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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 February 2013 (28.02.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/028912 A2

(51) **International Patent Classification**:
C12N 15/09 (2006.01)

(21) **International Application Number**:
PCT/US2012/052143

(22) **International Filing Date**:
23 August 2012 (23.08.2012)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/526,809 24 August 2011 (24.08.2011) US

(71) **Applicant** *(for all designated States except US)*: **NOVOZYMES, INC.** [US/US]; 1445 Drew Avenue, Davis, California 95618 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **YAVER, Debbie** [US/US]; 5632 Hoag Place, Davis, California 95618 (US). **JIN, Qiming** [CN/US]; 1770 Irongate Way, Sacramento, California 95835 (US).

(74) **Agent**: **STARNES, Robert**; Novozymes, Inc., 1445 Drew Avenue, Davis, California 95618 (US).

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

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

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/028912 A2

(54) **Title**: METHODS FOR PRODUCING MULTIPLE RECOMBINANT POLYPEPTIDES IN A FILAMENTOUS FUNGAL HOST CELL

(57) **Abstract**: The present invention relates to methods for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity. The present invention also relates to methods for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain. The present invention also relates to filamentous fungal strains expressing multiple recombinant polypeptides having biological activity.

METHODS FOR PRODUCING MULTIPLE RECOMBINANT POLYPEPTIDES IN A FILAMENTOUS FUNGAL HOST CELL

Reference to a Sequence Listing

This application contains a Sequence Listing in computer readable form, which is incorporated herein by reference.

Background of the Invention

Field of the Invention

The present invention relates to methods for producing multiple recombinant polypeptides in a filamentous fungal host cell.

Description of the Related Art

Recombinant production of a polypeptide in a filamentous fungal host cell may provide for a more desirable vehicle for producing the polypeptide in commercially relevant quantities. The recombinant production of a polypeptide is generally accomplished by constructing an expression cassette in which the DNA coding for the polypeptide is placed under the expression control of a promoter from a regulated gene. The expression cassette is introduced into the host cell, usually by plasmid-mediated transformation. Production of the polypeptide is then achieved by culturing the transformed host cell under inducing conditions necessary for the proper functioning of the promoter contained on the expression cassette.

Filamentous fungal cells may be transformed with a vector by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Transformation of a filamentous fungal host cell with two or more vectors, alone or together (co-transformation) is very inefficient and limted by the availability of useful selectable markers.

There is a need in the art for methods of constructing filamentous fungal strains capable of expressing multiple recombinant polypeptides by targeting tandem expression constructs to one or more (*e.g.*, several) specific genomic loci to achieve desired expression levels of all interested polypeptides.

The present invention provides improved methods for producing multiple recombinant polypeptides in a filamentous fungal host cell.

Summary of the Invention

The present invention relates to methods for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising:

(a) replacing an endogenous first gene by targeted integration by introducing into the filamentous fungal strain a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof;

(b) replacing an endogenous second gene by targeted integration by introducing into the filamentous fungal strain a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or

(c) a combination of (a) and (b).

The present invention also relates to filamentous fungal strains, comprising:

(a) an endogenous first gene replaced by targeted integration by introducing into the filamentous fungal strain a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof;

(b) an endogenous second gene replaced by targeted integration by introducing into the filamentous fungal strain a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological

activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or

(c) a combination of (a) and (b).

The present invention also relates to methods for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising:

(A) cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b); and optionally

(B) recovering the multiple recombinant polypeptides.

The present invention also relates to tandem constructs comprising (i) a homologous 5' region of a gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

The present invention also relates to methods for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising:

(a) inserting into an endogenous first locus by targeted integration a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking

region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof;

(b) inserting into an endogenous second locus by targeted integration a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or

(c) a combination of (a) and (b).

The present invention also relates to filamentous fungal strains, comprising: (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

The present invention also relates to methods for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising:

(A) cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first locus modified by insertion by targeted integration with a first tandem

construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b); and

(B) recovering the multiple recombinant polypeptides.

The present invention also relates to tandem constructs comprising (i) a homologous 5' region of a locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

Brief Description of the Figures

Figure 1 shows a restriction map of plasmid pAG43.

Figure 2 shows a restriction map of plasmid pSMai214.

Figure 3 shows a restriction map of plasmid pDM287.

Figure 4 shows a comparison of positive transformants for beta-glucosidase activity: between 45 transformants of pDM287 and 45 transformants of pEJG107 + pSMai214.

Figure 5 shows a restriction map of plasmid pDM286.

Figure 6 shows a restriction map of plasmid pDM290.

Figure 7 shows a restriction map of plasmid pJfyS142.

Figure 8 shows a restriction map of plasmid pJfyS144.

Figure 9 shows a restriction map of plasmid pJfyS139.

Figure 10 shows a restriction map of plasmid pQM18.

Figure 11 shows a restriction map of plasmid pQM21.

Figure 12 shows a restriction map of plasmid pAG121.

Figure 13 shows a restriction map of plasmid pRRAB01.

Figure 14 shows a restriction map of plasmid pDFng113-3.

Figure 15 shows a restriction map of plasmid pAmFs074.

Figure 16 shows a restriction map of plasmid pQM22.

Definitions

Acetylxylan esterase: The term "acetylxylan esterase" means a carboxylesterase (EC 3.1.1.72) that catalyzes the hydrolysis of acetyl groups from polymeric xylan, acetylated xylose, acetylated glucose, alpha-napthyl acetate, and *p*-nitrophenyl acetate. For purposes of the present invention, acetylxylan esterase activity is determined using 0.5 mM *p*-nitrophenylacetate as substrate in 50 mM sodium acetate pH 5.0 containing 0.01% TWEEN™ 20 (polyoxyethylene sorbitan monolaurate). One unit of acetylxylan esterase is defined as the amount of enzyme capable of releasing 1 µmole of *p*-nitrophenolate anion per minute at pH 5, 25°C.

Allelic variant: The term "allelic variant" means any of two or more (*e.g.*, several) alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

Alpha-L-arabinofuranosidase: The term "alpha-L-arabinofuranosidase" means an alpha-L-arabinofuranoside arabinofuranohydrolase (EC 3.2.1.55) that catalyzes the hydrolysis of terminal non-reducing alpha-L-arabinofuranoside residues in alpha-L-arabinosides. The enzyme acts on alpha-L-arabinofuranosides, alpha-L-arabinans containing (1,3)- and/or (1,5)-linkages, arabinoxylans, and arabinogalactans. Alpha-L-arabinofuranosidase is also known as arabinosidase, alpha-arabinosidase, alpha-L-arabinosidase, alpha-arabinofuranosidase, polysaccharide alpha-L-arabinofuranosidase, alpha-L-arabinofuranoside hydrolase, L-arabinosidase, or alpha-L-arabinanase. For purposes of the present invention, alpha-L-arabinofuranosidase activity is determined using 5 mg of medium viscosity wheat arabinoxylan (Megazyme International Ireland, Ltd., Bray, Co. Wicklow, Ireland) per ml of 100 mM sodium acetate pH 5 in a total volume of 200 µl for 30 minutes at 40°C followed by arabinose analysis by AMINEX® HPX-87H column chromatography (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

Alpha-glucuronidase: The term "alpha-glucuronidase" means an alpha-D-glucosiduronate glucuronohydrolase (EC 3.2.1.139) that catalyzes the hydrolysis of an alpha-D-glucuronoside to D-glucuronate and an alcohol. For purposes of the present invention, alpha-glucuronidase activity is determined according to de Vries, 1998, *J. Bacteriol.* 180: 243-249. One unit of alpha-glucuronidase equals the amount of enzyme capable of releasing 1 µmole of glucuronic or 4-O-methylglucuronic acid per minute at pH 5, 40°C.

Aspartic protease: The term "aspartic protease" means a protease that uses an aspartate residue(s) for catalyzing the hydrolysis of peptide bonds in peptides and proteins. Aspartic proteases are a family of protease enzymes that use an aspartate residue for catalytic hydrolysis of their peptide substrates. In general, they have two highly-conserved aspartates in the active site and are optimally active at acidic pH (Szecsi, 1992, *Scand. J. Clin. Lab. In vest.* Suppl. 210: 5–22). For purposes of the present invention, aspartic protease activity is determined according to the procedure described by Aikawa *et al.*, 2001, *J. Biochem.* 129: 791-794.

