



- (51) **International Patent Classification:**
C12N 15/80 (2006.01) C12N 9/22 (2006.01)
- (21) **International Application Number:**
PCT/US2015/066195
- (22) **International Filing Date:**
16 December 2015 (16.12.2015)
- (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]; 25 Page Mill Road, Palo Alto, California 94304 (US).
- (72) **Inventors:** 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,

[Continued on next page]

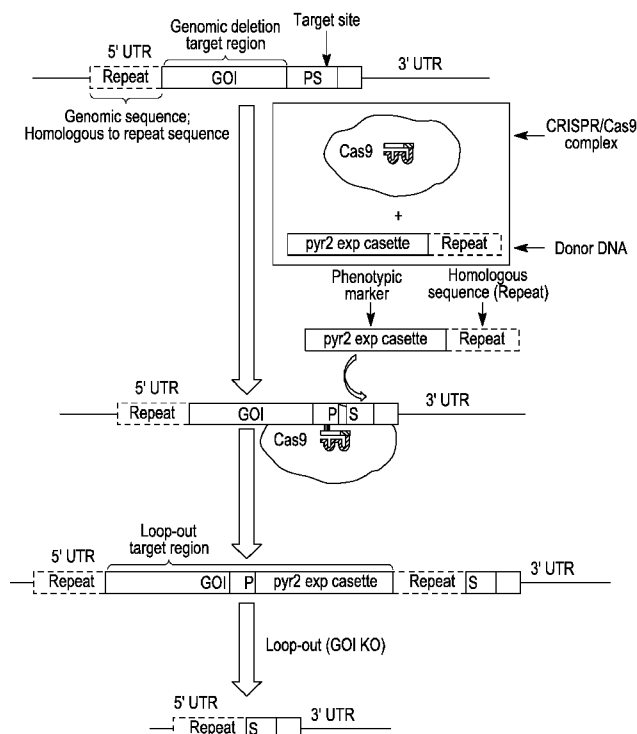
(54) **Title:** FUNGAL GENOME MODIFICATION SYSTEMS AND METHODS OF USE

FIG. 1

(57) **Abstract:** Compositions and methods are provided for genome modification at a target site in the genome of a fungal cell. Aspects of methods and compositions are drawn to a guide polynucleotide/Cas endonuclease system for promoting insertion of a donor DNA at a desired target site in a fungal host cell genome.



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.

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(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,

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. PCT/CN2014/093916, PCT/CN2014/093914, 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), is incorporated herein by reference. The sequence listing text file submitted via EFS contains the file "40532-WO-PCT-5(2015-831)_ST25.txt" created on December 11, 2015, which is 146 kilobytes in size.

BACKGROUND

[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 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 identified in a number of bacterial species, including but not limited to *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena*, and *Mycobacterium tuberculosis* (Groenen et al. (1993) *Mol. Microbiol.* 10:1057-1065; Hoe et al. (1999) *Emerg. Infect. Dis.* 5:254-263; Masepohl et al. (1996) *Biochim. Biophys. Acta* 1307:26-30; Mojica et al. (1995) *Mol. Microbiol.* 17:85-93).

[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 that relate to employing a guide RNA /Cas endonuclease system for inserting a donor DNA at a target site in the genome of a fungal cell, e.g., a filamentous fungal cell.

[07] Aspects of the present disclosure are drawn to methods for inserting a donor DNA at a target site in the genome of 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, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to introduce a double-strand break at a target site in a genomic locus of the genome of the fungal cells; and b) identifying at least one fungal cell from the population in which insertion of the donor DNA at the target site in 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 certain embodiments, the insertion has not occurred via a homologous recombination between the donor DNA and the genome of the fungal cells.

[009] In certain embodiments, the donor DNA does not comprise a sequence homologous to a genomic sequence in the genomic locus. In some embodiments, the donor DNA does not comprise a sequence that is homologous to a genomic sequence over at least 150, 200, 250, 300, 350, 400, 450, or 500 nucleotides length. In some embodiments, the donor DNA does not comprise a sequence that is homologous to a genomic sequence over at least 200 nucleotides length.

[010] In certain embodiments, the insertion of the donor DNA interrupts the expression or function of the genomic locus. In certain other embodiments, the insertion does not interrupt the expression or function of the genomic locus.

[011] In some embodiments of the method, the donor DNA comprises a gene of interest. In certain embodiments, the donor DNA comprises an expression cassette encoding a gene product of interest.

[012] In some embodiments, the gene of interest or the expression cassette encodes a protein of interest. In certain embodiments, the protein of interest 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.

[013] In certain embodiments, the gene of interest or the expression cassette encodes a phenotypic marker, e.g., a detectable marker, a selectable marker, a dominant heterologous selectable marker, a reporter gene, an auxotrophic marker, an antibiotic resistance marker, etc. (see description below). Any convenient phenotypic marker may be used.

[014] In some embodiments of the method, the donor DNA comprises, or further comprises (e.g., in the embodiments where the donor DNA comprises a gene of interest or an expression cassette), a sequence homologous to a genomic sequence in

the genomic locus (sometimes referred to herein as a “repeat sequence”), but the repeat sequence is not used for insertion of the donor DNA at the target site in the genomic locus. In some embodiments, the repeat sequence is at least about 150, 200, 300, 400, or 500 nucleotides long. In certain embodiments, the genomic sequence (i.e., the sequence to which the repeat sequence in the donor DNA is homologous) and the target site flank a genomic deletion target region. The genomic deletion target region is one defined by the user. In certain embodiments, the insertion of the donor DNA results in the genomic sequence and the sequence homologous to the genomic sequence (comprised in the donor DNA) flanking a loop-out target region comprising the genomic deletion target region. The genomic sequence and the sequence homologous to the genomic sequence are sometimes both referred to as the “repeat sequences” herein. In some embodiments where the donor DNA comprises an expression cassette encoding a phenotypic marker, the genomic sequence and the sequence homologous to the genomic sequence flank a loop-out target region that includes the genomic deletion target region and the phenotypic marker, e.g., a selectable marker. (See FIG. 1 for a schematic diagram showing examples of the donor DNA and genomic locus structural features).

[015] In certain embodiments, the method is one that results in the deletion of a genomic sequence (a genomic deletion target region) from the genome of the fungal cell. In such aspects of the present disclosure, the method further includes: c) culturing the fungal cell having the donor DNA inserted at the target site under conditions that promote or allow loop-out of the loop-out target region (i.e., the region between the genomic sequence and the repeat sequence in the donor DNA that is homologous to the genomic sequence), and d) identifying at least one fungal cell in the culture in which loop-out of the loop-out target region has occurred. This can be achieved by culturing the fungal cell under conditions in which only fungal cells that have lost the selectable marker can grow because the selectable marker is part of the loop-out target region.

[016] Another aspect of the present disclosure is drawn to a method for deleting a target region in the genome of a fungal cell, the method comprising: a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA, wherein the Cas endonuclease and guide RNA 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 fungal cells and allowing the donor DNA to be inserted at the target cite, wherein the donor DNA comprises a sequence homologous to a genomic sequence of

the fungal cells, and wherein the genomic sequence and the target site flank the target region in the fungal cell genome; b) culturing the population of fungal cells under conditions that allow homologous recombination between the genomic sequence and the sequence homologous to the genomic sequence; and c) identifying at least one fungal cell in the culture in which deletion of the target region has occurred; wherein the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells. The sequence on the donor DNA homologous to the genomic sequence of the fungal cells is sometimes referred to herein as a “repeat sequence”. In some embodiments, the repeat sequence is not used for insertion of the donor DNA at the target site in the genomic locus. In some embodiments, the repeat sequence is at least about 150, 200, 300, 400, or 500 nucleotides long.

[017] In certain embodiments of the method of deleting a target region in the genome of a fungal cell, the method further comprises, between steps a) and b), a step of identifying at least one fungal cell from the population in which insertion of the donor DNA at the target site has occurred. In some embodiments of the method, the donor DNA is not inserted at the target site via a homologous recombination between the donor DNA and the fungal cell genome.

[018] In certain embodiments of the methods described herein, the Cas endonuclease is a Type II Cas9 endonuclease or variant thereof. In some embodiments, 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. meningitidis*, *Francisella sp.*, *F. novicida*, *Pasteurella sp.*, and *P. multocida*. In specific embodiments, Cas9 endonucleases or variants thereof containing an amino acid sequence that has at least 70% identity to any one of SEQ ID NOs:1 to 7 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:1 to 7. In other embodiments, the Cas endonuclease or variant thereof is a Cpf1 endonuclease of the Type II CRISPR-Cas system.

[019] 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.

[020] 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).

[021] In certain embodiments 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:8 (encoding Cas9 from *S. pyogenes*; SEQ ID NO:1).

[022] 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:9 and SEQ ID NO:10 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.

[023] In certain embodiments, the expression cassette for the guide RNA comprises a DNA polymerase III dependent promoter functional in a *Euscomycete* or *Pezizomycete*, the promoter operably linked to the DNA encoding the guide RNA. In some instances, the promoter is derived from a *Trichoderma* U6 snRNA gene. In some 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: 11 or 12. In some embodiments, the promoter comprises the sequence of SEQ ID NO: 11 or 12. In certain embodiments, the DNA encoding the guide RNA comprises an intron sequence from a *Trichoderma* U6 snRNA gene.

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

[025] 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 a transcription factor, a repressor, protein 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.

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

[027] 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.

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

BRIEF DESCRIPTION OF THE DRAWINGS

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

[030] FIG. 1. Work flow of the application of pure SpyCas9 mediated DNA insertion for gene deletion in *T. reesei*.

[031] FIG. 2. Plasmid map of pET30a- SpyCas9.

[032] FIG. 3. Plasmid maps of pMD18T (T7-TrGA_Sth_sgR2).

[033] FIG. 4. SpyCas9 nuclease assay. Lane 1, DNA ladder; lane 2 and lane 3, SpyCas9 assay in the presence of water and TrGA_Sth_sgR2, respectively.

[034] FIG. 5. The morphology of transformants in Vogel-starch (glucose free) plate assays. Transformants 1 to 14 were picked from the plates with SpyCas9/sgrRNA treatment while transformants D1 through D5 were randomly selected from control plates. Panel A, regular Vogel agar plate. Panel B, Vogel-starch (glucose free) plate.

[035] FIGS. 6A-6C. Insertion-orientation-independent deletion cassette design (donor DNA) based on the position of target site (or protospacer, "PS") in a gene of interest (GOI). The genomic sequence is shown at the top of each of FIGS. 6A-6C, while the donor DNA is shown at the bottom of each of these figures. Donor DNA designs are shown for: PS near the C-terminus/3' end of the GOI (FIG. 6A); PS in or near the middle of the GOI (FIG. 6B); PS near the N-terminus/5' end of the GOI (FIG. 6C). UTR = untranslated; GOI = gene of interest; R1 = repeat sequence 1; R2 = repeat sequence 2; pyr2 exp cassette = expression cassette for pyr2 gene.

DETAILED DESCRIPTION

[036] The present disclosure includes compositions and methods that find use in inserting a donor DNA at a target site in the genome of 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 at the site, which thereby allows insertion of a donor DNA at the target site.

[037] 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.

[038] 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.

[039] 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.

[040] 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.

[041] 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.

[042] 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.

[043] 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.

[044] 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.

[045] 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.

[046] 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

[047] 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. The Cas endonucleases employed in the donor DNA insertion methods detailed herein are endonucleases that introduce double-strand breaks into the DNA 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 embodiments, the Cas endonuclease or variant thereof is a Cas9 endonuclease of the Type II CRISPR-Cas system. 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:1 to 7 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:1 to 7. 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.

[048] 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”.

[049] 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 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.

[050] 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.

[051] 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.

[052] 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.

[053] 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.

[054] 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, 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.

[055] 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 tracking, a modification or sequence that provides a binding site for proteins, a Locked Nucleic Acid (LNA), a 5-methyl dC nucleotide, a 2,6-Diaminopurine nucleotide, a 2'-Fluoro A nucleotide, a 2'-Fluoro U nucleotide; a 2'-O-Methyl RNA nucleotide, a phosphorothioate bond, linkage to a cholesterol molecule, linkage to a polyethylene glycol molecule, linkage to a spacer 18 molecule, a 5' to 3' covalent linkage, or any combination thereof. These modifications can result in at least one additional beneficial feature, wherein the additional beneficial feature is selected from the group of a modified or regulated stability, a subcellular targeting, tracking, a fluorescent label, a binding site for a protein or protein complex, modified binding affinity to complementary target sequence, modified resistance to cellular degradation, and increased cellular permeability.

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

[057] The terms “functional fragment”, “fragment that is functionally equivalent”, “functionally equivalent fragment”, and the like, are used interchangeably and refer to a 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.

[058] 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 with a guide polynucleotide. It is noted here

that a functional variant may have altered quantitative enzymatic activity as compared to the parent polypeptide.

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

[060] 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.

[061] 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.

[062] “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.

[063] “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.

[064] “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.

[065] 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).

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

[067] “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.

[068] 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).

[069] 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.

[070] 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.

[071] “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.

[072] 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.

[073] “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.

[074] “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.

[075] “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 diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, *Kluyveromyces lactis*, and *Yarrowia lipolytica* cell.

[076] 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*, *Corynascus*, *Chaetomium*, *Emericella*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Myceliophthora*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Hypocrea*, and *Trichoderma*.

[077] Suitable cells of filamentous fungal species include, but are not limited to, cells of *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *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 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 terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[078] 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 cleavage is desired to promote a genome modification, e.g., insertion of a donor DNA and subsequent deletion of a genomic region of interest. 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 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 that is to be deleted from the genome. 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.

[079] 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.

[080] 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.

[081] 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).

[082] 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.

[083] Intermediate and high stringency hybridization conditions are well known in the art. For example, intermediate stringency hybridizations may be carried out with an 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 $37 - 50^\circ\text{C}$. High stringency hybridization conditions may be hybridization at 65°C and 0.1X SSC (where 1X SSC = 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 $\mu\text{g/mL}$ 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 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.

[084] The phrase “substantially similar” or “substantially identical,” in the context of at least two nucleic acids or polypeptides, means that a polynucleotide or polypeptide 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.

[085] “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.

[086] The term “percentage of sequence identity” refers to the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the results by 100 to yield the percentage of sequence identity. Useful examples of percent sequence identities include, but are not limited to, 50%, 55%, 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.

[087] 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.

[088] 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.

[089] 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.

[090] 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.

[091] 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%.

[092] “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.

[093] A “mutated gene” is a gene that has been altered through human intervention. Such a “mutated gene” has a sequence that differs from the sequence of the corresponding non-mutated gene by at least one nucleotide addition, deletion, or substitution. In certain embodiments of the disclosure, the mutated gene 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.

[094] 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.

[095] 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 a target site in the genome of a fungal cell, generally in conjunction with the activity of a Cas/guide polynucleotide complex (where the guide polynucleotide defines the target site, as detailed above). In certain embodiments, the

donor DNA construct further comprises a sequence homologous to a genomic sequence in the genomic locus (also called a repeat sequence). By “homologous” is meant DNA sequences that are similar. For example, a “region homologous to a genomic sequence” that is found on the donor DNA is a region of DNA that has a similar sequence to a given “genomic sequence” in the fungal cell genome.

Collectively, the sequence homologous to a genomic sequence in the genomic locus and the genomic sequence itself are sometimes referred to herein as “the repeat sequences”. A homologous region can be of any length that is sufficient to promote or allow looping-out of the loop-out target region via homologous recombination between the repeat sequence and the homologous genomic sequence (which can be selected for under selective culture conditions). For example, the repeat sequence can comprise at least 50-55, 50-60, 50-65, 50-70, 50-75, 50-80, 50-85, 50-90, 50-95, 50-100, 50-200, 50-300, 50-400, 50-500, 50-600, 50-700, 50-800, 50-900, 50-1000, 50-1100, 50-1200, 50-1300, 50-1400, 50-1500, 50-1600, 50-1700, 50-1800, 50-1900, 50-2000, 50-2100, 50-2200, 50-2300, 50-2400, 50-2500, 50-2600, 50-2700, 50-2800, 50-2900, 50-3000, 50-3100 or more bases in length. “Sufficient homology” indicates that two polynucleotide sequences (e.g., direct repeat sequences in the donor DNA and the genome of fungal cell) have sufficient structural similarity to loop-out the sequence in between the repeat sequences, e.g., under appropriate selective culture conditions. The structural similarity includes overall length of each polynucleotide fragment, as well as the sequence similarity of the polynucleotides. Sequence similarity can be described by the percent sequence identity over the whole length of the sequences, and/or by conserved regions comprising localized similarities such as contiguous nucleotides having 100% sequence identity, and percent sequence identity over a portion of the length of the sequences.

[096] As used herein, a “genomic region” or “genomic locus” is a segment of a chromosome in the genome of a fungal cell that is present on either side of the target site (e.g., including the genomic deletion target and the genomic repeat sequence that is homologous to the repeat sequence in a donor DNA) or, alternatively, also comprises a portion of the target site. The genomic region can comprise at least 50-55, 50-60, 50-65, 50-70, 50-75, 50-80, 50-85, 50-90, 50-95, 50-100, 50-200, 50-300, 50-400, 50-500, 50-600, 50-700, 50-800, 50-900, 50-1000, 50-1100, 50-1200, 50-1300, 50-1400, 50-1500, 50-1600, 50-1700, 50-1800, 50-1900, 50-2000, 50-2100, 50-2200, 50-2300, 50-2400, 50-2500, 50-2600, 50-2700, 50-2800, 50-2900, 50-3000, 50-3100 or more bases.

[097] A “genomic deletion target” and equivalents is the sequence in the fungal genome that a user wants to delete according to aspects of the present disclosure (see FIG. 1). A “loop-out target region” and equivalents is the region between direct repeats (e.g., the genomic repeat sequence and the repeat sequence in the donor DNA that is homologous to the genomic repeat sequence) that is looped-out by homologous recombination between the direct repeats in the fungal genome. In certain embodiments, the loop-out target region includes the genomic deletion target and the selectable marker on the donor DNA inserted at the target site in the fungal genome. 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, select for, or screen 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.

[098] 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.

Methods and compositions for Modifying a Fungal Cell Genome

[0099] Methods are provided employing a guide RNA /Cas endonuclease system for inserting a donor DNA at a target site in the genome of a fungal cell, e.g., a filamentous fungal cell.

[0100] Aspects of the present disclosure include methods for donor DNA insertion at a target site in the genome of a fungal cell by transiently introducing a Cas endonuclease/guide polynucleotide complex into the cell along with a donor DNA. The Cas endonuclease/ guide polynucleotide complex is capable of introducing a double-strand break at the target site in the genome of the fungal cell

[0101] 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. 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). This set up may even be advantageous as the user can generate a 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 endonuclease into the cell, but rather only need introduce a guide polynucleotide into the Cas expressing cell.

[0102] 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:11 and functional variants thereof, e.g., SEQ ID NO:12).

[0103] 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).

[0104] 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.

[0105] This process can be exploited to provide an efficient mechanism to insert an entire donor DNA into a target site without the need for any homologous region.

