Title: GENE INACTIVATED MUTANTS WITH ALTERED PROTEIN PRODUCTION

(57) Abstract: A recombinant filamentous fungal cell (e.g., Aspergillus) having one or more inactivated chromosomal genes is provided. The chromosomal genes in some embodiments correspond to derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF and combinations thereof. The recombinant fungal cells may include further inactivated chromosomal genes which correspond to pepA, pepB, pepC, and pepD. The recombinant filamentous fungal cells may include a heterologous nucleic acid encoding a protein of interest. Also provided are methods of producing a protein of interest in said recombinant filamentous fungal cell.

PCR analysis of the mnn9 deletion strain
GENE INACTIVATED MUTANTS WITH ALTERED PROTEIN PRODUCTION

FIELD OF THE INVENTION
The present invention provides cells that have been genetically manipulated to have an altered capacity to produce expressed proteins. In particular, the present invention relates to filamentous fungal microorganisms, such as *Aspergillus* species wherein one or more chromosomal genes have been inactivated, and preferably, wherein one or more chromosomal genes have been deleted from the *Aspergillus* chromosome.

BACKGROUND OF THE INVENTION
Genetic engineering has allowed the improvement of microorganisms used as industrial bioreactors, cell factories and in food fermentations. In particular, filamentous fungi (e.g. *Aspergillus* and *Trichoderma* species) and certain bacteria (e.g., *Bacillus* species) produce and secrete a large number of useful proteins and metabolites (Bio/Technol. 5: 369 - 376, 713 - 719 and 1301 -1304 [1987] and Zukowski, "Production of commercially valuable products," In: Doi and McGlouglín (eds.) Biology of Bacilli: Applications to Industry, Butterworth-Heinemann, Stoneham. Mass pp 311-337 [1992]). Important production enzymes include glucoamylases, α-amylases, cellulases, neutral proteases, and alkaline (or serine) proteases, and important production proteins include hormones and antibodies. However, the occurrence of protein degradation and modification in some of these host cells provides a major hurdle for protein production, and in spite of advances in the understanding of production of proteins in filamentous fungal host cells, there remains a need for methods to increase expression of important proteins.

Accordingly, an object of the present invention is to provide an *Aspergillus* strain defective in protein degrading genes and protein modification genes, which can be used for more efficient production of heterologous or homologous proteins of interest.

SUMMARY OF THE INVENTION
The present invention is concerned with the inactivation of genes, which may be involved in protein degradation and modification (e.g., protease genes, endoplasmic reticulum (ER) degradation pathway genes and glycosylation genes). In some embodiments, the gene inactivation is a non-revertable inactivation that results in a genetically engineered microbial cell referred to as an inactivated mutant. In some embodiments, the inactivated mutant has an altered capacity to produce an expressed protein of interest.
In one aspect, the invention relates to an *Aspergillus* inactivated mutant comprising one or more non-revertible inactivated chromosomal genes selected from the group consisting of *derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd*, and *pepF* combinations thereof and homologous sequences thereto. In some embodiments, the inactivated mutant will further include a non-revertible inactivated chromosomal gene selected from the group consisting of *pepB, pepC, pepD*, combinations thereof and homologous sequences thereto. In other embodiments, the *Aspergillus* inactivated mutant is an *A. niger* inactivated mutant. In further embodiments, the inactivated mutant further comprises a polynucleotide encoding a heterologous protein of interest. In additional embodiments, the protein of interest is an enzyme, a protease inhibitor or an antibody or fragment thereof. In yet other embodiments, the *Aspergillus* inactivated mutant has an enhanced level of expression of the protein of interest compared to a corresponding parent *Aspergillus* strain when said inactivated mutant and parent strain are cultured under essentially the same growth conditions. In yet further embodiments, the one or more inactivated chromosomal genes have been deleted or the one or more inactivated chromosomal genes have been disrupted in the protein-coding region.

In a second aspect, the invention relates to a method for producing a protein of interest in an *Aspergillus* inactivated mutant comprising a) obtaining an *Aspergillus* inactivated mutant capable of producing a protein of interest, wherein said *Aspergillus* inactivated mutant has at least one non-revertible inactivated chromosomal gene selected from the group consisting of *derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd*, and *pepF* gene fragments thereof, and homologous sequences thereto; b) growing said *Aspergillus* inactivated mutant under conditions such that said protein of interest is expressed; and c) recovering the protein of interest. In some embodiments, the expression of said protein of interest in the inactivated mutant is enhanced compared to the expression of said protein of interest in a corresponding parent *Aspergillus*. In some embodiments, two chromosomal genes are inactivated. In other embodiments, the *Aspergillus* inactivated mutant further comprises inactivated chromosomal genes selected from the group consisting of *pepB, pepC, pepD* and combinations thereof and homologous sequences thereto. In additional embodiments, the protein of interest is an enzyme, a protease inhibitor or an antibody or fragments thereof. In some preferred embodiments, the protein of interest is a heterologous protein and in other embodiments the protein of interest is a homologous protein.

In a third aspect, the invention relates to a DNA sequence encoding the protein sequences of DER A, DER B, HTMA, MNN9, MNN10, OCHA, DPP4, Dpp5, PEPAa, PEPAb, PEPAc and PEPAd and functionally homologous sequence thereto.

In a fourth aspect, the invention relates to the DNA sequences comprising the genes of *derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc and pepAd*. 
In a fifth aspect, the invention relates to a method of making a recombinant filamentous fungal cell comprising introducing into a filamentous fungal cell a DNA construct that recombines with a chromosomal gene selected from the group of derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF or functionally homologous sequences thereto wherein the chromosomal gene is inactivated. In one embodiment, the inactivated gene is deleted and in another embodiment, the inactivated gene is disrupted.

**BRIEF DESCRIPTION OF THE DRAWINGS**

FIGS. 1A- B set forth a genomic Aspergillus derA DNA sequence (SEQ ID NO: 1).

FIG. 2 sets forth the putative protein sequence of DERA (SEQ ID NO: 2).

FIGS. 3A - B set forth a genomic Aspergillus derB DNA sequence (SEQ ID NO: 3).

FIG. 4 sets forth the putative protein sequence of DERB (SEQ ID NO: 4).

FIGS. 5A - E set forth a genomic Aspergillus htmA DNA sequence (SEQ ID NO: 5).

FIG. 6 sets forth the putative protein sequence of HTMA (SEQ ID NO: 6).

FIGS. 7A - D set forth a genomic Aspergillus mnn9 DNA sequence (SEQ ID NO: 7).

FIG. 8 sets forth the putative protein sequence of MNN9 (SEQ ID NO: 8).

FIGS. 9A - C set forth a genomic Aspergillus mnn10 DNA sequence (SEQ ID NO: 9).

FIG. 10 sets forth the putative protein sequence of MNN10 (SEQ ID NO: 10).

FIGS. 11A - E set forth a genomic Aspergillus ochA DNA sequence (SEQ ID NO: 11).

FIG. 12 sets forth the putative protein sequence of OCHA (SEQ ID NO: 12).

FIGS. 13A - C set forth a genomic Aspergillus dpp4 DNA sequence (SEQ ID NO: 13).

FIG. 14 sets forth the putative protein sequence of DPP4 (SEQ ID NO: 14).

FIGS. 15A - B set forth a genomic Aspergillus dpp5 DNA sequence (SEQ ID NO: 15).

FIG. 16 sets forth the putative protein sequence of DPP5 (SEQ ID NO: 16).

FIGS. 17A - B set forth a genomic Aspergillus pepAa DNA sequence (SEQ ID NO: 17).

FIG. 18 sets forth the putative protein sequence of PEPAa (SEQ ID NO: 18).

FIGS. 19A - C set forth a genomic Aspergillus pepAb DNA sequence (SEQ ID NO: 19).

FIG. 20 sets forth the putative protein sequence of PEPAb (SEQ ID NO: 20).

FIGS. 21A - B set forth a genomic Aspergillus pepAd DNA sequence (SEQ ID NO: 21).

FIG. 22 sets forth the putative protein sequence of PEPAd (SEQ ID NO: 22).

FIGS. 23A - C set forth a genomic Aspergillus pepF DNA sequence (SEQ ID NO: 23).

FIG. 24 sets forth the putative protein sequence of PEPF (SEQ ID NO: 24).
FIGS. 25A - B set forth a genomic Aspergillus pepB DNA sequence (SEQ ID NO: 25).

FIG. 26 sets forth the putative protein sequence of PEPB (SEQ ID NO: 26).

FIGS. 27A - D set forth a genomic Aspergillus pepC DNA sequence (SEQ ID NO: 27).

FIG. 28 sets forth the putative protein sequence of PEPC (SEQ ID NO: 28).

FIGS. 29A - B set forth a genomic Aspergillus pepD DNA sequence (SEQ ID NO: 29).

FIGS. 30 sets forth the putative protein sequence of PEPD (SEQ ID NO: 30).

FIGS. 31A - C set forth a genomic Aspergillus pepAc DNA sequence (SEQ ID NO: 31).

FIG. 32 sets forth the putative protein sequence of PEPAc (SEQ ID NO: 32).

FIG. 33 illustrates the general cloning strategy used for making inactivated mutants according to the invention. Fig. 33A illustrates the strategy for making a gene deletion using the vector pMW1-AderA to make a deletion of the derA gene. Further details are outlined in example 1a. Fig. 33B illustrates the strategy for making a disruption in the protein coding region of the gene using the vector PBS-disruption(ochA) as detailed for ochA in example 1f.

FIG. 34 depicts the analysis of the PCR fragment generated from total cellular DNA extracted from inactivated mutants of Aspergillus niger by fractionation on agarose gel. The gene \textit{mnr9} is representative of an inactivation by deletion (Fig. 34A), wherein lane 1 represents the DNA molecular weight marker, lane 3 represents a parent control which includes the \textit{mnr9} gene and lane 7 represents an inactivated strain with a \textit{mnr9} gene deletion. The gene \textit{ochA} is representative of an inactivation by disruption (Fig. 34B), wherein lane 1 represents the DNA molecular weight marker, lane 3 represents a parent control, which includes an \textit{ochA} gene and lane 7 represents an inactivated strain with an \textit{ochA} gene deletion. The genomic DNA was extracted from strains harboring either a gene deletion or a gene disruption. For gene deletions or disruptions two primers were designed; one primer was located on the coding region of a hydromycin gene (\textit{P}_{\text{hph}}, SEQ ID NO: 37) and one specific primer from each gene was used (See SEQ ID NOs: 38, 43, 48, 53, 58, 61, 67, 70, 73, 76, 79, 84, 89, 92 and 95). A specific PCR product was detected if the gene was deleted or disrupted. When DNA from the parent control strain was used as template PCR a band was not detected.

**DETAILED DESCRIPTION OF THE INVENTION**

The present invention provides recombinant fungal cells having one or more inactivated genes. In some embodiments, the fungal cells have been genetically manipulated to have an altered capacity to produce expressed proteins. In particular, the present invention relates to filamentous fungal cells, such as \textit{Aspergillus} cells having
enhanced expression of a protein of interest, wherein one or more chromosomal genes have been inactivated. In some preferred embodiments, the one or more chromosomal genes have been deleted from an *Aspergillus* chromosome and in other embodiments the one or more chromosome genes have been disrupted in the protein-coding region.

**Definitions**

All patents and publications, including all sequences disclosed within such patents and publications, referred to herein are expressly incorporated by reference. Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs (See *e.g.*, Singleton *et al.*, *Dictionary of Microbiology and Molecular Biology*, 2d Ed., John Wiley and Sons, New York [1994]; and Hale and Marham, *The Harper Collins Dictionary of Biology*, Harper Perennial, NY [1991], both of which provide one of skill with a general dictionary of many of the terms used herein). Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described. Numeric ranges are inclusive of the numbers defining the range. As used herein, the singular "a", "an" and "the" includes the plural reference unless the context clearly dictates otherwise. Thus, for example, reference to a "host cell" includes a plurality of such host cells.

Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxyl orientation, respectively. The headings provided herein are not limitations of the various aspects or embodiments of the invention that 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.

As used herein, "inactivated mutant" or "inactivated strain" (*e.g.*, an *Aspergillus* inactivated mutant) refers to genetically engineered recombinant host cells having one or more inactivated genes as encompassed by the invention. The term encompasses progeny thereof. In some embodiments, inactivation is the result of gene deletions and these inactivated mutants are sometimes referred to as deletion mutants. In other embodiments, inactivation is the result of disruption to the protein coding sequence and these inactivated mutants are sometimes referred to as disruption mutants. In some embodiments, the inactivation is non-reversible. In some embodiments, non-reversible refers to a strain, which will naturally revert back to the parental strain with a frequency of less than $10^{-7}$. In some embodiments, inactivation will result in a cell having no detectable activity for the gene or gene product corresponding to the inactivated gene.

A "corresponding parent strain" refers to the host strain (*e.g.*, the originating and/or wild-type strain) from which an inactivated mutant is derived.
The term "inactivation" includes any method that prevents the functional expression of one or more of the derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF, pepB, pepC, pepD genes, fragments or homologues thereof, wherein the gene or gene product is unable to exert its known function. Means of gene inactivation include deletions, disruptions of the protein-coding sequence, insertions, additions, mutations, gene silencing (e.g. RNAi genes antisense) and the like.

As used herein "protein-coding region" refers to the region of a gene that encodes the amino acid sequence of a protein.

As used herein "polypeptide" refers to peptide or protein sequences or portions thereof. The terms "protein", "peptide" and "polypeptide" are used interchangeably.

As used herein, the term "heterologous protein" refers to a protein or polypeptide that does not naturally occur in the host cell.

As used herein, "homologous protein" or "endogenous protein" refers to a protein or polypeptide native or naturally occurring in a cell.

As used herein, "host cell" or "host strain" refers to a cell that has the capacity to act as a host or expression vehicle for a newly introduced DNA sequence. In preferred embodiments of the present invention, the host cells are Aspergillus sp..

As used herein, the genus Aspergillus includes all species within the genus "Aspergillus," as known to those of skill in the art, including but not limited to A. niger, A. oryzae, A. awamori, A. kawachi and A. nidulans.

As used herein, "nucleic acid" refers to a nucleotide or polynucleotide sequence, and fragments or portions thereof, as well as to DNA, cDNA, and RNA of genomic or synthetic origin which may be double-stranded or single-stranded, whether representing the sense or antisense strand. It will be understood that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences may encode a given protein.

As used herein the term "gene" means a chromosomal segment of DNA involved in producing a polypeptide chain that may or may not include regions preceding and following the coding regions (e.g. promoter, terminator, 5' untranslated (5' UTR) or leader sequences and 3' untranslated (3' UTR) or trailer sequences, as well as intervening sequence (introns) between individual coding segments (exons)).

As used herein, the term "vector" refers to any nucleic acid that can be replicated in cells and can carry new genes or DNA segments into cells. Thus, the term refers to a nucleic acid construct designed for transfer between different host cells. An "expression vector" refers to a vector that has the ability to incorporate and express heterologous DNA fragments in a foreign cell. Many prokaryotic and eukaryotic expression vectors are commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.
As used herein, the terms "DNA construct," "expression cassette," and "expression vector," refer to a nucleic acid construct generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a target cell (i.e., these are vectors or vector elements, as described above). The recombinant expression cassette can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of an expression vector includes, among other sequences, a nucleic acid sequence to be transcribed, a promoter and a terminator. In some embodiments, DNA constructs also include a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a target cell. In some embodiments, a DNA construct of the invention comprises a selective marker.

As used herein, "transforming DNA," "transforming sequence," and "DNA construct" refer to DNA that is used to introduce sequences into a host cell or organism. The DNA may be generated in vitro by PCR or any other suitable techniques. In some embodiments, the transforming DNA comprises an incoming sequence, while in other embodiments it further comprises an incoming sequence flanked by homology boxes. In yet a further embodiment, the transforming DNA comprises other non-homologous sequences, added to the ends (i.e., stuffer sequences or flanks). The ends can be closed such that the transforming DNA forms a closed circle, such as, for example, insertion into a vector.

As used herein, the term "plasmid" refers to a circular double-stranded (ds) DNA construct used as a cloning vector, and which forms an extrachromosomal self-replicating genetic element in many bacteria and some eukaryotes. In some embodiments, plasmids become incorporated into the genome of the host cell.

As used herein, the terms "isolated" and "purified" refer to a nucleic acid or amino acid (or other component) that is removed from at least one component with which it is naturally associated.

As used herein, the term "enhanced expression" is broadly construed to include enhanced production of a protein of interest. Enhanced expression is that expression above the normal level of expression in the corresponding parent strain that has not been altered according to the teachings herein but has been grown under essentially the same growth conditions.

As used herein the term "expression" refers to a process by which a polypeptide is produced based on the nucleic acid sequence of a gene. The process includes both transcription and translation. In preferred embodiments, the process also includes secretion.

As used herein in the context of introducing a nucleic acid sequence into a cell, the term "introduced" refers to any method suitable for transferring the nucleic acid sequence into the cell.
As used herein, the terms “transformed” and "stably transformed" refers to a cell that has a non-native (heterologous) polynucleotide sequence integrated into its genome or as an episomal plasmid that is maintained for at least two generations.

As used herein “an incoming sequence” refers to a DNA sequence that is introduced into the host cell chromosome. In some embodiments, the incoming sequence is part of a DNA construct. In some embodiments, the incoming sequence encodes one or more proteins of interest. In other embodiments, the incoming sequence comprises a sequence that may or may not already be present in the genome of the cell to be transformed (i.e., it may be either a homologous or heterologous-sequence).—In some embodiments, the incoming sequence includes a functional or non-functional gene and/or a mutated or modified gene. In a preferred embodiment, the incoming sequence comprises a gene selected from the group consisting of derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF, pepB, pepC, pepD, fragments and homologous sequences thereof. In yet another embodiment, the incoming sequence includes a selective marker. In a further embodiment the incoming sequence includes two homology boxes. In some embodiments, the incoming sequence encodes at least one heterologous protein of interest.

As used herein, “homology box” refers to a nucleic acid sequence, which is homologous to a sequence in the Aspergillus chromosome. More specifically, a homology box is an upstream or downstream region having between about 80 and 100% sequence identity, between about 90 and 100% sequence identity, or between about 95 and 100% sequence identity with the immediate flanking coding region of a gene or part of a gene to be inactivated according to the invention. These sequences direct where in the chromosome a DNA construct or incoming sequence is integrated and directs what part of the chromosome is replaced by the DNA construct or incoming sequence. While not meant to limit the invention, a homology box may include about between 1 base pair (bp) to 200 kilobases (kb). Preferably, a homology box includes about between 1 bp and 10.0 kb; between 1 bp and 5.0 kb; between 1 bp and 2.5 kb; between 1 bp and 1.0 kb, and between 0.25 kb and 2.5 kb. A homology box may also include about 10.0 kb, 5.0 kb, 2.5 kb, 2.0 kb, 1.5 kb, 1.0 kb, 0.5 kb, 0.25 kb and 0.1 kb. In some embodiments, the 5' and 3' ends of a selective marker are flanked by a homology box wherein the homology box comprises nucleic acid sequences immediately flanking the coding region of the gene.

As used herein, the terms “selectable marker” and “selective marker” refer to a nucleic acid (e.g., a gene) capable of expression in host cell, which allows for ease of selection of those hosts containing the marker. Thus, the term “selectable marker” refers to genes that provide an indication that a host cell has taken up an incoming DNA of interest or some other reaction has occurred. Typically, selectable markers are genes that confer antimicrobial resistance or a metabolic advantage on the host cell to allow cells containing the exogenous DNA to be distinguished from cells that have not received any exogenous
sequence during the transformation. A "residing selectable marker" is one that is located on the chromosome of the microorganism to be transformed. A residing selectable marker encodes a gene that is different from the selectable marker on the transforming DNA construct. Selective markers are well known to those of skill in the art. As indicated above, preferably the marker is an antimicrobial resistant marker (e.g., amp\(^R\); phleo\(^R\); spec\(^R\); kan\(^R\); ery\(^R\); tet\(^R\); cmp\(^R\); hygro\(^R\) and neo\(^R\); See e.g., Guerot-Fleury, Gene, 167:335–337 [1995]; Palmeros et al., Gene 247:255-264 [2000]; and Trieu-Cuot et al., Gene, 23:331-341 [1983]). Other markers useful in accordance with the invention include, but are not limited to auxotrophic-markers, such as tryptophan, pyrG and amdsS; and detection markers, such as β-galactosidase.

As used herein, the term "promoter" refers to a nucleic acid sequence that functions to direct transcription of a downstream gene. In preferred embodiments, the promoter is appropriate to the host cell in which a desired gene is being expressed. The promoter, together with other transcriptional and translational regulatory nucleic acid sequences (also termed "control sequences") is necessary to express a given gene. In general, the transcriptional and translational regulatory sequences include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences.

A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA encoding a secretory leader (i.e., a signal peptide), is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

As used herein, "homologous genes" refers to a pair of genes from different, but usually related species, which correspond to each other and which are identical or very similar to each other. The term encompasses genes that are separated by speciation (i.e., the development of new species) (e.g., orthologous genes), as well as genes that have been separated by genetic duplication (e.g., paralogous genes). In preferred embodiments the homologous genes are functionally related.

As used herein, "ortholog" and "orthologous genes" refer to genes in different species that have evolved from a common ancestral gene (i.e., a homologous gene) by speciation. Typically, orthologs retain the same function in during the course of evolution. Identification
of orthologs finds use in the reliable prediction of gene function in newly sequenced genomes.

As used herein, "paralog" and "paralogous genes" refer to genes that are related by duplication within a genome. While orthologs retain the same function through the course of evolution, paralogs evolve new functions, even though some functions are often related to the original one. Examples of paralogous genes include, but are not limited to genes encoding trypsin, chymotrypsin, elastase, and thrombin, which are all serine proteinases and occur together within the same species.

As used herein, "homology" refers to sequence similarity or identity, with identity being preferred. This homology is determined using standard techniques known in the art (See e.g., Smith and Waterman, Adv. Appl. Math., 2:482 [1981]; Needleman and Wunsch, J. Mol. Biol., 48:443 [1970]; Pearson and Lipman, Proc. Natl. Acad. Sci. USA 85:2444 [1988]; programs such as GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package (Genetics Computer Group, Madison, WI); and Devereux et al., Nucl. Acid Res., 12:387-395 [1984]).

"Homologous sequences" as used herein means a nucleic acid or polypeptide sequence having at least 100%, at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 88%, at least 85%, at least 80%, at least 75%, at least 70% or at least 60% sequence identity to a subject nucleic acid or polypeptide sequence when optimally aligned for comparison. In some embodiments, homologous sequences have between 80% and 100% sequence identity, while in other embodiments between 90% and 100% sequence identity, and in more preferred embodiments, between 95% and 100% sequence identity. A functionally homologous sequence means the corresponding gene or protein functions in the same manner as the subject gene or protein.

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng and Doolittle (Feng and Doolittle, J. Mol. Evol., 35:351-360 [1987]). The method is similar to that described by Higgins and Sharp (Higgins and Sharp, CABIOS 5:151-153 [1989]). Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps.

Another example of a useful algorithm is the BLAST algorithm, described by Altschul et al., (Altschul et al., J. Mol. Biol., 215:403-410, [1990]; and Karlin et al., Proc. Natl. Acad. Sci. USA 90:5873-5878 [1993]). A particularly useful BLAST program is the WU-BLAST-2 program (See, Altschul et al., Meth. Enzymol., 266:460-480 [1996]). WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable
parameters are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. However, the values may be adjusted to increase sensitivity. A % amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the "longer" sequence in the aligned region. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored).

