccatcatcatacctagacagaagcttgtgggatactaagaccaacgtacaagctcgcactgta
cgctttgacttccatgtgaaaactcgatacggcgcgccctctaaattttatagctcaaccactccaatccaacctctgcacct
ctcactcgtctgatctactgttcaaatacagagaataaggacactatccaaatccaacagaatgggtaccacctcccagct
gcctgcctacaagcaggacttctcaaataccgccatcgacggcgcgctctcaagtttggcagcttcgagctcaagtcca
agcggatatccccctacttcttaacgcggggcgaattccacacggcgcgccctcgccggcgccatcgccctccgcctttgca
aagaccatcatcgaggcccaggagaaggccggcctagagttcgacatcgtcttcggccccggcctacaagggcatccc
gctgtgctccgccatcacatcaagctcggcgagctggcgccccagaaacctggaccgcgtctcctactcgtttgaccgca
aggaggccaaggaccacggcgaggcggaacatcgtcggcgcttcgctcaagggcaagagggtcctgattgtcga
cgacgtcatcaccqccqgcaccqccaagagggacqccattgagaagatcaccaagqagggcgcacgtcgcqccq

catcgtcgtggccctggaccgcatggagaagctccccgctgcggatggcgacgactccaagcctggaccgagtgccatt
ggcgagctgaggaaggagtagggcatccccatctttgccatcctcactctggatgacattatcgatggcatgaagggttt
gctacccctgaggatatcaagaacacggaggattaccgtgccaagtacaaggcgactgactgattgaggcggttcaatgt
cagaagggagagaaagactgaaaagggtggaaagaaggcacaattgttgttattattatttctatctcgaatcttctaga
5 tctgtcgtaaataaacaagcgtaactagctagcctccgtacaactgctgaattgatacccgatggaggggcagttatttat
ttgtttttcaagattttcattcgccgttgaactcgtctcacatcgcggtgattgcccgggtgcccatgtgttctcctactacccaa
gtccctcacgggtgtctcactttctttctctttatcctccctattttttcaagtcagcgacagagcagtcatatggggatacgt
gcaactgggactcacaacaggccatcttatggcctaatagccggcggttgatccactagtcattggttaaacagcacat
gcagtaacgccgactcggcgctcatttcgccacacccaatttgacctgagggatgctggaagctgctgagcagatcccg
10 taccgattcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaa
agtcccggtcattgtcgatgcagagatgcagatttcagcgggcgattctagggtagggcgactactactaataaccacct
agtcagtatgtatctagcacccggaggctaggcggttagtgacgggaacctgggtcattccatcgcaaccaggatcccg
acttcgttgcgcttctgccccacggggcgggagttggcagaggcagaatgcggagcagcccctgtctgccttgccg
gggcctgttgaagcaagcagacgagagcagagcgggtgagaagcgggtggtgacgcttgacggtacgaagacgagc
15 gagaatcccgtaagccgaggctgggctaattaattaatgaatcggcaaggggtagtactagtagacttgtagtctgccgg
attatgtattggagttggtcagtagaatgaaccacgggaatatcgggtcaccgggacatttgggatatagcgttctgagaag
ctgctggttgacgacatttgagaaggatgccctttacgacttataccgctatgccgggtatattaatttagccgttatgaaa
ctcaaagagacgatgataatgatgacgagtaattgttcgtttcaatttcgaaagctgactcccacgaagaatatgccaatg
accacggcatgaagcctgaactgggcgtgtgtaacactttaattgcctgacggcggacaaaacaaaggcggcagca
20 atgttgagaccgtgtgataaaccaaggttcccgaggggagagagagagagagagagagagagagagctaggtgaaaga
atgagtcgcctttgagtcacctgcgtctctctctccccctctctcactctctgtatcccatcaacctcttccctgttctctctat
cgcatccatgcgtttgcatcttccatttcattctttccccttgagccccatctatgcaaactcatcatccggcgccctgatggaat
ccttgaccttgatgagaatcgccgtcatccaaggctccagcctgctcgtgcgggtcgaactggaacagcagctcgctaaac
tcatctggctgtggttgcgacggcggtgcacaggctctcgagcagcttgacttgattgagaggagaactcggggtccttt
25 tggcggtaggactcgacggcgcgggcggtgcccaccatgtcgcccggtggcgaggtggcagatgccggccttgaagca
gtaggtcgagaggctccacttcatggtgccgttgcgatcatggtgttgatgatgcggctgtacgtctcgatggcgccgtagt
agtcgccgtcgaggggcgggcgaggctggcggtactgcgtccagagctt

SEQ ID NO:62

30 Fw1 (5'-cactactacatccatacgcagcaaacaatgg-3') and

SEQ ID NO:63

R3 (5'-ggtcaagaagcacatgccagagttcg -3')

SEQ ID NO: 64

TrGAF2 (5' gactgtctccaccatgtaatttttc 3')

SEQ ID NO:65

5 *E. coli* codon-optimized Cas9-D10A nickase gene (no stop codon)

atggataaaaaatacagcattggctcggccatcggaaccaacagcgttgggtgggcagtaataacagatgaatacaaaa
 gtgccgtcaaaaaaatttaaggttctgggaatacagatcgccacagcataaaaaagaatctgattggggcattgctgttt
 gattcgggtgagacagctgaggccacgcgtctgaaacgtacagcaagaagacgttacacacgtcgtaaaaatcgattt
 gctacttacaggaaatttttctaacgaaatggccaaggtagatgatagtttctccatcgtctcgaagaatctttctggtgag
 10 gaagataaaaaacacgaacgtcaccctatctttggcaatatcgtggatgaagtggcctatcatgaaaaataccctacgatt
 tatcatcttcgaagaagttggttgatagtagcgacaaagcggatctgcgtttaatctatcttgcgttagcgcacatgatcaa
 attcgtggtcattttctaattgaaggtgatctgaatcctgataactctgatgtggacaaattgtttatacaattagtgcacaccta
 taatcagctgttcgaggaaaacccattaatgcctctggagttgatgccaaagcgatttaagcgcgagactttctaagtcc
 cggcgtctggagaatctgatcgcccagttaccaggggaaaagaaaaatggctgtttggtaatctgattgccctcagtctgg
 15 ggcttaccggaacttcaaatacaatttgacctggctgaggacgcaaagctgcagctgagcaaagatacttatgatgatg
 acctcgacaatctgctcgcccagattggtgaccaatatgcggatctgtttctggcagcgaagaatcttccgatgctatcttgc
 tctcgatattctgcgtgttaataccgaaatcaccaaagcgctctgtctgcaagtatgatcaagagatacgcgagcacc
 accaggacctgactcttctaaggcactggtacgccaacagcttccggagaaatacaagaaatattcttcgaccagtcc
 aagaatggttacgcggttacatcgatggtggtgcatcacaggaagagttctataaatttataaaccaatccttgagaaa
 20 atggatggcacggaagagttactgttaaacttaaccggaagacttgcttagaaagcaacgtacattcgacaacggctc
 catccacaccagattcatttaggtgaacttcacgccatcttgcgcagacaagaagatttctatcccttctaaaagacaatc
 gggagaaaatcgagaagatcctgacgttccgcattccctattatgtcgggtcccctggcacgtggttaattctcggttgccctgg
 atgacgcgcaaaagtgaggaaaccatcaccccttgaactttgaagaagtcgtggataaagggtgctagcgcgcagtcttt
 tatagaaagaatgacgaacttcgataaaaacttgcceaacgaaaaagtcctgccaagcactctcttttatatgagtacttt
 25 actgtgtacaacgaactgactaaagtgaatacgttacggaaggtatgcgcaaacctgccttcttagtggcgagcagaa
 aaaagcaattgtcgatcttctttaaacgaatcgcaaggtaactgtaaaacagctgaaggaagattttcaaaaagat
 cgaatgctttgattctgtcgagatctcggtgtcgaagatcggttcaacgcttccctagggacctatcatgatttctgaagata
 ataaaagacaaagactttctcgacaatgaagaaaatgaagatattctggaggatattgtttgacctgaccttattcgaag
 atagagagatgatcgaggagcgcttaaaaacctatgcccacctgtttgatgacaaagtcataagcaattaaagcgccg
 30 cagatatacggggtggggccgcttgagccgcaagttgattaacggtattagagacaagcagagcggaaaaactatcct
 ggatttctcaaatctgacggatttgcgaaccgcaattttatgcagcttatacatgatgattcgcttacattcaaagaggatatt
 cagaaggctcaggtgtctgggcaaggtgattcactccacgaacatatagcaaatgtggccggctctctctgcgattaagaa
 ggggatcctgcaaacagttaaagttgtggatgaactgtaaaagtaatgggcccgcacaagccggagaaatcgtgata
 gaaatggcgcgcgagaatcaaacgacacaaaaagggtcaaaagaactcaagagagagaatgaagcgcattgagga