[0106] DNA taken up by transformation may integrate in a stable fashion in the genome or it may be transiently maintained. In some embodiments, donor DNA stably integrated into the genome is desired but integration of Cas endonuclease expression cassette or guide RNA expression cassette is not. In such embodiments, this goal can be achieved by direct introduction of Cas endonuclease and/or guide RNA transiently or transient introduction of Cas endonuclease expression cassette and/or guide RNA expression cassette. One can select for or screen for stable transformants with regard to the integration of donor DNA (e.g., using a gene product/marker encoded by the donor DNA) and unstable transformants with regard to the integration of Cas endonuclease expression cassette or guide RNA expression cassette (e.g., loss of different gene products/markers which are encoded on the DNA constructs comprising the Cas endonuclease expression cassette or guide RNA expression cassette). In

some other embodiments, especially in methods of deleting target sequences from host genome, even donor DNA may not be desired to be stably integrated into the genome, but just needs to be integrated transiently, until homologous recombination has occurred to loop-out the target region. In such cases, one can select for or screen for unstable transformants with regard to the integration of donor DNA (e.g., loss of the gene product/marker encoded by the donor DNA).

[0107] 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. We have demonstrated that it is possible to transiently express Cas endonuclease and/or guide RNA in fungal cells/unstable transformants.

[0108] 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).

[0109] 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 insertion of a donor DNA at a target site.

[0110] 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.

[0111] 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.

[0112] 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 and a donor DNA are introduced by transfection. 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, where the donor DNA is inserted. Subsequently, spores are recovered from the heterokaryon and subjected to selection or screening to recover the target strain with a donor DNA inserted at the target site. 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.

[0113] 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 presence of a functioning NHEJ pathway, non-homologous insertion of donor 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 having desired homologous recombination sites.

[0114] 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 or allow homologous recombination with their corresponding genomic 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.

[0115] 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%, 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.

[0116] The region of homology on the donor DNA can have homology to any sequence flanking the target site. While in some embodiments the regions of homology share significant sequence homology to the genomic sequence 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.

[0117] 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).

[0118] 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). 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:8, e.g., SEQ ID NO:9, as described below).

[0119] 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.

[0120] 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 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 needs to be made to express the fused guide RNA.

[0121] 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 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 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*. For example, the filamentous fungi *T. reesei* and *A. niger* find use in aspects of the disclosed methods.

[0122] 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, or PAM. In the case of the *S. pyogenes* Cas9, the PAM has the sequence NGG (5' to 3'; 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).

[0123] The length of the target site can vary, and includes, for example, target sites that are at least 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more nucleotides in length. It is further possible that the target site can be palindromic, that is, the sequence on one strand reads the same in the opposite direction on the

complementary strand. The 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.

[0124] 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.

[0125] 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.

[0126] 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.

[0127] 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, select for, or

screen 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 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.

[0128] 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.

[0129] 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, phenoxidases, phytases, polygalacturonases, proteases, rhamno-galacturonases, ribonucleases, transferases, transport proteins, transglutaminases, xylanases, hexose oxidases, and combinations thereof.

[0130] 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.

[0131] Non-limiting examples or embodiments of compositions and methods disclosed herein are as follows:

1. A method for inserting a donor DNA at a target site in the genome of a fungal cell, the method comprising:
 - a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to introduce a double-strand break at a target site in a genomic locus of the genome of the fungal cells; and
 - b) identifying at least one fungal cell from the population in which insertion of the donor DNA at the target site in 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 insertion has not occurred via a homologous recombination between the donor DNA and the genome of the fungal cells.
3. The method of embodiment 1 or 2, wherein the donor DNA does not comprise a sequence homologous to a genomic sequence in the genomic locus.
4. The method of any preceding embodiment, wherein the insertion of the donor DNA interrupts the expression or function of the genomic locus.
5. The method of any preceding embodiment, wherein the donor DNA comprises a gene of interest.
6. The method of any preceding embodiment, wherein the donor DNA comprises an expression cassette encoding a gene product of interest.
7. The method of embodiment 6, wherein the gene product of interest is a protein of interest.

8. The method of embodiment 6, wherein the gene product of interest is a phenotypic marker.
9. The method of embodiment 8, wherein the phenotypic marker is selected from the group consisting of an auxotrophic marker, an antibiotic resistance marker, a dominant heterologous selectable marker, and a reporter gene.
10. The method of any one of embodiments 1, 2, and 4-9, wherein the donor DNA comprises a sequence homologous to a genomic sequence in the genomic locus, wherein the genomic sequence and the target site flank a genomic deletion target region, and wherein the insertion of the donor DNA results in the genomic sequence and the sequence homologous to the genomic sequence flanking a loop-out target region comprising the genomic deletion target region.
11. The method of embodiment 10, further comprising:
 - c) culturing the at least one identified fungal cell under conditions that allow loop-out of the loop-out target region, and
 - d) identifying at least one fungal cell in the culture in which loop-out of the loop-out target region has occurred.
12. A method for deleting a target region in the genome of a fungal cell, the method comprising:
 - a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA, wherein the Cas endonuclease and guide RNA 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 fungal cells and allowing the donor DNA to be inserted at the target cite, wherein the donor DNA comprises a sequence homologous to a genomic sequence of the fungal cells, and wherein the genomic sequence and the target site flank the target region in the fungal cell genome;
 - b) culturing the population of fungal cells under conditions that allow homologous recombination between the genomic sequence and the sequence homologous to the genomic sequence; and

c) identifying at least one fungal cell in the culture in which deletion of the target region has occurred;
wherein the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells.

13. The method of embodiment 12, further comprising, between steps a) and b), a step of identifying at least one fungal cell from the population in which insertion of the donor DNA at the target site has occurred.

14. The method of embodiment 12 or 13, wherein the donor DNA is not inserted at the target site via a homologous recombination between the donor DNA and the fungal cell genome.

15. The method of any preceding embodiment, wherein the Cas endonuclease is a Type II Cas9 endonuclease or variant thereof.

16. The method of embodiment 15, 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. meningitidis*, *Francisella* sp., *F. novicida*, *Pasteurella* sp., and *P. multocida*.

17. The method of embodiment 16, 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:1 to 7.

18. 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.

19. 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.

20. The method of any one of embodiments 1 to 17 and 19, wherein the introducing step comprises directly introducing the Cas endonuclease into the fungal cells.
21. The method of any one of embodiments 1 to 18 and 20, wherein the introducing step comprises directly introducing the guide RNA into the fungal cells.
22. The method of embodiment 18, wherein the expression cassette for the Cas endonuclease comprises a Cas coding sequence that is optimized for expression in the fungal cell.
23. The method of embodiment 22, wherein the Cas coding sequence is a Cas9 coding sequence comprising a polynucleotide sequence that is at least 70% identical to SEQ ID NO:8.
24. The method of any preceding embodiment, wherein the Cas endonuclease is operably linked to a nuclear localization signal.
25. The method of any preceding embodiment, wherein the fungal cell is a filamentous fungal cell.
26. The method of any preceding embodiment, wherein the fungal cell is a Eumycotina or Pezizomycotina fungal cell.
27. 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.
28. 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.

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

EXAMPLES

[0132] In the following Examples, unless otherwise stated, parts and percentages are by weight and degrees are Celsius. It should be understood that these Examples, while indicating embodiments of the disclosure, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can make various changes and modifications of the disclosure to adapt it to various usages and conditions. Such modifications are also intended to fall within the scope of the appended claims.

Example 1: Heterologous expression of CRISPR SpyCas9 in E.coli

[0133] E. coli codon-optimized Streptococcus pyogenes Cas9 (SpyCas9) 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 (FIG. 2). As indicated in the plasmid map in FIG. 2, 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:13; includes a start codon methionine), a sequence encoding an SV40 nuclear localization signal (SEQ ID NO:14), a sequence encoding the SpyCas9 (SEQ ID NO:15), and a sequence encoding the BLR nuclear localization signal (SEQ ID NO:16) all in operable linkage. This entire coding sequence is shown in SEQ ID NO:17. The amino acid sequence of the N-terminal His6 tag / thrombin / S•Tag™ / enterokinase region encoded by SEQ ID NO:13 is shown in SEQ ID NO:18 (including the methionine at position 1), the amino acid sequence of the SV40 nuclear localization signal encoded by SEQ ID NO:14 is shown in SEQ ID NO:19, the amino acid sequence of the SpyCas9 encoded by SEQ ID NO:15 is shown in SEQ ID NO:1, and the amino acid sequence of the BLR nuclear localization signal encoded by SEQ ID NO:16 is shown in SEQ ID NO:20. The amino acid sequence encoded by SEQ ID NO:17 is shown in SEQ ID NO:21.

[0134] The pET30a-SpyCas9 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 cultivated for 24 hours in a 250ml shake flask with 25 ml of the Invitrogen MagicMedia™ E.coli Expression Medium (Thermo Fisher Scientific Inc., Grand Island, NY).

Example 2: Purification of SpyCas9

[0135] For purification of SpyCas9, a combination of affinity, hydrophobic interaction and size exclusion chromatographic steps were applied. Briefly, SpyCas9 expressing E. coli cells (Rosetta2 (De3)plysS, as described above) were cultured in a 250ml shake flask with 25 ml MagicMedia™ for 24 hours and harvested by centrifugation. Cells (approximately 40 grams) were pelleted and resuspended in 400 ml 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 ultra-sonicator (35% power, 20 min, 2s on/3s off) (SCIENT2-II D, Ningbo Scientz Biotechnology Co., LTD). The lysate was cleared by centrifugation at 20000g for 40 min.

[0136] Approximately 400 ml of clarified lysate was incubated with 5 ml Ni-NTA resin (GE Healthcare) overnight at 4° C with shaking at 30 rpm/min using 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 being washed extensively with equilibration buffer (20 mM HEPES, pH 7.5, 300 mM NaCl, 0.1% Triton X-100) followed by wash buffer (25 mM imidazole in equilibration buffer), the target protein was eluted with 250 mM imidazole in equilibration buffer.

[0137] To the active fraction collected from the affinity step, ammonium sulfate was added to a final concentration of 0.8 M and loaded onto a 20 ml phenyl-Sepharose HP column (GE Healthcare). The column was eluted with a gradient of 0.8 M to 0.0 M ammonium sulfate in 50 mM HEPES buffer pH 7.5 and the flow through was collected.

[0138] 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 fraction with the highest purity were pooled and concentrated via

Amicon 30 KDa membrane filter (Millipore). The final protein sample was stored at -20°C freezer in the 40% glycerol until use.

Example 3: In vitro DNA Cleavage assay

Preparation of substrate DNA fragment for in vitro SpyCas9 DNA cleavage assays

[0139] Genomic DNA was extracted from a *Trichoderma reesei* strain derived from RL-P37 and having the cellobiohydrolase 1, cellobiohydrolase 2, endoglucanase 1, and endoglucanase 2 genes deleted ($\Delta cbh1$, $\Delta cbh2$, $\Delta eg1$, and $\Delta eg2$ strain; also called “quad-delete strain”; see WO 92/06184 and WO 05/001036)) using the ZF Fungal/Bacterial DNA miniprep kit from Zymo (Cat No. D6005). With 1 ng of extracted genomic DNA, DNA fragment containing the *Trichoderma reesei* glucoamylase (TrGA) gene (Gene ID: 18483895) and its partial 5'-UTR (SEQ ID NO:22) 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:23) and 5'-ggcagactacaagtctactagtactac-3' (SEQ ID NO:24). PCR products were purified and concentrated with the DNA Clean & ConcentratorTM-5 kit from Zymo (Cat No. D4013(50)), and its DNA concentration was determined with NanoDropTM (Thermo Fisher).

[0140] SEQ ID NO:22 (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. The selected VT domain, TrGA_*Sth_sgR2*, is shown in bold (SEQ ID NO:25) and the 500bp fragment applied for further loop-out experiment is shown in underlined (SEQ ID NO:26).

gactgtctccaccatgtaattttccctgcgactccatataacgccggatcgtaaaattttctctttctttccttcttcaacaa
acaacggatctgtgctttgcgggtcccctgcgttcacgcgtcagggctgactgctctgcagctcgataactccatggagccat
caacttgctatggtgtcaatcatcctatcgacaggtccaagaacaagccggcctccggctgcctcattcgctgtcgcaaga
cggcttgagtgttggtgaggattcgggggccccatattccaaccctttttccaaggccgctcggccggtgaggttgag
gaaaaccatgggtgcctacatattatcgatgctggtgttgtagtagcaatgttgcggtggcagtttgagccgagcctcgt
cttgggcttctgacccaggcaacgccatctgactagctgcgccgaaggaaggatgattcattgtacgacgccagtcattg
gaatcttcaagtaaaagccccgacgaaccgacatgtcagatatcagaattctcctggctggtggggttggtggagactgc
ttacggagtcgatgcctcgtgactgtcatggccgctccagcctcctgggactctgtccgatattatgacacgagtaaagcc
tgcatgatgtcagtttgctgcgtctcatgtcgagaacaacacacacctgggtgctacataggcaatactacctcgtagcttcaaa

gttgactgttttgcttgatgtctttgatcatgccatccatcccctgtcttgacgtgcatgtggatctctacgtccagacggggag
aaagcttgtctgtgataaagtacgatgatgcattgatgcctgtggctacggccctttatccccatcgatcatctctatatta
atccaggagactctcctcctggcatgggtgagtacaagtgacgaggacatgtagaagcagagccacgcaacgtcttga
catctgtacctattttgggccaaaaatcgagacccaccagctcgtcctaccttacatgtgaagatcttagcccacaatcctac
tgctttactagtattactgcacagctgtcatcacgagtcctcgggtgcttgtaaaccagctcagctcctgagcacatgcagt
aacgccgactcggcgtcatttcgccacacccaatttgacctgagggatgctggaagctgctgagcagatcccgttaccg
attcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcc
cggcattgtcgatgcagagatgcagatttcagcgggagattctagggtagggcgactactactactaataccacctagtc
gtatgtatctagcaccggaggctaggcggtagtgagcgggaacctggctcattccatcgcaaccaggatcccgcacttcgt
tgcgcttctgccccacggggcgggagttggcagaggcagaatgcggagcagccccctgtctgccctggccggggcct
gttgaagcaagcagacgagagcagagcgggtgagaagcgggtggtgacgcttgacggtacgaagacgagcagagaat
cccgttaagccgaggctgggctccccccccgcgtcatcatatgccatcctgctctccagcccactcgtctccctgcctcgt
cgctccccctccccgattagctgcgcatgttctcctgacagcgtgactaatgacgcttgccagcccattcgctga
cgcatcccggcatctgagctagctcgtcacgctggcaatcttggcccaggcagagcagcaagacggcgggcatgattg
ggcgtgccctggcgggcatcagctggccatccgctgccacccgagaccgcatcaccgacttgcggatctctccgagc
agcaggaggctgatcctggccggcgagacgattgaaaagggctgccggggccggagcaggacagcggcgagagc
gagcagagagagaggaaaagaagaaggctgactgtcttattttcagccagccccggctcaacagaagcagaggagaa
ggcgaacgacgtcaacgacgacgacgacgacgaagacgggtgaagtccgttagttgaagatccttgccgtcaca
caccatctcgtggatattgtttccccctgccgttgcttgccacctgttccctcttctcccccttcttctcattccgagcgt
actggttctactccgcagccttcgggttgcccttctcttctgacacattgcaccgcccgtcgcggcacttggggccggag
aattcggcccttctgcagcattttggccctcagttccccatggggacgggtccacacttctctcttggccctgcagacctttgt
cgtcgggtccgagtcggaagaagctcagcttgagcgttgagtagcatctacgcgcgaatcactggacaaagtgcgcaa
gacgaagccgtcgtcgtcgtcgtcgtgttactgcgacaggcgctccgactgggggcatcggcataataaaaagat
gcccgccttcgcatggacctggccatgagccactcggcacgtctctctctcaacgcttctctcacacatcctcctcat
tccgcccacATGCACGTCCTGTCGACTGCGGTGCTGCTCGGCTCCGTTGCCGTTCAA
AAGGTCCTGGGAAGACCAGGATCAAGCGGTCTGTCCGACGTCACCAAGAGGTCT
GTTGACGACTTCATCAGCACCGAGACGCCTATTGCACTGAACAATCTTCTTTGCAAT
GTTGGTCCTGATGGATGCCGTGCATTGGGCACATCAGCTGGTGCGGTGATTGCAT
CTCCCAGCACAATTGACCCGGAAGTGTAAAGTTGGCCTTGATGAACCATATCATATATC
GCCGAGAAGTGGACCGCGTGCTGAGACTGAGACAGACTATTACATGTGGACGCGA
GATAGCGCTCTTGTCTTCAAGAACCTCATCGACCGCTTCACCGAAACGTACGATGC
GGGCCTGCAGCGCCGCATCGAGCAGTACATTACTGCCCAGGTCACTCTCCAGGGC
CTCTCTAACCCCTCGGGCTCCCTCGCGGACGGCTCTGGTCTCGGCGAGCCCAAG
TTTGAGTTGACCCTGAAGCCTTTCACCGGCAACTGGGGTCGACCGCAGCGGGATG

GCCCAGCTCTGCGAGCCATTGCCTTGATTGGATACTCAAAGTGGCTCATCAACAAC
 AACTATCAGTCGACTGTGTCCAACGTCATCTGGCCTATTGTGCGCAACGACCTCAA
 CTATGTTGCCCAGTACTGGTCAGTGCTTGCTTGCTCTTGAATTACGTCTTTGCTTGT
 GTGTCTAATGCCTCCACCACAGGAACCAAACCGGCTTTGACCTCTGGGAAGAAGT
 CAATGGGAGCTCATTCTTTACTGTTGCCAACCAGCACCGAGGTATGAAGCAAATCC
 TCGACATTGCTGCTACTGCACATGAGCATTGTTACTGACCAGCTCTACAGCACTT
 GTCGAGGGGCGCCACTCTTGCTGCCACTCTTGCCAGTCGGGAAGCGCTTATTCAT
 CTGTTGCTCCCCAGGTTTTGTGCTTTCTCCAACGATTCTGGGTGTCGTCTGGTGGA
 TACGTGCGACTCCAACAGTATGTCTTTTCACTGTTTATATGAGATTGGCCAATACTGAT
 AGCTCGCCTCTAGTCAACACCAACGAGGGCAGGACTGGCAAGGATGTCAACTCCG
 TCCTGACTTCCATCCACACCTTCGATCCCAACCTTGGCTGTGACGCAGGCACCTTC
 CAGCCATGCAGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTGTCGACTCCTTCC
 GCTCCATCTACGGCGTGAACAAGGGCATTCTGCCGGTGCTGCCGTGCCATTGG
 CCGGTATGCAGAGGATGTGTACTACAACGGCAACCCTTGGTATCTTGCTACATTTGC
 TGCTGCCGAGCAGCTGTACGATGCCATCTACGTCTGGAAGAAGACGGGGCTCCATC
 ACGGTGACCGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGTTCTGGCGTGACG
 GCCGGGACCTACTCCAGCAGCTCTTCGACCTTTACCAACATCATCAACGCCGTCTC
 GACATACGCCGATGGCTTCCTCAGCGAGGCTGCCAAGTACGTCCCCGCCGACGGT
 TCGCTGGCCGAGCAGTTTGACCGCAACAGCGGCACTCCGCTGTCTGCGCTTCAC
 CTGACGTGGTCGTACGCCTCGTTCTTGACAGCCACGGCCCGTCGGGCTGGCATC
 GTGCCCCCCTCGTGGGCCAACAGCAGCGCTAGCACGATCCCCTCGACGTGCTCC
 GCGCGTCCGTGGTCCGATCCTACTCGCGTCCCACCGCCACGTCATTCCCTCCGT
 CGCAGACGCCCAAGCCTGGCGTGCCCTCCGGTACTCCCTACACGCCCTGCCCCT
 GCGCGACCCCAACCTCCGTGGCCGTCACCTTCCACGAGCTCGTGTGACACAGT
 TTGGCCAGACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTGGGCAACTGGAGC
 ACGAGCGCCGCGGTGGCTCTGGACGCCGTCAACTATGCCGATAACCACCCCCTGT
 GGATTGGGACGGTCAACCTCGAGGCTGGAGACGTGCTGGAGTACAAGTACATCAA
 TGTGGGCCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCAACCACACTTACACG
 GTTCCTGCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGACACCTGGCAGTCG
 TAAatgaatcggcaaggggtagtactagtagactttagtctgcc (SEQ ID NO:22)

***In vitro* transcription and SpyCas9 DNA cleavage assays**

[0141] One VT domain in TrGA gene, TrGA_Sth_sgR2 (SEQ ID NO:25), and its specific PAM, were identified for downstream in vitro assay and transformation experiments.