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.

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 (Tm) of the nucleic acid binding complex or probe. For example, "maximum stringency" typically occurs at about Tm-5°C (5°C below the Tm of the probe); "high stringency" at about 5-10°C below the Tm; "intermediate stringency" at about 10-20°C below the Tm of the probe; and "low stringency" at about 20-25°C below the Tm. Functionally, maximum stringency conditions may be used to identify sequences having strict identity or near-strict identity with the hybridization probe; while an intermediate or low stringency hybridization can be used to identify or detect polynucleotide sequence homologs.

Moderate and high stringency hybridization conditions are well known in the art. An example of high stringency conditions includes hybridization at about 42°C in 50% formamide, 5X SSC, 5X Denhardt's solution, 0.5% SDS and 100 µg/ml denatured carrier DNA followed by washing two times in 2X SSC and 0.5% SDS at room temperature and two additional times in 0.1X SSC and 0.5% SDS at 42°C. An example of moderate stringent conditions include an overnight incubation at 37°C in a solution comprising 20% formamide, 5 x SSC (150mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1X SSC at about 37 - 50°C. 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.

As used herein, "recombinant" includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid sequence or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found in identical form within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, underexpressed, overexpressed or not expressed at
all as a result of deliberate human intervention. "Recombination," "recombining," or generating a "recombined" nucleic acid is generally the assembly of two or more nucleic acid fragments wherein the assembly gives rise to a chimeric gene.

In an alternative embodiment, the transforming DNA sequence comprises homology boxes without the presence of an incoming sequence. In this embodiment, it is desired to delete the endogenous DNA sequence between the two homology boxes. Furthermore, in some embodiments, the transforming sequences are wild-type, while in other embodiments, they are mutant or modified sequences. In addition, in some embodiments, the transforming sequences are homologous; while in other embodiments, they are heterologous.

As used herein, the term "target sequence" refers to a DNA sequence in the host cell that encodes the sequence where it is desired for the incoming sequence to be inserted into the host cell genome. In some embodiments, the target sequence encodes a functional wild-type gene or operon, while in other embodiments the target sequence encodes a functional mutant gene or operon, or a non-functional gene or operon.

As used herein, a "flanking sequence" refers to any sequence that is either upstream or downstream of the sequence being discussed (e.g., for genes A-B-C, gene B is flanked by the A and C gene sequences). In a preferred embodiment, the incoming sequence is flanked by a homology box on each side. In another embodiment, the incoming sequence and the homology boxes comprise a unit that is flanked by stuffer sequence on each side. In some embodiments, a flanking sequence is present on only a single side (either 3' or 5'), but in preferred embodiments, it is on each side of the sequence being flanked. The sequence of each homology box is homologous to a sequence in the Aspergillus chromosome. These sequences direct where in the Aspergillus chromosome the new construct gets integrated and what part of the Aspergillus chromosome will be replaced by the incoming sequence. In some embodiments these sequences direct where in the Aspergillus chromosome the new construct gets integrated without any part of the chromosome being replaced by the incoming sequence. In a preferred embodiment, the 5' and 3' ends of a selective marker are flanked by a polynucleotide sequence comprising a section of the inactivating chromosomal segment. In some embodiments, a flanking sequence is present on only a single side (either 3' or 5'), while in preferred embodiments, it is present on each side of the sequence being flanked.

As used herein, the term "primer" refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, (i.e., in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH). The primer is preferably single stranded for maximum efficiency in amplification. Preferably, the primer is an oligodeoxyribonucleotide.
As used herein, the term "polymerase chain reaction" ("PCR") refers to the methods of U.S. Patent Nos. 4,683,195, 4,683,202, and 4,965,188, hereby incorporated by reference. As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to bacterial enzymes, each of which cut double-stranded DNA at or near a specific nucleotide sequence.

A "restriction site" refers to a nucleotide sequence recognized and cleaved by a given restriction endonuclease and is frequently the site for insertion of DNA fragments. In certain embodiments of the invention restriction sites are engineered into the selective marker and into 5' and 3' ends of the DNA construct.

As used herein, "strain viability" refers to reproductive viability. In preferred embodiments, the inactivation of a chromosomal gene does not deleteriously affect division and survival of the inactivated mutant under laboratory conditions.

As used herein, the term "chromosomal integration" refers to the process whereby an incoming sequence is introduced into the chromosome of a host cell (e.g., Aspergillus). The homologous regions of the introduced (transforming) DNA align with homologous regions of the chromosome. Subsequently, the sequence between the homology boxes is replaced by the incoming sequence in a double crossover (i.e., homologous recombination).

"Homologous recombination" means the exchange of DNA fragments between two DNA molecules or paired chromosomes at the site of identical or nearly identical nucleotide sequences. In a preferred embodiment, chromosomal integration is by homologous recombination.

Preferred Embodiments

The present invention provides inactivated mutants (e.g., deletion mutants and disruption mutants) that are capable of producing a protein of interest. In particular, the present invention relates to recombinant filamentous fungal microorganisms, such as Aspergillus species having altered expression of a protein of interest, wherein one or more chromosomal genes have been inactivated, and preferably wherein one or more chromosomal genes have been deleted from the Aspergillus chromosome or wherein the protein-coding region of one or more chromosomal genes has been disrupted. Indeed, the present invention provides means for deletion of single or multiple genes. In preferred embodiments, such deletions provide advantages such as improved production of a protein of interest.

Inactivated Genes -

As indicated above, the present invention includes embodiments that involve single or multiple gene inactivations. In some embodiments, the gene inactivations are gene deletions or gene disruptions. In some embodiments the inactivations are non-revertable.
Genes to be inactivated according to the invention include but are not limited to those involved in protein degradation or protein modification, such as proteins in the ER degradation pathway, proteases genes, such as secreted serine and aspartic protease genes, glycosylation genes and glycoprotein degradation genes. In some embodiments, the chromosomal gene to be inactivated includes one or more of the following genes derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepF, pepAa, pepAb, pepAc and pepAd, or functionally homologous sequences thereto having at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%; at least 88%, at least 85%; at least 80%; at least 70% or at least 60% sequence identity therewith.

With respect to the genes to be inactivated according to the invention derA and derB genes are believed to function in the ER degradation pathway. ER degradation pathway enzymes include ER resident proteins such as those involved in the translocation of misfolded protein from the ER to the cytosol, and non ER resident proteins such as ubiquitin conjugating enzymes which target the misfolded protein for proteasomal degradation (Bonifacino and Weissman [1998] Ann. Rev. Cell. Biol.14:19 – 57). The htmA, mnn9, mnn10 and ochA genes are believed to function in glycoprotein modification. Glycoprotein modifying enzymes are enzymes that modify oligosaccharide molecules, which have been added to amino acid residues on a protein. The dpp4, dpp5, pepF, pepAa, pepAb, pepAc, pepAd, pepB, pepC and pepD genes are believed to be proteases. Proteases are protein-degrading enzymes, which catalyze the hydrolytic cleavage of proteins. More specifically, proteases are enzymes that cleave peptide bonds. In some embodiments, the protease genes are aspartic protease genes (e.g. pepAa, pepAb, pepAc, pepAd and pepB). Enzymatically active aspartic proteases are those enzymes or fragments thereof that contain aspartic acid residues at their active site. (Kosta, V (Ed) ASPARTIC PROTEINASES AND THEIR INHIBITORS, Walter de Gruyter, NY pp 27 – 40; 151 – 161 and 163 – 177). In other embodiments, the protease genes are dipeptidyl peptidases (e.g. dpp4 and dpp5). In other embodiments, the protease genes are serine carboxylypeptidase (e.g., pepF), and in further embodiments, the protease genes are serine proteases (e.g. pepC and pepD).

In some embodiments, inactivated genes will include two or more (e.g. two, three or four) inactivated genes according to the invention. In other embodiments, the inactivated genes will include at least one of the above-enumerated genes and a gene selected from the group consisting of pepB, pepC, pepD, combinations thereof and functionally homologous sequences thereto having at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94% at least 93%, at least 92%, at least 91%, at least 90%, at least 88%, at least 85%, at least 80%, at least 70% or at least 60% sequence identity therewith.

In other embodiments, inactivated genes will include any one of the above-enumerated genes and an inactivated pepA gene or homologous sequence thereto, such as

While not meant to limit the invention in any manner, genes to be inactivated include the following combinations and functionally homologous genes thereto: (a) mnn9 and mnn10; (b) mnn9 and ochA; (c) mnn9, mnn10 and ochA; (d) dpp4 and dpp5; (e) dpp4, dpp5 and pepA; (f) pepAa and pepAb; (g) pepAa, pepAb and pepAc; (h) pepAa and pepAc; (i) pepAa, pepAb, pepAc and pepB; (j) pepAa, pepAb, pepAC and pepC; (k) pepAa, pepAb, pepAC and pepD; (l) pepB, pepC, pepD and pepF; (m) pepAa, pepAb, and pepAc; and n) -dpp4,-dpp5 and mnn9. Further embodiments include any one of the above-mentioned combinations (a – n) and an inactivated pepA gene or homologous gene thereto.

In some embodiments, the DNA coding sequences of these genes from Aspergillus are provided in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, and SEQ ID NO: 31. As indicated above, it is contemplated that functionally homologous genes found in filamentous fungal cells will find use in the present invention. In some embodiments, the functionally homologous genes will have at least 80% sequence identity to any one of the above enumerated sequences.

Methods for determining homologous sequences from host cells are known in the art and include using a nucleic acid sequence disclosed herein to construct an oligonucleotide probe, said probe corresponding to about 6 to 20 amino acids of the encoded protein. The probe may then be used to clone the homologous protein degradation gene. The filamentous fungal host genomic DNA is isolated and digested with appropriate restriction enzymes. The fragments are separated and probed with the oligonucleotide probe prepared from the protein degradation sequences by standard methods. A fragment corresponding to the DNA segment identified by hybridization to the oligonucleotide probe is isolated, ligated to an appropriate vector and then transformed into a host to produce DNA clones.

In other embodiments, the DNA encodes the protein sequences selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 30 and SEQ ID NO: 32 and functionally homologous sequence thereto. In some embodiments, a functionally homologous sequence will be a protein found in a filamentous fungal cell (i.e. Aspergillus) and have at least 95% sequence identity to any one of the above enumerated sequences. In some embodiments, the functionally homologous sequence will be found in an Aspergillus niger or Aspergillus oryzae and will have at least 90% or also at least 95% sequence identity to and one of the above enumerated sequences.

In other embodiments, a protein sequence will differ from any one of the above enumerated protein sequences by one or more conservative
Methods of Inactivation and General Construction of DNA Constructs to be used to Inactivate Chromosomal Genes -

In some embodiments, the present invention includes a DNA construct comprising an incoming sequence. The DNA construct is assembled in vitro, followed by direct cloning of the construct into a competent host (e.g. an Aspergillus host), such that the DNA construct becomes integrated into the host chromosome. For example, PCR fusion and/or ligation can be employed to assemble a DNA construct in vitro. In some embodiments, the DNA construct is a non-plasmid construct, while in other embodiments it is incorporated into a vector (e.g., a plasmid). In some embodiments, circular plasmids are used. In preferred embodiments, circular plasmids are designed to use an appropriate restriction enzyme (i.e., one that does not disrupt the DNA construct). Thus, linear plasmids find use in the present invention.

In some embodiments, the incoming sequence comprises a derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF, pepC, pepB, pepD gene, gene fragments thereof, homologous sequences thereto; or immediate chromosomal coding region flanking sequences. A homologous sequence is a nucleic acid sequence having functional similarity to one of the above enumerated sequences and having at least 99%, 98%, 97%, 96%, 95%, 94% 93%, 92%, 91%, 90%, 88%, 85%, 80%, 70% or 60% sequence identity to a derA, derB, htmA, mnn9, mnn10, ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF, pepB, pepC or pepD gene or gene fragment thereof.

In some embodiments, wherein the genomic DNA is already known the 5' flanking fragment and the 3' flanking fragment of the gene to be deleted is cloned by two PCR reactions, and in embodiments wherein the gene is disrupted, the DNA fragment is cloned by one PCR reaction.

In some embodiments, the coding region flanking sequences include a range of about 1bp to 2500 bp; about 1bp to 1500 bp, about 1 bp to 1000 bp, about 1 bp to 500 bp, and 1 bp to 250 bp. The number of nucleic acid sequences comprising the coding region flanking sequence may be different on each end of the gene coding sequence. For example, in some embodiments, the 5' end of the coding sequence includes less than 25 bp and the 3' end of the coding sequence includes more than 100 bp.

In some embodiments, the incoming sequence comprises a selective marker flanked on the 5' and 3' ends with a fragment of the gene sequence. In other embodiments, when
the DNA construct comprising the selective marker and gene, gene fragment or homologous sequence thereto is transformed into a host cell, the location of the selective marker renders the gene non-functional for its intended purpose. In some embodiments, the incoming sequence comprises the selective marker located in the promoter region of the gene. In other embodiments, the incoming sequence comprises the selective marker located after the promoter region of gene. In yet other embodiments, the incoming sequence comprises the selective marker located in the coding region of the gene. In further embodiments, the incoming sequence comprises a selective marker flanked by a homology box on both ends. In still further embodiments, the incoming sequence includes a sequence that interrupts the transcription and/or translation of the coding sequence. In yet additional embodiments, the DNA construct includes restriction sites engineered at the upstream and downstream ends of the construct.

Whether the DNA construct is incorporated into a vector or used without the presence of plasmid DNA, it is used to transform a microorganism, which results in an inactivated mutant, preferably having a stable and non-reverting inactivation of the chromosomal gene. Methods used to ligate the DNA construct and to insert them into a suitable vector are well known in the art. Linking is generally accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide linkers are used in accordance with conventional practice. (See, Sambrook (1989) supra, and Bennett and Lasure, MORE GENE MANIPULATIONS IN FUNGI, Academic Press, San Diego (1991) pp 70 – 76.). Additionally, vectors can be constructed using known recombination techniques (e.g., Invitrogen Life Technologies, Gateway Technology). Examples of suitable expression and/or integration vectors that may be used in the practice of the invention are provided in Sambrook et al., (1989) supra, Ausubel (1987) supra, van den Hondel et al. (1991) in Bennett and Lasure (Eds.) MORE GENE MANIPULATIONS IN FUNGI, Academic Press pp. 396-428 and U.S. Patent No. 5,874,276. Particularly useful vectors include pFB6, pBR322, pUC18, pUC100 and pENTR/D.

In some embodiments, at least one copy of a DNA construct is integrated into the host chromosome. In some embodiments, one or more DNA constructs of the invention are used to transform host cells. For example, one DNA construct may be used to inactivate a derA gene and another construct may be used to inactivate a derB gene. Of course, additional combinations are contemplated and provided by the present invention.

Inactivation occurs via any suitable means, including deletions, substitutions (e.g., mutations), interruptions, and/or insertions in the nucleic acid gene sequence and gene silencing mechanisms, such as RNA interference (RNAi). In one embodiment, the expression product of an inactivated gene is a truncated protein with a corresponding change in the biological activity of the protein. In preferred embodiments, the inactivation results in a loss of biological activity of the gene. In some embodiments, the biological
activity of the inactivated gene in a recombinant fungal cell will be less than 25% (e.g. 20%, 15%, 10%, 5% and 2%) compared to the biological activity of the same or functionally homologous gene in a corresponding parent strain.

In some preferred embodiments, inactivation is achieved by deletion and in other preferred embodiments inactivation is achieved by disruption of the protein-coding region of the gene. In some embodiments, the gene is inactivated by homologous recombination. As used herein, “deletion” of a gene refers to deletion of the entire coding sequence, deletion of part of the coding sequence, or deletion of the coding sequence including flanking regions. The deletion may be partial as long as the sequences left in the chromosome render the gene functionally inactive. In preferred embodiments, a deletion mutant comprises deletion of one or more genes that results in a stable and non-reverting deletion. Flanking regions of the coding sequence may include from about 1 bp to about 500 bp at the 5' and 3' ends. The flanking region may be larger than 500 bp but will preferably not include other genes in the region which may be inactivated or deleted according to the invention. The end result is that the deleted gene is effectively non-functional. In simple terms, a “deletion” is defined as a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, have been removed (i.e., are absent). While not meant to limit the methods used for inactivation in some embodiments, derA, derB, htmA, mnn9, mnn10, pepC, pepB and functionally homologous genes may be inactivated by deletion.

A “disruption” is a change in a nucleotide or amino acid sequence, which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as compared to the parent or naturally occurring sequence. In some embodiments, the disruption may be by insertion of a marker gene into the protein-coding region in vitro through a restriction enzyme site. Flanking regions of the coding sequence may include about 1 bp to about 500 bp at the 5' and 3' ends. The flanking region may be larger than 500 bp, but will preferably not include other genes in the region. The DNA constrict aligns with the homologous sequence of the host chromosome and in a double crossover event the translation or transcription of the gene is disrupted. For example, ochA chromosomal gene is aligned with a plasmid comprising the gene or part of the gene coding sequence and a selective marker. In some embodiments, the selective marker is located within the gene coding sequence or on a part of the plasmid separate from the gene. The vector is integrated into the host chromosome, and the gene is inactivated by the insertion of the marker in the coding sequence. While not meant to limit the methods used for inactivation, in some embodiments ochA, dpp4, dpp5, pepAa, pepAb, pepAc, pepAd, pepF, pepD and functionally homologous sequences may be inactivated by this method.

An “insertion” or “addition” is a change in a nucleic acid or amino acid sequence in which one or more nucleotides or amino acid residues have been added as compared to the endogenous chromosomal sequence or protein product. In some embodiments inactivation
is by insertion in a single crossover event with a plasmid as the vector. For example, the
vector is integrated into the host cell chromosome and the gene is inactivated by the
insertion of the vector in the protein-coding sequence of the gene or in the regulatory region
of the gene.

In alternative embodiments, inactivation results due to mutation of the gene.

Methods of mutating genes are well known in the art and include but are not limited to site-
directed mutation, generation of random mutations, and gapped-duplex approaches (See
e.g., U.S. Pat. 4,760,025; Moring et al., Biotech. 2:646 [1984]; and Kramer et al., Nucleic
Acids Res.; 12:9441-[1984]).--

**Host Cells**

In the present invention, the host cell is preferably a filamentous fungal cell (See,
Alexopoulos, C. J. (1962), INTRODUCTORY MYCOLOGY, Wiley, New York) preferred filamentous
fungal cells include *Aspergillus* sp., (*e.g.*, *A. oryzae*, *A. niger*, *A. awamori*, *A. nidulans*, *A. sojae,
*A. japonicus*, *A. kawachi* and *A. aculeatus*); *Rhizopus* sp., *Trichoderma* sp. (*e.g.*, *Trichoderma
reesei* (previously classified as *T. longibrachiatum* and currently also known as *Hypocrea
jecorina*), *Trichoderma viride*, *Trichoderma koningii*, and *Trichoderma harzianum*) and *Mucor*
sp. (*e.g.*, *M. miehei* and *M. pusillus*). Most preferred host cells are *Aspergillus niger* cells. In
some embodiments, particular strains of *Aspergillus niger* include ATCC 22342 (NRRL 3112),
ATCC 44733, and ATCC 14331 and strains derived there from. In some embodiments, the host
cell is one that is capable of expressing a heterologous gene. The host cell may be a
recombinant cell, which includes a heterologous protein. In other embodiments, the host is one
that overexpresses a protein that has been introduced into the cell. In some embodiments, the
host strain is a mutant strain deficient in one or more genes such as genes corresponding to
protease genes other than the protease genes disclosed herein. For example a preferred host
is an *Aspergillus niger* in which a gene encoding the major secreted aspartyl protease, such as
aspergillopepsin has been deleted (U.S. Pat. Nos. 5,840,570 and 6,509,171).

**Methods of Determining Gene Inactivations**

One skilled in the art may use various methods to determine if a gene has been
inactivated. While not meant to limit the invention one method which can be used is the
phenol/chloroform method described in Zhu (Zhu et al., *Acad Mycologica Sinica* 13:34-40
[1994]). Briefly, in this method the genomic DNA is used as a template for PCR reactions.
Primers are designed so that one primer anneals to a selectable marker gene (*e.g.*, *a
hygromycin resistant marker gene, hph*) and a second primer anneals to a sequence further
3' from the DNA homologous fragment at the 3' end of the gene. An inactivated mutant will
produce a specific PCR product when its genomic DNA is used as a PCR reaction template
as opposed to the corresponding parent strain (having an non-inactivated gene) which will
not generate PCR fragments when its genomic DNA is used as a template. In addition the
PCR fragment from the inactivated mutant may be subjected to DNA sequencing to confirm the identity if the inactivated gene. Other useful methods include Southern analysis and reference is made to Sambrook (1989) supra.

Proteins of Interest -

In some embodiments an inactivated mutant encompassed by the invention will overexpress a homologous protein of interest and in other embodiments an inactivated mutant encompassed by the invention will express a heterologous protein of interest.

In some embodiments, the protein of interest is intracellular while in other embodiments, the protein of interest is a secreted polypeptide. In addition the protein of interest may be a fusion or hybrid protein. Preferred polypeptides include enzymes, including, but not limited to those selected from amylytic enzymes, proteolytic enzymes, cellulolytic enzymes, oxidoreductase enzymes and plant cell-wall degrading enzymes. More particularly, these enzyme include, but are not limited to amylases, glucoamylases, proteases, xylanases, lipases, laccases, phenol oxidases, oxidases, cutinases, cellulases, hemicellulases, esterases, peroxidases, catalases, glucose oxidases, phytases, pectinases, glucosidases, isomerases, transferases, galactosidases and chitinases. Particularly preferred enzymes include but are not limited to amylases, glucoamylases, proteases, phenol oxidases, cellulases, hemicellulases, glucose oxidases and phytases. In some particularly preferred embodiments of the present invention, the polypeptide of interest is a protease, cellulase, glucoamylase or amylase. These enzymes are well known in the art.

In some embodiments, the protein of interest is a secreted polypeptide, which is fused to a signal peptide (i.e., an amino-terminal extension on a protein to be secreted). Nearly all secreted proteins use an amino-terminal protein extension, which plays a crucial role in the targeting to and translocation of precursor proteins across the membrane. This extension is proteolytically removed by a signal peptidase during or immediately following membrane transfer.

In some embodiments of the present invention, the polypeptide of interest is a protein such as a protease inhibitor, which inhibits the action of proteases. Protease inhibitors are known in the art, for example the protease inhibitors belonging to the family of serine proteases inhibitors which are known to inhibit trypsin, cathepsinG, thrombin and tissue kallikrein. Particularly preferred protease inhibitors include Bowman-Birk inhibitors and soybean trypsin inhibitors (See, Birk, Int. J. Pept. Protein Res. 25:113-131 [1985]; Kennedy, Am. J. Clin. Neutr. 68:1406S-1412S [1998] and Billings et al., Proc. Natl. Acad. Sci. 89:3120 – 3124 [1992]).

In some embodiments of the present invention, the polypeptide of interest is selected from hormones, antibodies, growth factors, receptors, cytokines, etc. Hormones encompassed by the present invention include but are not limited to, follicle-stimulating hormone, luteinizing hormone, corticotropin-releasing factor, somatostatin, gonadotropin hormone, vasopressin,
oxytocin, erythropoietin, insulin and the like. Growth factors include, but are not limited to platelet-derived growth factor, insulin-like growth factors, epidermal growth factor, nerve growth factor, fibroblast growth factor, transforming growth factors, cytokines, such as interleukins (e.g., IL-1 through IL-13), interferons, colony stimulating factors, and the like. Antibodies include but are not limited to immunoglobulins obtained directly from any species from which it is desirable to produce antibodies. In addition, the present invention encompasses modified antibodies. Polyclonal and monoclonal antibodies are also encompassed by the present invention. In particularly preferred embodiments, the antibodies or fragments thereof are humanized antibodies, such as anti-p185 \textsuperscript{HER2} and HulD10.