Beta-glucosidase: The term "beta-glucosidase" means a beta-D-glucoside glucohydrolase (E.C. 3.2.1.21) that catalyzes the hydrolysis of terminal non-reducing beta-D-glucose residues with the release of beta-D-glucose. For purposes of the present invention, beta-glucosidase activity is determined using *p*-nitrophenyl-beta-D-glucopyranoside as substrate according to the procedure of Venturi *et al.*, 2002, Extracellular beta-D-glucosidase from *Chaetomium thermophilum* var. *coprophilum*: production, purification and some biochemical properties, *J. Basic Microbiol.* 42: 55-66. One unit of beta-glucosidase is defined as 1.0 µmole of *p*-nitrophenolate anion produced per minute at 25°C, pH 4.8 from 1 mM *p*-nitrophenyl-beta-D-glucopyranoside as substrate in 50 mM sodium citrate containing 0.01% TWEEN® 20.

Beta-xylosidase: The term "beta-xylosidase" means a beta-D-xyloside xylohydrolase (E.C. 3.2.1.37) that catalyzes the exo-hydrolysis of short beta-(1→4)-xylooligosaccharides to remove successive D-xylose residues from non-reducing termini. For purposes of the present invention, one unit of beta-xylosidase is defined as 1.0 µmole of *p*-nitrophenolate anion produced per minute at 40°C, pH 5 from 1 mM *p*-nitrophenyl-beta-D-xyloside as substrate in 100 mM sodium citrate containing 0.01% TWEEN® 20.

cDNA: The term "cDNA" means a DNA molecule that can be prepared by reverse transcription from a mature, spliced, mRNA molecule obtained from a eukaryotic or prokaryotic cell. cDNA lacks intron sequences that may be present in the corresponding genomic DNA. The initial, primary RNA transcript is a precursor to mRNA that is processed through a series of steps, including splicing, before appearing as mature spliced mRNA.

Cellobiohydrolase: The term "cellobiohydrolase" means a 1,4-beta-D-glucan cellobiohydrolase (E.C. 3.2.1.91 and E.C. 3.2.1.176) that catalyzes the hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, cellooligosaccharides, or any beta-1,4-linked glucose containing polymer, releasing cellobiose from the reducing or non-reducing ends of the chain (Teeri, 1997, Crystalline cellulose degradation: New insight into the function of cellobiohydrolases, *Trends in Biotechnology* 15: 160-167; Teeri *et al.*, 1998, *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose?, *Biochem. Soc. Trans.* 26: 173-178). Cellobiohydrolase activity is determined according to the procedures described by Lever *et al.*, 1972, *Anal. Biochem.* 47: 273-279; van Tilbeurgh *et al.*, 1982, *FEBS Letters,* 149: 152-156; van Tilbeurgh and Claeyssens, 1985, *FEBS Letters,* 187: 283-288; and Tomme *et al.*, 1988, *Eur. J. Biochem.* 170: 575-581. In the present invention, the Tomme *et al.* method can be used to determine cellobiohydrolase activity.

Cellulolytic enzyme or cellulase: The term "cellulolytic enzyme" or "cellulase" means one or more (*e.g.*, several) enzymes that hydrolyze a cellulosic material. Such enzymes include endoglucanase(s), cellobiohydrolase(s), beta-glucosidase(s), or combinations thereof. The two basic approaches for measuring cellulolytic activity include: (1) measuring the total cellulolytic activity, and (2) measuring the individual cellulolytic activities (endoglucanases, cellobiohydrolases, and beta-glucosidases) as reviewed in Zhang *et al.*, Outlook for cellulase improvement: Screening and selection strategies, 2006, *Biotechnology Advances* 24: 452-481. Total cellulolytic activity is usually measured using insoluble substrates, including Whatman №1 filter paper, microcrystalline cellulose, bacterial cellulose, algal cellulose, cotton, pretreated lignocellulose, *etc*. The most common total cellulolytic activity assay is the filter paper assay using Whatman №1 filter paper as the substrate. The assay was established by the International Union of Pure and Applied Chemistry (IUPAC) (Ghose, 1987, Measurement of cellulase activities, *Pure Appl. Chem*. 59: 257-68).

For purposes of the present invention, cellulolytic enzyme activity is determined by measuring the increase in hydrolysis of a cellulosic material by cellulolytic enzyme(s) under the following conditions: 1-50 mg of cellulolytic enzyme protein/g of cellulose in PCS (or other pretreated cellulosic material) for 3-7 days at a suitable temperature, *e.g.*, 50°C, 55°C, or 60°C, compared to a control hydrolysis without addition of cellulolytic enzyme protein. Typical conditions are 1 ml reactions, washed or unwashed PCS, 5% insoluble solids, 50 mM sodium acetate pH 5, 1 mM $MnSO_4$, 50°C, 55°C, or 60°C, 72 hours, sugar analysis by AMINEX® HPX-87H column (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

Cellulosic material: The term "cellulosic material" means any material containing cellulose. The predominant polysaccharide in the primary cell wall of biomass is cellulose, the second most abundant is hemicellulose, and the third is pectin. The secondary cell wall,

produced after the cell has stopped growing, also contains polysaccharides and is strengthened by polymeric lignin covalently cross-linked to hemicellulose. Cellulose is a homopolymer of anhydrocellobiose and thus a linear beta-(1-4)-D-glucan, while hemicelluloses include a variety of compounds, such as xylans, xyloglucans, arabinoxylans, and mannans in complex branched structures with a spectrum of substituents. Although generally polymorphous, cellulose is found in plant tissue primarily as an insoluble crystalline matrix of parallel glucan chains. Hemicelluloses usually hydrogen bond to cellulose, as well as to other hemicelluloses, which help stabilize the cell wall matrix.

Cellulose is generally found, for example, in the stems, leaves, hulls, husks, and cobs of plants or leaves, branches, and wood of trees. The cellulosic material can be, but is not limited to, agricultural residue, herbaceous material (including energy crops), municipal solid waste, pulp and paper mill residue, waste paper, and wood (including forestry residue) (see, for example, Wiselogel *et al.*, 1995, in Handbook on Bioethanol (Charles E. Wyman, editor), pp.105-118, Taylor & Francis, Washington D.C.; Wyman, 1994, *Bioresource Technology* 50: 3-16; Lynd, 1990, *Applied Biochemistry and Biotechnology* 24/25: 695-719; Mosier *et al.*, 1999, Recent Progress in Bioconversion of Lignocellulosics, in *Advances in Biochemical Engineering/Biotechnology*, T. Scheper, managing editor, Volume 65, pp.23-40, Springer-Verlag, New York). It is understood herein that the cellulose may be in the form of lignocellulose, a plant cell wall material containing lignin, cellulose, and hemicellulose in a mixed matrix. In a preferred aspect, the cellulosic material is any biomass material. In another preferred aspect, the cellulosic material is lignocellulose, which comprises cellulose, hemicelluloses, and lignin.

In one aspect, the cellulosic material is agricultural residue. In another aspect, the cellulosic material is herbaceous material (including energy crops). In another aspect, the cellulosic material is municipal solid waste. In another aspect, the cellulosic material is pulp and paper mill residue. In another aspect, the cellulosic material is waste paper. In another aspect, the cellulosic material is wood (including forestry residue).

In another aspect, the cellulosic material is arundo. In another aspect, the cellulosic material is bagasse. In another aspect, the cellulosic material is bamboo. In another aspect, the cellulosic material is corn cob. In another aspect, the cellulosic material is corn fiber. In another aspect, the cellulosic material is corn stover. In another aspect, the cellulosic material is miscanthus. In another aspect, the cellulosic material is orange peel. In another aspect, the cellulosic material is rice straw. In another aspect, the cellulosic material is switchgrass. In another aspect, the cellulosic material is wheat straw.

In another aspect, the cellulosic material is aspen. In another aspect, the cellulosic material is eucalyptus. In another aspect, the cellulosic material is fir. In another aspect, the cellulosic material is pine. In another aspect, the cellulosic material is poplar. In another

aspect, the cellulosic material is spruce. In another aspect, the cellulosic material is willow.

In another aspect, the cellulosic material is algal cellulose. In another aspect, the cellulosic material is bacterial cellulose. In another aspect, the cellulosic material is cotton linter. In another aspect, the cellulosic material is filter paper. In another aspect, the cellulosic material is microcrystalline cellulose. In another aspect, the cellulosic material is phosphoric-acid treated cellulose.

In another aspect, the cellulosic material is an aquatic biomass. As used herein the term "aquatic biomass" means biomass produced in an aquatic environment by a photosynthesis process. The aquatic biomass can be algae, emergent plants, floating-leaf plants, or submerged plants.

The cellulosic material may be used as is or may be subjected to pretreatment, using conventional methods known in the art, as described herein. In a preferred aspect, the cellulosic material is pretreated.

Coding sequence: The term “coding sequence” means a polynucleotide, which directly specifies the amino acid sequence of a polypeptide. The boundaries of the coding sequence are generally determined by an open reading frame, which begins with a start codon such as ATG, GTG, or TTG and ends with a stop codon such as TAA, TAG, or TGA. The coding sequence may be a genomic DNA, cDNA, synthetic DNA, or a combination thereof.