ggggataaaggaacttggatctcaaattctgaaagaacatccagttgaaaacactcagctgcaaaatgaaaaattgtac
 ctgtactacctgcagaatggaagagacatgtacgtggatcaggaattggatatcaatagactctcggactatgacgtagat
 cacattgtccctcagagcttctcaaggatgattctatagataataaagtacttacgagatcggacaaaaatcgcggtaaat
 cggataacgtcccatcggaggaagtcgttaaaaagatgaaaaactattggcgtcaactgctgaacgccaagctgatcac
 5 acagcgtaagtttgataatctgactaaagccgaacgcgggtggcttagtgaactcgataaagcaggatttataaacggc
 agttagtagaaacgcgccaattacgaaacacgtggctcagatcctcgattctagaatgaatacaaagtacgatgaaaa
 cgataaactgatccgtgaagtaaaagtcattaccttaaaaactaaactgtgtccgatttccgcaaagatttcagttttaca
 ggtccgggaaatcaataactatcacatgcacatgatgcataatttaaatgcgggtgtaggcacggcccttattaagaaatac
 cctaaactcgaaagtgaagttgtttatggggattataaagtgtatgacgttcgcaaaatgatcgcgaaatcagaacaggaa
 10 atcggtaaggctaccgctaaatactttttattccaacattatgaatttttaagaccgaaataactctcgcgaaatggtgaaat
 ccgtaaacggcctcttatagaaaccaatggtgaaacgggagaaatcggttgggataaaggctgtgactttgccaccggttcg
 taaagtcctctcaatgccgcaagttaacattgtcaagaagacggaagttcaaacagggggatttccaaagaatctatcct
 gccgaagcgtaacagtataaactattgccagaaaaaaagattgggatccaaaaaaatacggagggttattccct
 accgtcgcgtatagtgtgctggtggtgctaaagtcgagaaagggaaaagcaagaaattgaaatcagttaaagaactgc
 15 tgggtattacaattatggaaagatcgctccttgagaaaaatccgatcgacttttagaggccaaggggtataaggaagtga
 aaaaagatctcatcatcaaattaccgaagtatagtcttttgagctggaaaacggcagaaaaagaatgctggcctccgcg
 ggcgagttacagaaggggaaatgagctggcgctgccttccaaatatgttaattttctgtaccttgccagtcattatgaaaaact
 gaagggcagccccgaagataacgaacagaaacaattatcggtgaacagcataagcactatttagatgaaattataga
 gcaaattagtgaattttctaagcgcgttatcctcgcggtatgctaatttagacaaagtactgtcagcttataataaacatcggg
 20 ataagccgattagagaacaggccgaaaatatcattcatttgtttacctaaccaaccttgagaccagctgccttcaaata
 ttctgataccacaattgatcgtaaacggtatacaagtacaaaagaagtcttgacgcaaccttcattcatcaatctattactg
 gattatatgagacacgcattgatctttcacagctggcgaggagac

SEQ ID NO:66

25 Amino acid sequence of the Cas9-D10A nickase from *Streptococcus pyogenes*;
 encoded by SEQ ID NO:65

mdkkysiglaigtnsvgwavitdeykvpskkfkvlgntrhsikknliallfdsgetaeatrlkrtarrytrrknricylqeifsn
 emakvddsffhrleesflveedkkherhpifgnivdevayhekyptiyhlrkkldstdkadlrliylalahmikfrghfliegdl
 npdnsdvdklfiqlvtynqlfeenpinasgvdakailsarlsksrrlenliaqlpgekknglfgnlialslgltpnfksnfdlaed
 30 aklqlskdtydddldnllaigdqyadlflaaknlsdaillsdilrvnteitkapsasmikrydehqdltilkalvrqqpkeyke
 iffdqskngyagyidggasqeefykfikpilekmdgteellvklredllrkqrtdngsiphqihlgelhaulrrqedfypflkdn
 rekiekiltfripyyvgplargnsrfawmtrksetitpwnfeevvdkgasaqsfiermtnfdknlpnekvlpkhslllyeyftvy
 neltkvkyvtegmrkpafslsgeqkkaivdlflktnrkvtvkqlkedyfkkiectdsveisgvedrfnaslgythdllkiikdkdfld
 neenediledivltlftedremieerlktyahlfddkvmkqlkrtrytgwgrlsrklingirdkqsgktildflksdganrnmqli

hddsltfkediqaqvsgqgdsIhehianlagspaikkqilqtkvvdelvkvmgrhkpeniviemarenqttqkgqknsr
 ermkrriegikelgsqilkehpventqlqneklylylqngdmyvdqeldinrlsdydvdhivpqsfllkddsidnkvltrsd
 nrgksdnvpseevvkkmknywrqllnaklitqrkfdnltaergglseIdkagfkrqlvetrqitkhvaqildsrmtkyden
 dklirevkvitlksklvsdfrkdfqfykvreinnyhahdaylnavvgtalikkypklesefvygdykvydvrmiaakseig
 5 katakyffysnimnffkteitlangeirkprietngetgeivwdkgrdfatvrkvlsmqpvnivkktevqggfskesilpk
 dklarkkdwdpkkyggfdsptvaysvlvakvekgskklksvkellgitimerssfeknpidfleakgykevkdliiklpk
 yslfelengrkrmlasagelqkgnelalpskyvnflylashyeklkgspedneqkqlfveqhkhyldiieqisefskrvilad
 andkvlsaynkhrrdkpireqaeniihlftlnlgapaafkyfddtidrkrytstkevldatlihqsitglyetridlsqllggd

10 SEQ ID NO:67

N-terminal His6 tag / thrombin / S•Tag™ / enterokinase region amino acid sequence
 (with start methionine)

mhhhhhssglvprgsgmketaaakferqhmdspdlgtdddddka

15 SEQ ID NO:68

N-terminal His6 tag / thrombin / S•Tag™ / enterokinase region polynucleotide sequence
 (with start codon); encodes SEQ ID NO:67

atgcaccatcatcatcatcattcttctggtctggtgccacgcggttctggtatgaaagaaaccgctgctgctaaattcgaacg
 ccagcacatggacagcccagatctgggtaccgacgacgacgacaaggccatggcc

20

SEQ ID NO:69

SV40 NLS coding sequence (encodes SEQ ID NO:46)

ccaaaaaagaaacgcaaggtt

25 SEQ ID NO:70

BLR nuclear localization signal coding sequence (encodes SEQ ID NO:47)

Aagaagaaaaaactgaaactg

SEQ ID NO:71

30 tgatgacggtgaaaacctc

SEQ ID NO:72

aaaagcaccgactcgg

SEQ ID NO:73

Gla1rep Primer 1556F

ATGCGCAAATTTAAAGCGCTGATgtgtgtctaatgcctccaccac

5 SEQ ID NO:74

Gla1rep Primer 1557R

ATATGGATCTGCGCGCGATCGATgatcgtgctagcgctgctgttg

SEQ ID NO:75

10 Sequencing primer 1538F

CCACCACAGGAACCAAACC

SEQ ID NO:76

Sequencing primer 1539R

15 CTGCGACGGAGGGAATGACG

SEQ ID NO:77

Sequencing primer 1540F

GGGCAGGACTGGCAAGGATGT

20

SEQ ID NO:78

Sequencing primer 1541R

GCCGTCACGCCAGGAACAAG

25 SEQ ID NO:79

TrGA Primer KOF1

gaacaatcttcttgcaatgttggtc

SEQ ID NO:80

30 Pyr2 Primer R1

gaggaagtcctgctgtaggcaggc

SEQ ID NO:81

Pyr2 Primer F4

5 cgacagagcagtcatatggggatacg

THAT WHICH IS CLAIMED:

1. A method for homologous recombination of a donor DNA with a genomic locus in a filamentous fungal cell, the method comprising:
 - a) introducing into a population of filamentous fungal cells a Cas endonuclease, a guide RNA, and a donor DNA comprising a domain with homology to a genomic locus of the fungal cell, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to act at a target site in or near the genomic locus of the fungal cells; and
 - b) identifying at least one fungal cell from the population in which homologous recombination of the donor DNA with the genomic locus has occurred,
wherein the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells.
2. The method of claim 1, wherein the non-homologous end joining (NHEJ) mechanism at the target site in the fungal cells is not activated, non-functional, or reduced.
3. The method of claim 2, wherein the non-homologous end joining (NHEJ) pathway in the fungal cells comprises one or more non-functional or reduced-activity components.
4. The method of claim 3, wherein the one or more non-functional or reduced-activity components are selected from the group consisting of ku80, ku70, rad50, mre11, xrs2, lig4, xrs, and combinations thereof.
5. The method of claim 4, wherein the one or more non-functional or reduced-activity components is ku80.
6. The method of any preceding claim, wherein the Cas endonuclease is a Cas nickase.