In a further embodiment, the nucleic acid encoding the protein of interest will be operably linked to a suitable promoter, which shows transcriptional activity in a fungal host cell. The promoter may be derived from genes encoding proteins either homologous or heterologous to the host cell. The promoter may be a truncated or hybrid promoter. Further the promoter may be an inducible promoter. Preferably, the promoter is useful in a \textit{Trichoderma} host or an \textit{Aspergillus} host. Suitable nonlimiting examples of promoters include \textit{cbh1}, \textit{cbh2}, \textit{egl1}, \textit{egl2}, \textit{xyn1} and \textit{amy}. In one embodiment, the promoter is one that is native to the host cell. Other examples of useful promoters include promoters from the genes of \textit{A. awamori} and \textit{A. niger} glucoamylase genes (\textit{glaA}) (Nunberg \textit{et al.}, (1984) \textit{Mol. Cell Biol.} 4:2306-2315 and Boel \textit{et al.}, (1984) \textit{EMBO J.} 3:1581-1585); \textit{Aspergillus oryzae} TAKA amylase; \textit{Rhizomucor miehei} aspartic proteinase; \textit{Aspergillus niger} neutral alpha-amylase; \textit{Aspergillus niger} acid stable alpha-amylase; \textit{Trichoderma reesei} \textit{xln1} and the \textit{cellubiohydrolase 1} gene promoter (EPA 137280A1) and mutant, truncated and hybrid promoters thereof.

In some preferred embodiments, the polypeptide coding sequence is operably linked to a signal sequence which directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence may naturally contain a signal sequence naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. The DNA encoding the signal sequence is preferably that which is naturally associated with the polypeptide to be expressed. Preferably, the signal sequence is encoded by an \textit{Aspergillus niger} alpha-amylase, \textit{Aspergillus niger} neutral amylase or \textit{Aspergillus niger} glucoamylase. In some embodiments, the signal sequence is the \textit{Trichoderma cdh1} signal sequence which is operably linked to a \textit{cdh1} promoter.

**Transformation of Fungal Cells**

Introduction of a DNA construct or vector into a host cell includes techniques such as transformation; electroporation; nuclear microinjection; transduction; transfection, (e.g., lipofection mediated and DEAE-Dextrin mediated transfection); incubation with calcium phosphate DNA precipitate; high velocity bombardment with DNA-coated microprojectiles;

Cell Culture -

The fungal cells may be grown in conventional culture medium. The culture media for transformed cells may be modified as appropriate for activating promoters and selecting transformants. The specific culture conditions, such as temperature, pH and the like will be apparent to those skilled in the art. Preferred culture conditions may be found in the scientific literature such as Sambrook, (1982) supra, and from the American Type Culture Collection. Additionally, fermentation procedures for production of heterologous proteins are known per se in the art. For example, proteins can be produced either by solid or submerged culture, including batch, fed-batch and continuous-flow processes. Fermentation temperature can vary somewhat, but for filamentous fungi such as Aspergillus niger the temperature generally will be within the range of about 20°C to 40°C, generally preferably in the range of about 28°C to 37°C, depending on the strain of microorganism chosen. The pH range in the aqueous microbial ferment (fermentation admixture) should be in the exemplary range of about 2.0 to 8.0. With filamentous fungi, the pH normally is within the range of about 2.5 to 8.0; with Aspergillus niger the pH normally is within the range of about 4.0 to 6.0, and preferably in the range of about 4.5 to 5.5. While the average retention time of the fermentation admixture in the fermentor can vary considerably, depending in part on the fermentation temperature and culture employed, generally it will be within the range of about 24 to 500 hours, preferably presently about 24 to 400 hours. The type of fermentor employed is not critical, though presently preferred is operation under 15L Biolafitte (Saint-Germain-en-Laye, France).

Methods for Determining Expressed Protein Activity -

Various assays are known to those of ordinary skill in the art for detecting and measuring activity of intracellularly and extracellularly expressed polypeptides. Means for
determining the levels of secretion of a protein of interest in a host cell and detecting expressed proteins include the use of immunoassays with either polyclonal or monoclonal antibodies specific for the protein. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), fluorescence immunoassay (FIA), and fluorescent activated cell sorting (FACS). However, other methods are known to those in the art and find use in assessing the protein of interest (See e.g., Hampton et al., SEROLOGICAL METHODS, A LABORATORY MANUAL, APS Press, St. Paul, MN [1990]; and Maddox et al., J. Exp. Med., 158:1211 [1983]). In some preferred embodiments, the expression and/or secretion of a protein of interest are enhanced in an inactivated mutant. In some embodiments the production of the protein of interest is at least 100%, at least 95%, at least 90%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40%, at least 30%, at least 20%, at least 15%, at least 10%, at least 5% and at least 2% greater in the inactivated mutant as compared to the corresponding parent strain.

Protein Recovery –

Once the desired protein is expressed and, optionally, secreted the protein of interest may be recovered and further purified. The recovery and purification of the protein of interest from a fermentation broth can be done by procedures known per se in the art. The fermentation broth will generally contain cellular debris, including cells, various suspended solids and other biomass contaminants, as well as the desired protein product.

Suitable processes for such removal include conventional solid-liquid separation techniques such as, e.g., centrifugation, filtration, dialysis, microfiltration, rotary vacuum filtration, or other known processes, to produce a cell-free filtrate. It may be preferable to further concentrate the fermentation broth or the cell-free filtrate prior to crystallization using techniques such as ultrafiltration, evaporation or precipitation.

Precipitating the proteinaceous components of the supernatant or filtrate may be accomplished by means of a salt, followed by purification by a variety of chromatographic procedures, e.g., ion exchange chromatography, affinity chromatography or similar art recognized procedures. When the expressed desired polypeptide is secreted the polypeptide may be purified from the growth media. Preferably the expression host cells are removed from the media before purification of the polypeptide (e.g. by centrifugation).

When the expressed recombinant desired polypeptide is not secreted from the host cell, the host cell is preferably disrupted and the polypeptide released into an aqueous "extract" which is the first stage of purification. Preferably the expression host cells are collected from the media before the cell disruption (e.g. by centrifugation).

The manner and method of carrying out the present invention may be more fully understood by those of skill in the art by reference to the following examples, which examples
are not intended in any manner to limit the scope of the present invention or of the claims directed thereto.

**EXPERIMENTAL**

The following Examples are provided in order to demonstrate and further illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the experimental disclosure which follows, the following abbreviations apply: °C (degrees Centigrade); H₂O (water); dH₂O (deionized water); (HCl (hydrochloric acid); aa (amino acid); bp (base pair); kb (kilobase pair); kD (kilodaltons); gm (grams); µg (micrograms); mg (milligrams); µl (microliters); ml (milliliters); mm (millimeters); µm (micrometer); M (molar); mM (millimolar); µM (micromolar); MW (molecular weight); sec (seconds); min(s) (minute/minutes); hr(s) (hour/hours); NaCl (sodium chloride); PBS (phosphate buffered saline [150 mM NaCl, 10 mM sodium phosphate buffer, pH 7.2]); PCR (polymerase chain reaction); SDS (sodium dodecyl sulfate); w/v (weight to volume); v/v (volume to volume); ATCC (American Type Culture Collection, Rockville, MD); BD BioSciences (Previously CLONTECH Laboratories, Palo Alto, CA); Invitrogen (Invitrogen Corp., San Diego, CA); and Sigma (Sigma Chemical Co., St. Louis, MO).

Table 1 below illustrates the primers (and their sequence identification) used in the examples to make the corresponding gene inactivations.

TABLE 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Sequence (5' to 3' )</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>derA</td>
<td>P1a  TAGTTAACTCGTGCTCTCCTGGCGGC</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>P2a  AGGTCGACGAAGTATAGGAAGGTTGTGAACAG</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>T1a  AGGGATCCACGTCTGGTACTTCTTTCAACG</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>T2a  TCTCGCGATTGGATCAAACCATAAGATAC</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Prenh GAGGCGAAAGGAATAGAGTAG</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>PrenhOUT CTCAAGCAGAGGATTTGTC</td>
<td>38</td>
</tr>
<tr>
<td>derB</td>
<td>P1b  CGGTGTTACAGATTGGATTGTCTAATAAGCAG</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>P2b  TGGTCGACGGAGGACATTGGATTG</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>T1b  AGGGATCCCTAAAGATTATCCGCTTAGTCC</td>
<td>41</td>
</tr>
<tr>
<td>Gene</td>
<td>Sequence</td>
<td>Length</td>
</tr>
<tr>
<td>------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>T2b</td>
<td>AAGATATCCATCCAAGCTATGCGACATTTTACTCC</td>
<td>42</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>TAGAAGTGCGCATCAAATAG</td>
<td>43</td>
</tr>
<tr>
<td>htmA</td>
<td>P1c</td>
<td>CGGTTAACATATCATATTCGCGATTTGAGTTAC</td>
</tr>
<tr>
<td>P2c</td>
<td>TATCTCGAGCAGCGAACGAGATGACGAGGAGGAGAAG</td>
<td>45</td>
</tr>
<tr>
<td>T1c</td>
<td>ATGGATCCTAAAGTGCAAGTGTGTCAGACGCGGTG</td>
<td>46</td>
</tr>
<tr>
<td>T2c</td>
<td>AATGATATCCCGCAGTACCATCTCTCC</td>
<td>47</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>TCTTGGGCTAAATTAGAGGAGG</td>
<td>48</td>
</tr>
<tr>
<td>mnn9</td>
<td>P1d</td>
<td>TAGTTAACAGCCGAGAGTCAGAAAG</td>
</tr>
<tr>
<td>P2d</td>
<td>AGGTCGACAAGAGAGATGAGGAGGAAG</td>
<td>50</td>
</tr>
<tr>
<td>T1d</td>
<td>TTGGATCCGTCTACGGGCTGTGATTAC</td>
<td>51</td>
</tr>
<tr>
<td>T2d</td>
<td>ACCTCGCGACTTCTACCTCACAAACATTACC</td>
<td>52</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>CCGACAAGGAGCGACGAGAAG</td>
<td>53</td>
</tr>
<tr>
<td>mnn10</td>
<td>P1e</td>
<td>TCATGCTATTCTCTTTCCGTC</td>
</tr>
<tr>
<td>P2e</td>
<td>AGGCATGCACAAGATGTCAGT</td>
<td>55</td>
</tr>
<tr>
<td>T1e</td>
<td>AGGAGTCCGAGATGGAATTTGATA</td>
<td>56</td>
</tr>
<tr>
<td>T2e</td>
<td>TGGTTAGGATGATGTGGCTGAC</td>
<td>57</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>TGAATGATACGGGTGATGATGTCAG</td>
<td>58</td>
</tr>
<tr>
<td>ochA</td>
<td>Pf</td>
<td>TAGTTAACACAGCTGTCTGCAG</td>
</tr>
<tr>
<td>Tf</td>
<td>AGGTTAACATATGTCAAGAGATCAAAGTGC</td>
<td>60</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>ACAGCAAGATGTTGTGTGTCGT</td>
<td>61</td>
</tr>
<tr>
<td>dpp4</td>
<td>Pg</td>
<td>AGTCGCGAGATGTAGAAGAGGAGAAG</td>
</tr>
<tr>
<td>Tg</td>
<td>AGTCGCGAGCCTGTTTGAAGATGTG</td>
<td>63</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>TCTGGATAGAATGCAATTGCAGT</td>
<td>64</td>
</tr>
<tr>
<td>dpp5</td>
<td>Ph</td>
<td>TGCCAGGTCAGCCTTACAAAGAAG</td>
</tr>
<tr>
<td>Th</td>
<td>ACGATATCAGCAGCAGCCACACCCCATATAC</td>
<td>66</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>TCGTTATAGCTTTGCAGAATG</td>
<td>67</td>
</tr>
<tr>
<td>pepAa</td>
<td>Pi</td>
<td>GCACCTTCTTCCCTTTTTTTTTGTAC</td>
</tr>
<tr>
<td>Ti</td>
<td>AGGTTAACATTTGTAGATACAGGCCAC</td>
<td>69</td>
</tr>
<tr>
<td>P_T-OUT</td>
<td>TCGATGAGGGGTTGGAAGATG</td>
<td>70</td>
</tr>
<tr>
<td>pepAb</td>
<td>Pj</td>
<td>ACGTTAACCATATCACAGCTATATCCC 71</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td></td>
<td>Tj</td>
<td>ACGTTAACGCCAGGTCTCCTCTTCTGC 72</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>GGAGAGATAGGACGTAACCTCAGTG 73</td>
</tr>
<tr>
<td>pepAd</td>
<td>Pk</td>
<td>TGGTTAACTGTAATAGTGATTGGCTGTAC 74</td>
</tr>
<tr>
<td></td>
<td>Tk</td>
<td>ATGTTAACCAGGCTTGCTGCTTG 75</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>AGGAGCAAGGAAATACTGAGGAG 76</td>
</tr>
<tr>
<td>pepF</td>
<td>Pl</td>
<td>AGGTTAACTTGCTGCGAGGAAGCAAACCT 77</td>
</tr>
<tr>
<td></td>
<td>Ti</td>
<td>AGGTTAACATCAGCCGCGGTCAAAGT 78</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>TCGAGGCGGACGACAGTCTAT 79</td>
</tr>
<tr>
<td>pepB</td>
<td>P1m</td>
<td>CCGTTAACCTCCACGTATTTCAATATCC 80</td>
</tr>
<tr>
<td></td>
<td>P2m</td>
<td>AAGTCGACACAGTCTGGAGAATAGCGG 81</td>
</tr>
<tr>
<td></td>
<td>T1m</td>
<td>CCGGATCCCTGAGGGTGATCTTTGCAGGAGACCAC 82</td>
</tr>
<tr>
<td></td>
<td>T2m</td>
<td>GGTTAATAGCGTACTATCCTGCTGAG 83</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>TCGTTATAGCTTCGTAACACAAT 84</td>
</tr>
<tr>
<td>pepC</td>
<td>P1n</td>
<td>AAGTTAACCGTTCGGTAGGCATTGC 85</td>
</tr>
<tr>
<td></td>
<td>P2n</td>
<td>TCGTCGACAGTGATGTCGCTGACCATTGCC 86</td>
</tr>
<tr>
<td></td>
<td>T1n</td>
<td>CTGGATCCAAGCTGAAGAAAGACATCATGC 87</td>
</tr>
<tr>
<td></td>
<td>T2n</td>
<td>TAGATATCTGTCTATTCTATATGAGCCCTC 88</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>ATACAGCAGCAGTCTATCAATATGAG 89</td>
</tr>
<tr>
<td>pepD</td>
<td>Po</td>
<td>TAAAGCCTAGCAAGCAACATTG 90</td>
</tr>
<tr>
<td></td>
<td>To</td>
<td>AACAGAAAAAGCGAATAACAAACAGG 91</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>ACAAGAACCTGTCTCCAGTATGAG 92</td>
</tr>
<tr>
<td>pepAc</td>
<td>Pp</td>
<td>TGGTTAAGGAGGATTTGCTCTATTG 93</td>
</tr>
<tr>
<td></td>
<td>Tp</td>
<td>TGGTTAAGGGCTAGCTATGCTATG 94</td>
</tr>
<tr>
<td></td>
<td>P_{OUT}</td>
<td>TCGCTCGGGGTGGTTGTG 95</td>
</tr>
</tbody>
</table>

**EXAMPLE 1**

*Creation of Aspergillus Deletion Constructs and Strains*

a. Deletion of the derA gene.

Fig. 1 (SEQ ID NO: 1) sets forth the 2400 bp genomic DNA sequence of the Aspergillus derA gene and Fig. 2 (SEQ ID NO: 2) sets forth the 246 amino acids sequence translated from the derA genomic DNA of Fig. 1.

To construct the deletion plasmid, two pairs of PCR primers were designed. The first pair of PCR primers amplify the promoter region of the gene and they are indicated in Table 1 as SEQ ID NO: 33 (P1a) and SEQ ID NO: 34 (P2a). The second pair of PCR primers amplifies the terminator region of gene and they are indicated in Table 1 as SEQ ID NO: 35 (T1a) and SEQ ID NO: 36 (T2a). The terminator fragment DNA sequence (T1) was amplified in PCR using the following conditions: the PCR tube was heated at 94°C for 3 minutes to denature template DNA. Then, the PCR reaction was run at 94°C for 1 minute, 57°C for 1 minute and 72°C for 1 minute 30 seconds and this cycle was repeated 30 times. Finally, PCR reaction was extended at 72°C for 10 minutes before the tube was incubated at 4°C. The ends of the T1 fragment were then filled in with T4 DNA polymerase and then cut with restriction enzyme (BamH1). This modified PCR fragment was then cloned to pMW1 (Ulrich Kuck et al., [1989] Appl. Microbiol. Biotechnol. 31:358-365) to construct plasmid pMW1-T1(derA).

The promoter DNA sequence, (the P1 fragment), was amplified in PCR reaction using the same condition as the T1 fragment with two primers (SEQ ID NO: 33 and SEQ ID NO: 34). The ends of the P1 fragment were then filled in with T4 DNA polymerase and cut with restriction enzyme Sall. This modified PCR fragment was cloned to pMW1-T1(derA) to generate pMW1-ΔderA. The plasmid was analyzed by restriction enzyme digestion to confirm its identity. The plasmid was linearized by two restriction enzymes digestion (HapI and NruI).
The digested DNA fragment was used to transform a derivative of an AP-4 *Aspergillus niger* strain (Ward et al. [1993] Appl. Microbiol. Biotechnol 39:738-743) comprising an expression plasmid expressing *Tramete versicolor* laccase under the glucoamylase promoter and terminator control.

Figure 33A illustrates the general strategy used to make the deletion plasmids used in the examples provided and as described in detail herein.

The transformation protocol utilized was a modification of the Campbell method (Campbell et al. 1989. *Curr. Genet.* 16:53-56) wherein the beta-D-glucinase G (InterSpex Products, Inc. San Mateo, CA) was used to produce protoplasts and pH was adjusted to 5.5. All solutions and media were either autoclaved or filter sterilized through a 0.2 micron filter. The DNA was extracted from transformants using a phenol/chloroform method (Zhu et al. 1993. *Nucleic Acid Res.* 21:5279-80). The deletion strain was detected by PCR using two primers SEQ ID NO: 37 (P_hpg) and SEQ ID NO: 38 (P_outA), which gave a specific PCR product of 1064 bp when the DNA from the deletion strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain (Fig. 34).

b. Deletion of the derB gene.

Fig. 3 (SEQ ID NO: 3) sets forth the 2673 bp genomic DNA sequence of the *Aspergillus derB* gene and Fig. 4 (SEQ ID NO: 4) sets forth the 166 amino acid sequence translated from the derA genomic DNA of Fig. 3. The deletion plasmids were constructed as described above for Example 1a with the following differences.

The first pair of PCR primers used to amplify the promoter region are designated in Table 1 as SEQ ID NO: 39 (P1b) and SEQ ID NO: 40 (P2b). The second pair of primers used to amplify the terminator region are designated in Table 1 as SEQ ID NO: 41 (T1b) and SEQ ID NO: 42 (T2b). In this example, the ends of the T2 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (BamHI). The modified PCR fragment was then cloned to pMW1 to construct plasmid pMW1-T2(derB). The P2 fragment was amplified in PCR reaction using the same conditions as the P1 fragment. The ends of the P2 fragment were then filled in with T4 DNA polymerase and cut with restriction enzyme SalI.

This modified PCR fragment was cloned to pMW1-T2(derB) to generate pMW1-ΔderB. The plasmid was analyzed by restriction enzyme as described above in Example 1a. The plasmid was linearized by two restriction enzymes digestion (Hpal and EcoRV).

The digested DNA fragment was used to transform *Aspergillus niger* and DNA was extracted from the transformants as described above for example 1A. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 43 (P_outb), which gave a specific PCR product of 694 bp when the DNA from the deletion strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain. However, the wild type band was also identified from the deletion strain.
c. Deletion of the htmA gene.

Fig. 5 (SEQ ID NO: 5) sets forth the 7000 bp genomic DNA sequence of the *Aspergillus htmA* gene and Fig. 6 (SEQ ID NO: 6) sets forth the 1076 amino acid sequence translated from the *htmA* genomic DNA of Fig. 5. The deletion plasmids were constructed as described above for Example 1a with the following differences.

The first pair of PCR primers used to amplify the promoter region are designated in Table 1 as SEQ ID NO: 44 (P1c) and SEQ ID NO: 45 (P2c). The second pair of primers used to amplify the terminator-region are designated in Table 1 as SEQ ID NO: 46 (T1c) and SEQ ID NO: 47 (T2c). In this example, the P3 and T3 fragments were amplified in PCR reactions using the same conditions as the P1 and T1 fragments. The ends of the T3 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (BamHI). The modified PCR fragment was then cloned to pMW1 to construct plasmid pMW1-T3(htmA). The ends of the P3 fragment were then filled in with T4 DNA polymerase and cut with restriction enzyme Xhol. This modified PCR fragment was cloned to pMW1-T3(htmA) to generate pMW1-DhtmA. The plasmid was analyzed by restriction enzyme as described above in Example 1a. The plasmid was linearized by two restriction enzymes digestion (HpaI and EcoRV).

The digested DNA fragment was used to transform *Aspergillus niger* GAP3-4 (Ward et al. [1993] Appl. Microbiol. Biotechnol. 39:738 – 743) and DNA was extracted from the transformants as described above for example 1A. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 48 (P1outc), which gave a specific PCR product of 1497bp when the DNA from the deletion strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

d. Deletion of the mnn9 gene.

Fig. 7 (SEQ ID NO: 7) sets forth the 4947 bp genomic DNA sequence of the *Aspergillus mnn9* gene and Fig. 8 (SEQ ID NO: 8) sets forth the 369 amino acid sequence translated from the *mnn9* genomic DNA of Fig. 7. The deletion plasmids were constructed as described above for Example 1a with the following differences.

The first pair of PCR primers used to amplify the promoter region are designated in Table 1 as SEQ ID NO: 49 (P1d) and SEQ ID NO: 50 (P2d). The second pair of primers used to amplify the terminator region are designated in Table 1 as SEQ ID NO: 51 (T1d) and SEQ ID NO: 52 (T2d).

In this example, the ends of the P4 fragment were filled with T4 DNA polymerase and then cut with restriction enzyme (SalI). The modified PCR fragment was cloned to pMW1 to construct plasmid pMW1-P4 (mnn9). The ends of the T4 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (BamHI). The modified PCR fragment was
then cloned to pMW1-P(mnn9) to generate plasmid pMW1-Δmnn9. The plasmid was analyzed by restriction enzyme as described above in Example 1a. The plasmid was linearized by two restriction enzymes (HpaI and NruI).

The digested DNA fragment was used to transform *Aspergillus niger* and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 53 (P<sub>out4</sub>), which gave a specific PCR product of 1330 bp when the DNA from the deletion strain was used as template for PCR amplification (Fig. 34A, lane 7). No band was seen when the DNA was from the parent strain (Fig. 34A, lane 3).

e. Deletion of the mnn10 gene.