Control sequences: The term “control sequences” means nucleic acid sequences necessary for expression of a polynucleotide encoding a polypeptide. Each control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a polypeptide.

Ectopic integration: The term “ectopic integration” means the insertion of a nucleic acid into the genome of a microorganism at a non-targeted site or at a site other than its usual chromosomal locus, *i.e.*, random integration.

Endoglucanase: The term “endoglucanase” means an endo-1,4-(1,3;1,4)-beta-D-glucan 4-glucanohydrolase (E.C. 3.2.1.4) that catalyzes endohydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, cellulose derivatives (such as carboxymethyl cellulose and hydroxyethyl cellulose), lichenin, beta-1,4 bonds in mixed beta-1,3 glucans such as cereal beta-D-glucans or xyloglucans, and other plant material containing cellulosic components.

Endoglucanase activity can be determined by measuring reduction in substrate viscosity or increase in reducing ends determined by a reducing sugar assay (Zhang *et al.*, 2006, *Biotechnology Advances* 24: 452-481). For purposes of the present invention, endoglucanase activity is determined using carboxymethyl cellulose (CMC) as substrate according to the procedure of Ghose, 1987, *Pure and Appl. Chem.* 59: 257-268, at pH 5, 40°C.

Expression: The term "expression" includes any step involved in the production of a polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

Expression vector: The term "expression vector" means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to control sequences that provide for its expression.

Family 61 glycoside hydrolase: The term "Family 61 glycoside hydrolase" or "Family GH61" or "GH61" means a polypeptide falling into the glycoside hydrolase Family 61 according to Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280: 309-316, and Henrissat B., and Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, *Biochem. J.* 316: 695-696. The enzymes in this family were originally classified as a glycoside hydrolase family based on measurement of very weak endo-1,4-beta-D-glucanase activity in one family member. The structure and mode of action of these enzymes are non-canonical and they cannot be considered as bona fide glycosidases. However, they are kept in the CAZy classification on the basis of their capacity to enhance the breakdown of lignocellulose when used in conjunction with a cellulase or a mixture of cellulases.

Feruloyl esterase: The term "feruloyl esterase" means a 4-hydroxy-3-methoxycinnamoyl-sugar hydrolase (EC 3.1.1.73) that catalyzes the hydrolysis of 4-hydroxy-3-methoxycinnamoyl (feruloyl) groups from esterified sugar, which is usually arabinose in natural biomass substrates, to produce ferulate (4-hydroxy-3-methoxycinnamate). Feruloyl esterase is also known as ferulic acid esterase, hydroxycinnamoyl esterase, FAE-III, cinnamoyl ester hydrolase, FAEA, cinnAE, FAE-I, or FAE-II. For purposes of the present invention, feruloyl esterase activity is determined using 0.5 mM *p*-nitrophenylferulate as substrate in 50 mM sodium acetate pH 5.0. One unit of feruloyl esterase equals the amount of enzyme capable of releasing 1 µmole of *p*-nitrophenolate anion per minute at pH 5, 25°C.

Flanking: The term "flanking" means DNA sequences extending on either side of a specific DNA sequence, locus, or gene. The flanking DNA is immediately adjacent to another DNA sequence, locus, or gene that is to be integrated into the genome of a filamentous fungal cell.

Fragment: The term "fragment" means a polypeptide having one or more (*e.g.*,

several) amino acids absent from the amino and/or carboxyl terminus of a mature polypeptide main; wherein the fragment has enzyme activity. In one aspect, a fragment contains at least 85%, *e.g.*, at least 90% or at least 95% of the amino acid residues of the mature polypeptide of an enzyme.

Hemicellulolytic enzyme or hemicellulase: The term "hemicellulolytic enzyme" or "hemicellulase" means one or more (*e.g.*, several) enzymes that hydrolyze a hemicellulosic material. See, for example, Shallom, D. and Shoham, Y. Microbial hemicellulases. *Current Opinion In Microbiology*, 2003, 6(3): 219-228). Hemicellulases are key components in the degradation of plant biomass. Examples of hemicellulases include, but are not limited to, an acetylmannan esterase, an acetylxylan esterase, an arabinanase, an arabinofuranosidase, a coumaric acid esterase, a feruloyl esterase, a galactosidase, a glucuronidase, a glucuronoyl esterase, a mannanase, a mannosidase, a xylanase, and a xylosidase. The substrates of these enzymes, the hemicelluloses, are a heterogeneous group of branched and linear polysaccharides that are bound via hydrogen bonds to the cellulose microfibrils in the plant cell wall, crosslinking them into a robust network. Hemicelluloses are also covalently attached to lignin, forming together with cellulose a highly complex structure. The variable structure and organization of hemicelluloses require the concerted action of many enzymes for its complete degradation. The catalytic modules of hemicellulases are either glycoside hydrolases (GHs) that hydrolyze glycosidic bonds, or carbohydrate esterases (CEs), which hydrolyze ester linkages of acetate or ferulic acid side groups. These catalytic modules, based on homology of their primary sequence, can be assigned into GH and CE families. Some families, with an overall similar fold, can be further grouped into clans, marked alphabetically (*e.g.*, GH-A). A most informative and updated classification of these and other carbohydrate active enzymes is available in the Carbohydrate-Active Enzymes (CAZy) database. Hemicellulolytic enzyme activities can be measured according to Ghose and Bisaria, 1987, *Pure & Appl. Chem.* 59: 1739-1752, at a suitable temperature, *e.g.*, 50°C, 55°C, or 60°C, and pH, *e.g.*, 5.0 or 5.5.

High stringency conditions: The term "high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 65°C.

Homologous 3' or 5' region: The term "homologous 3' region" means a fragment of DNA that is identical in sequence or has a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least

99% to a region in the genome and when combined with a homologous 5' region can target integration of a piece of DNA to a specific site in the genome by homologous recombination. The term "homologous 5' region" means a fragment of DNA that is identical in sequence to a region in the genome and when combined with a homologous 3' region can target integration of a piece of DNA to a specific site in the genome by homologous recombination. The homologous 5' and 3' regions must be linked in the genome which means they are on the same chromosome and within at least 200 kb of one another.

Homologous flanking region: The term "homologous flanking region" means a fragment of DNA that is identical or has a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to a region in the genome and is located immediately upstream or downstream of a specific site in the genome into which extracellular DNA is targeted for integration.

Homologous repeat: The term "homologous repeat" means a fragment of DNA that is repeated at least twice in the recombinant DNA introduced into a host cell and which can facilitate the loss of the DNA, *i.e.*, selectable marker that is inserted between two homologous repeats, by homologous recombination. A homologous repeat is also known as a direct repeat.

Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, or the like with a nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

Isolated: The term "isolated" means a substance in a form or environment that does not occur in nature. Non-limiting examples of isolated substances include (1) any non-naturally occurring substance, (2) any substance including, but not limited to, any enzyme, variant, nucleic acid, protein, peptide or cofactor, that is at least partially removed from one or more or all of the naturally occurring constituents with which it is associated in nature; (3) any substance modified by the hand of man relative to that substance found in nature; or (4) any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated (*e.g.*, recombinant production in a host cell; multiple copies of a gene encoding the substance; and use of a stronger promoter than the promoter naturally associated with the gene encoding the substance).

Low stringency conditions: The term "low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25%

formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 50°C.

Mature polypeptide: The term "mature polypeptide" means a polypeptide in its final form following translation and any post-translational modifications, such as N-terminal processing, C-terminal truncation, glycosylation, phosphorylation, etc. In one aspect, the mature polypeptide of a *T. reesei* cellobiohydrolase I is amino acids 18 to 514 of SEQ ID NO: 2 based on the SignalP program (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6) that predicts amino acids 1 to 17 of SEQ ID NO: 2 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* cellobiohydrolase II is amino acids 19 to 471 of SEQ ID NO: 4 based on the SignalP program that predicts amino acids 1 to 18 of SEQ ID NO: 4 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* endoglucanase I is amino acids 23 to 459 of SEQ ID NO: 6 based on the SignalP program that predicts amino acids 1 to 22 of SEQ ID NO: 6 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* endoglucanase II is amino acids 22 to 418 of SEQ ID NO: 8 based on the SignalP program that predicts amino acids 1 to 21 of SEQ ID NO: 8 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* beta-glucosidase is amino acids 20 to 744 of SEQ ID NO: 10 based on the SignalP program that predicts amino acids 1 to 19 of SEQ ID NO: 10 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* xylanase I is amino acids 20 to 229 of SEQ ID NO: 12 based on the SignalP program that predicts amino acids 1 to 19 of SEQ ID NO: 12 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* xylanase II is amino acids 20 to 223 of SEQ ID NO: 14 based on the SignalP program that predicts amino acids 1 to 19 of SEQ ID NO: 14 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* xylanase III is amino acids 17 to 347 of SEQ ID NO: 16 based on the SignalP program that predicts amino acids 1 to 16 of SEQ ID NO: 16 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* beta-xylosidase is amino acids 21 to 797 of SEQ ID NO: 18 based on the SignalP program that predicts amino acids 1 to 20 of SEQ ID NO: 18 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* swollenin is amino acids 19 to 493 of SEQ ID NO: 20 based on the SignalP program that predicts amino acids 1 to 18 of SEQ ID NO: 20 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* subtilisin-like serine protease is amino acids 20 to 882 of SEQ ID NO: 22 based on the SignalP program that predicts amino acids 1 to 19 of SEQ ID NO: 22 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* aspartic protease is amino acids 21 to 407 of SEQ ID NO: 24 based on the SignalP program that predicts amino acids 1 to 20 of SEQ ID NO: 24 are a signal peptide. In another aspect, the mature polypeptide of a *T. reesei* trypsin-like serine protease is amino acids 20 to 259 of SEQ ID NO: 26 based on the SignalP program that predicts amino acids 1 to 19 of SEQ ID NO: 26 are a signal peptide. In another