7. The method of any one of claims 1-5, wherein the Cas endonuclease is a Cas9 endonuclease or variant thereof.
8. The method of claim 7, wherein the Cas9 endonuclease or variant thereof comprises a full length Cas9 or a functional fragment thereof from a species selected from the group consisting of: *Streptococcus* sp., *S. pyogenes*, *S. mutans*, *S. thermophilus*, *Campylobacter* sp., *C. jejuni*, *Neisseria* sp., *N. meningitides*, *Francisella* sp., *F. novicida*, and *Pasteurella* sp., *P. multocida*.
9. The method of claim 8, wherein the Cas9 endonuclease or variant thereof comprises an amino acid sequence that has at least 70% identity to any one of SEQ ID NOs:45 and 48 to 53.
10. The method of any preceding claim, wherein the donor DNA comprises a polynucleotide sequence of interest, and wherein homologous recombination at the genomic locus results in the insertion of the polynucleotide sequence of interest in the genomic locus.
11. The method of any preceding claim, wherein the introducing step comprises introducing a DNA construct comprising an expression cassette for the Cas endonuclease into the fungal cells.
12. The method of any preceding claim, wherein the introducing step comprises introducing a DNA construct comprising an expression cassette for the guide RNA into the fungal cells.
13. The method of any preceding claim or any one of claims 48-54, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising a sequence encoding a selectable marker.
14. The method of claim 13, wherein the DNA construct comprises both the sequence encoding the selectable marker and the donor DNA.

15. The method of any preceding claim, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising: a sequence encoding the Cas endonuclease, a sequence encoding the guide RNA, a sequence encoding a selectable marker, and the donor DNA.
16. The method of any one of claims 11 to 15, wherein the DNA construct is a linear DNA construct.
17. The method of any one of claims 11 to 15, wherein the DNA construct is a circular DNA construct.
18. The method of any one of claims 11 and 15-17, wherein the expression cassette for the Cas endonuclease or the sequence encoding the Cas endonuclease comprises a Cas coding sequence that is optimized for expression in the filamentous fungal cell.
19. The method of claim 18, wherein the Cas coding sequence is a Cas9 coding sequence comprising a polynucleotide sequence that is at least 70% identical to SEQ ID NO:44.
20. The method of any one of claims 1 to 10, 12 to 14, and 16-17, wherein the introducing step comprises directly introducing the Cas endonuclease into the fungal cells.
21. The method of any one of claims 1 to 11, 13 to 14, and 16 to 20, wherein the introducing step comprises directly introducing the guide RNA into the fungal cells.
22. The method of any preceding claim, wherein the Cas endonuclease is operably linked to a nuclear localization signal.
23. The method of any preceding claim, wherein the filamentous fungal cell is a Eumycotina or Pezizomycotina fungal cell.

24. The method of any preceding claim, wherein the fungal cell is selected from the group consisting of *Trichoderma*, *Penicillium*, *Aspergillus*, *Humicola*, *Chrysosporium*, *Fusarium*, *Myceliophthora*, *Neurospora*, *Hypocrea*, and *Emericella*.
25. The method of any preceding claim, wherein the target site is located within a region of a gene of interest selected from the group consisting of an open reading frame, a promoter, a regulatory sequence, a terminator sequence, a regulatory element sequence, a splice site, a coding sequence, a polyubiquitination site, an intron site, and an intron enhancing motif.
26. The method of any preceding claim, wherein the homologous recombination results in a modification of the DNA sequence at or near the target site, wherein the modification is selected from the group consisting of a deletion of one or more nucleotides, an insertion of one or more nucleotides, insertion of an expression cassette encoding a protein of interest, a substitution of one or more nucleotides, and any combination thereof.
27. The method of any preceding claim, wherein the identifying step comprises culturing the population of cells from step (a) under conditions to select for or screen for the homologous recombination or the modification.
28. The method of any preceding claim, wherein the identifying step comprises culturing the population of cells from step (a) under conditions to screen for unstable transformants.
29. The method of any preceding claim, wherein the introducing step comprises introducing into the fungal cells a DNA construct comprising a sequence encoding a selectable marker and the donor DNA, and wherein the identifying step comprises culturing the population of cells from step (a) under conditions to screen for unstable transformants that have lost the selectable marker yet retained the donor DNA.
30. A recombinant filamentous fungal cell produced by the method of any preceding claim.

31. A recombinant filamentous fungal cell comprising a first recombinant DNA construct comprising an expression cassette for a Cas endonuclease.
32. The recombinant fungal cell of claim 30 or 31, wherein the recombinant filamentous fungal cell comprises one or more non-functional or reduced activity component in the NHEJ pathway.
33. The recombinant fungal cell of claim 32, wherein the one or more components of the NHEJ pathway are selected from the group consisting of ku80, ku70, rad50, mre11, xrs2, lig4, and xrs.
34. The recombinant fungal cell of any one of claims 30 to 33 further comprising a donor DNA comprising a polynucleotide of interest.

2/12

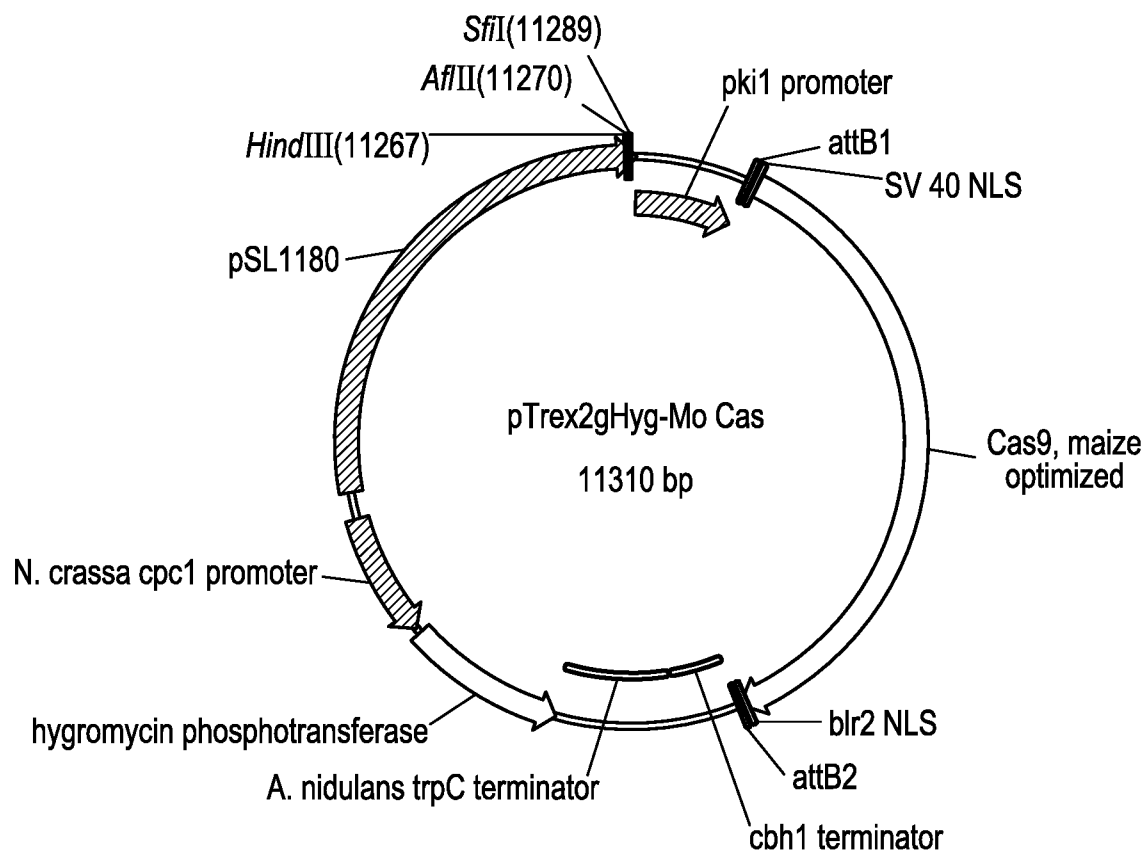
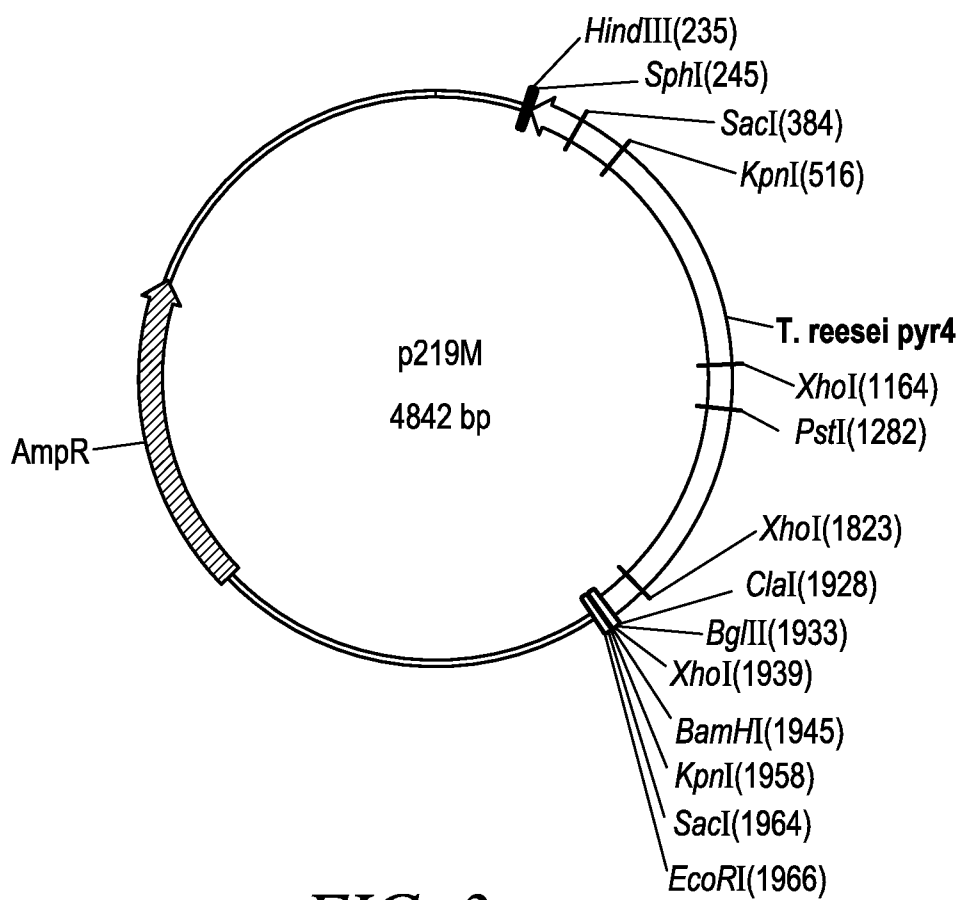