Fig. 9 (SEQ ID NO: 9) sets forth the 4524 bp genomic DNA sequence of the *Aspergillus mnn10* gene and Fig. 10 (SEQ ID NO: 10) sets forth the 466 amino acid sequence translated from the mnn10 genomic DNA of Fig. 9. The deletion plasmids were constructed as described above for Example 1a with the following differences.

The first pair of PCR primers used to amplify the promoter region are designated in Table 1 as SEQ ID NO: 54 (P1e) and SEQ ID NO: 55 (P2e). The second pair of primers used to amplify the terminator region are designated in Table 1 as SEQ ID NO: 56 (T1e) and SEQ ID NO: 57 (T2e).

In this example, the ends of the P5 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (SphiI). The modified PCR fragment was then cloned to pMW1 to construct plasmid pMW1-P5 (mnn10). The ends of the T5 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (BamHI). The modified PCR fragment was cloned to pMW1-P5(mnn10) to generate plasmid pMW1-Δmnn10. The plasmid was analyzed by restriction enzyme as described above in Example 1a. The plasmid was linearized by two restriction enzymes (NruI and EcoRV).

The digested DNA fragment was used to transform *Aspergillus niger* and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 53 (P<sub>out4</sub>), which gave a specific PCR product of 1295 bp when the DNA from the deletion strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

f. Disruption of the ochA gene.

Fig. 11 (SEQ ID NO: 11) sets forth the 6724 bp genomic DNA sequence of the *Aspergillus ochA* gene and Fig. 12 (SEQ ID NO: 12) sets forth the 380 amino acid sequence translated from the ochA genomic DNA of Fig. 11. The disruption plasmids were constructed as described above for Example 1a with the following differences.
The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 59 (Pf) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 60 (Ti). Using these primers, the coding region of the ochA gene including the promoter region of 80 bp and terminator region of 624 bp was amplified. The DNA sequence, named the W6 fragment, was amplified in a PCR reaction using the following conditions: The PCR tube was heated at 94°C for 4 min to denature template DNA, the PCR reaction was then run at 94°C for 1 min, 55°C for 2 min, and 72°C for 1 min 30 sec and this cycle was repeated 30 times. The PCR reaction was extended at 72°C for 10 min before the tube was incubated at 4°C. The-produced 1787 bp PCR fragment W6 was cloned to pBS-T, a TA vector derived from pBlue-script (Tian Wei Biotech. Co. Ltd) to construct plasmid pBS-W6(ochA). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the ochA gene at the EcoRV site to generate pBS-disruption ochA. The plasmid was linearized by restriction enzyme (Hpal) digestion.

Figure 33B illustrates the general strategy used to make the disruption plasmids used in the examples provided and as described in detail herein.

The digested DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 61 (P_t-outf), which gave a specific PCR product of 1336 bp when the DNA from the disruption strain was used as template for PCR amplification (Fig. 34B, lane 7), while no band was seen when the DNA was from the parent strain (Fig. 34B, lane 3).

g. Disruption of the dpp4 gene.

Fig. 13 (SEQ ID NO: 13) sets forth the 3989 bp genomic DNA sequence of the Aspergillus dpp4 gene and Fig. 14 (SEQ ID NO: 14) sets forth the 915 amino acid sequence translated from the dpp4 genomic DNA of Fig. 13. The disruption plasmids were constructed as described above for Example 1f with the following differences.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 62 (Pg) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 63 (Tg). Using these primers, the 950 – 3356 bp region of the coding region (817 – 3663) of the dpp4 gene was amplified.

The produced 2407 bp PCR fragment W7 was cloned to pBS-T to construct plasmid pBS-W7(dpp4). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the dpp4 gene at the EcoRI-EcoRI (2175 – 2257 bp) site to generate pBS-disruption dpp4. The plasmid was linearized by restriction enzyme digestion (NruI).

The digested DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 64 (P_t-outg), which
gave a specific PCR product of 1191 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

**h. Disruption of the dpp5 gene.**

Fig. 15 (SEQ ID NO: 15) sets forth the 2647 bp genomic DNA sequence of the *Aspergillus dpp5* gene and Fig. 16 (SEQ ID NO: 16) sets forth the 726 amino acid sequence translated from the *dpp5* genomic DNA of Fig. 15. The disruption plasmids were constructed as described above for Example 1f with the following differences.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 65 (Ph) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 66 (Th).

Using these primers, the 195 – 2490 bp region of the coding region (1 – 2647bp) of the dpp5 gene was amplified. The produced 2295 bp PCR fragment W8 was cloned to pBS-T to construct plasmid pBS-W8(dpp5). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the dpp5 gene at the BglII site to generate pBS-disruption dpp5. The plasmid was linearized by restriction enzyme (EcoRV) digestion.

The digested DNA fragment was used to transform *Aspergillus niger* and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 67 (P_{out}1), which gave a specific PCR product of 1282 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

**i. Disruption of the pepAa gene.**

Fig. 17 (SEQ ID NO: 17) sets forth the 2777 bp genomic DNA sequence of the *Aspergillus pepAa* gene and Fig. 18 (SEQ ID NO: 18) sets forth the 394 amino acid sequence translated from the *pepAa* genomic DNA of Fig. 17.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 68 (Pl) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 69 (Tl).

Using these primers, the coding region of the pepAa gene and some promoter region (355 bp) and terminator region (326 bp) was amplified. The DNA sequence, named as the W9 fragment was amplified in a PCR reaction as described above for Example 1f with the following differences.

The produced 1920 bp PCR fragment W9 was cloned to pBS-T to construct plasmid pBS-W9(pepAa). The DNA fragment containing the hygromycin resistant gene was inserted
into the coding region of the pepAa gene at the BstBI site to generate pBS-disruption pepAa. The plasmid was linearized by restriction enzyme (HpaI) digestion.

The digested DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 70 (P_out), which gave a specific PCR product of 1140 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

j. Disruption of the pepAb gene.

Fig. 19 (SEQ ID NO: 19) sets forth the 3854 bp genomic DNA sequence of the Aspergillus pepAb gene and Fig. 20 (SEQ ID NO: 20) sets forth the 417 amino acid sequence translated from the pepAb genomic DNA of Fig. 19.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 71 (Pj) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 72 (Tj).

Using these primers, the coding region of the pepAb gene and some promoter region (1025 bp) was amplified. The DNA sequence, named as the W10 fragment was amplified in a PCR reaction as described above for Example 1f with the following differences.

The produced 2170 bp PCR fragment W10 was cloned to pBS-T to construct plasmid pBS-W10(pepAb). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the pepAb gene at the Eco47III site to generate pBS-disruption pepAb. The plasmid was linearized by restriction enzyme (HpaI) digestion.

The digested DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 73 (P_out), which gave a specific PCR product of 1191 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

k. Disruption of the pepAd gene.

Fig. 21 (SEQ ID NO: 21) sets forth the 2411 bp genomic DNA sequence of the Aspergillus pepAd gene and Fig. 22 (SEQ ID NO: 22) sets forth the 480 amino acid sequence translated from the pepAd genomic DNA of Fig. 21.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 74 (Pk) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 75 (Tk).
Using these primers, the 1201 bp coding region of the 1443 bp pepAd gene and some promoter region (858 bp) was amplified. The DNA sequence, named as the W11 fragment was amplified in a PCR reaction as described above for Example 1f with the following differences.

The produced 2059 bp (23 – 2081 bp) PCR fragment W11 was cloned to pBS-T to construct plasmid pBS-W11(pepAd). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the pepAd gene at the Aaul site to generate pBS-disruption pepAd. The plasmid was linearized by restriction enzyme (HpaI) digestion.

The digested-DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 76 (P_coul), which gave a specific PCR product of 1086 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

I. Disruption of the pepF gene.

Fig. 23 (SEQ ID NO: 23) sets forth the 3525 bp genomic DNA sequence of the Aspergillus pepF gene and Fig. 24 (SEQ ID NO: 24) sets forth the 531 amino acid sequence translated from the pepF genomic DNA of Fig. 23. The disruption plasmids were constructed as described above for Example 1f with the following differences.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 77 (Pl) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 78 (Tl).

Using these primers, the coding region of the pepF gene and some promoter region (1058 bp) was amplified. The DNA sequence, named as the W12 fragment was amplified in a PCR reaction as described above for Example 1f.

The produced 2350 bp PCR fragment W12 was cloned to pBS-T to construct plasmid pBS-W12(pepF). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the pepF gene at the Nrl site to generate pBS-disruption pepF.

The plasmid was linearized by restriction enzyme (HpaI) digestion.

The digested DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 79(P_coul), which gave a specific PCR product of 1231 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

m. Deletion of the pepB gene.
Fig. 25 (SEQ ID NO: 25) sets forth the 3000 bp genomic DNA sequence of the *Aspergillus pepB* gene and Fig. 26 (SEQ ID NO: 26) sets forth the 282 amino acid sequence translated from the *pepB* genomic DNA of Fig. 25. The deletion plasmids were constructed as described above for Example 1a with the following differences.

The first pair of PCR primers used to amplify the promoter region are designated in Table 1 as SEQ ID NO: 80 (P1m) and SEQ ID NO: 81 (P2m). The second pair of primers used to amplify the terminator region are designated in Table 1 as SEQ ID NO: 82 (T1m) and SEQ ID NO: 83 (T2m).

In this example, the ends of the P13 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (SalI). The modified PCR fragment was then cloned to pMW1 to construct plasmid pMW1-P13 (pepB). The ends of the T13 fragment were filled in with T4 DNA polymerase. The modified PCR fragment was then cloned to pMW1-P13(pepB) to generate plasmid pMW1-ΔpepB. The plasmid was analyzed by restriction enzyme as described above in Example 1a. The plasmid was linearized by restriction enzyme (HpaI) digestion.

The digested DNA fragment was used to transform *Aspergillus niger* and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 84 (P,70um), which gave a specific PCR product of 1357 bp when the DNA from the deletion strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

**n. Deletion of the pepC gene.**

Fig. 27 (SEQ ID NO: 27) sets forth the 3220 bp genomic DNA sequence of the *Aspergillus pepC* gene and Fig. 28 (SEQ ID NO: 28) sets forth the 533 amino acid sequence translated from the *pepC* genomic DNA of Fig. 27. The deletion plasmids were constructed as described above for Example 1a with the following differences.

The first pair of PCR primers used to amplify the promoter region are designated in Table 1 as SEQ ID NO: 85 (P1n) and SEQ ID NO: 86 (P2n). The second pair of primers used to amplify the terminator region are designated in Table 1 as SEQ ID NO: 87 (T1n) and SEQ ID NO: 88 (T2n).

In this example, the ends of the T14 fragment were filled in with T4 DNA polymerase and then cut with restriction enzyme (BamHI). The modified PCR fragment was then cloned to pMW1 to construct plasmid pMW1-T14 (pepC). The ends of the P14 fragment were filled in with T4 DNA polymerase and cut with restriction (SalI). The modified PCR fragment was then cloned to pMW1-P14(pepC) to generate plasmid pMW1-ΔpepC. The plasmid was analyzed by restriction enzyme as described above in Example 1a. The plasmid was linearized by two restriction enzymes (HpaI and EcoRV) digestion.
The digested DNA fragment was used to transform *Aspergillus niger* strain GAP3-4 (Ward et al. [1993] Appl. Microbiol. Biotechnol. 39:738-743) and DNA was extracted from the transformants as described above for example 1a. The deletion strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 89 (P_doa4), which gave a specific PCR product of 1054 bp when the DNA from the deletion strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

**o. Disruption of the pepD gene.**

-Fig: 29 (SEQ ID NO: 29) sets forth the 2993 bp genomic DNA sequence of the *Aspergillus pepD* gene and Fig. 30 (SEQ ID NO: 30) sets forth the 416 amino acid sequence translated from the *pepD* genomic DNA of Fig. 29. The disruption plasmids were constructed as described above for Example 1f with the following differences.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 90 (Po) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 91 (To).

Using these primers, the coding region of the pepD gene and some promoter region (392 bp) and terminator region (521 bp) were amplified. The DNA sequence, named as the W15 fragment was amplified in a PCR reaction as described above for Example 1f. The produced 2317 bp PCR fragment W15 was cloned to pBS-T to construct plasmid pBS-W15(pepD). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the pepD gene at the BstBI site to generate pBS-disruption pepD. The plasmid was linearized by restriction enzyme digestion (StuI).

The digested DNA fragment was used to transform *Aspergillus niger* and DNA was extracted from the transformants as described above for example 1f. The disruption strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 92 (P_doa4), which gave a specific PCR product of 1344 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

**p. Disruption of the pepAc gene.**

Fig. 31 (SEQ ID NO: 31) sets forth the 4531 bp genomic DNA sequence of the *Aspergillus pepAc* gene and Fig. 32 (SEQ ID NO: 32) sets forth the 453 amino acid sequence translated from the *pepAc* genomic DNA of Fig. 31. The disruption plasmids were constructed as described above for Example 1f with the following differences.

The PCR primer used to amplify the promoter region is designated in Table 1 as SEQ ID NO: 93 (Pp) and the primer used to amplify the terminator region is designated in Table 1 as SEQ ID NO: 94 (Tp).
Using these primers, the coding region of the pepAc gene, some promoter region (789 bp) and some terminator region (509) were amplified. The DNA sequence, named the W16 fragment was amplified in a PCR reaction.

The produced 2753 bp PCR fragment W16 was cloned to pBS-T to construct plasmid pBS-W16(pepAc). The DNA fragment containing the hygromycin resistant gene was inserted into the coding region of the pepAc gene at the EcoRV site to generate pBS-disruption pepAc. The plasmid was linearized by restriction enzyme (Hpal) digestion.

The digested DNA fragment was used to transform Aspergillus niger and DNA was extracted from the transformants as described above for example 1f. The disruption strain was detected by PCR using two primers SEQ ID NO: 37 and SEQ ID NO: 95 (P_outP), which would give a specific PCR product of 1520 bp when the DNA from the disruption strain was used as template for PCR amplification while no band was seen when the DNA was from the parent strain.

EXAMPLE 2

Inactivated Double Deletion Mutants

a. Disruption of dpp4 and dpp5.

To construct the dpp4(amdS) deletion plasmid, the 2.7kb DNA fragment containing the amdS gene was inserted into the coding region (position 950 to 3356) of the dpp4 gene at the EcoRV-EcoRV site (position 2175 to 2256) in plasmid pBS-W7 (dpp4) to generate plasmid pBS-disruption dpp4(amdS). The plasmid was analyzed by restriction enzyme digestion to confirm its identity. The plasmid was linearized by restriction enzyme digestion (NruI). The digested DNA fragment was used to transform A. niger strain (dpp5-19) which expresses a Tramete laccase under the glucoamylase promoter and terminator control and carrying the disrupted dpp5 gene (as described in Example 1h). The double deletion strain was detected by PCR using two pairs of primers. The two primers of the first pair each respectively annealing to the amdS gene and 3’ downstream the W7 fragment on the chromosomal DNA which gave a specific PCR product of 1224 bp when the DNA from dpp4 deletion strain was used as a template for PCR amplification while no band was seen when the DNA was from the recipient strain.

Primers:

\[ P_{\text{out(dpp4)}} \quad 5'---\text{TCTGGATAGAAATGCAAATCGTAG}---3' \quad \text{SEQ ID NO: 64} \]

\[ P_{\text{emds}} \quad 5'---\text{TTTCCAGTCTAGACACGTATAACGGC}---3' \quad \text{SEQ ID NO: 96} \]

The second pair of primers was the same as the originally used primers for detection of the single dpp5 deletion strain (SEQ ID NOs: 37 and 67). The double deletion strain and its control strains were used for production of laccase and total protein production.
b. Disruption of mnn9 and ochA.

To construct the mnn9 (amdS) deletion plasmid, the 2.7kb DNA fragment containing the amdS gene was inserted into the pMW1-Δmnn9 (the amdS fragment directly replacing the hph fragment) to generate plasmid pMW1-disruption mnn9(amdS). The plasmid was analyzed by restriction enzyme digestion to confirm its identity. The plasmid was linearized by restriction enzyme digestion (AsuII-NruI). The digested DNA fragment was used to transform A. niger strain (ΔochA-23) which expresses a Tramete laccase under the glucoamylase promoter and terminator control and carrying the disrupted ochA gene as described in Example 1f. The double deletion strain was detected by PCR using two pairs of primers. The two primers of the first pair each respectively annealing to the amdS gene and 3' downstream the T4 fragment on the chromosomal DNA which gave a specific PCR product of 1380 bp when the DNA from mnn9 deletion strain was used as a template for PCR amplification while no band was seen when the DNA was from the recipient strain.

Primers:

\[ P_{\text{out(mnn9)}} \ 5'--GATATCAACCTCAGCGTCAATTGG--3' \quad \text{SEQ ID NO: 53} \]

\[ P_{\text{amdS}} \ 5'--TTTCC AGTCT AGACA CGTAT AACGCGC--3' \quad \text{SEQ ID NO: 97} \]

The second pair of primers was the same as the originally used primers for detection of the single ochA deletion strain (SEQ ID NOs: 37 and 61). The double deletion strain and its control strains were then used for production of the laccase and total protein production.

EXAMPLE 3

Use of Inactivated Mutants for the Production of a Heterologous Protein

To illustrate the advantages of using the inactivated mutant according to the invention, production of laccase in the parent (wild-type) was compared to the production of laccase in inactivated mutants as described above in examples 1 and 2.

Assays were performed in shake flask cultures using 250 ml baffled flasks containing 50 mL of growth media (Promosoy) suitable for laccase production as known in the art. The strains were grown in shake flasks for 120 hrs. Laccase activity was measured following a standard assay procedure based on the oxidation of ABTS; 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonate by oxygen in sodium acetate buffer (pH 4.6). The culture broths were incubated with ABTS in sodium acetate buffer (SIGMA) at 37°C for 30 min and color formation was measured at an optical density of OD 420 nM. The level of laccase produced by the inactivated strain was measured relative to the corresponding parent strain.

Results are illustrated in Table 2A and 2B. Total extracellular protein was measured using the Folin phenol method as described in Lowry, et al., [1951] J. Biol. Chem. 193:265-275 and results are illustrated in Table 2A.
### TABLE 2A - Single Inactivations

<table>
<thead>
<tr>
<th>Inactivated Mutant Strain (Δ)</th>
<th>Production of Laccase (% increase in OD420)</th>
<th>Total Protein % (compared to Parent (Wild-Type))</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔderA</td>
<td>-80</td>
<td>106.4</td>
</tr>
<tr>
<td>ΔderB</td>
<td>15.7</td>
<td>104.3</td>
</tr>
<tr>
<td>ΔhtmA</td>
<td></td>
<td>101.1</td>
</tr>
<tr>
<td>Δmnn9</td>
<td>14.6</td>
<td>99.6</td>
</tr>
<tr>
<td>Δmnn10</td>
<td>12.7</td>
<td>102.6</td>
</tr>
<tr>
<td>ΔochA</td>
<td>7.2</td>
<td>102.3</td>
</tr>
<tr>
<td>Δdpp4</td>
<td>6.0</td>
<td>102.7</td>
</tr>
<tr>
<td>Δdpp5</td>
<td>15.4</td>
<td>99.4</td>
</tr>
<tr>
<td>ΔpepB</td>
<td>8.6</td>
<td>99.3</td>
</tr>
<tr>
<td>ΔpepC</td>
<td></td>
<td>100.0</td>
</tr>
<tr>
<td>ΔpepD</td>
<td>4.8</td>
<td>102</td>
</tr>
<tr>
<td>ΔpepF</td>
<td>5.3</td>
<td>99.8</td>
</tr>
<tr>
<td>ΔpepAa</td>
<td>0.5</td>
<td>100.5</td>
</tr>
<tr>
<td>ΔpepAb</td>
<td>13.4</td>
<td>96.5</td>
</tr>
<tr>
<td>ΔpepAd</td>
<td>2.7</td>
<td>96.5</td>
</tr>
</tbody>
</table>

### TABLE 2B - Multiple Inactivations

<table>
<thead>
<tr>
<th>Inactivated Mutant strain (Δ)</th>
<th>Production of Laccase (% Increase in OD 420)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δdpp4</td>
<td>11.8</td>
</tr>
<tr>
<td>Δdpp5</td>
<td>15.3</td>
</tr>
<tr>
<td>Δdpp4/ Δdpp5</td>
<td>26.6</td>
</tr>
<tr>
<td>Δmnn9</td>
<td>13.0</td>
</tr>
<tr>
<td>ΔochA</td>
<td>8.5</td>
</tr>
<tr>
<td>Δmnn9/ ΔochA</td>
<td>16.8</td>
</tr>
</tbody>
</table>
IT IS CLAIMED:

1. A recombinant filamentous fungal cell comprising one or more inactivated chromosomal genes selected from derA (SEQ ID NO:1), derB (SEQ ID NO:3), htmA (SEQ ID NO:5), mnn9 (SEQ ID NO:7), mnn10 (SEQ ID NO:9), ochA (SEQ ID NO:11), dpp4 (SEQ ID NO:13), dpp5 (SEQ ID NO:15), pepAa (SEQ ID NO:17), pepAb (SEQ ID NO:19), pepAc (SEQ ID NO:31), pepAd (SEQ ID NO:21), pepF (SEQ ID NO:23), functionally homologous sequences thereto and combinations thereof.

2. The recombinant filamentous fungal cell of claim 1 comprising two inactivated chromosomal genes.

3. The recombinant filamentous fungal cell of claim 2, wherein the two inactivated chromosomal genes are a) dpp4 (SEQ ID NO:13) and dpp5 (SEQ ID NO:15); b) mnn9 (SEQ ID NO:7) and ochA (SEQ ID NO:11), or c) homologous sequences thereto having at least 95% sequence identity thereto.

4. The recombinant filamentous fungal cell of claim 1, wherein the inactivated chromosomal gene is dpp5 (SEQ ID NO:15) and homologous sequences thereof.

5. The recombinant filamentous fungal cell of claim 4 further comprising a second inactivated chromosomal gene.

6. The recombinant filamentous fungal cell of claim 1, wherein the one or more inactivated chromosomal genes are selected from derA (SEQ ID NO:1), derB (SEQ ID NO:3), htmA (SEQ ID NO:5), mnn9 (SEQ ID NO:7), mnn10 (SEQ ID NO:9), ochA (SEQ ID NO:11), dpp4 (SEQ ID NO:13), dpp5 (SEQ ID NO:15), and functionally homologous sequences thereof having at least 95% sequence identity thereto.

7. The recombinant filamentous fungal cell of claim 6, further comprising an inactivated chromosomal gene selected from pepAa (SEQ ID NO:17), pepAb (SEQ ID NO:19), pepAc (SEQ ID NO:31), pepAd (SEQ ID NO:21), pepF (SEQ ID NO:23); and functionally homologous sequences thereto having at least 95% sequence identity.

8. The recombinant filamentous fungal cell of claim 1, wherein said filamentous fungal cell is an Aspergillus cell.
9. The recombinant filamentous fungal cell of claim 1, further comprising an inactivated pepA.

10. The recombinant filamentous fungal cell of claim 1, wherein the inactivated gene has been deleted.

11. The recombinant fungal cell of claim 1, wherein the inactivated gene has been disrupted.

12. The recombinant filamentous fungal cell of claim 1, further comprising an inactivated gene selected from pepB (SEQ ID NO:25), pepC (SEQ ID NO:27), pepD (SEQ ID NO:29) and homologous sequences and combinations thereof.

13. The recombinant filamentous fungal cell of claim 1 further comprising an introduced nucleic acid which encodes a protein of interest.