Oligonucleotides were inserted into the pMD18T vector by Generay, resulting in pMD18T(T7-TrGA_Sth_sgR2) (FIG. 3) (see SEQ ID NO:27 below for the T7 promoter, CER domain, and the VT domain TrGA_Sth_sgR2 sequences). DNA fragment for the in vitro transcription were amplified from pMD18T (T7-TrGA_Sth_sgR2) by PCR with 0.4 μ M of each forward and reverse primers: 5'-ctttttacggttcctggc-3' (SEQ ID NO:28) and 5'-aaaagcaccgactcgg-3' (SEQ ID NO:29). PCR products were purified and concentrated with the DNA Clean & ConcentratorTM-5 kit from Zymo (Cat No. D4013); and its DNA concentration was determined.

[0142] With the above specific PCR product as template, RNA for VT domain TrGA_Sth_sgR2 was generated by in vitro transcription using MEGAshortscriptTM T7 transcription kit from Thermo Fisher Scientific Inc. according to the manufacturer's instructions. Transcribed RNAs were purified using MEGAclearTM Transcription Clean-Up kit from Thermo Fisher Scientific Inc. The RNA concentration was measured with NanoDropTM.

[0143] SpyCas9 in vitro DNA cleavage assays were performed to confirm the activity of the synthesized single-guide RNA. To initiate the assay, 1 μ g of purified SpyCas9, 200 ng of substrate DNA fragment, and 200 ng of single-guide 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 Celsius 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 min. The reaction results were analyzed using 0.8% agarose gel, running at 140 volts for 30 min.

[0144] As shown in FIG. 4, in the presence of specific single-guide RNA, SpyCas9 can successfully cut substrate DNA fragment into the desired sizes (lane 3), confirming the function of the synthesized RNA. In the absence of the guide RNA (TrGA_Sth_sgR2), no cutting of the substrate DNA is observed (Lane 2).

[0145] SEQ ID NO:27 shows the template sequence for transcription consisting of the T7 promoter, CER domain, and the VT domain TrGA_Sth_sgR2. The VT domain was shown in uppercase, while the T7 promoter and CER domain region were shown in bold and lowercase, respectively.

taatacgcactcactatagGGTGTGGATGGAAGTCAGGAgttttagagctagaaatagcaagttaaaataa
ggctagtagccgttatcaactgaaaaagtggcaccgagtcggtgc (SEQ ID NO:27)

Example 4: In vivo SpyCas9/sgRNA uptake experiment

Protoplast preparation

[0146] For protoplast preparation, 5×10^8 spores of a quad-delete strain of *T. reesei* (described above) with an additional alpha-amylase deletion (grown on a PDA plate for 5 days at 30 °C) were inoculated 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 MgSO₄ -10 mM Na-phosphate buffer and resuspended in 15 ml lysing enzyme buffer. Lysing Enzyme from *Trichoderma harzianum* (Sigma catalog #L1412)) was dissolved in 1.2 M MgSO₄ – 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 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 5 minutes at 4000 rpm, and washed two times with 1.2M sorbitol-10mM Tris-HCl buffer. The pellet was resuspended into at least 1ml 1.2 M Sorbitol - 10 mM Tris-HCl pH 7.5 - 10 mM CaCl₂ 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 CaCl₂ and 1 part of 25% PEG6000 – 50 mM CaCl₂ – 10mM Tris-HCl until 5×10^8 per ml for the future transformation.

Preparation of deletion cassette

[0147] The TrGA deletion cassette contained a pyr2 (orotate phosphoribosyltransferase) expression cassette including the pyr2 promotor, pyr2 CDS and pyr2 terminator, followed by a 500 bp repeat sequence for the further loop out. The nucleotide sequence of the TrGA knockout cassette is depicted as SEQ ID NO:30.

[0148] SEQ ID NO:30 below shows the nucleotide sequence of the TrGA knockout cassette. The pyr2 promotor (SEQ ID NO:31), pyr2 CDS (SEQ ID NO:32), pyr2

terminator (SEQ ID NO:33), and the 500bp repeat sequence (SEQ ID NO:34) are shown in lowercase, italic, bold and underlined, respectively.

ctcgagtttataagtgacaacatgctctcaaagcgctcatggctggcacaagcctggaaagaaccaacacaaagcata
 ctgcagcaaatcagctgaattcgtcaccaattaagtgaacatcaacctgaaggcagagtatgaggccagaagcacatct
 ggatcgcagatcatggattgcccctctgttgaagatgagaatctagaaagatggcggggatgagataagagcgatgg
 gggggcacatcatctccaagacaaacaacctttgcagagtcaggcaattttcgtataagagcaggaggaggaggagtcc
 agtcatttcacagcggtaaaatcactctagacaatctcaagatgagttctgccttgggtgacttatagccatcatcacct
 agacagaagcttgtgggatactaagaccaacgtacaagctcgactgtacgctttgacttccatgtgaaaactcgatacg
 gcgcgcctctaaattttatagctcaaccactccaatccaacctctgcatccctctcactcgtcctgatctactgttcaaactcag
 agaataaggacactatccaaatccaacagaaatgggtaccacctcccagctgcctgcctacaagcaggacttctcaaact
 ccgccaatcgacggcgggcgctcctcaagtttggcagcttcgagctcaagtccaagcggatatccccctacttctcaacgcgg
 gcgaattccacacggcgcgccctcgccggcgccatcgctccgcctttgcaaagaccatcatcgaggcccaggagaag
 gccggccttagagttcgacatcgtcttcggcccgccctacaagggcacccgctgtgctccgccatcaccatcaagctcgg
 cgagctggcgccccagaacctggaccgcgtctcctactcgtttgaccgcaaggaggccaaggaccaggcgaggggcg
 gcaacatcgtcggcgcttcgctcaagggcaagagggtcctgattgtcgacgacgtcatcaccgcccggcaccgccaaga
 gggacgccattgagaagatcaccaaggaggggcgcatcgtcgccggcatcgtcgtggccctggaccgcatggagaa
 gctccccgctcgggatggcgacgactccaagcctggaccgagtgccattggcgagctgaggaaggagtagggcatcc
 ccatctttgccatcctcactctggatgacattatcgatggcatgaagggtttgctaccctgaggatatcaagaacacgga
 ggattaccgtgccaagtacaaggcgactgactgattgagggcgttcaatgtcagaagggagagaaaagactgaaaag
gtggaaagaagaggcaaatgtgttattattatttctatctcgaatcttctagatcttctcgtaaataaacaagcg
taactagctagcctccgtacaactgcttgaatttgatacccgatggagggcagttattttttttttcaagatttt
ccattcgccgttgaactcgtctcacatcgcggtgattgcccggttgcccatgtgttctcctactacccaagtcctt
cacgggttgtctcactttcttctcctttatcctccctattttttcaagtcagcgacagagcagtcatatggggatac
gtgcaactgggactcacaacaggccatcttatggcctaatagccggcggttgatccactagtcattgagcacat
gcagtaacgccgactcggcgctatttcgccacaccaatttgacctgagggatgctggaagctgctgagcagatcccgt
taccgattcatggcactactacatccatacgcaagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaa
agtcccggcattgtcgatgcagagatgcagatttcagcgggcgattctagggtagggcgactactactaataaccacct
agtcagtatgtatctagcaccggaggctagggcggttagtgacgggaacctgggtcattccatcgcaaccaggatccccgc
acttcgttgcgcttctgccccacggggcgaggattggcagaggcagaatgcggagcagccccttgtctgccttggccg
ggccctgttgaagcaagcagacgagagcagagcggttgagaagcggtggtgacgctgacggtacgaagacgagc
gagaatcccgttaagccgaggctgggc (SEQ ID NO:30)

Transformation

[0149] To initiate the uptake experiment, 20µg Spycas9 protein was mixed with 16µg sgRNA (TrGA_Sth_sgR2, described in Example 3) and 2 µl of NEB buffer#3 (New England Biolabs) and the final volume was adjusted to 20 µl. After 30 min incubation at room temperature, the SpyCas9/sgRNA premixer (or 2µl of NEB buffer#3 dissolved in 18 µl nuclease-free water as control) was mixed with 10 µg deletion cassette to form a premixer solution with a final volume of 30ul. The premixer was added to 200 µL protoplast (1×10⁸) and kept on ice for 30 min. After incubation, protoplasts were added to cooled molten sorbitol/Vogel agar (1.1 M sorbitol of minimal Vogel agar) to be as the top layer of the minimal Vogel plate (Davis et al., (1970) *Methods in Enzymology* 17A, pp. 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).

[0150] Compared to the control plates (i.e., with no SpyCas9/sgRNA premix added) that have hundreds of transformants, only 14 transformants were obtained from the protoplast with the SpyCas9/sgRNA premixer treatment. Among those 14 transformants, 13 (>90%) displayed TrGA knock-out phenotype based on the Vogel-starch (glucose free) plate assays (FIG. 5) (Colonies with TrGA knockout phenotype will grow on regular Vogel agar plate (Panel A; all clones grew) but not on glucose free Vogel-starch plate (Panel B; clones 1-4 and 6-14 from the SpyCas9/sgRNA premixer treatment did not grow, demonstrating that they are TrGA deficient)).

[0151] All 13 transformants (1 to 4, 6 to 14, FIG. 5) displaying the TrGA knock-out phenotype were transferred and grown on a new Vogel plate for the downstream loop-out experiment. After 7 days growth, all the spores were collected and diluted to desired concentrations (Table 1) and subsequently spread on the Vogel agar plate supplemented with 1.2g/L FOA to select for loop-out of the pyr2 expression cassette. The randomly selected transformants (D1 to D5, FIG. 4) from control plate were processed similarly. After 7 days growth on the Vogel-FOA plate, colonies were observed for the transformants with SpyCas9/sgRNA treatment, but none were seen for transformants from the controls (no SpyCas9/sgRNA treatment; see Table 1). This indicates that the pyr2 expression cassette was looped-out via a recombination event between the repeat sequence present in the genome and in the TrGA knockout cassette (SEQ ID NO:34) rather than merely spontaneous mutation of the pyr2

expression cassette. If spontaneous mutations were the underlying cause, both experimental and control samples would have FOA resistant colonies.

Table 1. Results of loop-out experiment using Vogel-FOA agar plate

Starting spore concentration (cells/mL)	Colony number	
	With SpyCas9/sgRNA treatment	Control
10^8	>100	0
10^7	41	0
10^6	12	0
10^5	2	0

Loop-out Strain verification

[0152] 32 colonies from the Vogel-FOA plates were randomly selected and subjected to PCR confirmation with 0.4 μ M of each of forward and reverse primers: 5'-gggtgttgtagtagcaatg -3' (SEQ NO:35) and 5'-ggcagactacaagtctactagtactac-3' (SEQ ID NO:36). After sequencing each PCR product, 3 colonies displaying the expected loop-out sequences (SEQ ID NO:37) were confirmed, demonstrating the success of target gene deletion in *T. reesei* using the combination of SpyCas9, specific sgRNA and deletion cassette.

[0153] SEQ ID NO:37 shows the expected nucleotide sequences of the PCR product of loop-out strains. The upstream and downstream UTR sequences are shown in lowercase (SEQ ID NO:38 and 39, respectively) while the partial TrGA ORF fragment is shown in uppercase (SEQ ID NO:40). The 500bp fragment retained after the loop-out experiment is underlined (SEQ ID NO:41), which is identical to the repeat sequence present in the genome and in the TrGA knockout cassette (SEQ ID NO:34).

gggtgttgtagtagcaatgtttgcggtggcagtttgagccgagcctcgtcttgggcttctgacccaggcaacgccatctgac
tagctgcgccgaaggaaggatgattcattgtacgacgccagtcgaatggaatctcaagtaaaagcccgacgaaccgac
catgtcagatatacagaattctcctggctgggtggggtggtggagactgcttacggagtcgatgcctcgtgactgtcatggcc
gcgtccagcctcctgggactctgtccgatattatgacacgagtaaaagcctgcatgatgtcagtttgctgcgtctcatgtcgag
aacaacacacctggtgctacataggcaatactacctcgtagcttcaaagttgactgttttgcttgatgtctttgatcatgccca
tccatcccttgcttgcagtgcatgtggatctctacgtccagacggggagaaagcttgctgtgataaagtacgatgatgcatt
gatgcctgtggctacggccctttatcccatcgtcatgcatctctatattaatccaggagactctcctcctggcatgggtgagt
acaagtgacgaggacatgtagaagcagagccacgcaacgtcttgacatctgtacctatttgggccccaaaatcgagacc
caccagctcgtcctaccttacatgtgaagatcttagcccacaatcctactgttttactagtattactgcacagctgtcatcacg
agtcctcgggtgctgtgaaaccagctcagctcctgagcacatgcagtaacgccgactcggcgctcatttcgccacaccca

atttgacctgagggatgctggaagctgctgagcagatcccgttaccgattcatggcactactacatccatacgcagcaa
acatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcccggcattgtcgatgcagagatgcagatttcag
cgggcgattctagggtagggcgactactactaataccacctagtcagtatgtatctagcaccggaggctaggcgggta
gtggacgggaacctggtcattccatcgcaaccaggatcccgcacttcgttgcgcttctgccccacggggcgggagttgg
cagaggcagaatgcggagcagccccctgtctgccctggccggggcctgttgaagcaagcagacgagagcagagcgg
ttgagaagcgggtggtgacgcttgacgggtacgaagacgagcagagaatcccgttaagccgagggtgggcTGACTTC
 CATCCACACCTTCGATCCCAACCTTGGCTGTGACGCAGGCACCTTCCAGCCATGC
 AGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTGTGCGACTCCTTCCGCTCCATCTA
 CGGCGTGAACAAGGGCATTCTGCGCGTGCTGCCGTGCGCATTGGCCGGTATGC
 AGAGGATGTGTACTACAACGGCAACCCTTGGTATCTTGCTACATTTGCTGCTGCCG
 AGCAGCTGTACGATGCCATCTACGTCTGGAAGAAGACGGGCTCCATCACGGTGAC
 CGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGTTCTGCGTGACGGCCGGGAC
 CTACTCCAGCAGCTCTTCGACCTTTACCAACATCATCAACGCCGTCTCGACATACG
 CCGATGGCTTCCTCAGCGAGGCTGCCAAGTACGTCCCCGCCGACGGTTCGCTGG
 CCGAGCAGTTTGACCGCAACAGCGGCACTCCGCTGTCTGCGCTTCACCTGACGT
 GGTCTGACGCCTCGTTCTTGACAGCCACGGCCCGTCGGGCTGGCATCGTGCCCC
 CCTCGTGGGCCAACAGCAGCGCTAGCACGATCCCCTCGACGTGCTCCGGCGCGT
 CCGTGGTTCGGATCCTACTCGCGTCCCACCGCCACGTCATTCCCTCCGTGCGCAGAC
 GCCCAAGCCTGGCGTGCCCTTCGGTACTCCCTACACGCCCTGCCCTGCGCGAC
 CCCAACCTCCGTGGCCGTACCTTCCACGAGCTCGTGTGACACAGTTTGGCCAG
 ACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTGGGCAACTGGAGCACGAGCGC
 CGCCGTGGCTCTGGACGCCGTCAACTATGCCGATAACCACCCCCTGTGGATTGG
 GACGGTCAACCTCGAGGCTGGAGACGTGCTGGAGTACAAGTACATCAATGTGGG
 CCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCAACCACACTTACACGGTTCCT
 GCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGACACCTGGCAGTCGTAAatgaat
 cggcaaggggtagtactagtagactttagtctgcc (SEQ ID NO:37)

Additional Embodiments

[0154] It is noted here that in the Examples above, the donor DNA (SEQ ID NO:30) was designed to function in the loop-out reaction when inserted under only one orientation. Considering the fact that the in vivo DNA fragment insertion could occur in either orientation, one could design donor DNAs that would function in both conditions. FIGS. 6A-6C provide three alternative donor DNA designs that would function in the loop-out reaction regardless of the orientation of its insertion at the target site (indicated as “PS”

in each of FIGS. 6A-6C). These three examples of donor DNA configuration are based on the position of target site (or protospacer, PS).

[0155] In FIG. 6A, the target site is near the 3' end of the gene of interest (GOI). As such, the donor DNA includes two different repeat sequences (R1 and R2) derived from genomic sequences that are upstream (5') of the GOI (the directions of the arrows indicate the 5' to 3' orientation of the repeat sequences). The R1 and R2 repeat sequences in the donor DNA flank the pyr2 expression cassette and are oriented in a head to head configuration, with the pyr2 expression cassette in between. (It is noted that any desired detectable/selectable marker can be employed. Also, while the general orientation of the elements shown in FIGS. 6A-6C is important, these elements need not be in the precise locations with respect to a GOI. For example, the elements can be present in a non-coding region, e.g., an enhancer element.)

[0156] In FIG. 6B, the target site is near the center of the GOI. As such, the donor DNA includes two different repeat sequences (R1 and R2) oriented at the 3' end of the donor DNA in a tail to tail configuration. The genomic R1 site is present in the upstream of the GOI and the genomic R2 sequence is downstream of the GOI.

[0157] In FIG. 6C, the target site is near the 5' end of the gene of interest (GOI). As such, the donor DNA includes two different repeat sequences (R1 and R2) derived from genomic sequences that are downstream (3') of the GOI. The R1 and R2 repeat sequences in the donor DNA flank the pyr2 expression cassette and are oriented in a tail to tail configuration, with the pyr2 expression cassette in between.

[0158] In each of the scenarios in FIGS. 6A-6C, insertion of the donor DNA in either orientation will allow for loop-out of the pyr2 expression cassette and a significant region of the targeted GOI. Specifically, insertion of the donor DNA in either orientation will generate direct repeat sequences, either R1:R1 or R2:R2, that will function to loop-out the desired region.

[0159] In the present disclosure, the application of SpyCas9 mediated DNA fragment insertion at a desired target site, followed by the downstream loop-out via a recombination event between repeat sequences, successfully deleted the TrGA gene in *T. reesei*. While the method described above uses purified SpyCas9 enzyme and in vitro synthesized sgRNA to significantly reduce their continuous function, methods that employ either recombinant DNA encoded Cas and/or guide RNAs under transient transformation conditions can also be used (i.e., where non-stable transformants are

selected for). Application of the teachings of the present disclosure enable highly efficient and sequence specific genome modification that can be employed for a wide range of desired outcomes.