FIG. 2

3/12**FIG. 3**

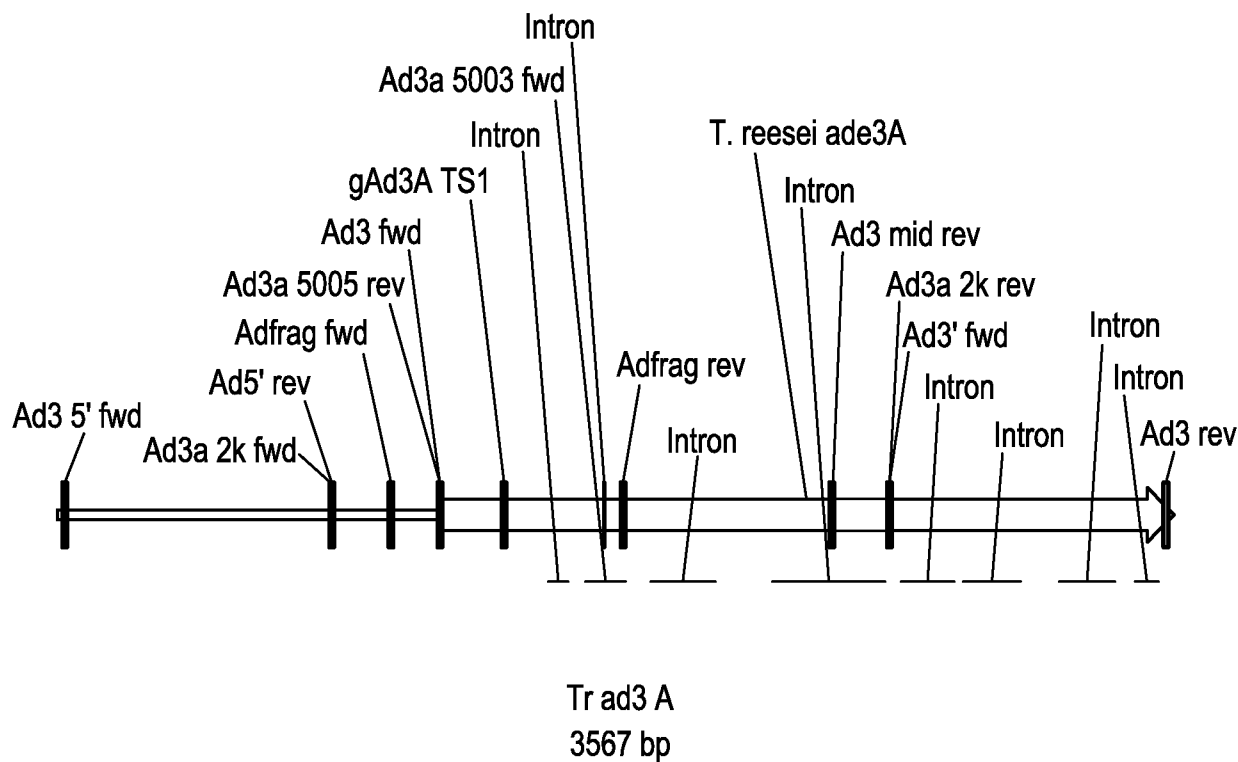


FIG. 4

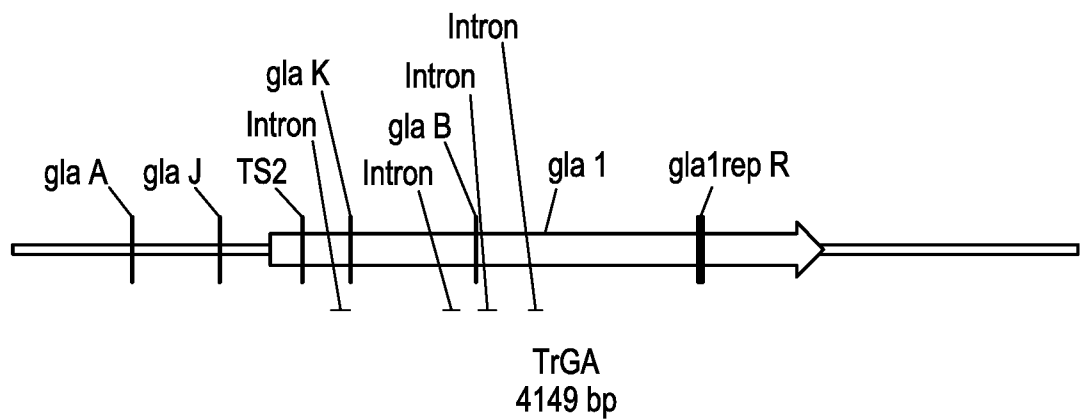


FIG. 5

5/12

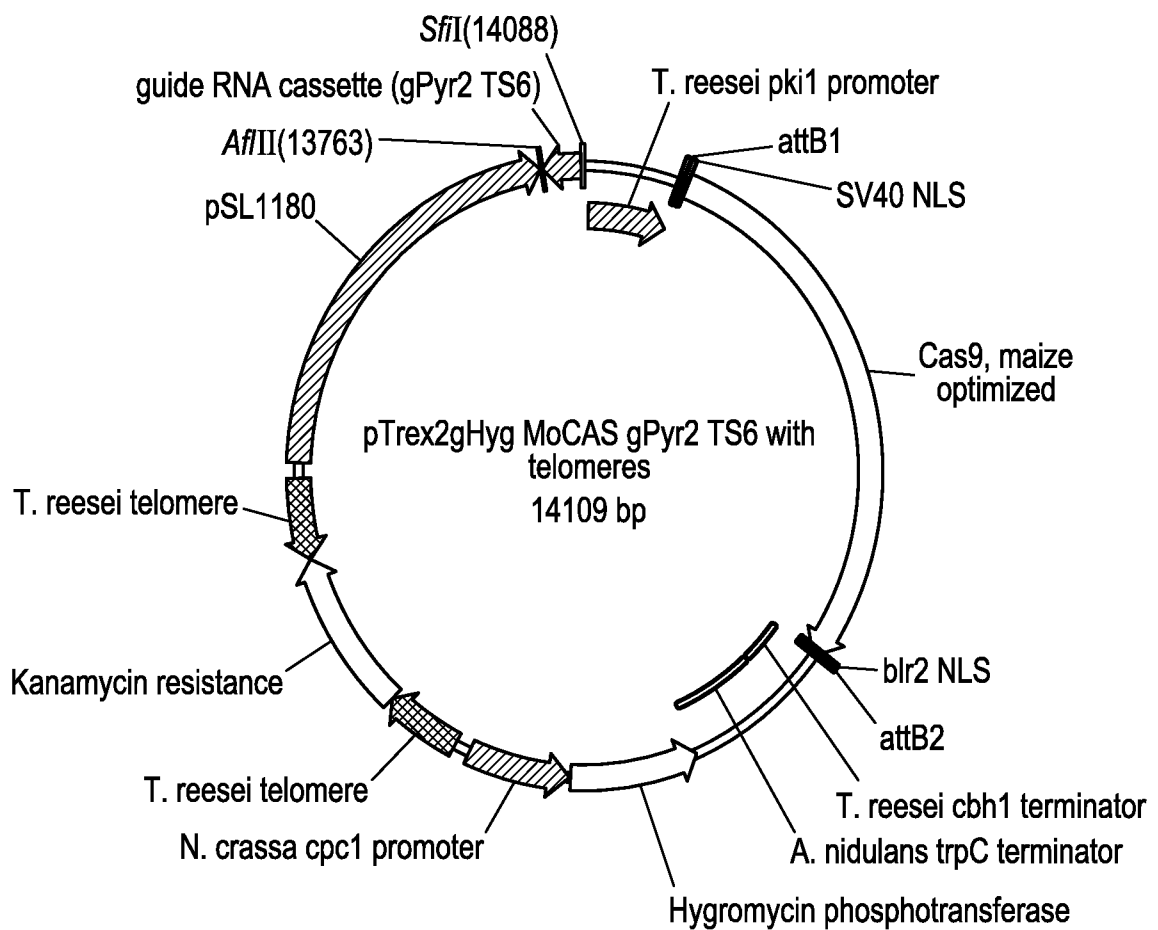