14. The recombinant filamentous fungal cell of claim 13, wherein the protein of interest is an enzyme.

15. The recombinant filamentous fungal cell of claim 13, wherein the protein of interest is a protease inhibitor.

16. The recombinant filamentous fungal cell of claim 13, wherein the protein of interest is an antibody of fragment thereof.

17. A method for producing a protein of interest in a filamentous fungal cell comprising
   a) obtaining a recombinant filamentous fungal cell according to claim 1,
   b) introducing a nucleic acid sequence which encodes a protein of interest into said recombinant fungal cell, and
   c) growing the recombinant fungal cell under suitable culture conditions for the expression and production of said protein of interest.

18. The method according to claim 17 further comprising recovering the protein of interest.

19. The method according to claim 18, wherein the expression of the protein of interest is enhanced compared to the expression of said protein of interest in a corresponding parent cell.
20. The method according to claim 19, wherein the filamentous fungal cell is *Aspergillus* cell.

21. A method for obtaining a protein of interest from an *Aspergillus* strain comprising
   a) obtaining a recombinant *Aspergillus* cell according to claim 8;
   b) transforming the recombinant *Aspergillus* cell with a nucleic acid sequence encoding
      a protein of interest;
   c) growing the transformed *Aspergillus* cell under suitable growth conditions to allow
      expression of the protein of interest, and
   d) recovering said protein.

22. The method according to claim 21, wherein the protein is a protease inhibitor.

23. The method according to claim 21, wherein the protein is an antibody or fragment
    thereof.

24. The method according to claim 21, wherein the protein is an enzyme.

25. The method according to claim 21, wherein the *Aspergillus* is *A. niger*.

26. An isolated nucleic acid encoding a protein having an amino acid sequence selected
    from the group of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10,
    SEQ ID NO:12 and functionally homologous sequences having at least 95% sequence identity
    thereto.

27. The isolated nucleic acid sequence of claim 26, wherein the nucleic acid sequence is
    selected from the group of sequences corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID
    NO:5, SEQ ID NO:7, SEQ ID NO:9 and SEQ ID NO:11.
FIG. 1A

CCCGGGCGAGTCATGACGCCTTTGAATTTAATGTGTTGAGGTGTCGCACTTCG
CGTCTCCTGGGCGGCTAGTTGAATTTAAATGAGCTCTGGCGGCATCGACTGGC
AGTGCAAGAAATAGCTCAAGGGCCAAACGCAGAGGAGGAGCTTTTGCAGGGAT
GCATGCGAGAGAGGCGCCCATCGGAGCCGGAAGCAGTAAGGTATACTGAAAT
AGCCATACGAGATCTCTGTCGACGAGCGAGGAGCTTTTGCGCCGCTGAAGAGCG
CGGGAGCGACACGTGACATATGCGCAGGGAGTTGACAGGTGCACTGAGGCG
CCAGCTGGAGTCGCGGATTACCGGGGAACTAAAACGTGAGTTATTAGGTGATTC
GCTCATCCCCTTATGCTACTAATGCTGGGACTGACGGCCACTTCCTCTCTCTGAGCTGA
TGACGACTGTCTGTCATCTCCTCTTCTCTTCTGCTCTACTCCCTACTGCGACC
CGGGCAATCTGGGGTGACCCGCGGAGCGAGCTGGCCAGTTCCCGCTGGGAGCAATGG
ATGAAATGCCCCCTGTAACGTGATGTTAGCAGACAGCCAGCTTTTGACCTCTAGTC
GTCAATGTCAGCTTCTACCAACTCAAGCTTGGTTTGTATTAGCTTCCCGGCGAGTCTATGT
TAAGGCTCTAGGCTCAGGCTAGTACCTCTCCTGGCTCACACTGTATAGGTACGAAATC
AGCCGTAACATCGATAGTATGCGCTGTCTACAAACCTTTCTATATCTTCCGAGC
TCGAATTCGACTTACTATTTCATGTTGCTCTCTTCTTGGCAGGAGATACGCGCCCTTGGAG
GAAATCATCGGGCGATGCGCCGGCCACTTTCTGTGGCTTTCTTGCTACGCGCATCGCCCT
TCTCCTCGTCTCTCGCAACTTTCTCTCCTCTCCTCGGACGCGCTCTCTCTCCA
GTCTGCTTACATCTGGAGTCTGCGCAACCGGGAAACCTGCTACCTACCTTTCTAGGAAATG
CTGCATTCACCCGCCCCATCTCTCTCCCTGTTCTCAGGCTTCTGCTGCGCCACGTCTGTACTTCT
TTCAACGATGTTTACCCTTTGCTCTCACTGGTCTGCAAGTCTCCTTTGCAGTCTCCCTATGTG
TGGGTAGCTCTGTTAGTGCTAGGGCCCGGAAAGGAGGACCAGCAGCGCTGCAACGTA
GGGAATTCGCGCTGTCTGCAACCCGAAAGTTGCGTAGCTATTTGTTGCAACCCACT
GGGGCAATTTACTGCGAAGTGAGCGCAGAAAGATCGTGCCGTAATCTGCGTCTGGGAAACGTC
GCATATATGTGTGACTGCCAACTACTCGCCGCCAACGGCGGAAGCATGAGAAGTTAA
TACTGTGACGATATAAGCAAGATACGCGCCGACAGACTTTGAGGAGATTTCTTATCTGTG
GCTGCTTTGCTCAGGGAGAGTGACTATTGATCTGCCCAACCCCGCCTGCGAGTGATTCG
FIG. 1B

GCTCTGGCCTCTCTATTGAAATGCAGAAGGAGAAGGAGCGTACAGAGGTCCCCGAGTTCA
ATAATCGGAGAGGAGAGAATCAAATTTTTGGCATCCAGCAAGGTGAGGAGTTGGGATCA
CTTGTGGGAGAGATGTCAGAAAAGGAGAGAGAGATCATCAAGGCAATAGCGGGGCG
CTACTGTCTGGGAGTTAGTAACAGGTGTGTGCTGTAAGAGAGACAGACTATCATGCTGCA
TCAATCAGACTAGTTGCTATCATGAAATACCAGGCAATTTTTTAGCTGATAATTTACA
TTTTCGAGAGGGAGGGAGGGAGGGCAGGACGATGGAAGGCGTCAGCCTCTCACA
TATGCAAGTAGAAATATCGTAGCATCCCAATTACCTGAGAGGGCCGGGCCCCAGTGTGCATC
CAGGTTGCAAGCTCAGAGTTTGGAATCTCCGCAATGTCTATCCCAAATAGCTGCTT
TCTACTTTCTCTTCTCCTCTTCTTCAATGCGCTACAGACTGCTATCCAGATTTCG
CCCTAGTTTCAAGCTCGATAGAGTTGGATCAGAATGATCGATCGATCTGACTTCTCTCTCTCTCC
CCGCGT5ACATAGCAACACGCGTATAGTTACCATGTAGAACAACCATGACAAATACTTCT
CTGCCTGAGCGTTGATCAACCAGTCAGAGACGACTTTGCGGATCAAAGACAGAT
AGTCCTATCTCTCTCACCATGAGACTAGCTATCTACACA
FIG. 2

MAAIWGNGGQAGQPLEQWFYEMPVPTRWTAATVATSVLVQCHVLTPFQLFYSFRAVY
VKSQYWRLFTTFLYPGPLNLDDLFFHVFFLQRYRLEESSGRPAPHFWSWLLFYHAMASLL
VLSPFLSPLFGLGTALSLSLVYIWSRRNPETRLSFLGMLVFTAPYPYLWPVLMAFSLVVG
VPKDEICGVVVGHVWYPNDVPVSLLHGHPFDPMPWVRLPESGPGERGTDAAANVNGE
FAAAAAPEVR
GGCTCAAGCACAAATGGGCTCAGTCTACACCGAGAAATACAACTCTTTACGTATACCTACTC
GCCAAATCCAGCGACAGAGATGAAACAGAGAAGCTACCGCTTTACACCAGATGGATTGT
CTAATAAGCAGCAAATGCGTGGCTCAATCATCTCAGAGTGCCCAAGAAGACCGGGCAG
TATCACCATTACGCTACTGTGAAAGACCTTCTCTGATTGTTTTGCTAGCCCTAC
GGCTGACAAACAGCTGACTCCGTCTGTAATCGCAACTTTGCGGGGTTTGGTTCGCGGGAAA
TCTTTCAAGGATAGGGGTCTATGCTCTCGAATGGAAGACTCTCTCTCTTCCCAGCGAG
TTCTCTGCAAAAGAAACCGTGCGGTACTGGGATGACATTCGACCCAGCCTGGCTGAGCT
GGCAACCAGCTGACATTGACTTTTGAGACTACTTTAGTGACCCATACAACTCAGGGTGAT
TTGGTTCATGCTCTCTGGAATCCCTGGAATTGTTGTTGATGGGAGCGCTGGTGGGCTGCT
GGCGCGCGAGCACAGCCCTGCTGACAGCAGCCCTCGCTGGGAGCGAGCAGCCAGCCTGCT
CCCGATAGCAAGAGCGACCACCTGTAGACCCACACCGACCGACGCTGGGAAACCCAGGGTCC
GGTATGGATGTCTGGGGGAGAGGACTCCGAGAAGAAGAGAAGCAGAAGCCGAGGAAACA
CATGGGCAAGAGAGGATCCGGGCGAGATGAGGCGACAGCCAGGCGGGGGAGAGGCGGAGGAGGAGGAT
GATTCCAGATTGTTTTTAAAGCGGTCTGATAACTCAAAAACAGGCACTAGATCAGGATGGTT
GACATTGTGAGACGGTCAGACATATATATATATTCCTCCCCATCTATCGAGGATAACGACGAC
AGAGCTCATGTAATCGGGGACTGAGAATACCCGTAACCGCTTCAATGACATGGCCCTG
CGGACTACTCCGACATGATCTGAGCGAGTAGACATCGACGCTATTCAATCTCGCCCTGC
TTGAAGTGCAAAAGCGGTTTTATGAGCAGCTGACTGAGATGATTAGCCCGATGTACCATAAC
GACAAATAAAATCTCCAAATAACTGTATAATATATCAACCGGAAAAATGAACACATACGCCC
AGAAATAAAACTCAAGATCATATTCTCCTTTATCATATGAGAAGAGCGGAGAAGACATATT
GCCGCTCAAGCACCACAAACATCCCCACCCGCTCAACCACATGAGATGGAAACCTCAAT
CAGCTAATCATATTGCTTCTCAGAGTGCCAACCTTTCTTTCTCTCTCCGATACATT
CCACCACCCTCTCCTGGAGTACAAAACGAGCTCTTAAAAAGTTTTTCTTATGTTCTCTATCT
CCTCAACGCCCCAGCGATACATAACCAACACCCACCTATATGAGTACCACCGGTGGTCCCTCA
CCCTCTCCGTTGGTCTCGACACTCTCTCTCCCCCTCTTTATCTAATCGCCAGATAGA
CAGCCAGACCCGACCCACCCAGCACCAAAAAACAAACCACATCTCTCTCCCACGAC
FIG. 3B

GCCGAGAATTAACCAATTAATACCAATACTTTCTAATAGAAAGAAGGCAGTG
ATCAATCAAAATGTCTCCCCAGCCAGAAACGCTCTTCACGTACAAAAACCTCTC
CACCAACCCCACAGGCTACCCCGCCCGGTACCCGAAAGATGACATGTGCTCCTC
GGGAAAGCAGTAAATCCAGGGGCTGAAAGTGACTCGGCTTCTCGAGGCGGCTGTTCGCGCC
GGTTGAGGTTCTCCATATGTAATACATCTCCTTCTTTTTTCCTCTTTCTCTGCTCTG
ATTTTCTTTGTCTTCAAGTCTTTCTGTTGATGCTTTGTTGAGCTTATGCTAACAGG
TTGGGACATACGTATAGTACCCCAACCGAAACGTGTGCAATTTCCATCCCTCCACC
CCGGTGACGCCCACATACGGATGCTTGAGAGGGCTGCTCTCCTATCCAGAGC
GTTGGAGAAGATTCTCATCTCCGTTATGATATGTTGTGGCGGACGAAATGAGTCTCTC
GGCGAACGTGGAGGCTCGAAGATGTGGAGGGAGGCGGAGGGGAGTATGAGCAGGAAGG
TGAGGGATGAGGTAGAAGGGGCTGCTGAGAACCCTCTCTCTCTTTAATTTGAGT
TGAAATGGTAAGGGGGCTGCTGCTGATAGACGCTGTGGTCGCTGTGGTGC
TCACTGTCTGTCTATACTCTGTGCTGAGAAGGAAATGTGCAGATGCTTTGAGATGTC
ATGGGTGCTTGGGTGGGCTTTGCTGTGGTCTTTCTCTTTCTCTCTCTCTCTCTACTTAT
ATATATTCTTTGATGCCACTTTCTAGGGGTAGACATGCGCCAGGAAATGCAATAG
CTTTTGTCCATATTATCGTCTTCTCTGTCTGTGATAAATTGTTAGGAGATGCTAATCT
CGAAATCGCCCTATAGTA
FIG. 4

MSSVAQKRLFIXIQXFXTNPPEGITAGPVTEEDMFHWEALIQGPEGTPFEGGVF
AAELKFPKDYPPLSPTMKFVGWHPNVYPNTVCISILHPPGDPNHYEHAS
ERWSPIQSVEKILISVMSMLAEPPNSANVEAAKMWREERRGEYERKVRDEVK
GLGL
FIG. 5A

GATATAAATGAACTGTCTGTGTCCACTCTCTCCGTGCTGAACCAGATCAACCCCTCCCTTCT
AACCAGCTGGCCTGTGGCTCCTCGCCCTGTTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT
CCCAATTTTCTCTCTACTCTCACTACCCCTCTGAGAGGTATCTACATGTCAAAATG
ATTGCTTTGGTGGTCGCCATCGCCGGGCTGACTCAGGGGTATCTCAGCTGCTC
TGATTGCTGCTCTCCTTGGCGTGTAAGGTCTCTCATTCAGCGGGGTGCTTGCATCGGTTCT
GCCTTCCCCTCAGCTTTATTTGTGTTGTTGTTCAACGAATCGGAAGTGGGCGATGACATA
TCAGCGAACAATCATATATATTCGGAATTACGGTTCTCGACAGCGACTATATATT
AATCAGTGTAGTCTCTGCGAACAACCTTTGGAAGTCGCGCCCTACGTCTGCAGATGTCATT
GTACGATTGGTCTTTTGAATGGAATCGGTTCTGCTATTCCGGGATCGAC
AACACATTTCCAGCTGATCCCCGGCATGTTCTCGCCCGATCGCTCTCGCATGTCAATA
CCAGTGCTCAATCGCGATACAGCCCTCGAGCTCCATACATCGGGAGGAGCGATTCC
ACATCATGGGAGAGTCGCCAGACTGCTGACCTTGAACAGGAATCTTGTGGAACGGGC
TGAGGGTGCAGCGATCGGTCTACCTGGTGGCAGCCAGGCGCTCTCGATGAACGACATGC
ACGCTGCTGACATCTCAATGGCAGGCAAGGGCGTATGTCTCTCAATTTGAGGTCAGCCTCC
GCAGCAGCAGGAAACGCTGCAGCCGGCTGCTGCTGAGTCGCTCAGAAGGAGAATGCAGA
TTACAGCACACTCCTCGCTGACCCCTGTAAGACCTCCCGCTGACATCGGAATCTCGTGTC
TATCAGGGCTGAGGAGAAGGGAGAGTCAGGGCGACTGATTACGCTGTTATATATCTGCA
TGTTTTTCATCTCTCGTTATCGCCATTATCTCTCGCGACTCGTCGCTTCCGTAGACATT
GGGCGAGAGCTCCAAATTACCTTCTCTCATCTCTCATGATCTTTAATCTTGGCAATTTTCT
GTCAAATCTCTCAAGCTCTTCCTGATTGGCAATTTACAGCCGGACTCGTACGTGATGCTGCT
ACGAATCGCATACCCAGTAGGGGACGCCAGCGCTAGCGCAACACGCGCGCTCCCAT
CCAGTTATTTGTAGCTTGAGATGAGACAGTTGGGACCGCGAGTCGAAATCGCGCTCGCG
AAAGGCTAGCAGGGCCCTCCTCGCGCATACGGTTTGGCAGAAAGTGTGTAAGTACA
TGATTTACAGGAATGGGGTATTTCTCTGATTGGCAAGAGTCCGCTAAATTTCGAGA
GGATTTACCCCGACTTGTTTGCTACTGGCCAGCGCATCAGAAAATACATGCGCAGCTACC
AAGGCTTTGGCAGGATGGAACATATTACGCGGATCCCCACGGCGCTAGCTACA
FIG. 5C

CGTTCTATACGGACACTTGGACTCCCTTACCCTCGAGTGAACCTTGAAGTACCGGGGTGCAA
CGGCAGCGCTACTCAGCACAACCTCACCAGCTTTTATGCACCCCCCTACGTGTGATAAGCCAA
TCTCTGAACTGCGCAAAGCCCTCCGCTCCTCGACCCTTGAAGATACGGAAAACCTGCA
GCGCTGCCGCTGAGCTCTAGCTCAGTCTCGAGTTTACAGTTTTTGAGCCGAGATTTACAGGGCGAT
GGACGATATGAGGACCTTGCAAGCAGAGCATTTCTGCGCCCTGGGAGAAGGAGGAGAGTGAA
TATTGGACGTATTGCTTGGCCTAGTGCAGTGAGCTTGGGCTTATCATTCTTCTATA
CCGGGTAGCTCAAATCAAGCAGTGGATTCAGTTGAGATGGCTACATCACTACGCAGTCCAG
ATCGGCGCAAGATTTGATAGCTTTTTCTGAAGATGCTTCTTCAAGTCTACGTACTACTCTC
GTCAAGGGGACGGCCCGCCGCGCAATCAGCAGCCGGTGGCATTGCCTGGGCACTATT
TCCTACCACTTTGAGAATACGCACCTCCGCGGGATGCCTTTCTGAAGTCTGGAGAGAG
TCTCATGCTCTCAAAAGCTCACCCTATACCCGGGAGAGGGCCATACAGCAGCGACAGTCT
GATGCAGGAGCAGATCTTACTCACCAGGAGGCACTCGGCTTTTTGAGATGCAAGATGCTTTAGCG
CCTTCTACCCGGGACCTTTACTTTAGCGGGGAAGTGAGATGAGCCATTTGGCATCCAT
CTTCTGACGACGGACAGTCTGGACATCGATTTCGCCTTCTGAGCAGTGAATTCGCC
CAGCCGGGACATTGACAGGCGCTTCTCTGATGTAGGCTGGCCCTGGAGATCGTGGGAAAT
CTACTTACTACCTACTACACGGAGCGACAAAGGGAACCCCTTGATATGCTGAGGAGATG
GTACTGGGGGATTGAAAGGCGGATGCTGGACCAAGTGCGGTTGGGTGATGATTACAGGA
CGTTGGGAATGGCGAGCTCAATGACGGCATGAGAGCTCTTCTCTGGTGAAACCTGCCA
AGTACATGTTTCTGCTGATCAGATTGATTTTGGATCATCCCTCTAAAAGCTAGACCCAGCTTCC
GTCTTCTCCACGAGGCCCACCCCTCTAATATCCCCAAGAACAAACAGCGGACCAGCGGC
TGAGCGCAAGCAGGTACAGTGGGTGGGAGAGCGAGGTTTGGAATGCCCAGAAGCACAC
CTCAGCTCTCAACGCTGGGGGTTCATCCACTGCGACGTTGCGATCTGTCTCCAGC
GCCAACCTGCGACGCCTACACTCTAGTGGCCTAGGATAGGGCTAGTGCGGAGGGCCCTCTTCT
GGATTACGCTGGGACCACCGAGCGATACGTGCTGGACTTGTGCTGCGCCACAACTC
ACACATTCTCTCCATGAGACATTGGCCCTCCAGAAGCTTGGCCATTAAAGCAACCCAGCGGC
CGATGCAAAATCCGCTCCTACGCTCGACATTCTTTTCCCGCGCTTTCCGCGTATGCTGGCT
GGGCTCTGGGATCTGGGAAAGGTGGCAGTTGGCATCTTTATTAAAGGCAATCGGAGGCCC
TACGACTAAGTATTGGTTCAGGATGTCCCCTGTGCAAGGGAATCCGCGAGGCGCAGAGATG
FIG. 5E

ATGTGTCCCTGATTGAGGAGGTGAGGACCTGGGCTGGAGGGTATATAATA
CCTGCTTTCTTTTCTTTTTTTTTATAAGGGTAGAGAGAGAGAGGATCTCTAGATCTGAA
TCAATAACGGAGATTTACTGTGATTTTGAATTATACATATACATATTTTGAGCTGGTTCT
GGTACTATATATCCGGAACCTCATTGAGTCCTATGATTTACTCATCTAACCCTTCTCTCTC
GAGTAATATATCTATTATAACAGTCCTATCCCCTCTACAATCTACTATTTACACTACAAACC
CACTACGAACCAACTCAATGCTATCCACACAGACCTCGAGCGAGACCACACACCCCCC
TCCACCTCTACATCAACATCACCACGGCCACCTCCCGACGCAACACACACACACCGCCAC
AACCACCTCACCGAATCTCACCTCCACAAACCCCGTCCC
FIG. 6

RTQVNLTSYLYIPGFAGIGADRLTPSIFRKETEHMFYHGFENYLEHAFPEDELRPLTCR
PLVRDEHPAHANLNDVGLGYSLLSLSLSSPQGPQKADYFQNQGRGRGFDM
DSKQVFETVIRGLGGLLSAHFLAVGDLFPTGYNPPETEANFARAWDKHSYPEGSRGIE
WKDGFEVYDGQQLRLAVDLANRLPAFYTDGLPYPRNLKYGVRQFPRYANSPFNAAPT
CDKANPESCQKPRRSSSTFETTETCAGSLVLEFTVLSRLTGDRGEYELAKRAWAMW
ARRSDIGLIGSGIDAEGERVWSYRTVQIKHVLHNLNHTTTSQIGAGIDSFPEYAFKS
YVLLSSGERPANTSSPSWHDPLSEYEHSDADAFLKWEKSHASIKRHLHRYEGH
QHPHRLQGDIIFGATRAFWIDSLSAFYPLLTLAGEVDEAIGIHLITTAWTRFSGLPE
RWNVATGDIEQGLSWYGRPEFVSTEPSTYLYRATKDPWYLHVEMVLRLDKRRCWTKCGW
AGIQDVRNGELNDRMESFLGETAKYMFLLYDFHDHPNKLQPDPFVFSTEGHPILIPKNS
TAQRAERKQVPVVESEGLTCPTAPQPPGLESTAAARCDFHAANLRLHLMPSRGLA
BGPLDDYARDHPSVSVDDLSSPNTYDFPWPTELVPFFNASAPMTIRPTLDISPAL
PGMVVGPGSLEERVDGIFIAIGGLRSMVQDVQPQGSGSAESDFRQVQINNVPLGK
DEKVYLLREETDVPDLTPFNPRTVDRAMIDVIDIVIPEIIIRRNDSDSHEPAAPRR
ANGAIHVSSSVDSKQGVADVSTSSMKTVLSLLVNTLTLLRDEVQGPQSLPPQKATSIL
RLLPAATTGLGSAPLPVEDATTVSITGDPQKRTWNSIYFADELCRNILREVAQ
NHQLVILKRGCNFSQKLDNIAAYPSRYALKLQIVIVSYDDEVVBEQREESDTITTPG
LAAVRAEPYLVPDLEETQMTAAAGVPRRQLSSVMMGQGQETYELLRQATVGKIRRYS
VRSQGVPINLNYIL
FIG. 7D