[0160] 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 certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

[0161] 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 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 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

Streptococcus pyogenes Cas9, no NLS (encoded by SEQ ID NO:8)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
 EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPI
 FGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN
 SDVDKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLF
 GNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS
 DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP EKYKEIFFDQSKN
 GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
 GELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYYVGPLARGNSRFAWMTRKSEETITP
 WNFEENVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKR
 RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQ
 VSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIAMARENQTTQ
 KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELD
 INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLL
 NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
 DKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLES
 EFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
 GETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
 DWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSF EKNPIDFLE
 AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFY LASHY
 EKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKH RDKPI
 REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDL SQ
 LGGD

SEQ ID NO:2

Streptococcus thermophilus LMD-9 Cas9

MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITA
 EGRRLLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSF LVPDDKRDSKYPI
 FGNLVEEKAYHDEFPTIYHLRKYLADSTKKADLR LVYLALAHMIKYRGHFLIEGEFNSK
 NNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIF

SEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDFLAKKLY
 DAILSGFLTVDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISKTYNEVFKDDTKN
 GYAGYIDGKTNQEDFYVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHL
 QEMRAILDKQAKFYPFLAKNKERIEKILTFRIPIYYVGPLARGNSDFAWSIRKRNEKITPW
 NFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRD
 YQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLL
 NIINDKEFLDDSSNEAIIIEIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWG
 KLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNI
 KEVVKSLPGSPAIAKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQQ
 RLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDIDRL
 SNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKRKTFWYQLLKSCLIS
 QRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRVR
 TVKIITLKSTLVSQFRKDFELYKVVREINDFHHAHDAYLNAVVASALLKKYPKLEPEFVYG
 DYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKE
 SDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAK
 EYLDPPKKYGGYAGISNSFTVLVKGTEIEGAKKKITNVLEFQGISILDRINYRKDKLNFLLE
 KGYKDIELIIELPKYSLFELSDGSRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAK
 RISNTINENHRKYVENHKKEFEELFYIILEFNENYVGAKKNGKLLNSAFQSWQNHSE
 LCSSFIGPTGSEKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGL
 YETRIDLAKLGEG

SEQ ID NO:3

Streptococcus mutans UA159 Cas9

MKKPYSIGLDIGTNSVGWAVVTDDYKVPAAKMKVLGNTDKSHIEKNLLGALLFDSGNT
 AEDRRLLKRTARRRYTRRRNRILYLQEIFSEEMGKVDDSFHRLSDSFLVTEDKRGERH
 PIFGNLEEEVKYHENFPTIYHLRQYLADNPEKVDLRLVYLALAHIIKFRGHFLIEGKFDTR
 NNDVQRLFQEFLAVYDNTFENSSLQEQNQVEEILTDKISKSAAKDRVLKLPNEKSN
 GRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEEELVLLAQIGDNYAELFLSAK
 KLYDSILLSGILTVTDVGTKAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSD
 VSKDGYAGYIDGKTNQEAIFYKYLKGLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPH
 QIHLQEMRAIIRQAIFYPFLADNQDRIEKLTFRIPIYYVGPLARGKSDFAWLSRKSAD
 KITPWNFDEIVDKESAEAFINRMTNYDLYLPNQKVLPHSLLYEKFTVYNELTKVKYK
 TEQGKTAFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKV
 FNASYGTYHDLCKILDKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRENYSDLLTKEQ

VKKLERRHYTGWGRLSAELIHGIRNKESRKTILDYLIDDGNSNRNFMQLINDDALSFKE
EIAKAQVIGETDNLNQVVSDIAGSPAIAKKGILQSLKIVDELVKIMGHQPENIVVEMAREN
QFTNQGRRNSQQRLKGLTDSIKEFGSQILKEHPVENSQQLQNDRLFLYYLQNGRDMYT
GEELDIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKSYW
SKLLSAKLITQRKFDNLTKAERGGLTDDDKAGFIKRQLVETRQITKHVARILDERFNTET
DENNKIRQVKIVTLKSNLVSNFRKEFELYKVBREINDYHHAHDAYLNAVIGKALLGVYP
QLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKDEHIS
NIKKVLSYPQVNVIVKKVEEQTGGFSKESILPKGNSDKLIPRKTCKFYWDTKKYGGFDSP
IVAYSILVIADIEKGKSKKLKTVKALVGVTIMEKMTFERDPVAFLEKGYRNVQEENIIL
PKYSLFKLENGRKRLASARELQKGNEIVLPNHLGTLTYHAKNIHKVDEPKHLDYVDKH
KDEFKELLDVVSNSFSKKYTLAEGNLEKIKELYAQNNGEDLKELASSFINLLTFTAIGAPA
TFKFFDKNIDRKRYTSTTEILNATLIHQSI TGLYETRIDLNKLGGD

SEQ ID NO:4

Campylobacter jejuni Cas9

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLAR
RKARLNHLKHLIANEFKLNIEDYQSFDESLAKAYKGSISPYLEFRALNELLSKQDFA
RVILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFENS
KEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFEEVLSVAFYKRALK
DFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLNKNTGILYTKDDLNALLNE
VLKNGTLTYKQTKLLGLSDDYEFKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAK
DITLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKALKLVTPLMLEGKKYDEAC
NELNLKVAINEDKKDFLPAFNETYYKDEVTPNPVVLRAIKEYRKVLNALLKKYGVHKINI
ELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEKLGLKINSKNILKLRLFKEQKEFC
AYSGEKIKISDLQDEKMLEIDHIYPYSRFSDDSYMKNVLVFTKQNQEKLNQTPFEAFGN
DSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQKNFKDRNLNDTRYIARLVNLYTKDYL
DFLPLSDDENTKLNLTQKGSKVHVEAKSGMLTSALRHTWGFSKDRNNHLHHAIDAVI
IAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQKVLDKIDEI
FVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVLKALELGKIRKVNGKIVKNGDMF
RVDIFKHKKTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCSLYK
DSLILIQTKDMQEPEFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKS
IGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK

SEQ ID NO:5

Neisseria meningitides Cas9

MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAIEVPKTGDSLAM
 ARRLARSVRRRLTRRRRAHRLLRTRRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAAL
 DRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVAGNAHALQTGDFRT
 PAELALNKFEEKESGHIRNQSRSDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGI
 ETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILEQGS
 ERPLTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEM
 KAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEAL
 LKHISFDKFVQISLKALRRIVPLMEQGGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEI
 RNPVVLRLALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDR
 EKAAAKFREYFPNFVGEPSKSKDILKLRLYEQQHKGKCLYSGKEINLGRNEKGYVEIDHA
 LPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFP
 RSKKQRILLQKFDEEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQI
 TNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTID
 KETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTLEKLRTLLAEKLSSR
 PEAVHEYVTPLFVSRAPNRKMMSGQGHMETVKSARKLDEGVSVLRVPLTQLKLKDLEK
 MVNREREPKLYEALKARLEAHKDDPAKAFAPFYKYDKAGNRTQQVKAVRVEQVQKT
 GVWVRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQKGDEED
 WQLIDDSFNFKFSLHPNDLVEVITKKARMFYGFASCHRGGTGNINIRIHDLDHKIGKNGIL
 EGIGVKTALSFOKYQIDELGKEIRPCRLKKRPPVR

SEQ ID NO:6

Francisella tularensis subsp. *novicida* Cas9

MNFKILPIAIDLGVKNTGVFSAFYQKGTSLERLDNKNKGKVEYELSKDSYTLMMNNRTARR
 HQRRGIDRKQLVKRLFKLIWTEQLNLEWDKDTQQAISFLFNRRGFSFITDGYSPEYLN
 VPEQVKAILMDIFDDYNGEDDLDSYLKLATEQESKISEIYNKLMQKILEFKLMKLCTDIKD
 DKVSTKTLKEITSYEFELLADYLANYESLKTQKFSYTDKQGNLKELSYYHHDKYNIQE
 FLKRHATINDRILDITLLTDDLDIWNFNFEKFDKNEEKLQNQEDKDHIQAHLHHFVFA
 VNLIKSEMASGGRHRSQYFQEITNVLDENNHQEGYLKNFCENLHNKKYSNLSVKNLV
 NLIGNLSNLELKPLRKYFNDKIHAKADHWDEQKFTETECHWILGEWRVGVKDQDKKD
 GAKYSYKDLCLNELKQKVTAGLVDFLELDPCRTIPPYLDNNNRKPPKCQSLILNPKFL
 DNQYPNWQQYLQELKKLQSIQNYLDSFETDLKVLKSSKDQPYFVEYKSSNQQIASGQ
 RDYKDLARILQFIFDRVKASDELLNEIYFQAKKLKQKASSELEKLESSKKLDEVIAN
 QLSQILKSQHTNGIFEQGTFLHLVCKYYKQRRQRARDSRLYIMPEYRYDKKLHKYNTG

RFDDDNQLLTYCNHKPRQKRYQLLNDLAGVLQVSPNFLKDKIGSDDDLFISKWLVEHI
RGFKKACEDSLKIQKDNRGLLNHKINIARNTKGKCEKEIFNLICKIEGSEDKKGNYKHGL
AYELGVLLFGEPNEASKPEFDRKIKKFNSIYSFAQIQQIAFAERKGNANTCAVCSADNA
HRMQQIKITEPVEDNKKIILSAKAQRLPAIPTRIVDGAVKKMATILAKNIVDDNWQNIQK
VLSAKHQLHIPIITESNAFEFEPALADVKGKSLKDRRKALERISPENIFKDKNNRIKEFA
KGISAYSGANLTDGDFDGAKEELDHIIPRSHKKYGTLNDEANLICVTRGDNKNKGNRIF
CLRDADNYKLKQFETDDLEIEKKIADTIWDANKKDFKFGNYRSFINLTPQEQKAFRH
ALFLADENPIKQAVIRAINNRNRTFVNGTQRYFAEVLANNIYLRAKKENLNTDKISFDYF
GIPTIGNGRGIAEIRQLYEKVDSDIQAYAKGDKPQASYSHLIDAMLAFCIAADEHRNDGS
IGLEIDKNYSLYPLDKNTGEVFTKDIFSQIKITDNEFSDKKLVRKKAIEGFNTHRQMTRD
GIYAENYLPILHKEKELNEVRKGYTWKNSEEIKIFKGKKYDIQQLNNLVYCLKFVDKPIDI
QISTLEELRNILTTNNIAATAEYYYINLKTQKLHEYYIENYNTALGYKKYSKEMEFLRSLA
YRSERVKIKSIDDVKQVLDKDSNFIIGKITLPFKKEWQRLYREWQNTTIKDDYEFLKSFF
NVKSITKLHKKVRKDFSLPISTNEGKFLVKKRKTWDNNFIYQILNDSRSDRGTKPFIPAF
DISKNEIVEAIDSFTSKNIFWLPKNIELQKVDNKNIFAIDTSKWFEVETPSDLRDIGIATIQ
YKIDNNSRPKVRVKLDYVIDDDSKINYFMNHSLLKSRYPDKVLEILKQSTIIEFESSGFNK
TIKEMLMGKLAGIYNETSNN

SEQ ID NO:7

Pasteurella multocida Cas9

MQTTNLSYILGLDLGASVGWAVVEINENEDPIGLIDVGVRIFERAEPKTGESLALSRR
LARSTRRLIRRRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEW
GAVLLHLIKHRGYLSKRKNESQTNNKELGALLSGVAQNHQLLQSSDYRTPAELALKKF
AKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQFGNPHCKEHIQQYMTTELLMWQ
KPALSGEAILKMLGKCTHEKNEFKAAKHTYSAERFVWLTKLNNLRILEDGAERALNEEE
RQLLINHPYEKSKLTYAQVRKLLGLSEQAIFKHLRYSKENAESATFMELKAWHAIKAL
ENQGLKDTWQDLAKKPDLLDEIGTAFSLYKTDEDIQQYLTKVPNSVINALLVSLNFDK
FIELSLKSLRKILPLMEQGKRYDQACREIYGHYGEANQKTSQLLPAIPAQEIRNPVVL
TSLQARKVINAIIRQYGSPARVHIETGRELGKSFKERREIQKQQEDNRTKRESAVQKFK
ELFSDFSSEPFSKDILKFRLEYEQQHKGKCLYSGKEINIHLNEKGYVEIDHALPFSRTWD
DSFNKVLVLASENQKGNQTPYEWLQGKINSERWKNFVALVLGSQCSAAKKQRLLT
QVIDDNKFIDRNLDTRYIARFLSNYIQENLLLVGKNKKNVFTPNGQITALLRSRWGLIK
ARENNNRHHALDAIVVACATPSMQQKITRFIRFKEVHPYKIENRYEMVDQESGEIISPH
FPEPWAYFRQEVNIRVFDNHPDVLKEMLPDRPQANHGFVQPLFVSRAPTRKMSGQ

GHMETIKSAKRLAEGISVLRIPLTQLKPNLLENMVNKEREPALYAGLKARLAEFNQDPA
KAFATPFYKQGGQQVKAIRVEQVQKSGVLVRENNGVADNASIVRTDVFIKNNKFFLVPI
YTWQVAKGILPNKAIVAHKNEDEWEEMDEGAKFKFSLFPNDLVELKTKKEYFFGYIIG
LDRATGNISLKEHDGEISKGKDG VYRVGVKLALSFEKYQVDELGKNRQICRPQQRQ
PVR

SEQ ID NO:8

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

atggacaagaagtacagcatcggcctcgacatcggcaccaactcggctgggctgggccgtcatcacggacgaataaa
gggtcccgtcgaagaagttcaaggtcctcggcaatacagaccgccacagcatcaagaaaaacttgatcggcgccctcct
gttcgatagcggcgagaccgcggaggcgaccaggctcaagaggaccgccaggagacgggtacactaggcgcaaga
acaggatctgctacctgcaggagatcttcagcaacgagatggcgaagggtggacgactccttctccaccgcctggagga
atcattcctggtggaggaggacaagaagcatgagcggcacccaatcttcggcaacatcgtcgacgaggtggcctacca
cgagaagtacccgacaatctaccacctccggaagaaactgggtggacagcacagacaaggcggacctccggctcatct
acctgcccctcgcgcatatgatcaagttccgcggccacttcctcatcgaggcgacctgaacccggacaactccgacgtg
gacaagctgttcatccagctcgtgcagacgtacaatcaactgttcgaggagaaccccataaacgctagcggcggtggacg
ccaaggccatcctctcggccagggtctcgaaatcaagaaggctggagaaccttatcgcgagttgccaggcgaaaaga
agaacggcctcttcggcaaccttattgcgctcagcctcggcctgacgccgaacttcaaatacaacttcgacctcgcgag
gacgccaaagctccagctctcaaaggacacctacgacgacgacctcgacaacctcctggcccagataggagaccagta
cgcggaacctcttctcgccgccaagaacctctccgacgctatcctgctcagcgacatccttcgggtcaacaccgaaattac
caaggcaccgctgtccgccagcatgattaaacgctacgacgagcaccatcaggacctcacgctgctcaaggcactcgt
ccgccagcagctccccgagaagtacaaggagatcttctcgaccaatcaaaaaacggctacgcgggatatatcgacgg
cgggtgccagccaggaagagttctacaagttcatcaaaccaatcctggagaagatggacggcaccgaggagttgctggt
caagctcaacaggaggacctcctcaggaagcagaggaccttcgacaacggctccatcccgcatcagatccacctgg
gcgaactgcatgccatcctcgggcgccaggaggacttctacccgttctgaaggataaccgggagaagatcgagaag
atcttgacgttcgcatcccatactacgtgggcccgtggctcgcggaactcccgggtcgctggatgacctcggaagtcg
gaggagaccatcacacctggaactttgaggaggtggctgataagggcgctagcgctcagagcttcatcgagcgcatg
accaacttcgataaaaacctgcccataatgaaaaagtcctcccaagcactcgctgctctacgagtacttcacctgtataa
cgagctcaccaagggtcaaatacgtcacccgagggtcatgcggaagccggcgttctgagcggcgagcagaagaaggc
gatagtggacctcctcttcaagaccaacaggaagggtgacctgaagcaattaaaagaggactacttcaagaaaataga
gtgcttcgactccgtggagatctcggcggtggaggatcgggtcaacgcctcactcggcacgtatcacgacctcctcaagat
cattaaagacaaggacttctcgacaacgaggagaacgaggacatcctcgaggacatcgctcctaccctgacctgttc
gaggaccgcgaaatgatcgaggagagggtgaagacctacgcgcacctgttcgacgacaagggtcatgaaacagctca

agaggcgccgtacactggttggggaaggctgtcccgcaagctcattaatggcatcagggacaagcagagcggcaag
 accatcctggacttctcaagtccgacgggtcgccaaccgcaactcatgcagctcattcacgacgactcgctcacgttc
 aaggaagacatccagaaggcacaggtgagcgggcagggtagctccctccacgaacacatcgccaacctggccggct
 cgccggccattaaaaaggcatcctgcagacgggtcaaggctgctgcagagctcgtgaaggatgggcccggcacaag
 cccgaaaatatcgtcatagagatggccagggagaaccagaccacccaaaaagggcagaagaactcgcgcgagcg
 gatgaaacggatcgaggagggcattaaagagctcgggtcccagatcctgaaggagcaccctggaaaatacccag
 ctccagaatgaaaagctctaccttactacctgcagaacggccgcgacatgtacgtggaccaggagctggacattaatc
 ggctatcggactacgacgtcgaccacatcgtgccgcagtcgttctcaaggacgatagcatcgacaacaagggtgtcac
 ccggtcggataaaaaatcggggcaagagcgacaacgtgccagcgaggaggtcgtgaagaagatgaaaaactactg
 gcgccagctcctcaacgcgaaaactgatcaccagcgcaagttcgacaacctgacgaaggcggaacgcggtggcttga
 gcgaactcgataaggcgggcttcataaaaaggcagctggctgcagacgcgccagatcacgaagcatgtcgcccagatc
 ctggacagccgcatgaataactaagtagatgaaaacgacaagctgatccgggaggtgaaggtagtcacgctgaagtcc
 aagctcgtgtcggacttccgcaaggacttccagttctacaagggtccgcgagatcaacaactaccaccacgcccacgacg
 cctacctgaatgcggtggctcgggaccgccctgatcaagaagtacccgaagctggagtcggagttcgtgtacggcgacta
 caaggctacgacgtgcgcaaaatgatcgccaagtccgagcaggagatcggcaaggccacggcaaaatacttcttcta
 ctgaaacatcatgaacttctcaagaccgagatcacctcgcgaacggcgagatccgcaagcgcccgtcatcgaaac
 caacggcgagacgggagatcgtctgggataaggccgggatttcgcgacgggtccgcaagggtgctctccatgccgca
 agtcaatatcgtgaaaaagacggaggtccagacgggaggggttcagcaaggagtcctcctccgaagcgcaactccg
 acaagctcatcgcgaggaagaaggattgggacccgaaaaaataatggcggcttcgacagcccagccgtcgcatacag
 cgtcctcgtcgtggcgaagggtggagaagggaagtcacaagaagctcaagtcctgaaggagctgctcgggatcacgat
 tatggagcgggtcctccttcgagaagaacccgatcgacttctagaggccaaggatataaggaggtcaagaaggacct
 gattattaaactgccgaagtactcgtcttcgagctggaaaacggccgcaagaggatgctcgctccgcaggcgagttgc
 agaagggaacgagctcgccctcccagcaaaatacgtcaatttctgtacctcgctagccactatgaaaagctcaaggg
 cagcccggaggacaacgagcagaagcagcttctcgtggagcagcacaagcattacctggacgagatcatcgagcag
 atcagcgagttctcgaagcgggtgatcctcgccgacgcgaacctggacaagggtgctgctggcatataacaagcaccgc
 gacaaaccaatacgcgagcaggccgaaaatatcatccacctcttcacctcaccaacctcggcgctccggcagccttca
 agtacttcgacaccacgattgaccggaagcggtacacgagcagcaaggaggtgctcgatgcgacgctgatccaccag
 agcatcacagggtctatgaaacacgcatcgacctgagccagctgggcggagac