FIG. 6

6/12

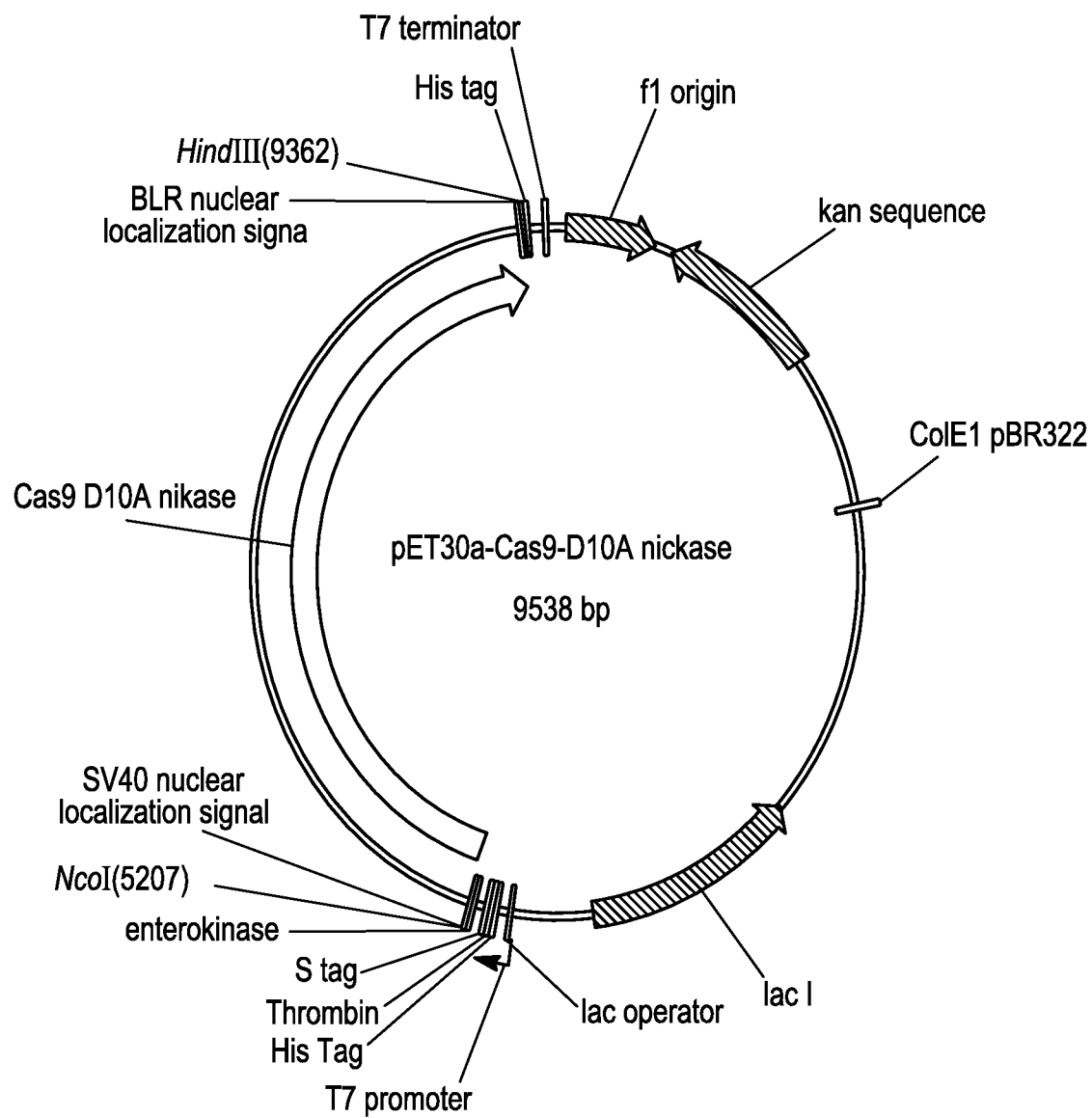
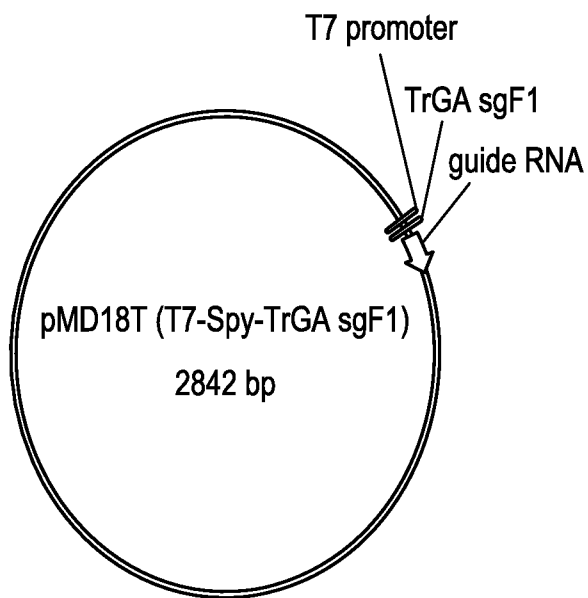
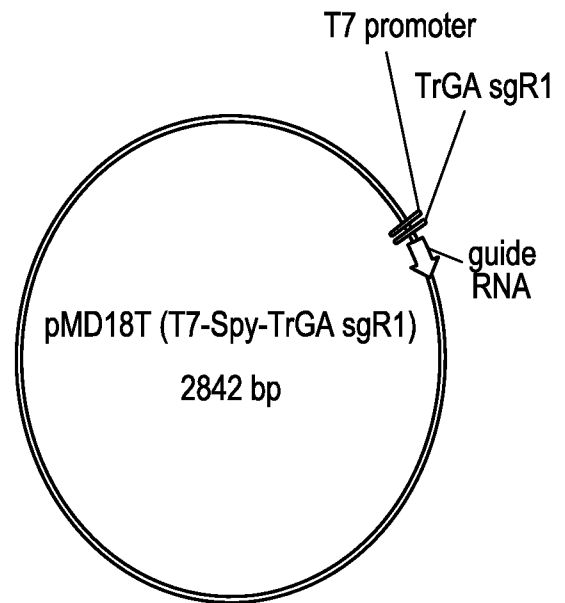
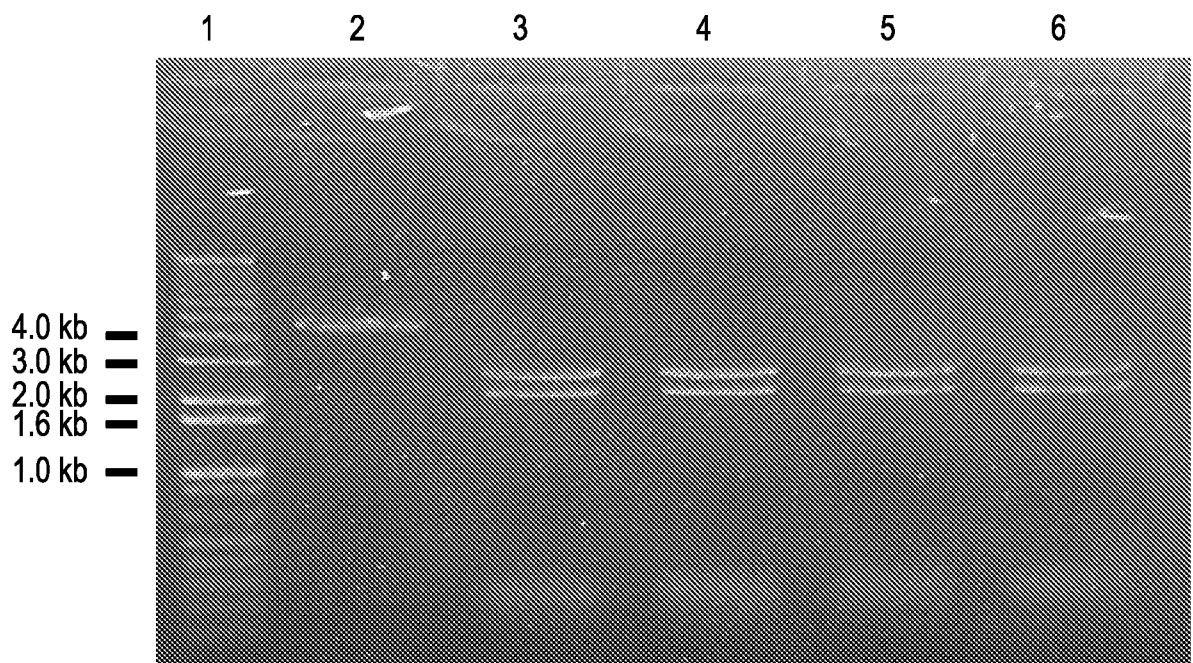