CTCGGGAGCAAGTGGAAGATGGTATCGCCTAGGAGGCAGCCTACTGGAACGCAGGCA
TGACGGCAAGGGAGGGGATCGATGTGGAGGACATAGCCAGGAGA
FIG. 8

LSDSFAFCSLPRPPPAPPPVAAPPVTDSSQLRREDAAEHPLSPPTKFLKSQAVREDGLKA
PPPVMHYNLNSSTSESIIKGSERVLLILTPARFQYQEFWDNVVKLSTYPSHELISIGFIVP
NTKDGHAATLAEISAISETQSGPIDSRFRHQHPSPGRLRPASRGREARAPQNVQPEST
SLLFITTLGPISTSWVLMLDSDIVGTATLQDLTAHNRPVIVPNCFQRYYNKDAKKMDVR
PYDFNSWIDSSTAEQLAETMGPDEISSREKLDCPPSGTEAHKAKFRGPRPSREIELD
GVGTTALLVKADVHRDGAMPPAPPPHYHLVETEGPAKMAKRLGYSVYGLPDDYFVRSPTIS
HPIEPTNAMATGVSL
FIG. 9A

CGAAATGCTGATAATGTTCCCGCTTTTGGCGACTGATCAGTTTTTATTCAAAGGACAT
GTCATGCTATTCCTCTCTCTGGTGAGCTCTGACTCTGGACCCATATTGATCTGATCT
TAATTATTTTCTAGCTCTTTTTAAATTTTCTGGGCGATACGATCTGTATCCTACT
AACGTATCTCTCTACACCTCGCTCCAAACCTCGTCTCTTTTTTTCCATCTCTGTCGCC
CTGTGTCCTCCCAAGTGGGCGGCGCCCTGTTTGAAAGAGACATCTCCCATGGACCTC
CACAGGGCGCTCTGCGAAGCAGAATCCCAATCAAGCGGCGCCTGCTGCTGCC
GCCCGTCATGCTCTTTGCTGCGATCGCCCTCTCTCTCCACCCCGGCGAGCGAGTGCT
GTCTGGACACTACTGCTCCAGTGGCTCTACACGCCTACAATGCTTCTCTGTCGCA
AATCCCATAGGCACCGACGCGATCTCTCTGGGCGCGCCGCCAAGCGGAAAGAGCGAG
ACGAGCTACCGCTCTCTCTGAGCAGAAGACAGCGATTTCTCTCGCGCTCAAGGCC
AGCTCTCCGCACCTTTTGGCGCGCTTGCTCTGGGTGGGCTTCGACATGGTTATGTC
GATAAGGATGAGTTGGCGGGGCGCGCGCTCTCTCTCAGACTCGGCTGGCTGCTG
GGTTTGGCGAGTGCTGGTCGCTCCGCGACAGAATCGGCTGGCTGCTGATCTTTT
TTGGCTTGGCTGATATATGTGGTTTGGGGCGGTGTAAGTGAAATCGCATACAAATGGGAAGA
ACGCTGCAGGTAGCTGACCTTGTTACAGCTCTCTCTCCAGAAGATCGTCCGGCTCTCGC
AGCCGGTGCGCGCAATCTCGTGTATCTGCGAATCCCAACATAGAAGCGGGCGGTGATG
AGTGGAAGGGAGCGCGCAATGGCGAGATCGCAGCGCAACAGCATAATGGAAAGCGAAGA
TATGTTGAACGAAGTGGGCTACGAGTTGGGAAACCCTCAATGTTGGCGGAAAGAGCGGTA
TTCACACGAAATGCCGCGAGCTGGGAGAAGGTGAAACCTATCCGGGAGACGATCGGA
AGCATCCCGATGCTAATGTGTTGCTGGTTGGATATTGATGGCTCCGGCTCATTGAGCTGACAT
CTTGTCGATTTTGGGTGGCTCGAACCTAGACACTGCTGATCATGGAAATCTCCTACTGT
TACAGGACCATATAATGGCAACCGCTTTGGAATGAAATCTCCGGGAGACGATCGGA
AACCATATGGAAATCGCACCAGCGGCAGCAGCTATCTGGACAGGTGTCTCTGTTT
GCCAAACGGGAAGGCGGACCATCTGCTACATGATTGTATTTGGTGGCAAGACTGTG
GGTCTTCAACCTGGGCTCTTTTCTCATGGCGGCTCCCTCTGGGCGACCGGCGCTGCTG
GGCTGTCGGAGACATGATGCGAAGGAAACATATGGAGTGGGAGCATAAAGAGCA
GGATCGGATGAAATACCTCTATGGCAGCGAGCGGCCTGGCTGCCAGCGCACTG


FIG. 9B

TCCCTCAACGCTATATCAACTCTGTAACCCTCGAGGGCTTGGGACGAGAAATGACCCG
AAATGTCACCTACAGAGATGAAGAGAAGACTCTTCCTGTGCAAATAGGGCTGGGCTAGTA
TGAGACCGACTCTGCTGCGGAGATGATCCAAGATCGTAAAGAATGCAAGCGATCTGAA
TGACATGTGGGAGCGATGAGAACAGATGAAACTCTACTCACTGTACGAGACTGATATCAGCT
GGCAGGAACGCAATGTAATGAAAGAAGACTCGTCAGCTGAGACTATGGAACATGAACTGAG
TTTGGCGAGGCGTTTGTGTTCGACGATGATGTGAGTTGAGCTGAGACTCGAGATACGTACG
ACCTCTTATCTGTCCTCCTGCAAATATCAAAAGACTTACGAAACTAAGTATATATAAATT
TAGTGTGCCACACACTACGTATAACCCGCTTTAG7AATTATCTCACCAGCCCACGTGAC
ACAATGAACTTAGACACCCGACCGCCCAACACCCCAACCAATCAAACACGGCAGTTT
GACATCTGCTACTCTCGGAGCTCCTGGGAACTTGAACTGTACATCGACTCGACTCTCTAC
GTACGTACTCTCCCTCCTGGCCTCCTCCACACACACCGACACTCTCCCTCCCCGAA
GCAACAAACCAAGGAAAGACCCGAATGCGCAGCCCTCCCAATCGCCTCCTCTCCTGGCA
AGAACACCCACGCCGCGCGACCTACGGCGCTCGACTTCGAGACAACGTGCGTGC
CACACGCGCCGCGACCGCATTATCCGTGAAACAATGGGTCCGCAGTATGATGCTCGTCT
GGTCCGGGAGAATTGGGAAGATGTTTAGTGCAGCCTGGGGAAGCTTTAATCATTGAGAAGT
GGTTTTTGAAGGGGTGAGTTTATACCTGTTCTTCTTTCTTTTCTTTTCTATCCCGGGAGCC
TTTGGGATGGTGCGGCTAGCGAGATAGGAGTTGAGAATATTGCGTAAATATGCTTC
TCTTTCCGGTGATAGAAAGTACCTCGAGTTGCTGGCGGAGCCGCAAGATCAGGGGTA
CTTGTCCAGGAAGAAACTACTTTGCTGGGGAGGGAACAAAGCTCGTGATAATTTGC
TCGGTGGATTGAAATCGAAATTGGGGTTGCGAGTTTCTCTGTGGTATGTATATGTAATAC
AAATATGCAATTGTGTTTTCTTTCTACTTCTTTTCTTTTCTTTCTTTCTGGTTTTTGAT
GGGAGTTAGGGTGAGTGCTGGTTTTGCAAGTCGCCCGGCTGTGATTGTGATGA
TATGGCTCCGAATGGGGTGGAGTTGCGCTGCTCGGTTTATATACTGGGTGTTGAGG
GCCAGTGAGGGCGTGCTGCTGCTGTTTATATCAGGTGATGAGTACTGGGAGTTGAGG
TGGTAAATATGCTTTGTGCTGATACATGTATAGGAAAGAAGCCTCGAAGCTCGAAGCTC
FIG. 9C
GAAGCTCGAAGCTGAGATCAATAATAGGACTGTCGTCCGTCTCCGGTACTGTCCCG
CGTATACACACGGCCCACTGGGTGCTTCTCTGCTACTGTTCCTTCGACATCATCTTCC
GTCCAAAGTCGTCCTCCACCGGCGCGCAAGCATGCGCACAGACTCGCTCTCATCCAGC
GAGGGAATCTTTGTATGGTGCTGCGCCAGAATGGAATCTGACCGGAGGGAAGTTGCG
AAGGAAGGCTCGTAAAGGGGTGGGATCAAGCTGTTGTGATGTGTTCTCGAGTCAACAT
CATCTCAAAAACATGACGTGAACTCACAACCAGTATCATATTACATCGGACACCCACTTGT
CTCCAACATCTCAATCAGTCTCAGTCTCTCCGGCTCCGGCTTTTCTCATCCAGGCTCTCC
GTCGCCCCTTCTCGTCTCGATCCGACTGCCGCATTTTCTCCACATTCCCGCTTGGCT
CTCTTTGTTTCCGCTGACGCTGTTCCAGCTCCCACGGACCCCAATGCCCCCTTGCTGCT
CTCTAGCCATGCAATGCCCTCCCGATTTCCAGACACATTCCGCAATCCACCCGAAAG
AACGGCGAGGCTCTCGCCCGCCGCAACCCGTCCGACTCCGATACCTACCACTGACGATC
ATCATCAGATTCCCAACGCCCCCTCTGAGCGGACGATCCCGAGGCCCAGCAGACGCTTGTT
CCGCTATACTCGCGACCCGCTATAACTCAATCTACGGAAGCTCGCCGCCGGCCTGGG
ATCTCCGGCGCCGAGCATCCAGTCTCGATGATGGATTGGCGCCTTGAATACAGGC
CCGAGCTAGGAATCGTCTCTCCAGAAGAGCGGAGGAGGTTGCTTCTCGAGGCGAGCAC
AGGACAGGGCGCTTGGGTGGCGGAGGTCAAGTCCTCGGAGGATGGATATTCTGG
TTATTTGAGATGAATGGCTGCTAGTTGCGAAGAGCGGAGGAGGAGGAGGAGGAGCAGT
TACGGCGCTGGCTGGAGGGATGTTGTTGCTGGCTGGAGGTGGATGTCGCGGCGCTGGCTT
CCGAGTCTGGCGCCGAGATTGTAAGTGGTGCGGAGGTACCGCAATCCGGCGAGGACGAG
TTAGGACACTAGCTGAGGGGTGGGAGGACAAAGGTCTACTCGGACGAGCGCTTCC
TGGATCGGCTGGTGGTGTCCGGCTCGTGTGAGGAGAGCGCGT
FIG. 11A

CTATATGCTGCTTACACTGATCTGCTTTGATCGTGCAGGAGCTTGAGCCAGAGACG
GCTGCGTCTACTACATAACACAGCTGCTCGCAAGCTCATTTGCTGCTGACATCCAC
CTAATTAGCGACATCTTCTCATAACACACGGAGAGCTGACCTAACCTCGGAAGCTGCTACTC
GGCGTCTGCTTTGATTACCTTTCATGCTCTACCGATCGTCGTCAACTCCTACACCAACGAGATTAGATCGTGA
CTTCCCGTCTCCGGGATAATTCTTCTCCGGCAGGACACCCCTCTACAAACAGCTGCCGCTGACC
CCGCAACGGAGGCGCCCTTGCGGAGTGGCTTGGCTACATCCCAGGGTTATCGTACGAGC
AGCAAGTTCCGCGGCTTCCATCTGGCACAGCTGGAAATATGCGGCGTCTAGATGTGCTT
TCAGCGAAAGCTGCTGAGTCCCCAGCTGCTGAGTCTGGGACACTTCATCCCAGGATTGTC
ACGGATGGTGTTCGGGACTCTGCTGCAAGGCTGACATGCTTGGCCATGACCTGGGGT
TCCTGATGTGTTGGAGGTTACGATGTGGCTCCGCTGTGAAATCTACAGCATCTGATAGCTC
TTGAAACGCGCGCTGCTGACTGGCGGACTGCAATCTGGGACTGGGACAGCTGGAGGGA
GTTGGCATTTGAGGCGATCTCCTGGAGGGCGGACCTGGGATGCTGGCCAGAGCTGGGA
TCAGCTTCTGCAATTGCACCACATCCAAGCCAAACCGGACACCCCACATGCGCGATATT
GCTTCCATTACATCAGGAGATCTTGGAGGATAGAAGAGCCGCTATTCTAAAGACTGG
CAAGATTGGCAAGACCGTGTATGGATTACACTGCGGGCCAGGGCGTGGGAGGATGCGGTTT
TCCGGTATTTCACAGTACACAGAAGTACCTTCAAACATGGGGCGCTGAGCTGGTAACAC
ACCTATGAGGACTTTACGGTGAGGATATAAGAAGGCTCAGAGATGGTGTTGCTTT
GCCCATCACCAGGTTCTCGAGCGGGAGTGCCACCATATGGGTTCCGGGAGATGTTGATGATC
CCATGCGATTCGTGGAAGCATATCAATTCTTGAGGAGTATGGGCGCCCTCAATTGCTCTATTGGTT
GACTCAAAAGCATACACGGCGCAACAGGAGATTTGGAAGGATGACTCCCTCTCTATAAGCGGT
CATATTAAACTGCTTTTACATTACACCTTACATAGTACGCTGGGTTGGAAATCCC
ACTGCTTTCGACAGGACTTGGCAACAGCGACGTCTTTAGACAGCTGGGATATAAGCATA
TAGCATAGGTTGCATATCATACGACTCTTTGCACTTTGCTGCTGGAACACGAGCAAGGGC
CCTTCTAGGGCACCCTACACAAACAACCCTCGCAGCATTCCACACACATTTGCCCTCTCAA
ACTCATTATGCCCCTTGGCTCCACCCAGCTAGCCATGTCGCTGTAACGAAAATGCTG
GCTTCCGGAGAGCTGGCTATTGGCTTGGCTACACTGACTCAGGGCTTCCAGCTTACGCT
FIG. 11C

GATCCTATCGACGAAATCCCTCATCCACGTCATATTTTGCGGCGACAAAATCTTTCATCGGA
GATGGGGCTAGCGGCCGTTGCTATTTCTGCGAGCAATGCCCAGTTGTGCAATGACACC
AGAGAGATCCCTAGGGCCTGCTGCGGCCCTCTGATACATCGTCGACTGTATCCGCGTCTTT
GCCACCTTCTCACCATGCAAGGCCGTGGTTGCGCAATTTGATTGAGAGGACTACCATCGGAT
ATGAAATCCCGTTACAGAAGGCCTCCTTCCAGCCTGCGCTGAGAGATCAGAGG
CCACACGATAGTCGAGAACGACCGCAGCAGCAGACATCGACAGAAAGTGCTGACGATCAGC
ATGCGCTGTTCCATCTATCTGAGCTTGACCGTGCTGACTTTGGTGTAGTGGACGCGAA
GAGCTAACCACCTTGGCAATGTCCTACCCCATCTAAGGGCTGCTCTAGAAATCGGAT
CGAGCGAGATAGCCGCTACTACCCCTAGGATAAAACTCATTGATAAACCCTGCTTTATT
ATCACAACCTTTGTGCCTCGAATTTTATGTCGATTTGTTAATATGGAAGCTTCTGGG
TCAGATACAAAGGGTGACCGAGACTCTTTGTCATTTACTATTGATTTTGGTAAACCGTGAAGCAT
ACTGCTGATGACGATCCTTGGTGTATATGGCTTTGGTAAACTATGATTTGAGATCGGTA
GAGTGAGAGTTGAGACATGCTTTTCTGCGCACAAGAAAGGCCTTGGACTGAGATCCTCT
GTTCACATCTGAAATCTGCTGTATGCAAATCATGATTAGATAATAGATCGGAT
GAAGCAAGAATGGCTGTTGTTGGCTGGAATCCTGAGTGAATATTGGGT
TTGGGGGGGGGGGAAAGAGGAGTTAGTATGATGAGAGCTTTGACTGGAGTTTTATAA
CCCGAGTACAGCAGAACCACTCCACCTCTGTCATACACTCTCTTGTGACCAAGCCAGGAG
TTTCATGTTATGGGAGAAGCTCCAGGCCTCGCGATTCCGGCTCGATAGCTCGACTGCTACGT
AGTCAATCACTACTTACTCTCCGCAGTCCGGGGTTTACTCCGATGCGCTCCACAGTACGC
CTTTATTGTCGCAAATTAATGTTGGTACCTCGGAAATTATGGCCACTTTTACCTCCG
CATATGATACCTTTGGTGTAATTTCACAAACAGAAGTGTCAATTGGCTGTGTCAGACTCAC
TCCAAAGCAGATCATATTGAGACTTGGCATTGTTGCTGATTTCAATTGCTGATAC
GCAACCGCGTCGTCGACCTGCGCTTCCGCTGATTAGAGACTCTCAAAGTCAAGATC
TCACATCACCCTTTGGCTTCCGGTAGGCATTTCTCTGCGCAGATAGCATTTTACTCCG
GTAGTGTCAGCCCTTTTGGAATTGCGTGAACAAAGTGGAGTATGATCTTTTTCGGCGACGACAC
CTGCACAGCTCGTGTCACCGAGATTAGCGGCCATTTGGAAATTATTTTAGTGAAGTCCAGT
FIG. 11E

CTTGGCATAGAATTGCGATGAAAGGGTGAATCCAGCTTGCTGCCCCATGAAAAGCATG
TTGAGAAATCTCAGCTGCATATATATCCATTTTGTTCACGATCTTCAAGCATGACTTTTTGCT
TATTGAAAAATCCGGACCCACAGTCTATCTAGGAGTTCAAGCATTTTCGAGACAATGGG
GTGCCGCCCTATTACAAGACGCGGTACCAGAGTACATCCTCGCTAACCTGTGCCCGATTC
TGTCGCACGAAAGATTTGGACGCAAGACATCGACATGGAGTCCATTATGAAGCGCTCC
TACGAACAACCTTTCTACCAGCAGAAATCGTATCATGACTTTAGATCTCGATCTA
FIG. 12

MLTFRKSLLLAAALLITFIVYLRSHTASSLPSPTSSAGHLNYQDYDGHADNERKGGTR
DTPQQLPLTPPSAPLARLRLYHPYPDLAEAKFPFAFIWQTWKYPSSMFLESRLDPESS
WSELHPGFVHEVVPPDDTQRHLIKLYGAVPDVFEAYDAMPLPVKADFFRYLILLARGG
IYSDIDTTALKPASDLPAELDLATVGAVVVGIEADPDMPDWHYARRIQFCQWTTIQAK
PGHPIMRDIVSYITEETLRKAGILKKTGMDKTVMEYTGPAWTDAVFRFNYFNDPEYFN
IEFGSTLNITYEDFTGQEGYKKVGDDVVLPLTSFSGVGQMGADVDPMAFVKHHFEG
MPFQLLLDDLTSKLTRQPGTWKDDSSL
FIG. 13B

CTTGCGTAGATTGTCTCGAAGGGAGGTTGCTCTATTACAAAAAGACGGCGGTTGAGA
TTTCTACGCGATTCGCGCTAGGGAAGGTTGCTATAGGAAAGGTTTTCTGGCAATTAGTGTA
ACATGGTGGTCTGGGAGACGGGAAATACGTGGTCTTTCTCGAAGGACAGAGCGGGCTGT
CCCTGAATTTTCCGCTGCAACTACTCTGTCAACGCGCATCTGGAACAGCCGACTCCCGAGG
GGCTGGAGGATTACCCAGAGTCAGGAGATCAAGTACCCCAAAGGCTGCGCTCCACAC
CCCGTGGTCAGTCTCGAGTTCCTCTGACGTTGAGAACAAGAGTCTTCTCTGATCGAGA
ACCGGATGATTTCGGAGATGACAGTCGCACTGCTATGAGGATGTGGCTGCCGAGA
GGAAGATATCTTTGTGCACGCAAACCACACAAACTTGTAATCTGAGATCGAGATTCGGCTGAC
TGCTGGTGGTAGAGGCTACGCAGACTACAGTTCTACCCAGAGATCAGCCCAAGTAGGGCC
GCCCTCATGATTGGATCTCGAATCTGATAGCAGCTCCGAGGGATTACGAGACACTCGGGTAC
TTCCAGGCCCTGGACAACTCAGAACCCATTCCTCTCACCAGGGTTGGTGAAGATGTT
GAGCGGCACACCGCGTTGGACTTGGCAAAAGCCATCTGTACCTTTCTACAAAGAGG
AAATCCCCACTGAGCCACACCTCAGTACCAGGTTGAAATGCAACGAGTCACACCTACGGCCT
CTAACAAGACACCTCCAAAGCGGGCTACTAGCAGTGATCTCTCTCTACCGGAACCGGCTA
CGCCCTGCTCAGCTACGGAGGTTCTTCCATTCATGGCAAGCAGTCTGACACCCGAGA
CCGACGAGCTCGAGTAGGGAGAGACCACGTAAGAACAACCGCTCTGGCAGGTATTGTT
GACTCTATACCCCTTCCACTGAGATCTACAGAAGCTAGCAGTCTCGGCTACGAGGCGGCAC
CGTTTTCTCTGCGGGCAAAACCGCGCTGACGTCCTCGGCGGCAACTAGGGCTACTACGAG
CTTACGACAAATCACCACGGCGAACCTTCTGGGGGAGAAGCTTACGCTGATGAAACC
CCGATGGCCATCTGGGCTGAGTACGGACCATCTAGACATGAACTAGATAGAAACA
AGATGCCGGGCAGCTTTCCCAGTACGCGCATGCCGCTACCCCTGTGACTGGCGAC
ATTATGGTAGCCCTCTTAAACCTCCTCCTCTCCTATAAACTCACACCTAATAAAT
AAATAGACTGGATCTACACGGAAGCCTGACATGCAACCAGCCCGACCACAAACCGGC
FIG 13C

TACGACAAACACCTCCATAACCGACATGACCAGCTTCTCCAACAAACCGTGCGATTCTTCGTT
CATCCACGGCGCCTCGGACAGAAGCTCCACATTCAAAACACGCTGTCCTCGTGAGATA
AACTGGAACCTGGGCGGGCGTGCGAGAACTACGATTDTGATTTCTATCCAGATTCAGATCAT
ATGATCAACTTTCCACAAATGCGATAGGATGTTTTATGAGCGTGAGCCCCCCTCCCTCC
CCGATCCCCGTGATTCTCAAGTAGGGGTGTAATTTGAGACATGTACTGATGATATTTGATA
ATAGGACTATCGAGCTGGCTCGTCAACGCTTTTCAACGATGAATGGCATCGCATAGCGGA
TCCGTCCCCAGATGACCTCAATGTTGGGAAGGTTGAAGGAGGTCGTTGCGGATCGTTGAT
ATGGAATTGAAATTGGATTGGTATAGTAGCTACATACATATATATATGATGTTCCGGGT
CATATCTAGTTCTACATACTACATAGCATGATACGTATGATGAGCTCAAAACGGG
TTTTCTATTTCATATAGTTACTCATCTACATACGGAAGGGAAGTACTTTAAATCGCATT
AAAGCATTACGTAAGTAGATGATTATTTTCATATCACCATGCAACTGAACAACAATCAA
CAAACATCCCAACATCTCTATGTGCTATGGATTTTCAGCTCAAACCAACATCAACACTC
AACACCAACATCTGTAATGGGCTATAGCAAG
FIG 15A