aspect, the mature polypeptide of another *T. reesei* subtilisin-like serine protease is amino acids 16 to 540 of SEQ ID NO: 28 based on the SignalP program that predicts amino acids 1 to 15 of SEQ ID NO: 28 are a signal peptide. In another aspect, the mature polypeptide of another *T. reesei* aspartic protease is amino acids 18 to 395 of SEQ ID NO: 30 based on the SignalP program that predicts amino acids 1 to 17 of SEQ ID NO: 30 are a signal peptide. It is known in the art that a host cell may produce a mixture of two of more different mature polypeptides (*i.e.*, with a different C-terminal and/or N-terminal amino acid) expressed by the same polynucleotide.

Mature polypeptide coding sequence: The term "mature polypeptide coding sequence" means a polynucleotide that encodes a mature polypeptide having enzyme activity. In one aspect, the mature polypeptide coding sequence of a *T. reesei* cellobiohydrolase I is nucleotides 52 to 1545 of SEQ ID NO: 1 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 51 of SEQ ID NO: 1 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* cellobiohydrolase II is nucleotides 55 to 1608 of SEQ ID NO: 3 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 54 of SEQ ID NO: 3 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* endoglucanase I is nucleotides 67 to 1374 of SEQ ID NO: 5 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 66 of SEQ ID NO: 5 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* endoglucanase II is nucleotides 64 to 1254 of SEQ ID NO: 7 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 63 of SEQ ID NO: 7 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* beta-glucosidase is nucleotides 58 to 2612 of SEQ ID NO: 9 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 57 of SEQ ID NO: 9 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* xylanase I is nucleotides 58 to 749 of SEQ ID NO: 11 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 57 of SEQ ID NO: 11 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* xylanase II is nucleotides 58 to 778 of SEQ ID NO: 13 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 57 of SEQ ID NO: 13 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* xylanase III is nucleotides 49 to 1349 of SEQ ID NO: 15 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 48 of SEQ ID NO: 15 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* beta-xylosidase is nucleotides 61 to 2391 of SEQ ID NO: 17 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 60 of SEQ ID NO: 17 encode a signal

peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* swollenin is nucleotides 55 to 2776 of SEQ ID NO: 19 or the cDNA sequence thereof based on the SignalP program that predicts nucleotides 1 to 54 of SEQ ID NO: 19 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* subtilisin-like serine protease is nucleotides 58 to 2774 of SEQ ID NO: 21 based on the SignalP program that predicts nucleotides 1 to 57 of SEQ ID NO: 21 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* aspartic protease is nucleotides 61 to 1299 of SEQ ID NO: 23 based on the SignalP program that predicts nucleotides 1 to 60 of SEQ ID NO: 23 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of a *T. reesei* trypsin-like protease is nucleotides 58 to 930 of SEQ ID NO: 25 based on the SignalP program that predicts nucleotides 1 to 57 of SEQ ID NO: 25 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of another *T. reesei* subtilisin-like serine protease is nucleotides 46 to 1681 of SEQ ID NO: 27 based on the SignalP program that predicts nucleotides 1 to 45 of SEQ ID NO: 27 encode a signal peptide. In another aspect, the mature polypeptide coding sequence of another *T. reesei* aspartic protease is nucleotides 52 to 1339 of SEQ ID NO: 29 based on the SignalP program that predicts nucleotides 1 to 51 of SEQ ID NO: 29 encode a signal peptide.

Medium stringency conditions: The term “medium stringency conditions” means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 55°C.

Medium-high stringency conditions: The term “medium-high stringency conditions” means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 60°C.

Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic, which comprises one or more (*e.g.*, several) control sequences.

Operably linked: The term “operably linked” means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a

polynucleotide such that the control sequence directs expression of the coding sequence.

Polypeptide having cellulolytic enhancing activity: The term "polypeptide having cellulolytic enhancing activity" means a GH61 polypeptide that catalyzes the enhancement of the hydrolysis of a cellulosic material by enzyme having cellulolytic activity. For purposes of the present invention, cellulolytic enhancing activity is determined by measuring the increase in reducing sugars or the increase of the total of cellobiose and glucose from the hydrolysis of a cellulosic material by cellulolytic enzyme under the following conditions: 1-50 mg of total protein/g of cellulose in PCS, wherein total protein is comprised of 50-99.5% w/w cellulolytic enzyme protein and 0.5-50% w/w protein of a GH61 polypeptide having cellulolytic enhancing activity for 1-7 days at a suitable temperature, *e.g.*, 50°C, 55°C, or 60°C, and pH, *e.g.*, 5.0 or 5.5, compared to a control hydrolysis with equal total protein loading without cellulolytic enhancing activity (1-50 mg of cellulolytic protein/g of cellulose in PCS). In a preferred aspect, a mixture of CELLUCLAST® 1.5L (Novozymes A/S, Bagsværd, Denmark) in the presence of 2-3% of total protein weight *Aspergillus oryzae* beta-glucosidase (recombinantly produced in *Aspergillus oryzae* according to WO 02/095014) or 2-3% of total protein weight *Aspergillus fumigatus* beta-glucosidase (recombinantly produced in *Aspergillus oryzae* as described in WO 2002/095014) of cellulase protein loading is used as the source of the cellulolytic activity.

The GH61 polypeptides having cellulolytic enhancing activity enhance the hydrolysis of a cellulosic material catalyzed by enzyme having cellulolytic activity by reducing the amount of cellulolytic enzyme required to reach the same degree of hydrolysis preferably at least 1.01-fold, *e.g.*, at least 1.05-fold, at least 1.10-fold, at least 1.25-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 20-fold.

Pretreated corn stover: The term "PCS" or "Pretreated Corn Stover" means a cellulosic material derived from corn stover by treatment with heat and dilute sulfuric acid, alkaline pretreatment, or neutral pretreatment.

Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *Trends Genet.* 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity"

(obtained using the –nobrief option) is used as the percent identity and is calculated as follows:

(Identical Residues x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the –nobrief option) is used as the percent identity and is calculated as follows:

(Identical Deoxyribonucleotides x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

Subsequence: The term "subsequence" means a polynucleotide having one or more (*e.g.*, several) nucleotides absent from the 5' and/or 3' end of a mature polypeptide coding sequence; wherein the subsequence encodes a fragment having enzyme activity. In one aspect, a subsequence contains at least 85%, *e.g.*, at least 90% or at least 95% of the nucleotides of the mature polypeptide coding sequence of an enzyme.

Subtilisin-like serine protease: The term "subtilisin-like serine protease" means a protease with a substrate specificity similar to subtilisin that uses a serine residue for catalyzing the hydrolysis of peptide bonds in peptides and proteins. Subtilisin-like proteases (subtilases) are serine proteases characterized by a catalytic triad of the three amino acids aspartate, histidine, and serine. The arrangement of these catalytic residues is shared with the prototypical subtilisin from *Bacillus licheniformis* (Siezen and Leunissen, 1997, *Protein Science* 6: 501-523). Subtilisin-like serine protease activity can be determined using a synthetic substrate, N-succinyl-L-Ala-L-Ala-L-Pro-L-Phe-p-nitroanilide (AAPF) (Bachem AG, Bubendorf, Switzerland) in 100 mM NaCl-100 mM MOPS pH 7.0 at 50°C for 3 hours and then the absorbance at 405 nm is measured.

Targeted integration: The term "targeted integration" means the stable integration of extracellular DNA at a defined genomic locus.

Transformant: The term "transformant" means a cell which has taken up extracellular DNA (foreign, artificial or modified) and expresses the gene(s) contained therein.