SEQ ID NO:9

Filamentous fungal cell codon optimized *Streptococcus pyogenes* Cas9-encoding gene;
 with N- and C-terminal NLS sequences

atggcaccgaagaagaagcgcaagggtgatggacaagaagtacagcatcggcctcgacatcggcaccaactcgggtg
 gctgggcccgtcatcacggacgaatataagggtcccgtcgaagaagttcaaggctcctcggaatacagaccgccacagca

tcaagaaaaacttgatcggcgccctcctgttcgatagcggcgagaccgcgaggcgaccaggctcaagaggaccgcc
aggagacggtacactaggcgcaagaacaggatctgctacctgcaggagatcttcagcaacgagatggcgaagggtgg
acgactccttctccaccgcctggaggaatcattcctggtggaggaggacaagaagcatgagcggcacccaatcttcgg
caacatcgtcgacgaggtggcctaccacgagaagtacccgacaatctaccacctccggaagaaactggtggacagca
cagacaaggcggacctccggctcatctaccttgccctcgcgcatatgatcaagttccgcgccacttctcatcgagggc
gacctgaacccggacaactccgacgtggacaagctgttcatccagctcgtgcagacgtacaatcaactgttcgaggaga
accccataaacgctagcggcgtggacgccaaggccatcctctcggccaggctctcgaaatcaagaaggctggagaac
cttatcgcgagttgccaggcgaaaagaagaacggcctcttcggcaaccttattgcgtcagcctcggcctgacgccga
acttcaaatcaaaacttcgacctcgcgaggagcgccaagctccagctctcaaaggacacctacgacgacacctcgaca
acctcctggcccagataggagaccagtacgcgacctcttctcgccgccaagaacctctccgacgctatctgtctcagc
gacatccttcgggtcaacaccgaaattaccaaggcaccgctgtccgccagcatgattaaacgctacgacgagcaccatc
aggacctcacgctgctcaaggcactcgtccgccagcagctccccgagaagtacaaggagatcttcttcgaccaatcaaa
aaacggctacgcgggatatacgacggcgggtgccagccaggaagagttctacaagttcatcaaaccaatcctggagaa
gatggacggcaccgaggagttgctggtcaagctcaacagggaggacctcctcaggaagcagaggaccttcgacaac
ggctccatcccgcatcagatccacctgggccaactgcatgccatcctcgggcgccaggaggacttctacccgttctgaa
ggataaccgggagaagatcgagaagatcttgacgttccgcatccatactacgtggggcccgctggctcgcggaactcc
cggttcgctggatgacccggaagtcggaggagaccatcacaccttggaactttgaggaggtggtcgataagggcgct
agcgtcagagcttcatcgagcgcatgaccaacttcgataaaaacctgccaatgaaaaagtcctccccaagcactcgc
tgctctacgagtacttcaccgtgtacaacgagctcaccaagggtcaaatacgtcaccgagggcatgcggaagccggcgtt
cctgagcggcgagcagaagaaggcgatagtggtacctcctcttcaagaccaacaggaagggtgacctgaagcaattaa
aagaggactacttcaagaaaatagagtgcctcgactccgtggagatctcgggcggtggaggatcggttcaacgcctcactc
ggcacgtatcacgacctcctcaagatcattaaagacaaggacttctcgacaacgaggagaacgaggacatcctcgag
gacatcgtcctcaccctgacctgttcgaggaccgcaaatgatcgaggagaggctgaagacctacgcgcacctgttcg
acgacaagggtcatgaaacagctcaagaggcgccgctacactgggtgggaaggctgtcccgcaagctcattaatggca
tcagggacaagcagagcggcaagaccatcctggacttctcaagtcggacgggttcgccaaccgcaacttcatgcagc
tcattcacgacgactcgctcacgttcaaggaagacatccagaaggcacagggtgagcgggcagggtgactccctccacg
aacacatcgccaacctggccggctcgccggccattaaaaagggtcctcgacaggtcaaggctcgtcgacgagctc
gtgaagggtgatgggcccgcacaagcccgaataatcgctcatagagatggccagggagaaccagaccacccaaaaa
gggcagaagaactcgcgcgagcggatgaaacggatcgaggagggtcattaaagagctcgggtcccagatcctgaag
gagcaccctcggtgaaaatacccgctccagaatgaaaagctctactctactacctgcagaacggccgacatgtac
gtggaccaggagctggacattaatcggtatcggaactacgacgtcgaccacatcgtgccgcagtcgttctcaaggacg
atagcatcgacaacaagggtgctcaccggctcgataaaaaatcggggcaagagcgacaacgtgccagcgaggaggt
cgtgaagaagatgaaaaactactggcgccagctcctcaacgcgaaactgatcaccagcgcaagttcgacaacctga
cgaaggcggaacgcgggtggcttgagcgaactcgataaggcggggttcataaaaaggcagctggtcgagacgcgcca

gatcacgaagcatgtcgcccagatcctggacagccgcatgaataactaagtacgatgaaaacgacaagctgatccggg
 aggtgaaggtgatcacgctgaagtccaagctcgtgtcggacttccgcaaggacttccagttctacaaggtccgcgagatc
 aacaactaccaccacgcccacgacgcctacctgaatgcggtggtcgggaccgccctgatcaagaagtacccgaagct
 ggagtcggagttcgtgtacggcgactacaaggtctacgacgtgcgcaaaatgatcgccaagtccgagcaggagatcgg
 caaggccacggcaaaatacttcttctactcgaacatcatgaacttctcaagaccgagatcacctcgcgaacggcgag
 atccgcaagcgccccgctcatcgaaaccaacggcgagacggggcgagatcgtctgggataagggccgggatttcgcgac
 ggtccgcaaggtgctctccatgccgcaagtcaatatcgtgaaaaagacggaggtccagacggggcggttcagcaagg
 agtccatcctcccgaagcgcaactccgacaagctcatcgcgaggaagaaggattgggacccgaaaaaatatggcggc
 ttcgacagcccgaccgtcgcatacagcgtcctcgtcgtggcgaaggtggagaagggcaagtcaaagaagctcaagtcc
 gtgaaggagctgctcgggatcacgattatggagcgggtcctcctcgagaagaacccgatcgacttctagaggccaagg
 gatataaggaggtcaagaaggacctgattattaaactgccgaagtactcgtcttcgagctggaaaacggccgcaaga
 ggatgctcgcctccgcaggcgagttgcagaagggcaacgagctcgccctcccgagcaaatacgtcaatttctgtacctc
 gctagccactatgaaaagctcaagggcagcccgaggacaacgagcagaagcagctcttcgtggagcagcacaag
 cattacctggacgagatcatcgagcagatcagcgagttctcgaagcgggtgatcctcgccgacgcgaacctggacaag
 gtgctgtcggcatataacaagcaccgcgacaaaccaatacgcgagcaggccgaaaatatcatccacctcttcacctca
 ccaacctcggcgctccggcagccttcaagtacttcgacaccacgattgaccggaagcgggtacacgagcacgaaggag
 gtgctcgatgcgacgctgatccaccagagcatcacagggctctatgaaacacgcatcgacctgagccagctgggcgga
 gacaagaagaagaagctcaagctctag

SEQ ID NO:10

Streptococcus pyogenes Cas9 with N- and C-terminal NLS sequences (encoded by
 SEQ ID NO:9)

MAPKKKRKVMDDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLVESFLVEE
 DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLI
 EGDLPNDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLP
 GEKKNGLFGNLIALLSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADL
 FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
 SIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRFAWMTRK
 SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
 KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
 FNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLFEDREMIEERLKTYAHLFDDKV
 MKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFK

EDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA
IKKYPKLESEFVYGDYKVYDVRKMIKSESEQEIGKATAKYFFYSNIMNFFKTEITLANGEI
RKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRN
SDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSF
EKNPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSA
YNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSITE
LYETRIDLSQLGGDKKKKLKL

SEQ ID NO:11

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

AAAAAACACTAGTAAGTACTTACTTATGTATTATTAAGTACTTTAGCTAACTTCTGCA
GTACTACCTAAGAGGCTAGGGGTAGTTTTATAGCAGACTTATAGCTATTATTTTTAT
TTAGTAAAGTGCTTTTAAAGTAAGGTCTTTTTTATAGCACTTTTTATTATTATAATAT
ATATTATATAATAATTTTAAGCCTGGAATAGTAAAGAGGCTTATATAATAATTTATAG
TAATAAAAGCTTAGCAGCTGTAATATAATTCCTAAAGAAACAGCATGAAATGGTATT
ATGTAAGAGCTATAGTCTAAAGGCACTCTGCTGGATAAAAATAGTGGCTATAAGTC
TGCTGCAAACTACCCCCAACCTCGTAGGTATATAAGTACTGTTTGATGGTAGTCT
ATC

SEQ ID NO:12

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

AATTCCTAAAGAAACAGCATGAAATGGTATTATGTAAGAGCTATAGTCTAAAGGCA
CTCTGCTGGATAAAAATAGTGGCTATAAGTCTGCTGCAAACTACCCCCAACCTCG
TAGGTATATAAGTACTGTTTGATGGTAGTCTATC

SEQ ID NO:13

N-terminal His6 tag / thrombin / S•Tag™ / enterokinase region polynucleotide sequence
(with start codon); encodes SEQ ID NO:18

atgcaccatcatcatcatcattctctggtctggtgccacgcggttctggtatgaaagaaaccgctgctgctaaattcgaacg
ccagcacatggacagcccagatctgggtaccgacgacgacgacaaggccatggcc

SEQ ID NO:14

SV40 NLS coding sequence (encodes SEQ ID NO:19)

ccaaaaaagaaacgcaagggt

SEQ ID NO:15

E. coli codon-optimized Cas9 gene (no stop codon)

atggataaaaaatacagcattggtctggatatcggaaccaacagcgtgggtgggcagtaataacagatgaatacaaa
gtgccgtcaaaaaaatttaaggttctgggaatacagatcgccacagcataaaaaagaatctgattggggcattgctgttt
gattcgggtgagacagctgaggccacgcgtctgaaacgtacagcaagaagacgttacacacgtcgtaaaaatcgattt
gctacttacaggaaatTTTTCTAACGAAATGGCCAAGGTAGATGATAGTTTCTCCATCGTCTCGAAGAATCTTTCTGGTTGAG
GAAGATAAAAAACACGAACGTCACCCTATCTTTGGCAATATCGTGGATGAAGTGGCCTATCATGAAAAATACCCTACGATT
TATCATCTTCGAAGAAGTTGGTTGATAGTACGGACAAAGCGGATCTCGTTAATCTATCTTGCCTTAGCGCACATGATCAA
ATTCTGTGTCATTTCTTAATTGAAGGTGATCTGAATCCTGATAACTCTGATGTGGACAAATTGTTATACAATTAGTGCAACCTA
TAATCAGCTGTTGAGGAAAACCCCTAATGCCTCTGGAGTTGATGCCAAAGCGATTTAAGCGGAGACTTTCTAAGTCC
CGGCGTCTGGAGAATCTGATCGCCAGTTACCAGGGGAAAAGAAAAATGGTCTGTTGGTAATCTGATTGCCCTCAGTCTGG
GGCTTACCCCGAACTCAAATCCAATTTGACCTGGCTGAGGACGCAAAGCTGCAGCTGAGCAAAGATACTTATGATGATG
ACCTCGACAATCTGCTCGCCAGATTGGTGACCAATATGCGGATCTGTTCTGGCAGCGAAGAATCTTCGGATGCTATCTTGC
TGTGGATATTCTGCGTGTAATACCGAAATCACCAAAGCGCCTCTGTCTGCAAGTATGATCAAGAGATACGACGAGCACC
ACCAGGACCTGACTCTCTTAAGGCCTGGTACGCCAACAGCTCCGGAGAAATACAAAGAAATATTCTCGACCAGTCC
AAGAATGGTTACGCGGGCTACATCGATGGTGGTGCATCACAGGAAGAGTTCTATAAATTTATAAACCAATCCTTGAGAAA
ATGGATGGCACGGAAGAGTTACTTGTTAACTTAACCGCGAAGACTTGCTAGAAAGCAACGTACATTCGACAACGGCTC
CATCCACACCAGATTCATTAGGTGAACCTCACGCCATCTGCGCAGACAAGAAGATTCTATCCCTCTTAAAAGACAATC
GGGAGAAAATCGAGAAGATCCTGACGTTCCGATTCCCTATTATGTCGGTCCCCTGGCACGTGTAATCTCGGTTTGCTGG
ATGACGCGCAAAAGTGAGGAAACCATCACCCCTTGAACCTTGAAGAAGTCGTGGATAAAGGTGCTAGCGCGCAGTCTTT
TATAGAAAGATGACGAACCTCGATAAAAACCTGCCCAACGAAAAAGTCTGCCAAGCACTCTCTTTATATGAGTACTTT
ACTGTGTACAACGAACGTACTAAAGTGAAATACGTACGGAAGGTATGCGCAAACCTGCCTTTCTAGTGGCGAGCAGAA
AAAAGCAATTGTGATCTTCTTTAAACGAATCGCAAGGTAACGTGAAAACAGCTGAAGGAAGATTATTTCAAAAAGAT
CGAATGCTTTGATTCTGTCGAGATCTCGGTTGTCGAAGATCGTTCAACGCTTCTTAGGGACCTATCATGATTGCTGAAGATA
ATAAAAGACAAAGACTTTCTCGACAATGAAGAAAATGAAGATATTCTGGAGGATATTGTTTGACCTTGACCTATTCTGAAG
ATAGAGAGATGATCGAGGAGCGCTTAAAACCTATGCCACCTGTTTGATGACAAAGTCATGAAGCAATTAAGCGCCG
CAGATATACGGGGTGGGGCCGCTGAGCCGCAAGTTGATTAACGGTATTAGAGACAAGCAGAGCGGAAAAACTATCCT

ggatttctcaaattcgacggatttgcgaaccgcaattttatgcagcttatacatgatgattcgcttacattcaaagaggatatt
 cagaaggctcaggtgtctgggcaagggtgattcactccacgaacatatagcaaatttggccggctctctcgcgattaagaa
 ggggatcctgcaaacagttaaagttgtggatgaacttgtaaagtaatgggccgccacaagccggagaatatcgtgata
 gaaatggcgcgcgagaatcaaacgacacaaaaagggtcaaaagaactcaagagagagaatgaagcgcattgagga
 ggggataaaggaacttgatctcaaattctgaaagaacatccagttgaaaacactcagctgcaaaatgaaaaattgtac
 ctgtactacctgcagaatggaagagacatgtacgtggatcaggaattggatatcaatagactctcgactatgacgtagat
 cacattgtccctcagagcttctcaaggatgattctatagataataaagtacttacgagatcggacaaaaatcgcggtaaat
 cggataacgtcccatcggaggaagtcgttaaaaagatgaaaaactattggcgtcaactgctgaacgccaaagctgatcac
 acagcgtaaagtttgataatctgactaaagccgaacgcgggtggtcttagtgaactcgataaagcaggatttataaacggc
 agttagtagaaacgcgccaaattacgaaacacgtggctcagatcctcgattctagaatgaatacaaagtacgatgaaaa
 cgataaactgatccgtgaagtaaaagtcattaccttaaaatctaaactgtgtccgatttccgcaaagatttcagttttaca
 ggtccgggaaatcaataactatcacatgcacatgatgcataatttaaatgcgggtgtaggcacggcccttattaagaaatac
 cctaaactcgaaagtgaagttgtttatggggattataaagtgtatgacgttcgcaaaatgatcgcgaaatcagaacaggaa
 atcggtaaggctaccgctaaatactttttattccaacattatgaatttttaagaccgaaataactctcgcgaaatggtgaaat
 ccgtaaacggcctcttatagaaaccaatggtgaaacggggagaaatcggttgggataaaggctgtgactttgccaccgttcg
 taaagtcctctcaatgccgcaagttaacattgtcaagaagacggaagttcaaacagggggatttccaaagaatctatcct
 gccgaagcgtaacagtataaacttatgccagaaaaaaagattgggatccaaaaaaatcaggaggctttgattccct
 accgtcgcgtatagtgtctggtggttgctaaagtcgagaaagggaaaagcaagaaattgaaatcagttaaagaactgc
 tgggtattacaattatggaaagatcgctccttgagaaaaatccgatcgacttttagaggccaaggggtataaggaagtga
 aaaaagatctcatcaaaattaccgaagtatagtcttttgagctggaaaacggcagaaaaagaatgctggcctccgcg
 ggcgagttacagaagggaaatgagctggcgtgccttccaaatatgttaattttctgtaccttgccagtcattatgagaaact
 gaagggcagccccgaagataacgaacagaaacaattatcggtggaacagcataagcactatttagatgaaattataga
 gcaaattagtgaattttctaagcgcgttatcctcgcggtatgctaatttagacaaagtactgtcagcttataataaacatcggg
 ataagccgattagagaacaggccgaaaatatcattcatttgtttacctaaccaaccttgagcaccagctgccttcaata
 tttcgataccacaattgatcgtaaacggtatacaagtacaaaagaagtcttgacgcaaccttcattcatcaatctattactg
 gattatatgagacacgcattgatctttcacagctgggcggagac

SEQ ID NO:16

BLR2 nuclear localization signal coding sequence (encodes SEQ ID NO:20)

aagaagaaaaaactgaaactg

SEQ ID NO:17

The nucleotide sequence of the SpyCas9 synthetic gene in plasmid pET30a- SpyCas9. The oligonucleotides encoding the N-terminal His6 tag, the SV40 nuclear localization

signal, and the BLR nuclear localization signal are shown in bold underline, italic underline, and underlined, respectively.