FIG. 7

7/12

*FIG. 8A**FIG. 8B**FIG. 9*

8/12

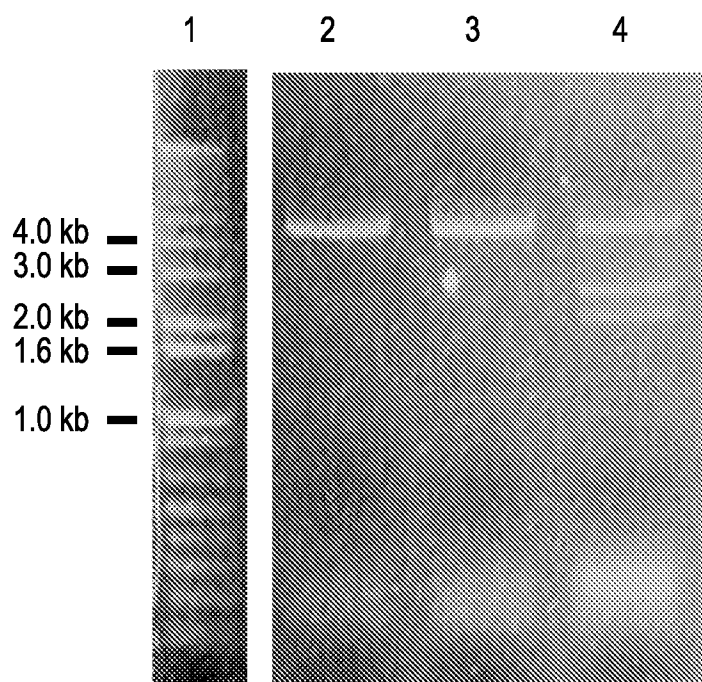


FIG. 10

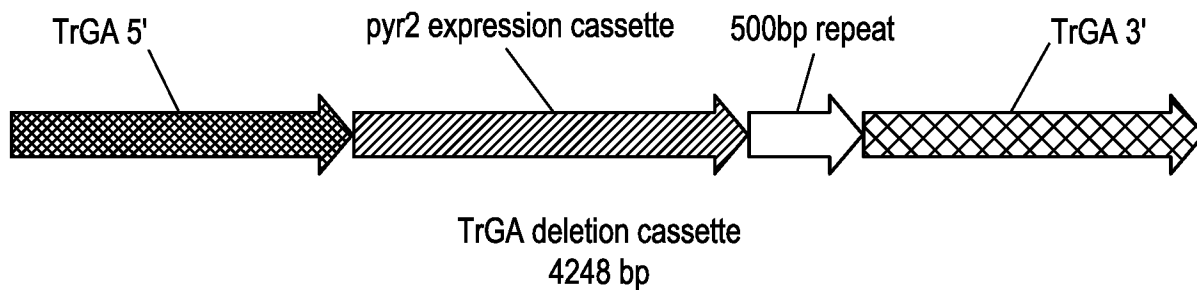


FIG. 11

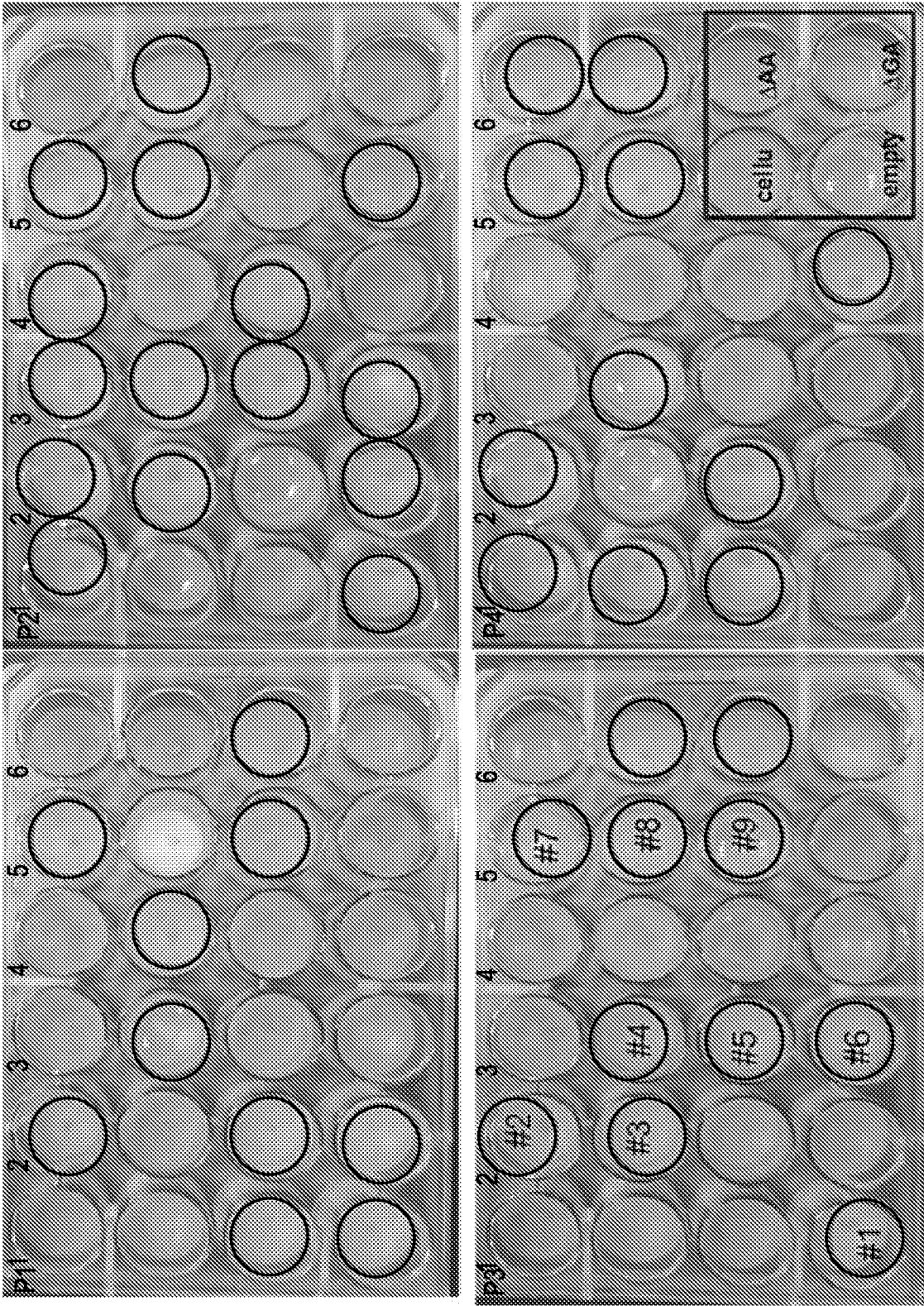


FIG. 12

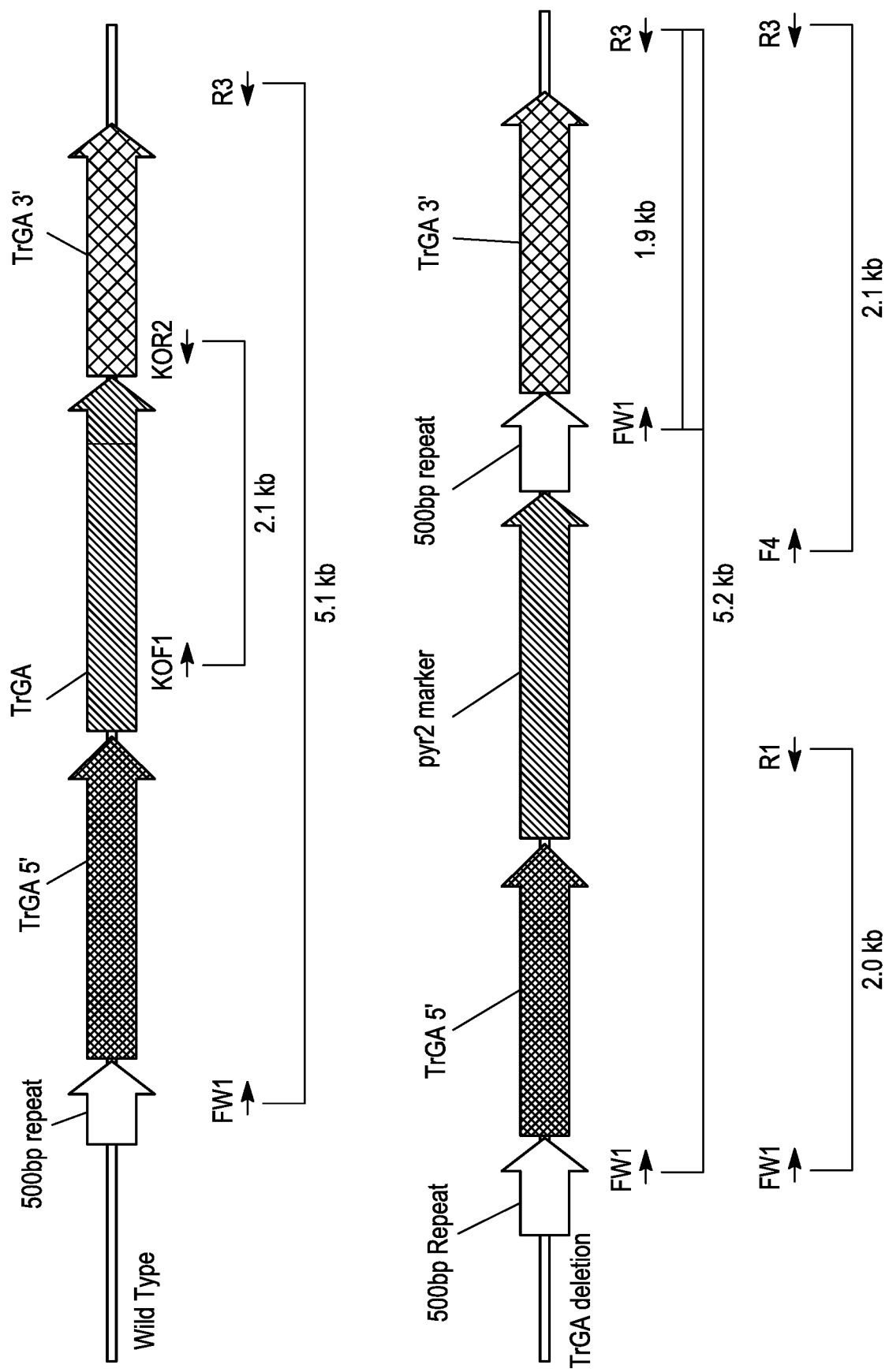


FIG. 13A

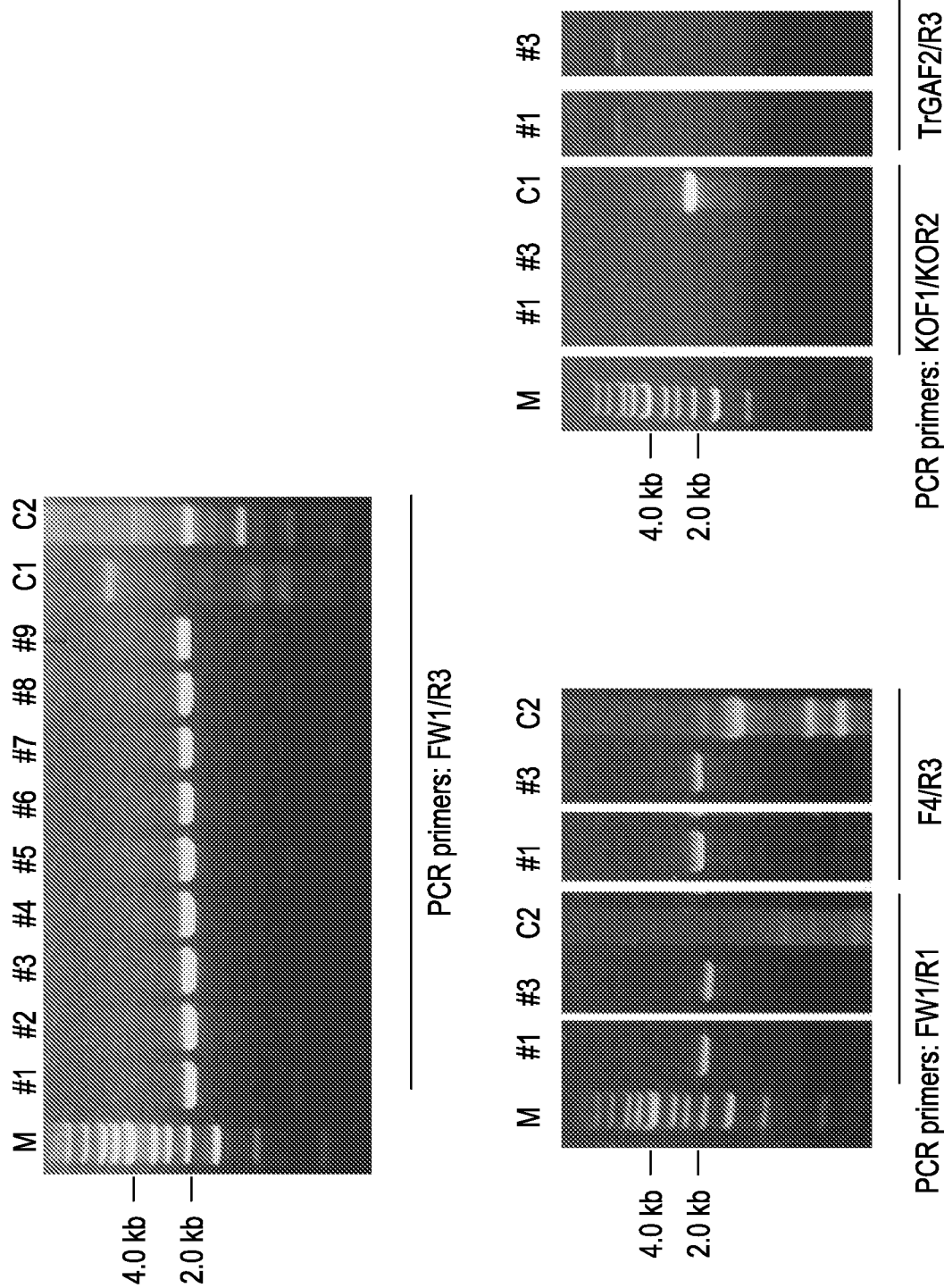


FIG. 13B

12/12

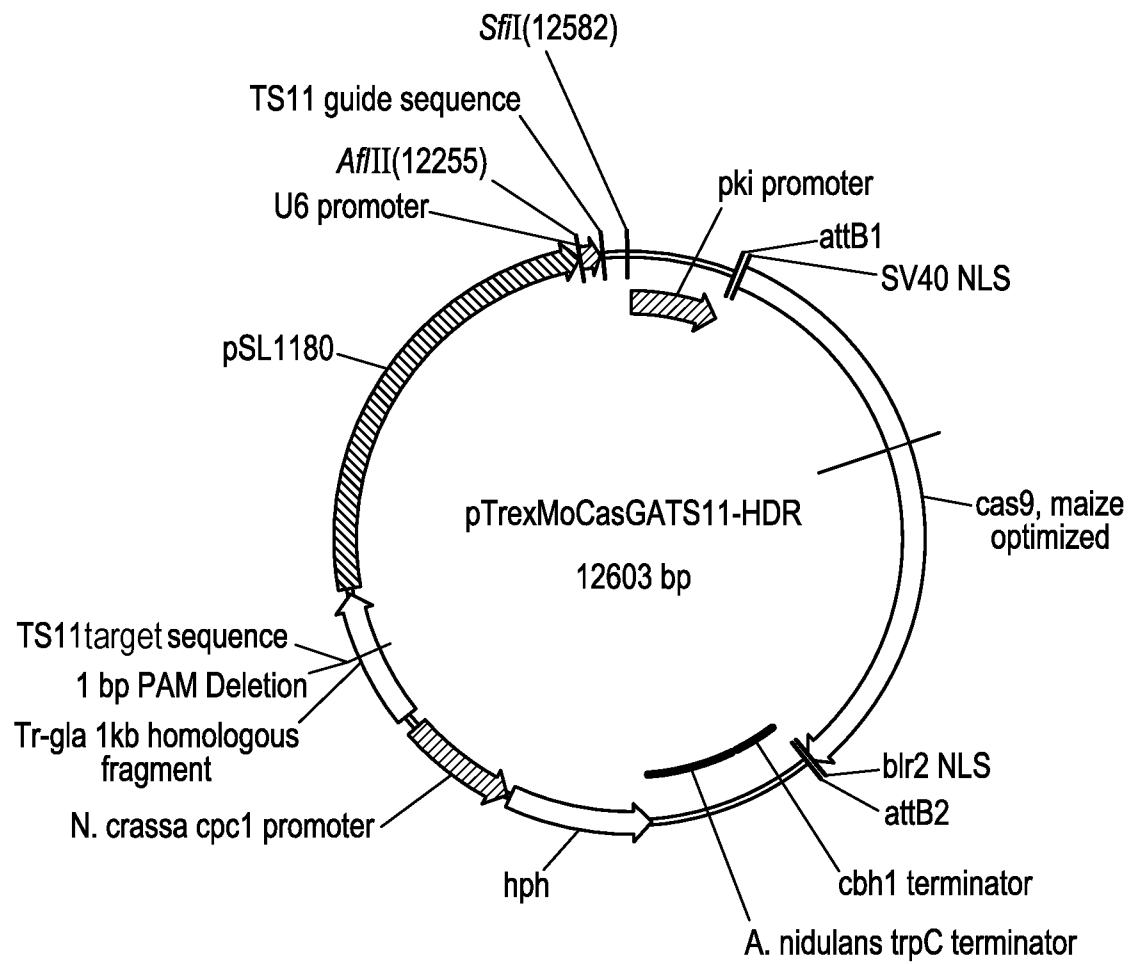


FIG. 14

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/065693

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/80 C12N9/22
ADD.

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, BIOSIS, EMBASE, MEDLINE, PAJ, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J. E. DICARLO ET AL: "Genome engineering in <i>Saccharomyces cerevisiae</i> using CRISPR-Cas systems", NUCLEIC ACIDS RESEARCH, vol. 41, no. 7, 4 March 2013 (2013-03-04), pages 4336-4343, XP55086617, ISSN: 0305-1048, DOI: 10.1093/nar/gkt135	1-34
Y	abstract page 4337, left-hand column page 4338, left-hand column ----- -/-	1-29, 31-34



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

10 March 2016

Date of mailing of the international search report

13/04/2016

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

Weinberg, Suzanna

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2015/065693

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TAKAYUKI ARAZOE ET AL: "Site-specific DNA double-strand break generated by I-SceI endonuclease enhances ectopic homologous recombination in <i>Pyricularia oryzae</i> ", FEMS MICROBIOLOGY LETTERS, vol. 352, no. 2, 26 February 2014 (2014-02-26), pages 221-229, XP055256149, GB ISSN: 0378-1097, DOI: 10.1111/1574-6968.12396	30