Transformation: The term "transformation" means the introduction of extracellular DNA into a cell, *i.e.*, the genetic alteration of a cell resulting from the direct uptake, incorporation and expression of exogenous genetic material (exogenous DNA) from its surroundings and taken up through the cell membrane(s).

Transformation efficiency: The term "transformation efficiency" means the efficiency by which cells can take up the extracellular DNA and express the gene(s) contained therein, which is calculated by dividing the number of positive transformants expressing the gene(s) by the amount of DNA used during a transformation procedure.

Trypsin-like serine protease: The term "trypsin-like serine protease" means a protease with a substrate specificity similar to trypsin that uses a serine residue for catalyzing the hydrolysis of peptide bonds in peptides and proteins. For purposes of the present invention, trypsin-like serine protease activity is determined according to the procedure described by Dienes *et al.*, 2007, *Enzyme and Microbial Technology* 40: 1087-1094.

Variant: The term "variant" means a polypeptide having enzyme activity comprising an alteration, *i.e.*, a substitution, insertion, and/or deletion, at one or more (*e.g.*, several) positions. A substitution means replacement of the amino acid occupying a position with a different amino acid; a deletion means removal of the amino acid occupying a position; and an insertion means adding an amino acid adjacent to and immediately following the amino acid occupying a position.

Very high stringency conditions: The term "very high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 70°C.

Very low stringency conditions: The term "very low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 45°C.

Xylan-containing material: The term "xylan-containing material" means any material comprising a plant cell wall polysaccharide containing a backbone of beta-(1-4)-linked xylose residues. Xylans of terrestrial plants are heteropolymers possessing a beta-(1-4)-D-xylopyranose backbone, which is branched by short carbohydrate chains. They comprise D-glucuronic acid or its 4-*O*-methyl ether, L-arabinose, and/or various oligosaccharides, composed of D-xylose, L-arabinose, D- or L-galactose, and D-glucose. Xylan-type polysaccharides can be divided into homoxylans and heteroxylans, which include glucuronoxylans, (arabino)glucuronoxylans, (glucurono)arabinoxylans, arabinoxylans, and

complex heteroxylans. See, for example, Ebringerova *et al.*, 2005, *Adv. Polym. Sci.* 186: 1–67.

In the processes of the present invention, any material containing xylan may be used. In a preferred aspect, the xylan-containing material is lignocellulose.

Xylan degrading activity or xylanolytic activity: The term "xylan degrading activity" or "xylanolytic activity" means a biological activity that hydrolyzes xylan-containing material. The two basic approaches for measuring xylanolytic activity include: (1) measuring the total xylanolytic activity, and (2) measuring the individual xylanolytic activities (*e.g.*, endoxylanases, beta-xylosidases, arabinofuranosidases, alpha-glucuronidases, acetylxylan esterases, feruloyl esterases, and alpha-glucuronyl esterases). Recent progress in assays of xylanolytic enzymes was summarized in several publications including Biely and Puchard, Recent progress in the assays of xylanolytic enzymes, 2006, *Journal of the Science of Food and Agriculture* 86(11): 1636-1647; Spanikova and Biely, 2006, Glucuronoyl esterase - Novel carbohydrate esterase produced by *Schizophyllum commune*, *FEBS Letters* 580(19): 4597-4601; Herrmann, Vrsanska, Jurickova, Hirsch, Biely, and Kubicek, 1997, The beta-D-xylosidase of *Trichoderma reesei* is a multifunctional beta-D-xylan xylohydrolase, *Biochemical Journal* 321: 375-381.

Total xylan degrading activity can be measured by determining the reducing sugars formed from various types of xylan, including, for example, oat spelt, beechwood, and larchwood xylans, or by photometric determination of dyed xylan fragments released from various covalently dyed xylans. The most common total xylanolytic activity assay is based on production of reducing sugars from polymeric 4-O-methyl glucuronoxylan as described in Bailey, Biely, Poutanen, 1992, Interlaboratory testing of methods for assay of xylanase activity, *Journal of Biotechnology* 23(3): 257-270. Xylanase activity can also be determined with 0.2% AZCL-arabinoxylan as substrate in 0.01% TRITON® X-100 (4-(1,1,3,3-tetramethylbutyl)phenyl-polyethylene glycol) and 200 mM sodium phosphate buffer pH 6 at 37°C. One unit of xylanase activity is defined as 1.0 μmole of azurine produced per minute at 37°C, pH 6 from 0.2% AZCL-arabinoxylan as substrate in 200 mM sodium phosphate pH 6 buffer.

For purposes of the present invention, xylan degrading activity is determined by measuring the increase in hydrolysis of birchwood xylan (Sigma Chemical Co., Inc., St. Louis, MO, USA) by xylan-degrading enzyme(s) under the following typical conditions: 1 ml reactions, 5 mg/ml substrate (total solids), 5 mg of xylanolytic protein/g of substrate, 50 mM sodium acetate pH 5, 50°C, 24 hours, sugar analysis using *p*-hydroxybenzoic acid hydrazide (PHBAH) assay as described by Lever, 1972, A new reaction for colorimetric determination of carbohydrates, *Anal. Biochem* 47: 273-279.

Xylanase: The term "xylanase" means a 1,4-beta-D-xylan-xylohydrolase (E.C. 3.2.1.8) that catalyzes the endohydrolysis of 1,4-beta-D-xylosidic linkages in xylans. For purposes of the present invention, xylanase activity is determined with 0.2% AZCL-arabinoxylan as substrate in 0.01% TRITON® X-100 and 200 mM sodium phosphate buffer pH 6 at 37°C. One unit of xylanase activity is defined as 1.0 μmole of azurine produced per minute at 37°C, pH 6 from 0.2% AZCL-arabinoxylan as substrate in 200 mM sodium phosphate pH 6 buffer.

Detailed Description of the Invention

The present invention relates to methods for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising: (a) replacing an endogenous first gene by targeted integration by introducing into the filamentous fungal strain a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) replacing an endogenous second gene by targeted integration by introducing into the filamentous fungal strain a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

The present invention also relates to tandem constructs comprising (i) a homologous 5' region of a gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

The present invention also relates to methods for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising: (a) inserting into an endogenous first locus by targeted integration by introducing into the filamentous fungal strain a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) inserting into an endogenous second locus by targeted integration by introducing into the filamentous fungal strain a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

The present invention also relates to tandem constructs comprising (i) a homologous 5' region of a locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

In the methods of the present invention, each of the tandem constructs integrates by double homologous recombination into a targeted site in the chromosome of the filamentous fungal strain. In one aspect, the homologous 5' region of the first gene or first locus, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp. In another aspect, the homologous 3' region of the first gene or first locus, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp. In another aspect, the homologous 5' region of the second gene or second locus, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800

bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp. In another aspect, the homologous 3' region of the second gene or second locus, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

The methods of the present invention may further comprise replacing one or more (*e.g.*, several) additional endogenous genes each by targeted integration with a corresponding tandem construct for each gene comprising (i) a homologous 5' region of the gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

The methods of the present invention may further comprise inserting into one or more (*e.g.*, several) additional endogenous loci each by targeted integration a corresponding tandem construct for each locus comprising (i) a homologous 5' region of the locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

In the methods of the present invention, the filamentous fungal host cell may be further transformed with a tandem construct comprising (i) one or more (*e.g.*, several) selectable markers, (ii) a fifth polynucleotide encoding a fifth polypeptide having biological activity operably linked to a fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

The present invention provides several advantages including improved methods for producing multiple recombinant polypeptides in a filamentous fungal host; methods allowing easy replacement and/or deletion of one or more recombinant polypeptides introduced into the filamentous fungal host (if one or more recombinant polypeptides are introduced at a single locus, one of more recombinant polypeptides can be replaced and or/deleted in a single step, instead of multiple steps); and method allowing easy modification (*i.e.*, introducing a variant/mutant gene) of existing one or more recombinant poly peptides in a filamentous fungal strain (if one or more recombinant polypeptides were introduced at a

single locus, one or more variant/mutant of those recombinant polypeptides can be easily modified in a single step, instead of multiple steps). An additional advantage of a constructed filamentous fungal strain of the present invention is the possibility of mutagenizing the strain and selecting for yield improved mutants. A tandem construct of such a mutant can then be replaced with a new tandem construct expressing new multiple recombinant polypeptides thereby taking advantage of the improved yield productivity of the mutant as a host cell.

Endogenous Genes

In the methods of the present invention, any gene endogenous to a filamentous fungal strain may be replaced or may be a locus. The gene may be native or foreign to the filamentous fungal strain. The term "endogenous gene" or variations thereof, *e.g.,* "endogenous first gene" or "endogenous second gene", will be understood to encompass one or more (*e.g.*, several) genes. Where more than one gene is replaced, the genes are preferably contiguous. Such multiple genes may be, for example, a metabolic pathway or portion thereof.