ATGGGAGCTCTTCAAGTGGCTGCTCATCAGGCTCTGGCTGCTGGCGCTCAGCAGTGTCTACGCTT
GACCCCGGAGTAAGGATATCTCAAATCTATCTGGAATTGCCATATATCGTGATAGCTAACCC
AGCTTACCTGATAGGATGATCGTGGCCCATCGGAGGACGAAACGAGTAGTAACTACCAAAACC
CCTCGGTTGTAATGCCCATTGCGCTGACCCTGTTGACCTACGATATATCTGCTGCTGAGACAC
GAGAAGGACACGGGTATATGATCGTGCTGGCTGAGACGATGTTGATGACTATCTTCTGAGAG
CACTGGTGAGCTGGTCGATCGACCTGAATCGGGCCAGAGCCACACCATCTCCTACCGATGATA
CGATATCGGAGGAGATACACTGCTATGCTGTGGGATGCTGATGATTGCAGACTCTTTCTCTGAC
AGCAGCAAAGCGGCGAGTTCCGTGATGCTGAGTCGATGTTGAGATGGATACATCAGCTACAC
TGCTAAATGCGATTCTGGTGTCGACGCTTTTAACCATGCTCTGAGAATACGTATCTACACCC
TGCTCAGTCTACAAGGCGACTCTCTCTGCTGCGCTCTCCTGCGACATCAAATCCAGGCTG
GACAGATTCCTCGGCACCTCGATTTTCATCTCTTGAGAAAGTCTATCCCAACGGACACG
CATACATTCACATCCTGAGAAGGTCTTTGCACATCCAGCAGCCCAGCATCTAGACAGCACTC
TTTTGCGGCGACTGGGCACTTTTACCTGAGACCACCGCCCTCCACGCTGGATTTCTCTGCGTAC
CTGCAAAGGTCTCAACGCGACCGAGCGATGTTCTAATATTTATCTCTTCTACGGGGGATTGA
CGAACCTGGTTACACCAATCGGAGGGTCGCCAAGACCCATTCCTCCTTTTTGAGGCAAC
GACGACTTGACCTGCTGCAAGGCTGAAAATGGTTTACCTTTCAACAGCAACACGCGCAGA
GCTGCTCTTGTCAACAAACCGGCTGCTATGCTATCCTGTCGCCAACACCGACGGGCTCTG
CGACTGCTTTTCTGCATACGCGCCGCTGATGCCTGCAACCCCCGAGGGAGTGGTGAAGGA
GAATCCATAATACCCGTTTCTTCCCTGTAAGGCACAAAAATACGCTAACCCAAAATGGC
AATCATTACATAAGGACTCGGACCAGGCGACATGCTATACGTATACTCCATACGGGATGACA
CTATACCCCCTTGGAAAGGACTGGGACCGACATCCATCGCTGGACATGGGTGGTGAT
GGAGCAAACCTCGTGCCTGGCAAGCCAAAGATCTAGGACACGAAACTTTTCGCCACATCC
AGGGGAGTGCAGGGGACATTTCAAGCCACGGAACCTTCACCAGGCGCGGGTGCTGTCG
TCAATACGTTCTACTCAACTCTCTCTTGGTCACTGAGGGCCTCTCCTGGAACAGCT
GAGCGCTACCGCCGGCAGCTTACGCGCTCTTTTCATGGGCTGATCAACACACTGCGCTCAAGGAC
GAGATCGACCCCGGACTTAGCGCCCTAGTTCTCCGACTTTGGAAGAGTTCTACTTTGA
CGCAACTGGAGATCCGTTAATCTATCCCCCTCCCTCCCACCCCAACCACATCACAAACAT
ACTAAACTCAACCGAGCTCCAAGGATGATGACCTAACCACCCCAAGACTTCGACTCATC
FIG. 17A

CTATGGACACTTCTTCTTTCTTCTTTCCCTATCCCTACCCCATCCCGCAGGTCAGAATGAA
GATGCGTTCTCCCTGGTTCTCTCCTGAGAGTCTCCAGAGTACTGAGTGGCTGATCTACAG
GATTGGTTGATGATCCACCCTGGGCTCTTGTAACCTAGGCTCTCAAGATGCGTCTCAGTGCTAG
CACCTGGACACCTTTCTGCTTCTACAATTTGCTTTCTGAGTCTCTTCTTCTTCTTCTTCT
TCCTTAGTCGTCGCTGGGCTCCAGTGGTCTGCTAGCCGCTCCTGTAACCTACGTAATA
CCAGGTCAAGTCAGGACTTTGCGACCCATACTGCTCCCCTCTTTTTATATCAATCAAT
CCTCGGAAGGAATCTATTTCTATTCTATAATCTAGATTATTTTATTGATGTTTTTAC
AAATTTTATTTTCTGAAACCTGGCGGCGCTATCTCTTTTTATAGTGACCCTGCTG
TGGCTGTCTATGTCAGCAACCGATCAGCGCTCCCTCCATGCTTCTGCTACTTCTGCTG
TGCTCTCAGCTACGGCCCTCTCTCTACTCCCTTCTTCTATTAGGCCGCTTACCGACACAC
GTTCCTCTTACAGGTAGACGGTCTGCTGGGCTAGGAGCGTATCGAGGCTGCTGCTG
CTCCGGAAAGCATCTATTTCTATAATCTAGATTATTTTATTGATGTTTTTAC
AGACTTTAACCTATACGACAACCCATCTGTGCTGCTGGGAGCGGATAGCTCCGCTG
AGTCTTCCTGCTGATGCGCCTCTCTGCTGATTAAGCTGACTCTAGCTAGTTGG
TGTTGCTGATGAGATCTCTCTCTCTGCTGATGAGATCTGACTCTAGCTAGTTGG
TCCTGACCTTCAAGAGATGGACGGAACACCTCTGCTGACGACTCTTCTGAGAGCTG
GTAGCCTCTGGCCCGTCCGGAAGCACTTCAACCTGGGACGCTCCCTCCGACAGG
AGCAAGCTTCTGTGTCCCGCGACAGTATCCCGACTGCTGTATGTGACGACTC
CACGTCTCTGCGCAAAAGCTGCTGACGAGGTGGCAGCGATCGCAAAAGCAAGAATCACTCGACCGGC
FIG. 17B

AACATTGCCAACATCGCGTGGAATCGATGACCTGCAACTGAGCGTTCTCCACCTCCCAA
GTACTCCTGGCGAGACGAGATGAACTGGGATGGTCTCTGCTGACACAGTTGGAACTGCAG
ACACCAGGTACCTCCCTTATGCTGCTGGATGAGACGTGGTTATCGTACTATGCGCAA
GTTCACCAACTCGTGTCATCGAGCTGACGGTGGTACATCTACCCCTGCAACACCAC
TCTTCCCCAGCTTTCTCAGGCTTGCTGCCGCGATGCGAAGCGCTGCCAGATCCCCGTAACC
TGATCAATTTCTCCAAGGTGCAACACCGACACACACCACCCGAGGCTGATTAGCTGCTC
CCCTTTTTTGGATGATTGCACTGATATTGTGTGCTGTTAGTGCTGTTTGGGGG
GCATTCAATCCAAAGGAAAACACCTCGCTGCGAGATTCTGCGGCGTATATTTTCTCAGAGGCG
TTTTTGTTGTCTTCGACATCGCGCGGCCCCCTCGCTTGGTGTGCTTCCATCCCAAGAACTA
GTTTCCTTTTCTCTGACTTTTTCCCGCGTGGTAATAAATATCGTCTGATTATTGGACTG
TCTCCTACGTTGGCAAGATGGATGATAGTTTTGCTACGTGACATTGCTTTATGGGTT
CTGTGACTGAGCCAGAAGGTCGCTGGTGTATCTCAATTGATACAGTGCCGACCG
TTATGGCTTACCACATCGAAAAACATAGACACTCTTTCTAACCCTATCCATGATACAA
GTATACTTCCGATGCATATTATGTTGTGCTATCAAGGCGCCATGTTTATTATCTAATG
AAACCAACGATGGTCTCATCTTCTACGTTGTGTTTAAAGTGCCGAGATATGAGAAG
ATAGATATAGAAGCACCAGCAAGTGCTAAGGGCTAATCAGGCGCGTTAACAGGTAACCT
CCGCGAAGGAACACGAGGTAGCGCAAATAGAAACTACACTGCACCTCTCCCGAGT
CCCCAAAGACAAAACACGACAAAATGCCGCCAAGAGACACAACACACAGCCGACACCAA
AAGCAGAAAAATATCTGTCCCTTGTACCTGGTACCAGGCCACTGCAAGACACACACAC
AGAGCCTGCAGCAAGAAAATGGCTTCCTGCGATCGTGCTGAGGCAATAATGCGCGGCGCGCG
FIG. 18

MQLLQSLIVAVCFSYGVLSLPHGPSNQHKARSFKVERVRRGTGALHGPAALRKAYRKYG
IAPSSFNIIDLADFKEITTHAAAGSEIAEBPDQTGAATSVNDAEFVSPVGLIGQKIV
MTFDTGSSDFWVFDTNLNETLGTHTEHYNPNSSTFPKMDGMTYFDDVSYGDSSYAGPGVT
DTVPIGGAIVKEQAPGVPDQVSQSFIEDINSNLVGGLFSSINTKPEAQDTFFANVAP
SLDEPVMTASLKADGVGEYEFGTIDDKKYQGNIANISVDSSNGYWQFSTPKYSVAGEL
KDGLSLNTSIADTGTSLMLLLDEDVTTAYAQVPNSVYVSSAGGYYIYPCNITLPSFSLVL
GESSLATIPGNLINFSKVGTNNTTQACKLPPFMCIEHD
FIG. 19A

CTTTCCCGAAAGTGACCCTGACGCTTAAGAAAACCATTATATACATGACATTTA
ATAAAAAATAGGGCTACTACAGAGGCCCTTTGCTCTCGCGCGGTTTTCTGAGATGCCCAT
AACCTCTGCACACTGCGACTCCCGGAGACGCCTACAGCTTGCTGTAACCGCATGGCCCG
GAGCAGACAAGCGCTAGGGCGGCGCTACGGGCTGGTGGCGGTGGTGCCTGCGCTTA
ACTATGCGCATCAGAGGAGATTGTACTGAGAGTAGCACCATAATGCGGCTGTAATACCG
CACAGATGCTAAAGGAGAAATACCCGATACAGGCGCCATTGCGCATTACAGGCGCAAAGGG
CTGTTGGAGGGGAGTCGGTGGGCTTGGCTGCCCTTTGCTATTCATACCGGCGGACGCAAGG
GATTGTCTGCAAGGGCATTAAAGTGTGCTAGTACGCATTCTTCTTATTATACCCCTCT
GCCCTCTCTTAAATGGTACTATAAAGTGAATGCTGTTGGTTATTCACTA
ATTGATTGCTATTCTTCGAGCGAGATAAAGGATAGTACGAATATTCTATATCATGCCATA
GTCAGATTATATCTATAGAATACGCAATCGAGTACCAGACGGTAGTTAG
TACGGTCTACCCGAGCGAGAGTAGGAGGATTAAAACTTATGAGGAGTATTCAGGATCAT
GTTTTGGCCTGCAATTATGACTATTACAGATGGCATTGCTGATCGACTACACAG
CTATATCCCTCTATTGCTCTGCTTTGTAGACATATTGATTTAGCAGATCGCTCATAGAGA
ATGCATTATAGGGAGAGATTGGTGTGGAGGACATAGGACATTCTTCGAGGCGCTACTCGG
CATTCTGCAGCTATCGCTCTCGGTAGGGAGGTGCAGAACCAGCTGAGACTCAGGCT
TCGGTGTATCTTGATACGAGCACTAGAATACTGCTCAATCTGTGATAGTACTCTGCTCTT
TGTTCTCTACACACTATGCTACTAGTATCTATCAAAAAAGCAGGAAGATGAGATCGAGGC
CACCCGGTTAGGGCGGCGGCGGGAGGAACAAAAATACGGGACAGAATCTCAGTG
ATTGGGAGAAGAGAGATGGCAGACTGCGAATACCATACACGACAGAAAATAAGCCGA
AACTAACAGATCAAATCATCAGTACATGAAAGACGCTCGTAAATGATGATAAGCAA
TCCACCAATATCATACATACATAGAAGACGCTCGTAAATGATGATAAGCAA
CCCAACGCTGACGATCACCGAGGCTTTGAGAACCAACAGACAAGATGACAGGCCCCG
CGCACACGGCCACTAATGCGCTCAGACAGAATGCGACGATGCTCCC
GTTCTGGCTCTTGCAAGTAAGAATGGGTATGCCTGATAGTCCGCTAGCCGGATCT
CGCGATCGTAAAGCATTATCATCAATCGATTGTGCAACTCGGACTGCGGAGATT
TCAGGGCCAGCTGAAAAGGTTGGCTGGCAGTTTCTGAGTGGATGATGATGTATAGC
GGAAGAGAGGGGGGAGATGACCTCTAACTTCTGTTGCTACTTTACTTTAGGGACTAGCCACC
TAGTATTTAGTGACTCTAGTGGAGATTGGTTAAAAAGCCTGTGGTTTTCTCTCTCTCTCC
TCGTCATTCTCTCTCATTACACTCCACTCCAANACTCCCTCCACTTTGAGTCAATA
TTTCACTCATGCTACAAAATTAATCAGACTTCATCCCCAATCTCAACTACAAGCGCTCAGG
CAACAAATCTCATCAGCGACTCCATCCAAAGACACCGGTCGTGGCTGACACTGAAAGACCCATC
CTACACTCTTCAGACTGCTATCCATCCCAACAGACCGAGGGTCGTGGCTGACACTGAAAGACCCATC
GGGGGTGCAGGCTCCCTGAGGTGCTGAGGCTAGAAAGAGGAGCGAAGAGCCACCCAGAAGGGT
TGCCGAGGTTCCGGCAGAAGATGTGCAAGAGACGAGTCCATTTGAGCAGCTTTGAGGCTCCTTTC
TGGAACCCCGGCGACAAACTGGATGTTGGACTTTGACACTGTTCAGCTGATCTTTGG
GTACACCCCATTTAGGAAAGACCTAAATGAGAAAGAGCGTCATGACAGAAGTCTGTACG
TGTTCCAAACAAAACCTCCCCCTACAACTACCTCTATCCAGAAGAACACCACATGCGACCTCA
CTCGTCGAAAATCGAGCAGCTTTGAAAGGTTGAAATCCTGGCAATTCTCCTACG
GAAGCGGATTGCTCGGTCTAGGATACCTGCAACATACATCCAAACCGCTACGAGCAAATG
CCGTGAAAAGCCGCCGTCAGAACATGATCCTGCAGGATGACATTTCCAAGGCTCGGTGAG
CTTCTACGGCCAGCTGGAGACTTGGGGGAGACTGAGGAGCAGGTCTGGTTTACACTTT
TGGCTTCTATTGACCAGGGATCTGAGAAGACCGCAGGTGAAAGGTTCTACACTACCCCTG
TCGATAAACGTCAAGGCTTTGCTGTATTTACCTCCGCGACGGTAAATGGAAAG
ACCTTAAACCGGGTGGTAACACCCCGCAGTTGTACCCGGATCAGGACCTGGGCTTTG
GGACGATGCAACGTTGGAGCCATTTATAGTGCAAATGCGCCTATTATGATTCAGG
AAAGTACAGGGCTGATCTATCCGACCAGATAGCGGCGAGATAAGCTGCCACTGTGGTCG
TTTGCCGTTGGAGAAAGGACGTTCTGAGTCCAGAGGGACCTGGCGTTTCGGAGGC
GAAGACGGCTAGTCTATGAGGAAATACCAGATGCTGTGGTGATATGACCATAGGACATCT
TTGGAGACACATTGTTGAGAGATATTTATGCTGTAAAGTGCATTTGCTGTGTCGGCTTAAAGG
GGTGATATCGAGCTACACTGGATTTGCAAGACCTTTGAGTGGTGAGGAGGACCTGCGGCTTT
GGAGCGGTCCAGCGCGGAGTGCTCGCGGACAGCTCGATTTCTGCCTTATTAGTGAGTCGTA
TTACAAATCTGCGCCGGTGAGTCTCGAGCTGGGAACCCCTGTGGCTTACCC
AACCCAATCGCCTTGACGCACATCCCAACCCCTTTTCCGCGAGCTGGAAGCTATGACGAAGGCC
CGCAACGGATCGGCCCCTCCCAACAGTGTCGCGAGCCTGAAATGCGGAAATGGCGGCTTGGCA
GTATTTCTCTTACCGATCTGTGCGGTATTTTCACACCGCATATGCTGACTCTCAGTA
CAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCCGACACCCCGCAAACACCGCTGAC
CGGCCCTGACGGGCTTGTCTGCTCCGATCAGCTTCAGTTACGCAAGCTTGGTGGGCTTTC
CGGAGCGGACTGTGTCAGAGGTGTTTCACCCGTCTCACCAGAAACCGGGAGACGGAAAGG
GCCTCGTGATACCGCCTATTTTTATAGGTTAATGTGATGATATAATTGTTTTCTTAGACG
TCAGGTTGCGACCTTTTCGGGAAATGTGCGCGAAACCCTATTGGTTTTTCTAAAT
ACATTCGAAATATGTATTCG
FIG. 21A

TGTGGCTTGATGCAATTATGATTATGAAACTCGTAAGTTAGGTACCTGTAATCTACCTGCTATATATATA
TGTACTGTTTTTCTCCGCGGCGTACAAGTGATCTTAATCTGACTGCTAAAGACCTATA
GTAGGAGGTGTGATACTAAGGATGGAATTGATGCTCAGGACTGTACTTTTGCGCTG
TCGCTCTCATTCTGTGCTCACGNGAAACACATGACCGGTGCTGCCGGCTCCCGGCA
CGCAAAAGGGCTGAGAACGGGCACTTGAGATGGGCTGGCCTTTTGGTTTGGTTGGCTCCAGATG
CCGTGGCAGCGCGAAAACCTGCGAGTTGGCGCTTTTAGGTCTTTTGCGCTCCAGAGTC
TAAGCATAATTTCTTCTGCTGATGTAATTCTCTTGTGGAGGCGTCACTTTAATT
CTTAAGCTGCAAGGACACCTGGGCTGGTCTATTTTCTTTTCTTTCTTTTCTCGTG
TTCATTTCACCTCCCTTGCTCTCTCTTCTTTTGCGACATTACAAATCAACTTTTCT
TATACACTTTCTTCTCTACCTGCTTTTCTTTCTATCTACTGACTGTTGACCTCAGGTT
ATCAAAACATTATTTTTTTCGTFCTTTTATCTACCTTGTTGGCGCCCTCTCCCA
CTAGCAATTCTATCGTGGTACCTTTTCTAGACATTGTCTTGGAGTGTCATTTTCCTGATT
GACCTCGGCTATGAAATTATTATTCTCAATCTGCTCTAAAAACCAATTCTACTCTATCATTT
ACACATTCTGATACCTGCTGAGATAGAGAGACTGAGCTACCTGTGCTGCA
TCTCCCACGCTGCTGTTACAGCAGGTCTTGGCTTCCCTGACCTGATCTTTTTCAGTTTATGCA
TCTCCACCAGGCTGCTGTTACAGCAGGTCTTGGCTTCCCTGACCTGATCTTTTTCAGTTTATGCA
CATACAGATTCTGCACTGCTGAGATAGAGAGACTGAGCTACCTGTGCTGCA
CTGTTTCTGCGGACTACACAAAACCAAGGCGTACCTGACTAGAGCTACCTGTGCCGCCAGTTTCTCCATCC
CAGCGATGCTCCCGTGACAATCTCTACTCTCCCTTGCTTACAGGTAATAACGTGCTGCTGACATTGATT
TCAAGATTGCTGATGCAGCTGAAAACCTCTCTCGCTCTAACACGGCGCTCTCGATACAGAT
GGTAGCGACATTCTACATCTCTGCTGTTAAGCTTTTGGGTGGTGATGAGAAATCTATGTA
CATGGTGTCTGCAACAGGGCGCTTGATACCTGCTTGGGTCTTGCTCTTGACCTGACCTCGCCA
CACCGCTGCACTGCAATTACCTTGGTTCTCGGAGCGATTCTTCTCGAACCTTTAGGAAATGCA
TCGAGAGATGGGTGCTTGGCTTAACTGAGGGCTGTGCTGACCGGCTTGGCTAGGAAAGA
CAAGCTCAGATTTGCAAATGTCAGTACGATGACTTGATTTTGCGAAGTTTCTCTCCAAGCGAT
CGATAACTCTGAGTCTACCAATGGAGGGACTTCTCTGCTCTCGTCTCGTCAACCAACAGAT
AGTTCTACGCAACCAACATCAATCTGAGTCGGCTGGGAGAAGTAAAGTGTTTTCTAAGTC
GAATATCGTGGCTTGCGCCTTCATCAGCTAGCTCGCCAGGGAGTGGCAGGCTAGGCTTTCG
FIG. 21B

GCACCTACTGACAAGGACAAGTACACCCGCGATATCACCTACACCCTTGATACCTGTATATCCTATGCTTCTTATAGTTGGCGGCTGACTTCATGCGATTT
CTCCAAACAAATCAGCCATCATCGATAACCGGAACCTTTTTATGCTATATGCTGCTTTTCCAAGCG
ACTCGAAAGACGCTGACGCTCATTCCCGGCGCCAAATCTTTCGGGGAGCTACCAACATT
ATTCCGTCGAACACAACACTACTAAGCTACAAGTGGCAATTTCTTCTGTGGTTGAATACACCAT
CTCGCCAGAGGTACCTACGTGGGAGCAACTTCAGGGTTCTGGATGCGTTTCAGAACATTATCA
GCTACGACTTTATTTGATGATGACATCCTGGCTCTGGGTGACACGTTCTCTCATAAAAAATGTT
TATGCTGTGTGTTGACTACGATGATGTACAGGTGGTGCTCGGATTTGCAAGCGGTTCCTCGAAC
CACCTCTGCAGCTCGAACACTCTACGAGCTCTGGAAACAAGCAGCACACCTGGGTCCAACCTACAA
CGGGCAGCTCAGCTACGAGCTCTGCTAGCTCTAGTAGTTGTATCTGATGCTGAA
TCAGGAAGTAGCATAGCATTCCCGCTCCTAGTATTTCTTCTCTGCTCCTGGCGATTC
TTCCCTCATGCTTTGGCTCTAGTTAAACCGCATCTACTCGACGCCCTGAAACCTCGGGAAA
CATATGCAATTTTACATGCTGCTATTTGTATTTGCATATATTCTTCG
FIG. 22

MHLPLLVTAAACLCASATAFIPTYTIKLDTSDDISARDLARRFLPVPNPSDLADDSTS
SASDELSLSNIKRPVRRNDFKIVVAETPSWSNTAALDQGSDISYISVVNIGSDEKKS
MYMLLDTGGSDTWFVGSNCTSTPCTMHNFTFGDS DSLTLEMTSEESVGYGTSVESGLLG
KDKLTIANVTVRMTFGLASNASDFESYPMGILGILGRTNDSYDNPTFMDAVAESNVF
KSNIVGFALSRSPAKDGTVSFGTDDKDKYTGDITYTDTVGS S YWRIVPD VVVGTVGTC
DFS N S AIIIDTGT SYAMLPDDS KTLHSLIPG AKSSG S YHIICN TTTLQVAFSGVNY
TISP KDV GATSSGCVSI ISYDLFGDWWL GDTFLKNVYAVFDYDELRV GFAERSS
NTTSA NS TSSGTS STSS GT TTTGSS SSSSSSSSA SSSSSSA ESGSSMTIPAPQYFFS ALA
IASF MLWL
FIG. 23B