Y	abstract page 222, right-hand column; figure 1 page 225	1-29, 31-34
A	----- DE BOER P ET AL: "Highly efficient gene targeting in <i>Penicillium chrysogenum</i> using the bi-partite approach in DELTAlig4 or DELTaku70 mutants", FUNGAL GENETICS AND BIOLOGY, SAN DIEGO, CA, US, vol. 47, no. 10, 1 October 2010 (2010-10-01), pages 839-846, XP027264914, ISSN: 1087-1845 [retrieved on 2010-07-24] abstract	2-5,32, 33
X,P	----- WO 2015/054507 A1 (PRONUTRIA INC [US]; BASU SUBHAYU [US]; CHEN YING-JA [US]; HARVEY CAITL) 16 April 2015 (2015-04-16) abstract; examples 7-9	1-34
X,P	----- TAKAYUKI ARAZOE ET AL: "Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus", BIOTECHNOLOGY AND BIOENGINEERING., vol. 112, no. 12, 14 December 2015 (2015-12-14), pages 2543-2549, XP055256237, US ISSN: 0006-3592, DOI: 10.1002/bit.25662 the whole document	1-34
X,P	----- CHRISTINA S. NØDVIG ET AL: "A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi", PLOS ONE, vol. 10, no. 7, 15 July 2015 (2015-07-15), page e0133085, XP055256394, DOI: 10.1371/journal.pone.0133085 the whole document ----- -/--	1-34

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2015/065693

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	RUI LIU, LING CHEN, YANPING JIANG, ZHIHUA ZHOU, GEN ZOU: "Efficient genome editing in filamentous fungus <i>Trichoderma reesei</i> using the CRISPR/Cas9 system", CELL DISCOVERY, vol. 1, 15007, 12 May 2015 (2015-05-12), pages 1-11, XP002755316, DOI: 10.1038/celldisc.2015.7 the whole document	1-34