The endogenous gene may encode a polypeptide selected from the group consisting of an antibody, an antigen, an antimicrobial peptide, an enzyme, a growth factor, a hormone, an immunodilator, a neurotransmitter, a receptor, a reporter protein, a structural protein, or a transcription factor.

In one aspect, the enzyme is selected from the group consisting of an oxidoreductase, a transferase, a hydrolase, a lyase, an isomerase, and a ligase. In another aspect, the enzyme is selected from the group consisting of an acetylmannan esterase, acetyxylan esterase, aminopeptidase, alpha-amylase, alpha-galactosidase, alpha-glucosidase, alpha-1,6-transglucosidase, arabinanase, arabinofuranosidase, beta-galactosidase, beta-glucosidase, beta-xylosidase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, coumaric acid esterase, cyclodextrin glycosyltransferase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, endoglucanase, esterase, feruloyl esterase, GH61 polypeptide having cellulolytic enhancing activity, glucocerebrosidase, glucose oxidase, glucuronidase, glucuronoyl esterase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, mannanase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phospholipase, phytase, phenoloxidase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, urokinase, and xylanase.

In another aspect, one or more (*e.g.*, several) endogenous protease genes are inactivated. In another aspect, the one or more (*e.g.*, several) endogenous protease genes

are subtilisin-like serine protease, aspartic protease, and trypsin-like serine protease genes as described in WO 2011/075677, which is incorporated herein by reference in its entirety.

In another aspect, the enzyme is a subtilisin-like serine protease. In another aspect, the enzyme is an aspartic protease. In another aspect, the enzyme is a trypsin-like serine protease.

In another aspect, the enzyme is an endoglucanase. In another aspect, the enzyme is a cellobiohydrolase. In another aspect, the enzyme is a beta-glucosidase. In another aspect, the enzyme is a GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme is a xylanase. In another aspect, the enzyme is a beta-xylosidase. In another aspect, the enzyme is a swollenin. In another aspect, the enzyme is an acetyxylan esterase. In another aspect, the enzyme is a feruloyl esterase. In another aspect, the enzyme is an arabinofuranosidase. In another aspect, the enzyme is a glucuronidase. In another aspect, the enzyme is an acetylmannan esterase. In another aspect, the enzyme is an arabinanase. In another aspect, the enzyme is a coumaric acid esterase. In another aspect, the enzyme is a galactosidase. In another aspect, the enzyme is a glucuronoyl esterase. In another aspect, the enzyme is a mannanase. In another aspect, the enzyme is a mannosidase.

In another aspect, the endogenous gene may be a cellobiohydrolase I gene. In another aspect, the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

In another aspect, the endogenous gene may be a cellobiohydrolase II gene. In another aspect, the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%,

e.g., at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

In another aspect, the endogenous gene may be an endoglucanase I gene. In another aspect, the endoglucanase I gene encodes an endoglucanase I selected from the group consisting of: (i) an endoglucanase I comprising the mature polypeptide of SEQ ID NO: 6; (ii) an endoglucanase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 6; (iii) an endoglucanase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 5; and (iv) an endoglucanase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 5 or the full-length complement thereof.

In another aspect, the endogenous gene may be an endoglucanase II gene. In another aspect, the endoglucanase II gene encodes an endoglucanase II selected from the group consisting of: (i) an endoglucanase II comprising the mature polypeptide of SEQ ID NO: 8; (ii) an endoglucanase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 8; (iii) an

endoglucanase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 7; and (iv) an endoglucanase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 7 or the full-length complement thereof.

In another aspect, the endogenous gene may be a beta-glucosidase gene. In another aspect, the beta-glucosidase gene encodes a beta-glucosidase selected from the group consisting of: (i) a beta-glucosidase comprising the mature polypeptide of SEQ ID NO: 10; (ii) a beta-glucosidase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 10; (iii) a beta-glucosidase encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 9; and (iv) a beta-glucosidase encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 9 or the full-length complement thereof.

In another aspect, the endogenous gene may be a xylanase I gene. In another aspect, the xylanase I gene encodes a xylanase I selected from the group consisting of: (i) a xylanase I comprising the mature polypeptide of SEQ ID NO: 12; (ii) a xylanase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 12; (iii) a xylanase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 11; and (iv) a xylanase I

encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 11 or the full-length complement thereof.

In another aspect, the endogenous gene may be a xylanase II gene. In another aspect, the xylanase II gene encodes a xylanase II selected from the group consisting of: (i) an xylanase II comprising the mature polypeptide of SEQ ID NO: 14; (ii) a xylanase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 14; (iii) a xylanase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 13; and (iv) a xylanase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 13 or the full-length complement thereof.

In another aspect, the endogenous gene may be a xylanase III gene. In another aspect, the xylanase III gene encodes a xylanase III selected from the group consisting of: (i) an xylanase III comprising the mature polypeptide of SEQ ID NO: 16; (ii) a xylanase III comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 16; (iii) a xylanase III encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 15; and (iv) a xylanase III encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 15 or the full-length complement thereof.

In another aspect, the endogenous gene may be a beta-xylosidase gene. In another aspect, the beta-xylosidase gene encodes a beta-xylosidase selected from the group

consisting of: (i) a beta-xylosidase comprising the mature polypeptide of SEQ ID NO: 18; (ii) a beta-xylosidase comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 18; (iii) a beta-xylosidase encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 17; and (iv) a beta-xylosidase encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 17 or the full-length complement thereof.

In another aspect, the endogenous gene may be a swollenin gene. In another aspect, the swollenin gene encodes a swollenin selected from the group consisting of: (i) a swollenin comprising the mature polypeptide of SEQ ID NO: 20; (ii) a swollenin comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 20; (iii) a swollenin encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 19; and (iv) a swollenin encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 19 or the full-length complement thereof.

In another aspect, the endogenous gene may be a subtilisin-like serine protease gene. In another aspect, the subtilisin-like serine protease gene encodes a subtilisin-like serine protease selected from the group consisting of: (i) a subtilisin-like serine protease comprising the mature polypeptide of SEQ ID NO: 22; (ii) a subtilisin-like serine protease comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least

94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 22; (iii) a subtilisin-like serine protease encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 21; and (iv) a subtilisin-like serine protease encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 21 or the full-length complement thereof.

In another aspect, the endogenous gene may be an aspartic protease gene. In another aspect, the aspartic protease gene encodes an aspartic protease selected from the group consisting of: (i) an aspartic protease comprising the mature polypeptide of SEQ ID NO: 24; (ii) an aspartic protease comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 24; (iii) an aspartic protease encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 23; and (iv) an aspartic protease encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 23 or the full-length complement thereof.

In another aspect, the endogenous gene may be a trypsin-like serine protease gene. In another aspect, the trypsin-like serine protease gene encodes a trypsin-like serine protease selected from the group consisting of: (i) a trypsin-like serine protease comprising the mature polypeptide of SEQ ID NO: 26; (ii) a trypsin-like serine protease comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 26; (iii) a trypsin-like serine protease encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at

least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 25; and (iv) a trypsin-like serine protease encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 25 or the full-length complement thereof.

In another aspect, the endogenous gene may be a subtilisin-like serine protease gene. In another aspect, the subtilisin-like serine protease gene encodes a subtilisin-like serine protease selected from the group consisting of: (i) a subtilisin-like serine protease comprising the mature polypeptide of SEQ ID NO: 28; (ii) a subtilisin-like serine protease comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 28; (iii) a subtilisin-like serine protease encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 27; and (iv) a subtilisin-like serine protease encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 27 or the full-length complement thereof.

In another aspect, the endogenous gene may be an aspartic protease gene. In another aspect, the aspartic protease gene encodes an aspartic protease selected from the group consisting of: (i) an aspartic protease comprising the mature polypeptide of SEQ ID NO: 30; (ii) an aspartic protease comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 30; (iii) an aspartic protease encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence

of SEQ ID NO: 29; and (iv) an aspartic protease encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 29 or the full-length complement thereof.

Tandem Constructs

The present invention also relates to tandem constructs comprising (i) a homologous 5' region of a gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

The present invention also relates to tandem constructs comprising (i) a homologous 5' region of a locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

The tandem constructs can be constructed by operably linking one or more (*e.g.*, several) control sequences to each polynucleotide of the construct that direct the expression of the coding sequence in a filamentous fungal host cell under conditions compatible with the control sequences. Manipulation of each polynucleotide prior to insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

The control sequence may be a promoter, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding a polypeptide. The promoter contains transcriptional control sequences that mediate the expression of the polypeptide. The promoter may be any polynucleotide that shows transcriptional activity in the host cell including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

In one aspect, the promoters are different promoters. In another aspect, two or more (*e.g.*, several) of the promoters are the same promoter.