atg**caccatcatcatcatcat**tcttctggtctggtgccacgcggttctggtatgaaagaaaccgctgctgctaaattcgaac
gccagcacatggacagcccagatctgggtaccgacgacgacgacaaggccatggccccaaaaaagaaacgcaag
gttatggataaaaaatacagcattggtctggatatcggaaccaacagcggtgggtgggcagtaataacagatgaataca
aagtgcggtcaaaaaaatttaaggttctggggaatacagatcgccacagcataaaaaagaatctgattggggcattgctg
tttgattcgggtgagacagctgaggccacgcgtctgaaacgtacagcaagaagacgttacacacgtcgtaaaaatcgtat
ttgctacttacaggaaatttttctaacgaaatggccaaggtagatgatagtttctccatcgtctcgaagaatcttttctggtga
ggaagataaaaaacacgaacgtcaccttatcttggcaatatcgtggatgaagtggcctatcatgaaaaataccctacga
tttatcatcttcgcaagaagttggttgatagtagcgacaaagcggatctgcgtttaatctatcttgcgttagcgcatgatca
aatttcgtggtcatttcttaattgaaggtgatctgaatcctgataactctgatgtggacaaattgtttatacaattagtgcaaacct
ataatcagctgttcgaggaaaaccccatatgcctctggagttgatgccaaagcgatttaagcgcgagactttctaagtc
ccggcgtctggagaatctgatcgccagttaccaggggaaaagaaaaatggtctgtttggaatctgattgccctcagtcg
gggcttaccggaacttcaaatccaattttgacctggctgaggacgcaaagctgcagctgagcaaagatacttatgatgat
gacctcgacaatctgctcgcccagattggtgaccaatatgcggatctgttctggcagcgaagaatcttcggatgctatctt
gctgtcggatattctgcgtgttaataccgaaatcaccaaagcgcctctgtctgcaagtatgatcaagagatacgacgagca
ccaccaggacctgactcttctaaggcactggtacgccaacagcttccggagaaatacaaaagaaatattctcgaccagt
ccaagaatggttacgcgggtacatcgatgggtgcatcacaggaagagttctataaatttataaaccaatccttgaga
aaatggatggcacggaagagttactgttaaaacttaaccggaagacttgcttagaaagcaacgtacattcgacaacgg
ctccatcccacaccagattcattaggtgaacttcacgccatcttgcgcagacaagaagatttctatcccttcttaaagaca
atcgggagaaaaatcgagaagatcctgacgttccgcattccctattatgtcgggtcccctggcacgtggtaattctcggttgcct
ggatgacgcgcaaaagtgaggaaaccatcaccccttgaactttgaagaagtcgtggataaagggtgctagcgcgcagt
cttttatgaaagaatgacgaacttcgataaaaacttgcccaacgaaaaagtcctgccaagcactctcttttatatgagta
ctttactgtgtacaacgaactgactaaagtgaataacgttacggaaggatgcgcaaacctgccttcttagtggcgagcag
aaaaagcaattgtcgatcttctttaaacgaatcgcaaggtaactgtaaaacagctgaaggaagattattcaaaaag
atcgaatgctttgattctgtcgagatctcgggtgtcgaagatcgtttcaacgcttcccttagggacctatcatgattgctgaagat
aataaaagacaaagactttctgacaatgaagaaaatgaagatattctggaggatattgtttgaccttgaccttattcgaa
gatagagagatgatcgaggagcgcttaaaaacctatgccacctgtttgatgacaaagtcataagcaattaaagcgcc
gcagatatacggggtggggccgcttgagccgcaagttgattaacggtattagagacaagcagagcggaaaaaactatcc
tggatttctcaaatctgacggatttgcaaccgcaattttatgcagcttatacatgatgattcgcttacattcaagaggatat
tcagaaggctcaggtgtctgggcaaggtgattcactccacgaacatatagcaaatttgccgggtctcctgcgattaagaa
ggggatcctgcaaacagttaaagttgtggatgaacttgtaaagtaatgggccgccacaagccggagaatatcgtgata
gaaatggcgcgcgagaatcaaacgacacaaaaagggtcaaaagaactcaagagagagaatgaagcgcattgagga
ggggataaagggaacttgatctcaattctgaaagaacatccagttgaaaacactcagctgcaaaatgaaaaattgtac

ctgtactacctgcagaatggaagagacatgtacgtggatcaggaattggatatcaatagactctcggactatgacgtagat
cacattgtccctcagagcttctcaaggatgattctatagataataaagtacttacgagatcggacaaaaatcgcggtaa
cggataacgtcccatcggaggaagtcgttaaaaagatgaaaaactattggcgtcaactgctgaacgccaagctgatcac
acagcgtgaagtttgataatctgactaaagccgaacgcgggtgggtcttagtgaactcgataaagcaggatttataaacggc
agttagtagaaacgcgccaattacgaaacacgtggctcagatcctcgattctagaatgaatacaaagtacgatgaaaa
cgataaactgatccgtgaagtaaaagtcattacctaataaacttaaaacttggtgctcgatttccgcaaagatttcagtttaca
gggtccgggaaatcaataactatcacatgcacatgatgcataatttaaatgcgggttaggcacggcccttattaagaaatac
cctaaactcgaaagtgagttgtttatggggattataaagtgtatgacgttcgcaaaatgatcgcgaaatcagaacaggaa
atcggtaaggctaccgctaataacttttttattccaacattatgaatttttaagaccgaaataactctcgcgaatggtgaaat
ccgtaaacggcctcttatagaaaccaatggtgaaacgggagaaaatcggttgggataaaggctcgtgactttgccaccgttcg
taaagtcctctcaatgccgcaagttaacattgtcaagaagacggaagttcaaacagggggatttccaaagaatctatcct
gccgaagcgtaacagtataaacttattgccagaaaaaaagattgggatccaaaaaaatacggaggctttgattccct
accgtcgcgtatagtgtgctgggttgctaaagtcgagaaagggaagcaagaaattgaaatcagttaaagaactgc
tgggtattacaattatggaaagatcgtccttgagaaaaatccgatcgacttttagaggccaaggggtataaggaagtga
aaaaagatctcatcatcaaattaccgaagtatagtcttttgagctggaaaacggcagaaaaagaatgctggcctccgcg
ggcgagttacagaagggaatgagctggcgctgccttccaatatgttaatttctgtaccttgccagtcattatgagaaact
gaagggcagccccgaagataacgaacagaaacaattatcgtggaacagcataagcactatttagatgaaattataga
gcaaattagtgaattttctaagcgcgttatcctcgcgatgctaatttagacaaagtactgtcagcttataataaacatcggg
ataagccgattagagaacaggccgaaaatatcattcattgtttaccttaaccaaccttgagcaccagctgccttcaaata
tttcgataaccacaattgatcgtaaacggtatacaagtacaaaagaagtcttgagcgaaccttattcatcaatctattactg
gattatatgagacacgcattgatctttcacagctgggcggagacaaagaagaaaaaactgaaactg

SEQ ID NO:18

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

Mhhhhhssglvprgsgmketaaakferqhmdspdlgtdddkama

SEQ ID NO:19

SV40 NLS

PKKKRKV

SEQ NO:20

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

KKKKLKL

SEQ ID NO:21

The amino acid sequence of the SpyCas9 protein expressed from plasmid pET30a-SpyCas9. The N-terminal His6 tag, the SV40 nuclear localization signal, and the BLR nuclear localization signal are shown in bold underline, italic underline, and underlined, respectively.

mhssglvprgsgmketaaakferqhmdspdlgtdddkkamadkkysigldigtnsvgwavit
deykvpskkfkvlgntrhsikknligallfdsgetaeatrkrtrrrytrrknricylqeifsnemakvddsffhrleesflveed
kkherhpifgnivdevayhekyptiyhrkkklvdstdkadrllylalahmikfrghfliegdlnpdnsdvdklfiqlvtynqlfe
enpinasgvdakailsarlsksrrlenliaqlpgekknglfgnlialsgltpnfksnfdlaedakqlskdyddldnllaqigd
qyadflaaknlsdaillsdilrvnteitkaplsasmikrydehhqdltlkalvrqqipekykeiffdqskngyagyidggasqe
efykfikpilekmdgteellvknredllrkqrtfdngsiphqihlgelhaulrrqedfypflkdnrekiekiltfripyyvgplargnsr
fawmtrkseetitpwnfeevvdkgasaqsfiermtnfdklnpnekvlpkhslyeyftvyneltkvkyvtegmrkpafisge
qkkaivdlflktnrkvtvkqlkedyfkkiectdsveisgvedrfnaslgtyhdllkiikdkdfldneenediledivltltfedremi
eerlkyahlfddkvmkqlkrryrgwgrlsrklingirdkqsgktildflksdgfanrnmqlihddsltfkediqaqvsgqgd
slhehianlagspaikkilqtkvvdelvkmgrhkpeniviemarenqttqkgqknsrermkrieegikelgsqilkehp
ventqlqneklylylqngrdmyvdqeldinrlsdydvdhivpqsfllkddsidnkvltksdnrgksdnvpseevvkkmn
ywrqllnaklitqrkfdnltkaergglseldkagfikrqlvetrqitkhvaqildsrmtkydendklirekvitlksklvsdrkdf
qfykvreinnyhhahdaylnavvgtalikkypklesefvygdykvydvrmkiakseigkatakyffysnimnffkteitla
ngeirkrplietngetgeivwdkgdrfatvrkvlsmqpvnivkktevqtggskesilpkrrnsdkliarkkdwdpkkyggfds
ptvaysvlvvakvekgskklksvkellgitimerssfeknpidfleakgykevkdliiklpkyslfelengrkrmlasagelq
kgnelalpskyvnflylashyeklkgspedneqqlfveqhkhyldieiieqisefskrviladanldkvlsaynkhrrdkpireq
aeniihlftltnlgapaafkydttidrkrytstkevldatlihqsitglyetridlsqlggd

kkkkkl

SEQ ID NO:22

The nucleotide sequences of the substrate DNA fragment. The UTR sequences are shown in lowercase while the TrGA gene is shown in uppercase. The selected VT domain, TrGA_Sth_sgR2, is shown in bold and the 500bp fragment applied for further loop-out experiment was shown in underlined.

gactgtctccaccatgtaattttccctgcgactccatataacgccggatcgtaaatcttcttcttcttctccttctctcaacaa
acaacggatctgtgctttgcgggtcccctgcgttcacgcgtcagggcgactgctctgcagctcgataactccatggagccat
caacttgctatggtgtcaatcatcctatcgacaggtccaagaacaagccggcctccggctgcctcattcgctgtcgcaaga
cggcttgagtgttggtgaggattcgggggccccatattccaaccctttttccaaggccgctcggccggtgaggttgag
gaaaaccatgggtgcctacatattatcgatgctggtgttgtagtagcaatgttgcggtggcagttgagccgagcctcgt

cttgggcttctgacccaggcaacgccatctgactagctgcgccgaaggaaggatgattcattgtacgacgccagtcattg
gaatcttcaagtaaaagccccgacgaaccgacatgtcagatatcagaattctcctggctgggtggggttgggtggagactgc
ttacggagtcgatgcctcgtgactgtcatggccgcgtccagcctcctgggactctgtccgatattatgacacgagtaaagcc
tgcatgatgtcagtttctgcgtctcatgtcgagaacaacacacctgggtgtacataggcaatactacctcgtagcttcaaa
gttgactgttttcttggatgtcttggatcatgccatccatcccttgtcttgcagtgcattgtggatctctacgtccagacggggag
aaagcttgtctgtgataaagtagcatgatgcattgatgcctgtgggtacggccctttatccccatcgtcatgcattctatatta
atccaggagactctcctcctggcatgggtgagtacaagtgacgaggacatgtagaagcagagccacgcaacgtcttga
catctgtacctattttgggccaaaaatcgagacccaccagctcgtcctaccttacatgtgaagatcttagcccacaatcctac
tgttttactagtattactgcacagctgtcatcacgagtcctcggttgcctgtgaaacccagctcagctcctgagcacatgcagt
aacgccgactcggcgtcatttcgccacacccaatttgacctgagggatgctggaagctgctgagcagatcccgttaccg
attcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcc
cggcattgtcgatgcagagatgcagatttcagcgggcgattctagggtagggcgactactactaataccacctagtca
gtatgtatctagcaccggaggctaggcggttagtgacgggaacctgggtcattccatcgcaaccaggatcccgcacttcgt
tgcgcttctgccccacggggcgagggttggcagaggcagaatgcggagcagccccctgtctgccctggccggggcct
gttgaagcaagcagacgagagcagagcgggtgagaagcgggtggtgacgcttgacggtacgaagacgagcagagaat
cccgttaagccgaggctgggctccccccccgcatcatcatgccatcctgctctccagcccactcgtctccctgcctcgt
cgctccccctcccccgattagctgcgcatgttctcctgacagcgtgactaatgacgcgttgcagccccattcgctga
cgcatcccggcatctgagtctagctcgtcacgctggcaatcttggcccaggcagagcagcaagacggcgggcatgattg
ggcgtgcccgtggcgggcatcagctggccatccgctgccacccgagaccgcatcaccgacttgcggatctctccgagc
agcaggaggctgatcctggccggcgagacgattgaaaagggctgccggggccggagcaggacagcggcgagagc
gagcgagagagaggaaaagaagaaggtcgactgtcttattttcagccagccccggctcaacagaagcagaggagaa
ggcgaacgacgtcaacgacgacgacgacgacgaagacgggtgaagtccgttagttgaagatccttgcggtcaca
caccatctcgtggatattgttccccctgccgttgcgttgccacctgttccctcttcttcccccttctcctcattccgagcgt
actgggttctactccgcagccttcgggttgcccttcttcttctgacgacttgcaccgcccgtcgcggcacttgggccccggag
aattcgcccttctgcagcattttggccctcagttcccatggggacggtccacacttctcttggccctgcagacctttgt
cgtcgggtccgagtcggaagaagctcagctttagcgtttagtagcatctacgcgcgaatcactggacaaagtcggcaa
gacgaagccgtcgtcgctgctgctgtgttactgcgacaggcgctccgactgggggcatcggcataataaaaagat
gcccgccttcgcatggacctggccatgagccactcggcacggtctctctctcaacgcttctctcacacatcctccttcat
tccgcccacATGCACGTCCTGTGCGACTGCGGTGCTGCTCGGCTCCGTTGCCGTTCAA
AAGGTCCTGGGAAGACCAGGATCAAGCGGTCTGTCCGACGTCACCAAGAGGTCT
GTTGACGACTTCATCAGCACCGAGACGCCTATTGCACTGAACAATCTTCTTTGCAA
TGTTGGTCCTGATGGATGCCGTGCATTGGGCACATCAGCTGGTGCGGTGATTGCA
TCTCCAGCACAAATTGACCCGGACTGTAAGTTGGCCTTGATGAACCATATCATATA
TCGCCGAGAAGTGGACCGCGTGCTGAGACTGAGACAGACTATTACATGTGGACG

CGAGATAGCGCTCTTGTCTTCAAGAACCTCATCGACCGCTTCACCGAAACGTACG
ATGCGGGCCTGCAGCGCCGCATCGAGCAGTACATTACTGCCCAGGTCACTCTCCA
GGGCCTCTCTAACCCCTCGGGCTCCCTCGCGGACGGCTCTGGTCTCGGCGAGCC
CAAGTTTGAGTTGACCCTGAAGCCTTTCACCGGCAACTGGGGTTCGACCGCAGCG
GGATGGCCCAGCTCTGCGAGCCATTGCCTTGATTGGATACTCAAAGTGGCTCATC
AACAACAACCTATCAGTCGACTGTGTCCAACGTCATCTGGCCTATTGTGCGCAACGA
CCTCAACTATGTTGCCAGTACTGGTCAGTGCTTGCTTGCTCTTGAATTACGTCTT
TGCTTGTGTGTCTAATGCCTCCACCACAGGAACCAAACCGGCTTTGACCTCTGGG
AAGAAGTCAATGGGAGCTCATTCTTTACTGTTGCCAACCAGCACCGAGGTATGAA
GCAAATCCTCGACATTCGCTGCTACTGCACATGAGCATTGTTACTGACCAGCTCTA
CAGCACTTGTGAGGGCGCCACTCTTGCTGCCACTCTTGCCAGTCGGGAAGCG
CTTATTCATCTGTTGCTCCCCAGGTTTTGTGCTTTCTCCAACGATTCTGGGTGTG
TCTGGTGGATACGTGCGACTCCAACAGTATGTCTTTTCACTGTTTATATGAGATTGG
CCAATACTGATAGCTCGCCTCTAGTCAACACCAACGAGGGCAGGACTGGCAAGGA
TGTCAACTCCG**TCCTGACTTCCATCCACACC**TTCGATCCCAACCTTGGCTGTGAC
GCAGGCACCTTCCAGCCATGCAGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTG
TCGACTCCTTCCGCTCCATCTACGGCGTGAACAAGGGCATTCCCTGCCGGTGCTGC
CGTCGCCATTGGCCGGTATGCAGAGGATGTGTACTACAACGGCAACCCTTGGTAT
CTTGCTACATTTGCTGCTGCCGAGCAGCTGTACGATGCCATCTACGTCTGGAAGA
AGACGGGCTCCATCACGGTGACCGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGT
TCCTGGCGTGACGGCCGGGACCTACTCCAGCAGCTCTTCGACCTTTACCAACATC
ATCAACGCCGTCTCGACATACGCCGATGGCTTCCTCAGCGAGGCTGCCAAGTACG
TCCCCGCCGACGGTTCGCTGGCCGAGCAGTTTGACCGCAACAGCGGCACTCCGC
TGTCTGCGCTTCACCTGACGTGGTCGTACGCCTCGTTCTTGACAGCCACGGCCCG
TCGGGCTGGCATCGTGCCCCCCTCGTGGGCCAACAGCAGCGCTAGCACGATCCC
CTCGACGTGCTCCGGCGCGTCCGTGGTTCGGATCCTACTCGCGTCCCACCGCCAC
GTCATTCCCTCCGTGCGCAGACGCCCAAGCCTGGCGTGCCTTCCGGTACTCCCTAC
ACGCCCCTGCCCTGCGCGACCCCAACCTCCGTGGCCGTACCTTCCACGAGCTC
GTGTCGACACAGTTTGGCCAGACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTG
GGCAACTGGAGCACGAGCGCCGCCGTGGCTCTGGACGCCGTCAACTATGCCGAT
AACCACCCCCTGTGGATTGGGACGGTCAACCTCGAGGCTGGAGACGTGCTGGAG
TACAAGTACATCAATGTGGGCCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCA
ACCACACTTACACGGTTCCTGCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGA
CACCTGGCAGTCGTAAAtgaatcggcaaggggtagtactagtagactttagtctgcc

SEQ ID NO:23

forward primer for SEQ ID NO:22:

5'- gactgtctccaccatgtaatttttc-3'

SEQ ID NO:24

reverse primer for SEQ ID NO:22:

5'-ggcagactacaagtctactagtagtactac-3'

SEQ ID NO:25

TrGA_Sth_sgR2 VT domain

TCCTGACTTCCATCCACACC

SEQ ID NO:26

500bp fragment applied for further loop-out experiment

gagcacatgcagtaacgccgactcggcgctcatctcgccacacccaatttgacctgagggatgctggaagctgctgagc
agatcccggttaccgattcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattg
cccgcaaaagtcccggttgcgatgcagagatgcagattcagcgggctgattctagggtagggcgactactactacta
ataccacctagtcagtatgtatctagcaccggaggctaggcggttagtgacgggaacctgggtcattccatcgcaaccag
gatcccgacttcggttgcgcttctgccccacggggcgagggttggcagaggcagaatgcggagcagcccctgtctgcc
ctggccggggcctgttgaagcaagcagacgagagcagagcggttgagaagcggtggtgacgcttgacggtacgaag
acgagcgagaatcccggttaagccgaggctgggc

SEQ ID NO:27

The template sequence for in vitro transcription consisting of the T7 promoter, CER domain, and the VT domain TrGA_Sth_sgR2. The VT domain is shown in uppercase, while the T7 promoter and CER domain region are shown in bold and lowercase, respectively.

taat**acgactcactat**agGGTGTGGATGGAAGTCAGGAgttttagagctagaaatagcaaggtaaaataa
 ggctagtccgttatcaactgaaaaagtggcaccgagtcggtgc

SEQ ID NO:28

Forward

5'-ctttttacggttcctggc-3'

SEQ ID NO:29

Reverse

5'-aaaagcaccgactcgg-3'

SEQ ID NO:30

The nucleotide sequence of the TrGA knockout cassette. The *pyr2* promotor, *pyr2* CDS, *pyr2* terminator and the 500bp repeat sequence was shown in lowercase, italic underline, bold, and underlined, respectively.