A,P	----- FULLER KEVIN K ET AL: "Development of the CRISPR/Cas9 System for Targeted Gene Disruption in <i>Aspergillus fumigatus</i> ", EUKARYOTIC CELL NOV 2015, vol. 14, no. 11, November 2015 (2015-11), pages 1073-1080, XP002755315, ISSN: 1535-9786 the whole document	1-34
A,P	----- CHI ZHANG ET AL: "Highly efficient CRISPR mutagenesis by microhomology-mediated end joining in <i>Aspergillus fumigatus</i> ", FUNGAL GENETICS AND BIOLOGY, vol. 86, 14 December 2015 (2015-12-14), pages 47-57, XP055256363, US ISSN: 1087-1845, DOI: 10.1016/j.fgb.2015.12.007 the whole document	1-34
A	----- PRASHANT MALI ET AL: "Cas9 as a versatile tool for engineering biology", NATURE METHODS, NATURE PUBLISHING GROUP, GB, vol. 10, no. 10, 1 October 2013 (2013-10-01), pages 957-963, XP002718606, ISSN: 1548-7105, DOI: 10.1038/NMETH.2649 [retrieved on 2013-09-27] the whole document	1-34
A	----- WO 2014/191521 A2 (CELLECTIS [FR]) 4 December 2014 (2014-12-04) abstract; figure 1; example 4	6
A	----- CHANDLER JULIE M ET AL: "Protein profiling of the dimorphic, pathogenic fungus, <i>Penicillium marneffei</i> ", PROTEOME SCIENCE, BIOMED CENTRAL, LONDON, GB, no. 1, 4 June 2008 (2008-06-04), page 17, XP021037177, ISSN: 1477-5956 page 2, left-hand column	1-34

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2015/065693

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2015054507	A1	16-04-2015	NONE

WO 2014191521	A2	04-12-2014	AU 2014273085 A1 10-12-2015
			CA 2913869 A1 04-12-2014
			WO 2014191521 A2 04-12-2014