Examples of suitable promoters for directing transcription of the constructs in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus nidulans*

acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TAKA amylase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor, as well as the NA2-tpi promoter (a modified promoter from an *Aspergillus* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus* triose phosphate isomerase gene; non-limiting examples include modified promoters from an *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus nidulans* or *Aspergillus oryzae* triose phosphate isomerase gene); and mutant, truncated, and hybrid promoters thereof. Other promoters are described in U.S. Patent No. 6,011,147, which is incorporated herein in its entirety.

The control sequence may also be a transcription terminator, which is recognized by a host cell to terminate transcription. The terminator is operably linked to the 3'-terminus of the polynucleotide encoding the polypeptide. Any terminator that is functional in the host cell may be used in the present invention.

In one aspect, the terminators are different terminators. In another aspect, two or more (*e.g.*, several) of the terminators are the same terminator.

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus nidulans* acetamidase (*amdS*), *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, *Fusarium oxysporum* trypsin-like protease, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor.

The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader is operably linked to the 5'-terminus of the polynucleotide encoding the polypeptide. Any leader that is functional in a host cell may

be used.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the polynucleotide and, when transcribed, is recognized by a host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase *Aspergillus oryzae* TAKA amylase, *Fusarium oxysporum* trypsin-like protease, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, and *Trichoderma reesei* endoglucanase V.

The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide into a cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the polypeptide. However, any signal peptide coding sequence that directs the expressed polypeptide into the secretory pathway of a host cell may be used.

Effective signal peptide coding sequences for filamentous fungal host cells are the signal peptide coding sequences obtained from the genes for *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, and *Trichoderma reesei* endoglucanase V.

The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may

be obtained from the genes for *Myceliophthora thermophila* laccase (WO 95/33836) and *Rhizomucor miehei* aspartic proteinase.

Where both signal peptide and propeptide sequences are present, the propeptide sequence is positioned next to the N-terminus of a polypeptide and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

It may also be desirable to add regulatory sequences that regulate expression of the polypeptide relative to the growth of a host cell. Examples of regulatory sequences are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory sequences in filamentous fungi include the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TAKA alpha-amylase promoter, *Aspergillus oryzae* glucoamylase promoter, *Trichoderma reesei* cellobiohydrolase I promoter, and *Trichoderma reesei* cellobiohydrolase II promoter. Other examples of regulatory sequences are those that allow for gene amplification. In these cases, the polynucleotide encoding the polypeptide would be operably linked to the regulatory sequence.

The tandem constructs of the present invention preferably contain one or more (*e.g.*, several) selectable markers that permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like. Examples of selectable markers for use in a filamentous fungal host cell include, but are not limited to, *adeA* (phosphoribosylaminoimidazole-succinocarboxamide synthase), *adeB* (phosphoribosylamino-imidazole synthase), *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae amdS* and *pyrG* genes and a *Streptomyces hygroscopicus bar* gene. Preferred for use in a *Trichoderma* cell are *adeA*, *adeB*, *amdS*, *hph*, and *pyrG* genes. Examples of bacterial selectable markers are markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin, or tetracycline resistance.

The one or more (*e.g.*, several) selectable markers may be a dual selectable marker system as described in WO 2010/039889 A2, which is incorporated herein by reference in its entirety. In one aspect, the one or more selectable markers is a *hph-tk* dual selectable marker system.

In each tandem construct of the present invention, the one or more (*e.g.*, several) selectable markers are different markers, unless a selectable marker is reused as described herein.

One or more (*e.g.*, several) of the selectable markers may be reused for replacing one or more (*e.g.*, several) additional endogenous genes or for inserting into one or more additional endogenous loci each by targeted integration with a corresponding tandem construct for each gene or locus. The one or more tandem constructs may further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers. Upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for replacing the one or more additional endogenous genes or for inserting into the one or more additional endogenous loci each by targeted integration with the corresponding tandem construct for each gene or locus.

In one aspect, the first and second homologous repeats are identical. In another aspect, the first and second homologous repeats have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other. In another aspect, the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp. The fragment containing one repeat may be longer than the fragment containing the other repeat.

The tandem constructs of the present invention may further comprise one or more (*e.g.*, several) additional polynucleotides encoding other polypeptides having biological activity. For example, a tandem construct may contain one additional polynucleotide, two additional polynucleotides, three additional polynucleotides, etc.

Polypeptides Having Biological Activity

The polypeptides may be any polypeptides having a biological activity of interest. The term "polypeptide" is not meant herein to refer to a specific length of the encoded product and, therefore, encompasses peptides, oligopeptides, and proteins. The term "polypeptide" also encompasses two or more (*e.g.*, several) polypeptides combined to form the encoded product. The polypeptides also include fusion polypeptides, which comprise a combination of partial or complete polypeptide sequences obtained from at least two different polypeptides wherein one or more (*e.g.*, several) may be heterologous to the filamentous fungal host strain. The polypeptides may further include naturally occurring allelic and engineered variations of the below-mentioned polypeptides and hybrid polypeptides.

In one aspect, the polypeptides are selected from the group consisting of an antibody, an antigen, an antimicrobial peptide, an enzyme, a growth factor, a hormone, an immunodilator, a neurotransmitter, a receptor, a reporter protein, a structural protein, or a transcription factor.

In another aspect, the enzyme is selected from the group consisting of an oxidoreductase, a transferase, a hydrolase, a lyase, an isomerase, and a ligase. In another aspect, the enzyme is selected from the group consisting of an acetylmannan esterase, acetyxylan esterase, aminopeptidase, alpha-amylase, alpha-galactosidase, alpha-glucosidase, alpha-1,6-transglucosidase, arabinanase, arabinofuranosidase, beta-galactosidase, beta-glucosidase, beta-xylosidase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, coumaric acid esterase, cyclodextrin glycosyltransferase, cutinase, deoxyribonuclease, endoglucanase, esterase, feruloyl esterase, GH61 polypeptide having cellulolytic enhancing activity, glucocerebrosidase, glucose oxidase, glucuronidase, glucuronoyl esterase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, mannanase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phospholipase, phytase, phenoloxidase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, urokinase, and xylanase.

In another aspect, the polypeptides are selected from the group consisting of an albumin, a collagen, a tropoelastin, an elastin, and a gelatin.

In another aspect, the polypeptides having biological activity may be different polypeptides. In another aspect, two or more (*e.g.*, several) of the polypeptides having biological activity are the same polypeptide.

In another aspect, the polypeptides comprise one or more enzymes selected from the group consisting of a cellulase, a cip1 protein, a GH61 polypeptide having cellulolytic enhancing activity, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin. In another aspect, the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase. In another aspect, the hemicellulase is one or more enzymes selected from the group consisting of a xylanase, an acetylxylan esterase, a feruloyl esterase, an arabinofuranosidase, a xylosidase, and a glucuronidase.

In another aspect, one of the polypeptides is a cellulase. In another aspect, one of the polypeptides is an endoglucanase. In another aspect, one of the polypeptides is a cellobiohydrolase. In another aspect, one of the polypeptides is a beta-glucosidase. In another aspect, one of the polypeptides is a GH61 polypeptide having cellulolytic enhancing activity. In another aspect, one of the polypeptides is a swollenin protein. In another aspect, one of the polypeptides is a cip1 protein. In another aspect, one of the polypeptides is an

esterase, In another aspect, one of the polypeptides is an expansin. In another aspect, one of the polypeptides is a laccase. In another aspect, one of the polypeptides is a ligninolytic enzyme. In another aspect, one of the polypeptides is a pectinase, In another aspect, one of the polypeptides is a peroxidase. In another aspect, one of the polypeptides is a protease. In another aspect, one of the polypeptides is a swollenin.

In another aspect, one of the polypeptides is a hemicellulase. In another aspect, one of the polypeptides is a xylanase. In another aspect, one of the polypeptides is a beta-xylosidase. In another aspect, one of the polypeptides is an acetyxylan esterase. In another aspect, one of the polypeptides is a feruloyl esterase. In another aspect, one of the polypeptides is an arabinofuranosidase. In another aspect, one of the polypeptides is a glucuronidase. In another aspect, one of the polypeptides is an acetylmannan esterase. In another aspect, one of the polypeptides is an arabinanase. In another aspect, one of the polypeptides is a coumaric acid esterase. In another aspect, one of the polypeptides is a galactosidase. In another aspect, one of the polypeptides is a glucuronoyl esterase. In another aspect, one of the polypeptides is a mannanase. In another aspect, one of the polypeptides is a mannosidase.