ctcgagtttataagtgacaacatgctctcaaagcgctcatggctggcacaagcctggaaagaaccaacacaaagcata
 ctgcagcaaatacagctgaattcgtcaccaattaagtgaacatcaacctgaaggcagagtatgaggccagaagcacatct
 ggatcgcagatcatggattgcccctcttggtgaagatgagaatctagaaagatggcggggatgagataagagcgatgg
 gggggcacatcatctccaagacaaacaacctttgcagagtcaggcaattttcgtataagagcaggaggaggaggtcc
 agtcatttcacagcggtaaaatcactctagacaatctcaagatgagttctgccttgggtgacttatagccatcatcatacct
 agacagaagcttgtgggatactaagaccaacgtacaagctgcactgtacgctttgacttccatgtgaaaactcgatacg
 gcgcgcctctaaattttatagctcaaccactccaatccaacctctgcacccctctcactcgtcctgatctactgttcaaactcg
 agaataaggacactatccaaatccaacagaatggctaccacctcccagctgcctgcctacaagcaggacttcctcaaat
ccgccatcgacggcgcgctcctcaagtttggcagcttcgagctcaagtccaagcggatatccccctacttctcaacgcgg
gcgaattccacacggcgcgctcgcgcctcgccggcgccatcgctccgcctttgcaaagaccatcatcgaggcccaggagaa
gcccggcctagagttcgacatcgtcttcggccccggcctacaagggcatcccgcctgtgctccgccatcaccatcaagctcgg
cgagctggcgccccagaaacctggaccgcgtctcctactcgtttgaccgcaaggaggccaaggaccacggcgaggggcg
gcaacatcgtcggcgcttcgctcaagggcaagagggctctgattgtcgacgacgtcatcaccgccggcaccgccaaga
gggacgccattgagaagatcaccaaggaggggcgccatcgctcgccggcatcgctgagccctggaccgcatggagaa
gtccccgcctgcggatggcgacgactccaagcctggaccgagtgccattggcgagctgaggaaggagtagggcatcc
ccatctttgccatcctcactctggatgacattatcgatggcatgaagggctttgctaccctgaggatatcaagaacacgga
ggattaccgtgccaagtacaaggcgactgactgattgagggcgttcaatgtcagaagggagagaaagactgaaaag
gtggaaagaagaggc**aaattgtt****gttattattatttctatctgaatcttctagatcttgtcgtaaataaacaagcg**
taactagctagcctccgtacaactgcttgaattgatacccgatggagggcagttat**ttttttgtttttcaagatttt**
ccattcgccgttgaactcgtctcacatcgcggtgattgcccgggtgcccattgttctcctactaccccaagtcctt
cacgggttgtctcactttcttctcctttatcctccctat**ttttttcaagtcagcgacagagcagtcatatggggatac**
gtgcaactgggactcacaacaggccatcttatggcctaatagcggcggttgatccactagtc**aattg**agcacat
gcagtaacgccgactcggcgctatttcgccacacccaatttgacctgagggatgctggaagctgctgagcagatcccgt
taccgattcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgcccgcaaa
agtcccgccattgtcgatgcagagatgcagatttcagcggggcgattctagggtagggcgactactactaataaccacct

agtcagtatgtatctagcaccggaggctaggcggttagtggacgggaacctgggtcattccatcgcaaccaggatcccg
acttcgttgcgcttctgccccacggggcgaggattggcagaggcagaatgcggagcagccccttgtctgccctggccg
gggcctgttgaagcaagcagacgagagcagagcggttgagaagcggtggtgacgcttgacggtacgaagacgagc
gagaatcccgtaagccgaggctgggc

SEQ ID NO:31

pyr2 promotor

ctcgagtttataagtgacaacatgctctcaaagcgctcatggctggcacaagcctggaaagaaccaacacaaagcata
 ctgcagcaaatcagctgaattcgctaccaattaagtgaacatcaacctgaaggcagagtatgaggccagaagcacatct
 ggatcgcagatcatggattgcccctcttgttgaagatgagaatctagaaagatggcggggtatgagataagagcgatgg
 gggggcacatcatctccaagacaaacaacctttgcagagtcaggcaattttcgtataagagcaggaggaggaggagtc
 agtcatttcacgagcggtaaaatcactctagacaatttcaagatgagttctgccttgggtgacttatagccatcatcatacct
 agacagaagcttgtgggatactaagaccaacgtacaagctcgactgtacgctttgacttccatgtgaaaactcgatacg
 gcgcgcctctaaattttatagctcaaccactccaatccaacctctgcatccctctcactcgtcctgatctactgttcaaactcag
 agaataaggacactatccaaatccaacaga

SEQ ID NO:32

pyr2 CDS

atggctaccacctcccagctgcctgcctacaagcaggacttctcaaattccgccatcgacggcggtcctcaagtttg
 cagcttcgagctcaagtccaagcggtatccccctacttctcaacgcgggcgaattccacacggcgcgctcgccggc
 gccatcgctccgctttgcaaagaccatcatcgaggccaggagaaggccggcctagagtcgacatcgcttctggcc
 cggcctacaagggtatcccgctgtgctccgccatccatcaagctcggcgagctggcgccccagaacctggaccgcg
 tctcctactcgttgaccgcaaggaggccaaggaccacggcgagggcggaacatcgtcggcgcttcgctcaaggga
 agagggctcgtgattgtcgacgacgtcatcaccgcccgcaccgccaagaggagcgcattgagaagatcaccaaggag
 ggcggcatcgtcggcgcatcgctggtggcctggaccgcatggagaagctccccgctgcggatggcgacgactccaag
 cctggaccgagtgccattggcgagctgaggaaggagtagggcatccccatcttggccatcctcactctggatgacattatc
 gatggcatgaagggtttgctaccctgaggatatcaagaacacggaggattaccgtgccaagtacaaggcgactgact
 ga

SEQ ID NO:33

pyr2 terminator

ttgaggcgttcaatgtcagaaggagagaaagactgaaaagggtgaaagaaggagcaaattgtgtattattattattct
 atctcgaatcttctagatcttctgtaataaacaagcgtaactagctagcctccgtacaactgcttgaatttgatacccgat
 ggagggcagttatttttttttcaagattttcattcgccgttgaactcgtctcacatcgcggtgattgcccgggtgcccattg
 gttctcctactacccaagtccctcacgggtgtctcactttcttctccttattctccttatttttttcaagtcagcgacagagca
 gtcatatggggatacgtgcaactgggactcacaacaggccatcttatggcctaataagccggcggttgatccactagtc

attg

SEQ ID NO:34

500bp repeat sequence

agcacatgcagtaacgccgactcggcgtcatttcgccacacccaatttgacctgagggatgctggaagctgctgagca
gatcccgttaccgattcatggcactactacatccatacgcagcaaacatgggcttgggcttggcttctcaatgcaaaattgc
ccgcaaaagtcccggcattgtcgatgcagagatgcagatttcagcgggcgattctagggtagggcgactactactactaa
taccacctagtcatgtatgtatctagcaccggaggctaggcgggttagtgagcgggaacctggcattccatcgcaaccagg
atcccgcacttcgttgcgcttctgccccacggggcgaggagttggcagaggcagaatgcggagcagccccttgtctgccc
tgcccggggcctgttgaagcaagcagacgagagcagagcgggtgagaagcgggtggtgacgcttgacggtacgaaga
cgagcgagaatcccgtaagccgaggctgggc

SEQ ID NO:35

ggtgttggtagtagcaatg

SEQ ID NO:36

ggcagactacaagtctactagtactac

SEQ ID NO:37

The expected nucleotide sequences of the PCR product of loop-out strains. The UTR sequences are shown in lowercase, the partial TrGA ORF fragment is shown in uppercase, and the 500bp fragment retained after the loop-out (repeat sequence) is underlined.

ggtgttggtagtagcaatgtttgcggtggcagtttgagccgagcctcgtcttgggcttctgaccaggcaacgccatctgac
tagctgcgccgaaggaaggatgattcattgtacgacgccagtcgaatggaatctcaagtaaaagcccgacgaaccgac
catgtcagatatcagaattctcctggctggtggggttggttgagactgcttacggagtcgatgcctcgtgactgtcatggcc
gcgtccagcctcctgggactctgtccgatattatgacacgagtaaacgctgcagatgacgatttgcgtctcatgtcag
aacaacacacctggtgctacataggcaatactacctcgtagcttcaaagttgactgtttgcttggatgtcttgatcatgcca
tccatcccttgtctgcagtgcatgtggatctctacgtccagacggggagaaagcttgtctgtgataaagtacgatgatgcatt
gatgcctgtggctacggccctttatcccatcgtcatgcatctctatattaatccaggagactctcctcctggcatgggtgagt
acaagtgacgaggacatgtagaagcagagccacgcaacgtcttgacatctgtacctatttggggccaaaatcgagacc
caccagctcgtcctacctacatgtgaagatcttagcccacaatcctactgtttactagtattactgcacagctgtcatcacg
agtcctcgggtgtgtgtgaaaccagctcagctcctgagcacatgcagtaacgccgactcggcgtcatttcgccacaccca
atttgacctgagggatgctggaagctgctgagcagatcccgttaccgattcatggcactactacatccatacgcagcaa
acatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcccggcattgtcgatgcagagatgcagatttcag
cgggcgattctagggtagggcgactactactactaataccacctagtcagtatgtatctagcaccggaggctaggcggtta
gtggacgggaacctggtcattccatcgcaaccaggatcccgcacttcgttgcgcttctgccccacggggcgaggatgg
cagaggcagaatgcggagcagccccttgtctgcccctggccggggcctgttgaagcaagcagacgagagcagagcgg

ttgagaagcgggtggttgacgcttgacgggtacgaagacgagcgagaatcccgttaagccgagggtgggcTGACTTC
 CATCCACACCTTCGATCCCAACCTTGGCTGTGACGCAGGCACCTTCCAGCCATGC
 AGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTGTGCGACTCCTTCCGCTCCATCTA
 CGGCGTGAACAAGGGCATTCTGCGGGTGCTGCCGTGCGCATTGGCCGGTATGC
 AGAGGATGTGTACTACAACGGCAACCCTTGGTATCTTGCTACATTTGCTGCTGCCG
 AGCAGCTGTACGATGCCATCTACGTCTGGAAGAAGACGGGGCTCCATCACGGTGAC
 CGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGTTCTGCGTGACGGCCGGGAC
 CTACTIONCAGCAGCTCTTCGACCTTTACCAACATCATCAACGCCGTCTCGACATACG
 CCGATGGCTTCCTCAGCGAGGCTGCCAAGTACGTCCCCGCCGACGGTTCGCTGG
 CCGAGCAGTTTGACCGCAACAGCGGCACTCCGCTGTCTGCGCTTCACCTGACGT
 GGTCGTACGCCTCGTTCTTGACAGCCACGGCCCGTCGGGCTGGCATCGTGCCCC
 CCTCGTGGGCCAACAGCAGCGCTAGCACGATCCCCTCGACGTGCTCCGGCGCGT
 CCGTGGTCGGATCCTACTCGCGTCCCACCGCCACGTCATTCCCTCCGTGCGCAGAC
 GCCCAAGCCTGGCGTGCCCTCCGGTACTCCCTACACGCCCTGCCCTGCGCGAC
 CCCAACCTCCGTGGCCGTACCTTCCACGAGCTCGTGTGACACAGTTTGGCCAG
 ACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTGGGCAACTGGAGCACGAGCGC
 CGCCGTGGCTCTGGACGCCGTCAACTATGCCGATAACCACCCCCTGTGGATTGG
 GACGGTCAACCTCGAGGCTGGAGACGTGCTGGAGTACAAGTACATCAATGTGGG
 CCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCAACCACACTTACACGGTTCCT
 GCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGACACCTGGCAGTCGTAAtgaat
 cggcaaggggtagtactagtagactttagtctgcc

SEQ ID NO:38

Upstream UTR sequence from SEQ ID NO:37

ggtgtttgtagtagcaatgtttgcggtggcagtttagccgagcctcgtcttgggcttctgaccaggcaacgccatctgac
 tagctgcgccgaaggaaggatgattcattgtacgacgccagtcgaatggaatctcaagtaaaagcccgacgaaccgac
 catgtcagatatcagaattctcctggctggtggggtggttgagactgcttacggagtcgatgcctcgtgactgtcatggcc
 gcgtccagcctcctgggactctgtccgatattatgacacgagtaaagcctgcatgatgtcagtttctgctctcatgtcgag
 aacaacacacctggtgctacataggcaatactacctcgtagcttcaaagttgactgttttcttgatgtcttgatcatgcca
 tccatcccttcttgcagtgcatgtggatctctacgtccagacggggagaaagcttctgtgataaagtacgatgatgcatt
 gatgcctgtggctacggccctttatcccatcgtcatgcatctctatattaatccaggagactctcctcctggcatgggtgagt
 acaagtgacgaggacatgtagaagcagagccacgcaacgtcttgacatctgtacctattttgggcaaaaatcgagacc
 caccagctcgtcctaccttacatgtaagatcttagccacaatcctactgtttactagtattactgcacagctgtcatcacg
 agtcctcggttctgtgaaacccagctcagctcctgagcacatgcagtaacgccgactcggcgctcatttcgccacacca
 atttgacctgagggatgctggaagctgctgagcagatcccgttaccgattcatggcactactacatccatacgcagcaa

acatgggcttgggcttggcttctcaatgcaaaattgcccgcaaaagtcccggcattgtcgatgcagagatgcagatttcag
cgggcgattctagggtagggcgactactactactaataccacctagtcagtatgtatctagcaccggaggctaggcggtta
gtggacgggaacctggcattccatcgcaaccaggatcccgcacttcgttgcgcttctgccccacggggcgaggattgg
cagaggcagaatgcggagcagcccctgtctgcctggccggggcctgtgaagcaagcagacgagagcagagcgg
ttgagaagcgggtggtgacgcttgacggtacgaagacgagcgagaatcccgttaagccgaggctgggc

SEQ ID NO:39

Downstream UTR Sequence

Tgaatcggcaaggggtagtactagtagactttagtctgcc from SEQ ID NO:37

SEQ ID NO:40

partial TrGA ORF fragment from SEQ ID NO:37

TGACTTCCATCCACACCTTCGATCCCAACCTTGCTGTGACGCAGGCACCTTCCA
GCCATGCAGTGACAAAGCGCTCTCCAACCTCAAGGTTGTTGTGCGACTCCTTCCGC
TCCATCTACGGCGTGAACAAGGGCATTCTGCGCGTGCTGCCGTGCGCATTGGC
CGGTATGCAGAGGATGTGTACTACAACGGCAACCCTTGGTATCTTGCTACATTTGC
TGCTGCCGAGCAGCTGTACGATGCCATCTACGTCTGGAAGAAGACGGGCTCCATC
ACGGTGACCGCCACCTCCCTGGCCTTCTTCCAGGAGCTTGTTCTGCGGTGACG
GCCGGGACCTACTCCAGCAGCTCTTCGACCTTTACCAACATCATCAACGCCGTCT
CGACATACGCCGATGGCTTCCTCAGCGAGGCTGCCAAGTACGTCCCCGCCGACG
GTTGCTGGCCGAGCAGTTTGACCGCAACAGCGGCACTCCGCTGTCTGCGCTTC
ACCTGACGTGGTCGTACGCCTCGTTCTTGACAGCCACGGCCCGTCCGGGCTGGCA
TCGTGCCCCCCTCGTGGGCCAACAGCAGCGCTAGCACGATCCCCTCGACGTGCT
CCGGCGCGTCCGTGGTCGGATCCTACTCGCGTCCCACCGCCACGTCATTCCCTC
CGTCGCAGACGCCCAAGCCTGGCGTGCCTTCCGGTACTCCCTACACGCCCTGC
CCTGCGCGACCCCAACCTCCGTGGCCGTACCTTCCACGAGCTCGTGTGACAC
AGTTTGGCCAGACGGTCAAGGTGGCGGGCAACGCCGCGGCCCTGGGCAACTGG
AGCACGAGCGCCGCCGTGGCTCTGGACGCCGTCAACTATGCCGATAACCACCCC
CTGTGGATTGGGACGGTCAACCTCGAGGCTGGAGACGTGCTGGAGTACAAGTAC
ATCAATGTGGGCCAAGATGGCTCCGTGACCTGGGAGAGTGATCCCAACCACACTT
ACACGGTTCCTGCGGTGGCTTGTGTGACGCAGGTTGTCAAGGAGGACACCTGGC
AGTCGTAA

SEQ ID NO:41

The 500bp fragment retained after the loop-out from SEQ ID NO:37

agcacatgcagtaacgccgactcggcgtcatttcgccacacccaatttgacctgagggatgctggaagctgctgagca
gatcccgttaccgattcatggcactactacatccatacgcagcaaacaatgggcttgggcttggcttctcaatgcaaaattgc
ccgcaaaagtcccggcattgtcgatgcagagatgcagattcagcgggcgattctagggtagggcgactactactactaa
taccacctagtcagtatgtatctagcacccggaggctaggcggtagtggaacggaacctggcattccatcgcaaccagg
atcccgcaacttcgttgcgcttctgccccacggggcgaggaggcagagggcagaatgcggagcagcccctgtctgccc
tgccgggggcctgttgaagcaagcagacgagagcagagcgggtgagaagcgggtggtgacgctgacggtacgaaga
cgagcgagaatcccgtaagccgaggctgggc

THAT WHICH IS CLAIMED:

1. A method for inserting a donor DNA at a target site in the genome of a fungal cell, the method comprising:
 - a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA, wherein the Cas endonuclease and guide RNA are capable of forming a complex that enables the Cas endonuclease to introduce a double-strand break at a target site in a genomic locus of the genome of the fungal cells; and
 - b) identifying at least one fungal cell from the population in which insertion of the donor DNA at the target site in 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 insertion has not occurred via a homologous recombination between the donor DNA and the genome of the fungal cells.
3. The method of claim 1 or 2, wherein the donor DNA does not comprise a sequence homologous to a genomic sequence in the genomic locus.
4. The method of any preceding claim, wherein the insertion of the donor DNA interrupts the expression or function of the genomic locus.
5. The method of any preceding claim, wherein the donor DNA comprises a gene of interest.
6. The method of any preceding claim, wherein the donor DNA comprises an expression cassette encoding a gene product of interest.
7. The method of claim 6, wherein the gene product of interest is a protein of interest.
8. The method of claim 6, wherein the gene product of interest is a phenotypic marker.

9. The method of claim 8, wherein the phenotypic marker is selected from the group consisting of an auxotrophic marker, an antibiotic resistance marker, a dominant heterologous selectable marker, and a reporter gene.
10. The method of any one of claims 1, 2, and 4-9, wherein the donor DNA comprises a sequence homologous to a genomic sequence in the genomic locus, wherein the genomic sequence and the target site flank a genomic deletion target region, and wherein the insertion of the donor DNA results in the genomic sequence and the sequence homologous to the genomic sequence flanking a loop-out target region comprising the genomic deletion target region.
11. The method of claim 10, further comprising:
- c) culturing the at least one identified fungal cell under conditions that allow loop-out of the loop-out target region, and
 - d) identifying at least one fungal cell in the culture in which loop-out of the loop-out target region has occurred.
12. A method for deleting a target region in the genome of a fungal cell, the method comprising:
- a) introducing into a population of fungal cells a Cas endonuclease, a guide RNA, and a donor DNA, wherein the Cas endonuclease and guide RNA 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 fungal cells and allowing the donor DNA to be inserted at the target cite, wherein the donor DNA comprises a sequence homologous to a genomic sequence of the fungal cells, and wherein the genomic sequence and the target site flank the target region in the fungal cell genome;
 - b) culturing the population of fungal cells under conditions that allow homologous recombination between the genomic sequence and the sequence homologous to the genomic sequence; and
 - c) identifying at least one fungal cell in the culture in which deletion of the target region has occurred;
- wherein the Cas endonuclease, the guide RNA, or both are introduced transiently into the population of fungal cells.