ACTCGTGGTGAAATCTACAACCAATGTGTCTGTTGTAAGTGTGATATACACTGACGCTAGTT
GAGTTTACATGGCGCGTATGATCCTAACCTATTTTTTGTAGGTGACCCAAACTCTTGGG
AAGGGGATTTCTCTCTGGGTGTCCTCAGCTGCGAGATTTTGCGGCTAGGAGATGCTGAGACT
TTGTGAAAGTTCTTCAAGAATGCTGAGCTCTTGGGATCAAAACTCAAGATCTAT
GTACTGGAGAAAATTATCGGCCTTTGATTTGATTTATCATCCTACCTCCATTTCTAGA
TCAGAATGATACAGACACTTCTACACTTCAAACAGTAAAGGTGAGTTATACACTTCAACAAAT
TAACAGGCTTGTTGACTAGTGCCAGGAGGAGCACCTGTTGTCTCCTTTGGCAGAAGAAATAT
GCCCTCTCTCAATTTCAATGCAAGCTTTTTGGCGGAGACTAGAGACCATCCTGACAGCAAT
TGGAATAGAAAGATTTACTAGACACTATCAGTCTTCTCCACAGATCTGCCTAGCAGGC
CAAGGCTATGAACTGGAGCGATCCACCTGCTGACTGTGTATAGTACATGTTAATACGC
GTCTTGAGTTCAACCCTGCTTTTCAACCCTACGAAAATCAACGAGATGTGCCCCATTCT
CTGGGAGCTCTTGGAAGTTCCCAAGCAGTACATCTCCCTGCGGGCGCCAGCATCT
ACCTTGACCGCGCTGATGTTAACGCTGACAGCAGCTCTAAACACCCCTCTCCATG
TGATCGGTAACCGTTATATGACATGGTCACCTTTACCAAGGTAACCCCTCTCTCTGAC
CAGAACATGCACTGAATGGAAATGCTCTGAGATCGACGCCGCCGCCAGACGACACCCTCA
CATGACATCCCTGACCTGATACAATGAGTGTCTTTCAAGGAGCCTGTACGACCCAC
AAGTTGGTCAGGGTGTCATGGGACCATCAGCACTATGACGGTCCTTATAGTGGGTGAG
ACCTTCCAGAGCGACACATGCAGCCCAATTCACAACCAGGTGCTATACCACCTTAC
TGATTGGCTCTGCTGGCGCGCTGATACCTCTGAAGTGGTGGTGACTGACCGGAC
GATGTACAGATGATGTAAGTTATGATCTTGAGATACGTGGTATGCGAATGTACAT
GAATGCTTTCTACTGGCACTCTTAAAGCAAAATTCATAGTAGAGTAGCTGCTCTACTTA
CCCCTACTCTCCTACCTTTCAAACCTGGAAGCGGAAGAATGTGAAATACCTAAGCAACTA
ACGTAGCTGATTGGAGCAGAACATACATCTACCTGACGACTTTCTACTTAATG
CGCTATTTGACTGCTAACTGGGTATAATCTGGAAGCTGCAGTCCAAATCTGACATTA
FIG. 23C

CTCAATGTGCTTGCCCAGGAAACGATATTTTGACTTTATATGTCTGAAAATGAACAAATT
GTCCCCGAGAGGAGAGGAGAGGAGGGACGGTGAAATACTTTAGGAGTCAGTCACCGCAGTA
TCTCCACCTAATGCGTAACACAGGAAATGGGACACGAATGGACAGCAGTGATATACGAG
TACACCTTTCCTAACAATGCTGATGTCTGTAAGCAAATGGCCTAAGTACGCTAGATAGA
GAATCTATTTACAAATCAGTAGTAAAGGAATGAGTCGCCAACCAGAA
FIG. 24

MLFRSLLSLSTAVLAVSLCTDNSAAKGRPGQKARDAMNIAKRSANAVKHSKIPVEDYQ
FLNNKTKPYRVESSLQFDFVDPQDMDLGEMYSGLVPIEKGNVSRSLPPQFQPTIEPVEDEIWL
NGPGSSSLEAFLQENFRWQPQGTPQVENMPYSWVNLTVLWNLDQPVGTGFSLGVPTA
TSEEIAEDFVKFFQWQIQFGGKKNFYVTGESYAGRYPYISAAFLQNDTEHFNLK
GALAYDCIGQFQDYOQEEAPVQFQKNNALFNASPLAELESIHEQCQKDFIDQYL
VFPSGQPPKAMNWSPTCDVYDIVNNALDPNPFYBINECMPIUWNVLGFPEV
DYLPAASFYFRAVQKRAMHAPNITWSECSVESVFGDGGEQGQEDYSANPIEHLP
QVIBGTRVLIGNQVYMILNTQILSINQMTWNGKLGFDTAAPSTPNIDIDPLMLYN
VFENGYPQGQQGQGMGQHYERGLMWAETFQSGHMQPQFRVSYRHEWLLGRDRL
FIG 25B

TCACCCACACCTTCAGCGCAGCTGTTAGGTGATCTTTTGAGACACGCCTGAGTTG
ATCGTGAGGACTTTCCAGTGTCGGTACCTCCCTTGTGTTTTGCTGACTTTCCGGC
TTACCTTCACCAATGTCAGGCCACCCGCGCGCTCCACTGCAGGCCTCTTGAGCT
ACCATATGGGACATTGAGCGAGTTCGAGACCCGCCAGTGTCCTGCTTGGAG
CAGCGTCACCTGACCTACTTTCAATTGCTCTCTATGCGATATGCCTGCTTCA
ATGCTTTCTCAGAGGAACAAACCCGAGGATGAAATGAAAATGAGGATGAGCTATC
CGGATTAGTCTGACTTGGTGTGATTGTAATTCGTTTTCTGTGATGGTTTTGAATGAT
TGTAACCTACTTTTAAGTGAGAAGAAATGGATGACGCGCTGCTGCTGAAATGGCTGTC
CTGCTTTATAATTGTAAGAGATCTTCCAGAAGCTGCTGCTCGGAGATCTGAGATCTGAGA
TCATCGAGATCTGCGAAGTCGGCGCTGCTGTAAAGTCTTTTCGCTCTGTCGATCATAACT
TTGATAAAAGCCTTTGTAGATAAGCTACACTACAGATTAATGCGGCTCAAATTTGAT
CTTACTGAGAATCCCATCGAATAGAGCTTCAGAGGCCTGGGTGGAATTGTCGGCCG
TGAGTTGATAGTAGGAAGCTGCTCACCACATGAAAGCACCGCCGCTGATATGGGAGCAG
TAGGGGAGAGCGCTGAAAAGACTTTCCCTATAGTTCTATAAGAGGCTTTCGCACTAGT
CAGGACTTGGAAATAGAATACATAGACGCGCTGCTTACCGTTCCCGGATAATAGTC
GCGAGCCTAGTACATACATAGCCATGTTCAACAACAGGAATCGAGCACACGAGATACCTATG
AGAAGCCCTCTCCATCACGAATTTGTCTCCAGAGAAGAGGGAGGTATTCTCAGATTAT
TTTGAATGTCAGGGCCCATAGTATGCTCGTGTAGCTCAGTTGGAGTATGGGATGAGGCC
ATATAGTTCAAGCTGGGGGAGACATGACGTGGGGAAGGTACACTCGTGATCCGTATAGG
CAGCAGTAGCAGCCATATTCTACTTTGTATCTAATGATAACGAGAAATTGGCGCGTG
CGTGAATAGTTCAGAGGCAAGAAGAAGCTTTACTACCTTCGTAATGCGTCTCTCTCT
CTATGGAGGAGCGAGATCGCTCTCAAGCTCCTCAGCCAGAGGATATGTCGACATTACCCC
CGACAACTGGATGACCTGCTTGGAAGCGCATATTGCTGACGAAGCTTAACGA
FIG 27A

GGATCCATCCATTCTACCTACGCTTTTCCTGTTGCTGAGCTGTGCTCTACCCAGGTCC
CAGTTTCTCCGACCGCCTGCTTATCGGGGCTACGACACAACGATGATTCTGCTTGTCAT
CCGGGCGTATGCGTGAAATTACCGCGATGCGGCTGCATCATTACACACTCTGCTGCTGCCAT
CTTGGCTGAAATCCGTCGCCACATCCACGTCTCCCTCCTCTCCCCTCCTCTCTCTCTCCTCAAC
CTCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCCTCTCTTCT
TCTCATCCTCTCAGTTGTCTTCACTTTGTACACGGCTACACTCCTCCTCTCTCTCTCTCTCTAC
GGCTGGCTGATTGGGGCCCTCTCTAACTCCACATCATGAAAGGCGTCTCCCTGCTCCTTTGGGACG
CCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
AAGAAGCCGCCTCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TGACTCTCCAGCCGAGCGGACTGAGATCGAAAGACGCGGCTGCTTCCGCTTTGGGACG
AGGTCTATCTGGGCTCCAGAAACAACCTTGTGACTTTGTGCTCTGTGCTTGTTACTCT
GGTCAGTCTCCAGGAGTGCTCATGAGCAGACTCCCAGACACCCCATGTGAGTACAC
CCCTCATCGCTACCCGTCCGTTACCTAAGATAAGCCTTAAACCACGCTGCTCAATTAG
GTGCTGATTGACATCGCGCGGATTCCGAAGATTCACATGCCATGGAAGGCAGCAGAAGAA
CGCCCCCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
ACAACTACCTGTATCGGCTCCAGGGGCTGAGGGCGCGGTGACGCTACACATGACG
GGTATCAACGTTGACCAAGCTGCCAGGCGGCTGCCACTGGGGGGCCGGAAGAAGAATCC
TACCAGACGATGAGATCTCGATGCAATGGTGCAAGGAACTACTCGTCTCCGGAACCATTGG
CTGGTAAGAGTGACCTGGTCTCGCTCCAGGAAGCAGCCAACCTCATATGCTGCAAGGCTCTCGGG
TCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TCATACAAAGAAGCAGTGCCAAAGAAGGCAAGGCTCAGATAGGAAAGGCCGAGCAGCAGC
TTGGCAAACATGAGTCTCGGTGCTGGCAAGTCTAAGACCTCGAGGATGCTGTTAACGCT
GGTGTCGAGGCTGCTCTCAGCTCCGCTGGCCGGTAAATGCAATGCTGAATGGTCTGG
CAACTCTTACTCTGCTGGTCGAAGGCAAGCCAACATCCAGGGTGTCCGGCTCAT
ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CATGGCCTCTTCAACATTGCTGCGCCTCTCCGCTAATTTGCTTCCTCCACGCGCCCTCCT
FIG. 27D

TGAGTCGCGATGACGAAATATCGTTTCTCGAATCCTCCACGCTGCTTT
CACGATCGATTTGTATTCTCGGGGCTCTTGCACAAATGACCAGGCGCTCAACGTTGG
CGGGGAAGAGATCCGGGGAATT
FIG. 28

MKGLGSLPLLTAASPVFVDSIHNEAAPILSATNAKEVPDSYIVVFKKHVTSELASA
HHSWVQDIHDSQSERTELKKRLPGLGDENVYLGKNTFQAGSLIGYSGHFEDVIEQV
RRHPDVYIERDSVEHTMEGATEKKNAPWGLARISHRDSLTFGQNFNYALYASEBEGGYDA
YTITDITIONDHFDRGATWGTKPTNDEDLDGNGHGTHCSGTMMAGKGYGVAKKANYA
VKVLRSSGTMSDVSQVESYAYAVQAHIKKADAKNGKVKGFKGSVANMSLGGGKSKTE
DAVNAVGEALHFAVAAAGNDNADACNYSPAABAEKAIEGASTLDLDERAYFSNYGECTDI
FAPGLNILSTWIGSNYATNISGTMSASPHELALAYFVSLQPSDSSAFVEELTPAKL
KIIDIAIAETGALTDPSNTPNVSHAAVGIYKRNETLQKFSSLPGTVVPRTPSTLAA
VTATSLPIRTASRTVLRSFTRPKSCSPRSLVPSTARSRMPSHRSELVLSRRSEDLV
FF
FIG. 29B

GGATTAACATGAGTCTTGATGTCGCCCTCTCTGGGATCTAAATGCGTATACCGT
GATGCAGGTGGAGCTACTCTTATGCGTTAAACACGATGTTGAACTGTCTGACTCTCA
GGTGGTCCTGGTGGTGGCCCTGGGAAATGGAATGTAATGCTTGGCAACTGTCCAA
CCATTGAGCTAAATTTGACTATGTTGTTGGACATTTTTGCGAGACAAAGGAATTCTCA
TTCTGACCCGACGCATTACTGGGCTATCAAGAGAAAGCAATGCGTACGCTTCTG
TCTCAACTACGCGCTCTGCTTGGATACATTTTGGGCGAAGGCAATGACTTTTCCTG
TGGACCCTGAAAACAGGCACTCAGCTGCTTGCATAGATGCTGCTCAGGACGGAGC
CCAATCTTCTGGCTCAACAATGGAAACAGCGCCGCTGTCACAAAAGGGGTAGCGATGG
AGATGGAGACTAGGTGCTAATAGGATATGTGCTTGAATAGTGGGATCGAGAG
GTAGACTAGTTTATAAGGCAATAATGCTGATGATGCTGCTGCTGGGCTGTTG
TCGGCATTTTTCACTTTTGGGACAGGCAATATGCTGCTGCTGGGCTGCTGCT
CAGCCAGCTCCAGGAGACTGAAGACTCAGATCCCAATGAGTGATGAAAGATGCTGCACA
TGGAAAATGCGCTCGCCTGCTAACCACTCCCAGCGGAGTGTACCGGACGGGACAGAGACAGA
TCATGGACGCTAATCCAGCAAGCAACTGTGGCTCTTACTGAAAGATCACCATTCCACA
GGTCATGCGCATGATGATGATGATGTCTTCTCCTCCATCAGAAACACTGAGCGTGT
GTGGAATGAATCTAGATCAAGGAGATCAAACCCGCTCCCAGAAGATCAGGGCTATGC
CCAATAATGGAACGGGTGACGTGTAACCCCTGTATTCAAACTCGGTTTTTGTATTGCTTCTTCT
GGTCTCCTGGATGCTGCTACTACCTACAGACCTTGTTGTGCTCTGCTTGGAGACTCTGGAGA
CAGGTTCCTTGTATAGCTGCTAGCCTAAGTGCCGGGTCTAGGAAAACAGATGTGCTGAGG
TCTTCTCGGATTTCTCAATGGAAGACTGTGCCCCGGGTCTTTACGCCCTCTGTTGAAGAG
CGAAAAAGGAGACGCCTTCTGGCCCTGCTTCCGCAAATAGCAGGCCTACCAGCCCGACG
GGATCC
FIG. 30

MAFLKRILPLALLLAPAVSFATEQVPHPTIQTIPGKYIVTFKSGIDNAKIESHAAWVTE
LHRRSLEGSRSTTEDDPAGIERTYRIANFAGYAGSFDEKTIEEIRKHNVAYVEQDVW
YLDTLVTERAPWGLGSISHRGASSTDYIYDSAGECTYAYVVDTGILATHNEFGGRAS
LAYNAAGGENEHDGVGSGTHVAGTIKGKTYGVSKNALLLSVKVFVGESSSTSVILDGDNW
AANDIVSKNRTSKAAINMSSGGGYSYAFNNAVENAPDEGVLSCVAAGNERDAARTSPA
SAPDAITVAAINRNARASFSNYGSVDIFAPGEQVLSANTGSNSATNTISGTMATPH
VTGLILYMGLRLATPAAATTELKRLATRANVTVAGSPNLLAYNGNSGVSNGSGDDG
DED
FIG. 31A

GCTACGGACCAACCCACCACATACATCATTACGACTCAGGAAGCCAGAG
CTCTCTCGCTCCGCGATTGCACAAGTTCACCAAACGCTGGAAGACCCACGAACCG
CTGTTCTGCCCTTCCGGCATGACATGTTCTCGAGTCTACGGGAGAGATGTA
TGAGATGTCTATCTGGAGCGGATGTGGAATCGAAATCGGGATGATTGAGCCCTAAT
GAACGGAGTTCTCTGGGAAAGCGGCGCGCCGGAGCGGCCGAGGTTGGAATTGTGTTT
GGTGTCCGCTAGATGGTTGTTAATGAGAAAGATGCGGTAGAGTGAACGCGAG
GGATGGAGGGGATTCGTTGCGGAAGCGGTTGTTGACATATTCTCCGCTAAGTC
TTGTCCAGCTCATCCATCCTCCTCACTGGTCGTGACGGAGATCTCTGCTAACG
AGGCAGGTCTCTGATTATGTGACGAGAAGTTCATGCACCCGAAACTTGAAGAG
TATGGCGGTGCGAGCGGCACTTTGCTACCCCTCTAATAACTGTCATGCTATTT
GGTGTACGAGCACTCCCCGGGAAACACGCACTGGTGTTGCTCGGCGGATG
CGAATCATATATCCACGGGATGGACGTTGCGGAGTAGTTGAGCGGTGGCAA
GGACGCTGGTAGACGATGTTGGAGCGTTTGCGGAGTGGATGAGCAGGCGGCT
GATTAGCTGGGACATTTGGAGTAAGCGGTGGTGTTGTCTCAACATTGGGTGAAG
GATGAGTGGTGTTGAGGAATTTGGGAAAGAAGGAGGGAAGGTGAACCGGCTT
GCTAGGGTGTGTTGAGGAATTTGGGAAAGAAGGAGGGAAGGTGAACCGGCTT
GATTGATTAAAGCCATTTAAAGGGGTATGGATTGTGAAACATGAATTTCTACAATCAT
TCACAATATACTATAGTGAATAATGGGGACATATCCATCGCTATATCGTAGG
GTCGGTTGGGAGGTGCGAGTTTTTGCTCCACGGAGCGAAAGCTTGGTCGCTGC
GTGGTGCGTGCAGCTTCCGGGAGATGATGCTGCAGGGGTGTGATGGAAGAGCTT
TCGGCAGTGTCAGTCATCCGGCAGGAAATGCCAGATACGGGGCGAGTAATGGAAG
CGCCCTTCTCGGTCCAGTGAATGCCTCGGCGGACGGGGAAGGGCGGCTCGATCCT
CAGCGAATGGGATTTGGATAGTGCTCTGAGGGTGGCTGCGGCGGCAATTGAGCTAC
TTAGAACCACATCGAGTTAAAGGGGTGCTGGAGTCAACGTACGATGTACTGGGATGTC
GTC
FIG. 31B

GCCACCCAGCGCTGGTCAGCTCTTTGCGCTCGAGTGACAGCCGGTCAATGGTTGATG
GTGTGAGAACGCTGCTGACCCTTGGAGATGCAGCCCTTCGATTTGGTGACCAGATTTGAGG
TTTGTGTGAAGTCTCCCAGAAGTCCACAGAAGGCGAGAACGGCAGAGGCGAGGATA
GCCAGCGCCAGCGGAAACATCGGCTCTACATGGCCAGAATCTAAACAGAGGGTGGTAC
AGGCCCTCTCTCCACATGGGTGATGATATTGCTGTAAGCATATCGTCGTCGG
TGTTGTGGAGAGCGCTCGACTGATACATGGATTGAGACTGACTACCTCAGTAA
TTGCGCCAGAATACATGACAGCCAGGTTTCGCTCCATACCATACGTCGCCAGGTAATCA
TCACCATACTCGATGCGCTGTATCTCCCTGCTCCTAATTTCTCTTGAAGACTCTCTCA
CAGTCCAGAGTGAGCAAGCAGAGCACTCGAGACTGGAACGAGTTTCGCGCCGCGGTGCTC
GAGGCTATACATGTGAAGCCGGCTCTAAACGACCCGATTACGACTCCTGGAACCGGTG
TCGAGGATACAAGTCAAAGGATACACAGCCTGAGTGGGATTAGGGTGATAGGCAGAACCTCT
CGCTTCTGAGAGGGAGATCAAGCTGAGCTGACCCTCGACTGTTGAGGAGCGCTTCAC
GTTAGGCAACACTACGAGGTTCTACATTTCTCTGAGACTGTTTCTGTCAGCTCG
AGGTACTGGGAGGGCCCTTTGATTCTGGACTCGAGCAAGAGCGGCGGGGGCTTGTTG
AGACCGGTGCTGCTGCGCTCTTGGAGGAATATTTTCTCTCCCTGGATGCGGGGAGGG
AGTGCGCGCAACACGAGCGACCGCCCCAGCAAGAGCGAGGCGAGCTGGGACAGGGTACC
ACGGGAGATAATACTGATGACACGTCATGGAGAATGTGACTTGAAGATGCTGATGCTGAC
TGACAGGAAGAAGCTCTCGTCGAGAAGAGGGGTTGTTGCTGGAGAATATAATAGTG
TTGGGAAATATTACATCGGGTATGCTGGACAAAGACCCAACAAATCGTACAATAAAC
ATAACCGAAATGAGACAGAATTACGCTGATGCGAAGTACATGCTGCTGTCGACATGC
ATAAGCTCCAGGCTCTGCGAAGTCGGGTTAGCAATAGACGCAGTGCAGTTGGTGAAG
AAGACAAGCCGCGACGTTATAGTGCTGCTGATGAGTGACAGACATGCGCCTCCA
CTGGATGGCTAAGGAAATAGCGCCGCGTGGGCAGCGACGATTTCGTGATAG
GGAGGCTTGGTCGCTGCGGCTCTGGATTGTTGCGAGAAGCCAGGCCACAGAGAGA
TAAGACGGAAGTCGGCGCTGCGAAGTGCCGACAGGGGAGGGTTGCGAAGGTTGAGACT
GAAAGAGATAGATAGATTGGATATATTGGGCGACGAGAAATATTTCTCTCAGAGCGTC
GAGTGTATCATTTGTTGCGCTTTTCTCTTCCTGTAATAGATGATACCGAGCCTTGGAA
TAAATTGTATCTCGGCGACCACTCTGGGTAACCCCTGGAAGGCTATTAGCGCCCGGCGA
AAT
FIG. 32

MYIPVGLATASLLAGAALAPTPSPLKGRLNIVRSGHHTVYKPAFAAPSHNKA
SSKYLELSKTSKGNVNRSAAYVRKSTSSGSSLISLFEGBEFATSITIGGDSF
DVIVDTGSSDTWVKTGFTCIDLDTGRETSEQSCDFGSTWTVEESSFKIEEGEEFA
IEYGDGEYLYGVMNGETVALADITVDQITIGVTRAAWEGDGTSGTGLAYPALT
SAYSTTTDEQIVSYNIITTMWEELIEPLFSLAERDVSGAAGYLALGGLPPVDF
VEDFTKTSILVTHNEGYSKAYDFYTINIDAVTLNGKLSAGGDIQYIMQVDSDG
TTLNYXPSIAEIEAFAFSPAATYSDDEGAYIVDCATPPHTGITISGTKTYINP
LDMLDAGTDEGNATICISGIVDGSDTSEDLYILGDFTQKNVTVFDIGATELR
FAARENYTSNDTY
FIG. 33A: General cloning strategy for gene deletion in Aspergillus.
FIG. 33B: General cloning strategy for gene disruption in Aspergillus.
FIG. 34A: PCR analysis of the mnn9 deletion strain

FIG. 34B: PCR analysis of ochA disruption strain