13. The method of claim 12, further comprising, between steps a) and b), a step of identifying at least one fungal cell from the population in which insertion of the donor DNA at the target site has occurred.
14. The method of claim 12 or 13, wherein the donor DNA is not inserted at the target site via a homologous recombination between the donor DNA and the fungal cell genome.
15. The method of any preceding claim, wherein the Cas endonuclease is a Type II Cas9 endonuclease or variant thereof.
16. The method of claim 15, 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*, *Pasteurella* sp., and *P. multocida*.
17. The method of claim 16, 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:1 to 7.
18. 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.
19. 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.
20. The method of any one of claims 1 to 17 and 19, wherein the introducing step comprises directly introducing the Cas endonuclease into the fungal cells.
21. The method of any one of claims 1 to 18 and 20, wherein the introducing step comprises directly introducing the guide RNA into the fungal cells.

22. The method of claim 18, wherein the expression cassette for the Cas endonuclease comprises a Cas coding sequence that is optimized for expression in the fungal cell.
23. The method of claim 22, wherein the Cas coding sequence is a Cas9 coding sequence comprising a polynucleotide sequence that is at least 70% identical to SEQ ID NO:8.
24. The method of any preceding claim, wherein the Cas endonuclease is operably linked to a nuclear localization signal.
25. The method of any preceding claim, wherein the fungal cell is a filamentous fungal cell.
26. The method of any preceding claim, wherein the fungal cell is a Eumycotina or Pezizomycotina fungal cell.
27. 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*.
28. 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.
29. A recombinant fungal cell produced by the method of any preceding claim.

1/5

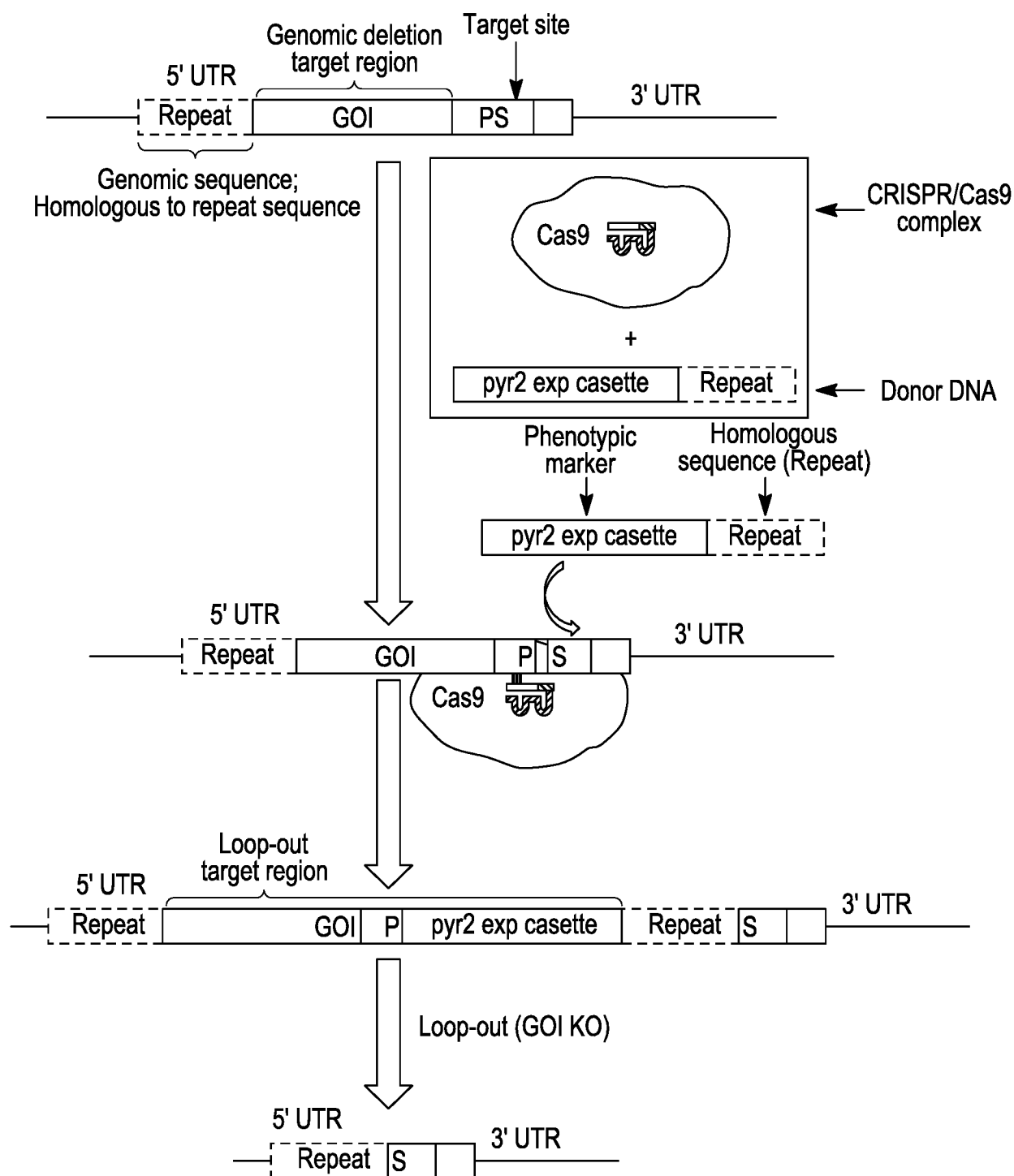


FIG. 1

2/5

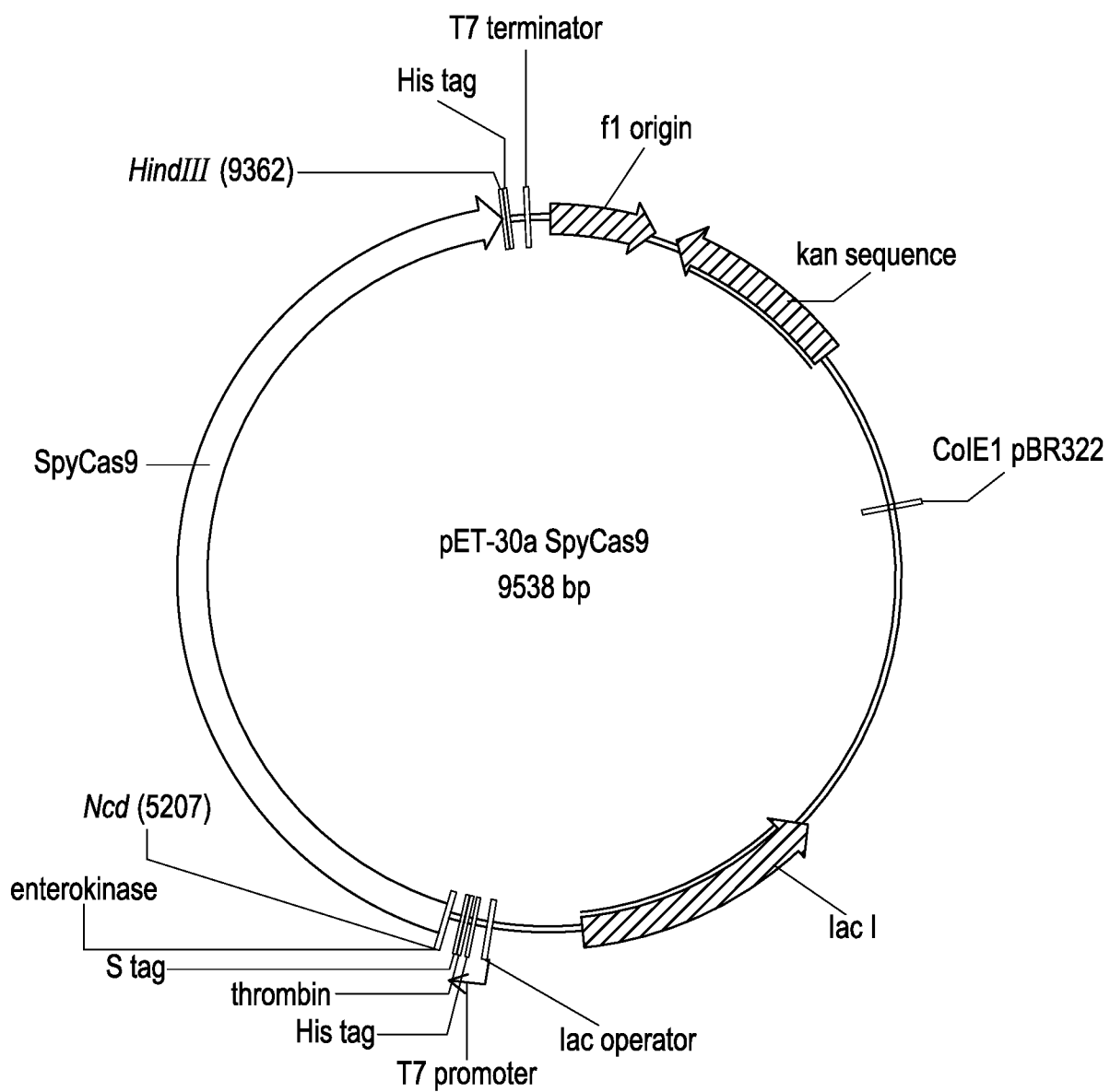


FIG. 2

3/5

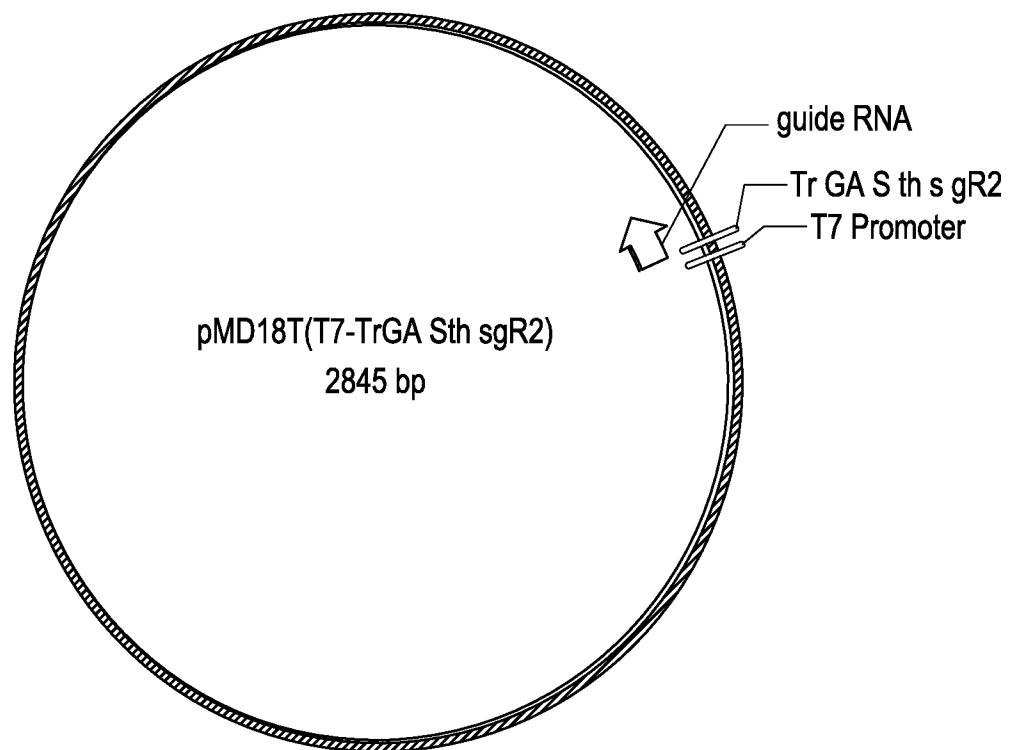


FIG. 3

4/5

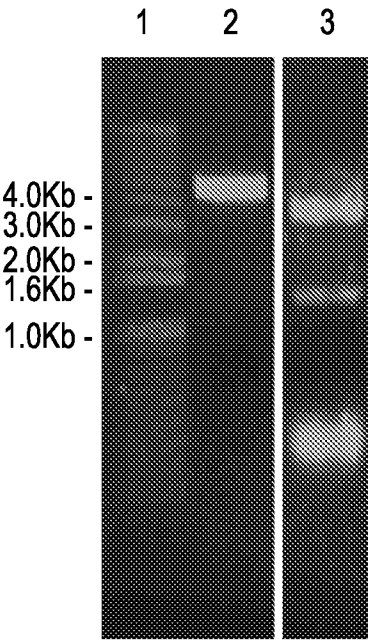


FIG. 4

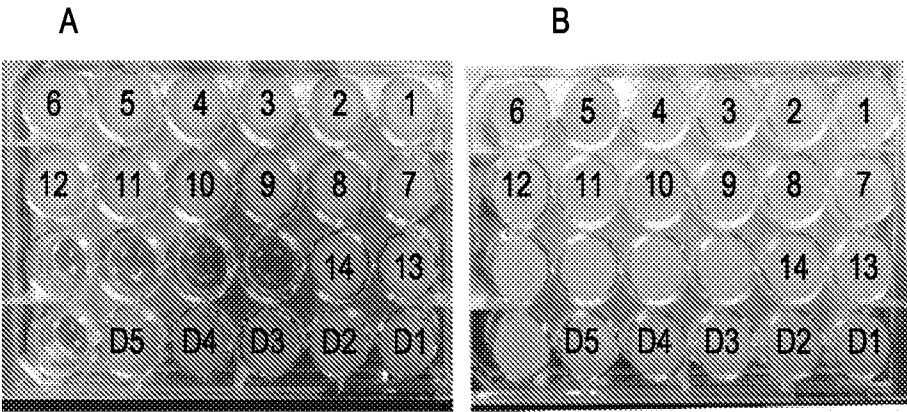
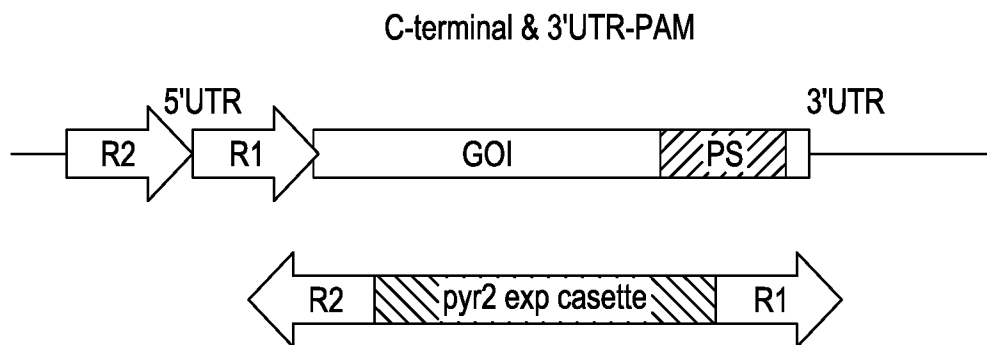
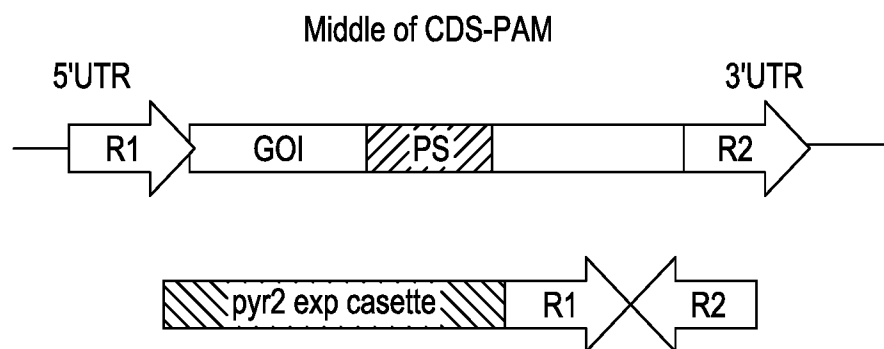
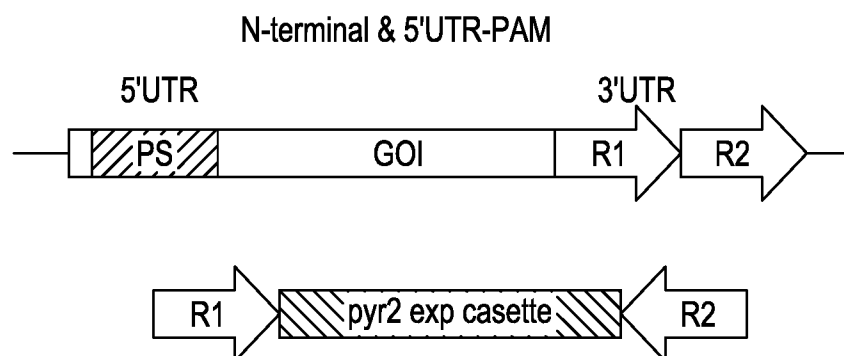


FIG. 5

5/5**FIG. 6A****FIG. 6B****FIG. 6C**

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/066195

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, XP055086617, ISSN: 0305-1048, DOI: 10.1093/nar/gkt135	1-29
Y	abstract page 4337, left-hand column page 4338, left-hand column ----- -/-	1-28

☒ 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

8 April 2016

Date of mailing of the international search report

18/04/2016

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/US2015/066195

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2015/066195

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	29
Y	abstract page 222, right-hand column; figure 1 page 225	1-28
X	----- DE BOER P ET AL: "Highly efficient gene targeting in <i>Penicillium chrysogenum</i> using the bi-partite approach in DELTA _{lig4} or DELTA _{ku70} 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]	29
A	abstract page 843, right-hand column	1-28
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	1-29
X,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 ----- -/--	1-29

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/066195

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	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	1-29
A	----- P. MALI ET AL: "RNA-Guided Human Genome Engineering via Cas9", SCIENCE, vol. 339, no. 6121, 3 January 2013 (2013-01-03), pages 823-826, XP055111247, ISSN: 0036-8075, DOI: 10.1126/science.1232033 page 823 the whole document -& P. MALI ET AL: "RNA-Guided Human Genome Engineering via Cas9", SCIENCE, vol. 339, no. 6121, 3 January 2013 (2013-01-03), pages 823-826, XP055111247, ISSN: 0036-8075, DOI: 10.1126/science.1232033	1-29
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	1-29
X,P	----- RUI LIU, LING CHEN, YANPING JIANG, ZHIHUA ZHOU, GEN ZOU: "Efficient genome editing in filamentous fungus Trichoderma reesei 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 ----- -/--	1-29

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/066195

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	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	1-29
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	1-29
A	----- DEWEI JIANG ET AL: "Molecular tools for functional genomics in filamentous fungi: Recent advances and new strategies", BIOTECHNOLOGY ADVANCES., vol. 31, no. 8, 1 December 2013 (2013-12-01), pages 1562-1574, XP055250956, GB ISSN: 0734-9750, DOI: 10.1016/j.biotechadv.2013.08.005 the whole document	1-29
A	----- BASAK ANINDITA ET AL: "A pseudouridine residue in the spliceosome core is part of the filamentous growth program in yeast.", CELL REPORTS 21 AUG 2014, vol. 8, no. 4, 21 August 2014 (2014-08-21), pages 966-973, XP002755520, ISSN: 2211-1247 abstract	1-29
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	1-29

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2015/066195

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