Examples of endoglucanases as the polypeptides having biological activity include, but are not limited to, a *Trichoderma reesei* endoglucanase I (Penttila *et al.*, 1986, *Gene* 45: 253-263, *Trichoderma reesei* Cel7B endoglucanase I (GENBANK™ accession no. M15665), *Trichoderma reesei* endoglucanase II (Saloheimo, *et al.*, 1988, Gene 63:11-22), *Trichoderma reesei* Cel5A endoglucanase II (GENBANK™ accession no. M19373), *Trichoderma reesei* endoglucanase III (Okada *et al.*, 1988, *Appl. Environ. Microbiol.* 64: 555-563, GENBANK™ accession no. AB003694), *Trichoderma reesei* endoglucanase V (Saloheimo *et al.*, 1994, *Molecular Microbiology* 13: 219-228, GENBANK™ accession no. Z33381), *Aspergillus aculeatus* endoglucanase (Ooi *et al.*, 1990, *Nucleic Acids Research* 18: 5884), *Aspergillus kawachii* endoglucanase (Sakamoto *et al.*, 1995, *Current Genetics* 27: 435-439), *Erwinia carotovara* endoglucanase (Saarilahti *et al.*, 1990, *Gene* 90: 9-14), *Fusarium oxysporum* endoglucanase (GENBANK™ accession no. L29381), *Humicola grisea* var. *thermoidea* endoglucanase (GENBANK™ accession no. AB003107), *Melanocarpus albomyces* endoglucanase (GENBANK™ accession no. MAL515703), *Neurospora crassa* endoglucanase (GENBANK™ accession no. XM_324477), *Humicola insolens* endoglucanase V, *Myceliophthora thermophila* CBS 117.65 endoglucanase, basidiomycete CBS 495.95 endoglucanase, basidiomycete CBS 494.95 endoglucanase, *Thielavia terrestris* NRRL 8126 CEL6B endoglucanase, *Thielavia terrestris* NRRL 8126 CEL6C endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7C endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7E endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7F endoglucanase,

Cladorrhinum foecundissimum ATCC 62373 CEL7A endoglucanase, and *Trichoderma reesei* strain No. VTT-D-80133 endoglucanase (GENBANK™ accession no. M15665).

Examples of cellobiohydrolases as the polypeptides having biological activity include, but are not limited to, *Aspergillus aculeatus* cellobiohydrolase II (WO 2011/059740), *Chaetomium thermophilum* cellobiohydrolase I, *Chaetomium thermophilum* cellobiohydrolase II, *Humicola insolens* cellobiohydrolase I, *Myceliophthora thermophila* cellobiohydrolase II (WO 2009/042871), *Thielavia hyrcanie* cellobiohydrolase II (WO 2010/141325), *Thielavia terrestris* cellobiohydrolase II (CEL6A, WO 2006/074435), *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, and *Trichophaea saccata* cellobiohydrolase II (WO 2010/057086).

Examples of beta-glucosidases as the polypeptides having biological activity include, but are not limited to, beta-glucosidases from *Aspergillus aculeatus* (Kawaguchi *et al.*, 1996, *Gene* 173: 287-288), *Aspergillus fumigatus* (WO 2005/047499), *Aspergillus niger* (Dan *et al.*, 2000, *J. Biol. Chem.* 275: 4973-4980), *Aspergillus oryzae* (WO 2002/095014), *Penicillium brasilianum* IBT 20888 (WO 2007/019442 and WO 2010/088387), *Thielavia terrestris* (WO 2011/035029), and *Trichophaea saccata* (WO 2007/019442).

The beta-glucosidase may also be a fusion protein. In one aspect, the beta-glucosidase is an *Aspergillus oryzae* beta-glucosidase variant BG fusion protein (WO 2008/057637) or an *Aspergillus oryzae* beta-glucosidase fusion protein (WO 2008/057637).

Examples of other endoglucanases, cellobiohydrolases, and beta-glucosidases are disclosed in numerous Glycosyl Hydrolase families using the classification according to Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280: 309-316, and Henrissat B., and Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, *Biochem. J.* 316: 695-696.

Other cellulolytic enzymes that may be used in the present invention are described in WO 98/13465, WO 98/015619, WO 98/015633, WO 99/06574, WO 99/10481, WO 99/025847, WO 99/031255, WO 2002/101078, WO 2003/027306, WO 2003/052054, WO 2003/052055, WO 2003/052056, WO 2003/052057, WO 2003/052118, WO 2004/016760, WO 2004/043980, WO 2004/048592, WO 2005/001065, WO 2005/028636, WO 2005/093050, WO 2005/093073, WO 2006/074005, WO 2006/117432, WO 2007/071818, WO 2007/071820, WO 2008/008070, WO 2008/008793, U.S. Patent No. 5,457,046, U.S. Patent No. 5,648,263, and U.S. Patent No. 5,686,593.

Examples of GH61 polypeptides having cellulolytic enhancing activity as the polypeptides having biological activity include, but are not limited to, GH61 polypeptides from *Thielavia terrestris* (WO 2005/074647, WO 2008/148131, and WO 2011/035027), *Thermoascus aurantiacus* (WO 2005/074656 and WO 2010/065830), *Trichoderma reesei* (WO 2007/089290), *Myceliophthora thermophila* (WO 2009/085935, WO 2009/085859, WO

2009/085864, WO 2009/085868), *Aspergillus fumigatus* (WO 2010/138754), GH61 polypeptides from *Penicillium pinophilum* (WO 2011/005867), *Thermoascus* sp. (WO 2011/039319), *Penicillium* sp. (WO 2011/041397), and *Thermoascus crustaceous* (WO 2011/041504).

Examples of xylanases as the polypeptides having biological activity include, but are not limited to, xylanases from *Aspergillus aculeatus* (GeneSeqP:AAR63790; WO 94/21785), *Aspergillus fumigatus* (WO 2006/078256), *Penicillium pinophilum* (WO 2011/041405), *Penicillium* sp. (WO 2010/126772), *Thielavia terrestris* NRRL 8126 (WO 2009/079210), and *Trichophaea saccata* GH10 (WO 2011/057083).

Examples of beta-xylosidases as the polypeptides having biological activity include, but are not limited to, beta-xylosidases from *Neurospora crassa* (SwissProt accession number Q7SOW4), *Trichoderma reesei* (UniProtKB/TrEMBL accession number Q92458), and *Talaromyces emersonii* (SwissProt accession number Q8X212).

Examples of acetylxylan esterases as the polypeptides having biological activity include, but are not limited to, acetylxylan esterases from *Aspergillus aculeatus* (WO 2010/108918), *Chaetomium globosum* (Uniprot accession number Q2GWX4), *Chaetomium gracile* (GeneSeqP accession number AAB82124), *Humicola insolens* DSM 1800 (WO 2009/073709), *Hypocrea jecorina* (WO 2005/001036), *Myceliophtera thermophila* (WO 2010/014880), *Neurospora crassa* (UniProt accession number q7s259), *Phaeosphaeria nodorum* (Uniprot accession number Q0UHJ1), and *Thielavia terrestris* NRRL 8126 (WO 2009/042846).

Examples of feruloyl esterases (ferulic acid esterases) as the polypeptides having biological activity include, but are not limited to, feruloyl esterases form *Humicola insolens* DSM 1800 (WO 2009/076122), *Neosartorya fischeri* (UniProt Accession number A1D9T4), *Neurospora crassa* (UniProt accession number Q9HGR3), *Penicillium aurantiogriseum* (WO 2009/127729), and *Thielavia terrestris* (WO 2010/053838 and WO 2010/065448).

Examples of arabinofuranosidases as the polypeptides having biological activity include, but are not limited to, arabinofuranosidases from *Aspergillus niger* (GeneSeqP accession number AAR94170), *Humicola insolens* DSM 1800 (WO 2006/114094 and WO 2009/073383), and *M. giganteus* (WO 2006/114094).

Examples of alpha-glucuronidases as the polypeptides having biological activity include, but are not limited to, alpha-glucuronidases from *Aspergillus clavatus* (UniProt accession number alcc12), *Aspergillus fumigatus* (SwissProt accession number Q4WW45), *Aspergillus niger* (Uniprot accession number Q96WX9), *Aspergillus terreus* (SwissProt accession number Q0CJP9), *Humicola insolens* (WO 2010/014706), *Penicillium aurantiogriseum* (WO 2009/068565), *Talaromyces emersonii* (UniProt accession number Q8X211), and *Trichoderma reesei* (Uniprot accession number Q99024).

The accession numbers are incorporated herein by reference in their entirety.

Expression Vectors

The present invention also relates to expression vectors comprising a tandem construct of the present invention. A tandem construct may be inserted into a vector or the various components of a tandem construct may be joined together to produce a recombinant expression vector. The vector may include one or more (*e.g.*, several) convenient restriction sites to allow for insertion of polynucleotides at such sites. In creating the expression vector, the coding sequences are located in the vector so that the coding sequences are operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the polynucleotides. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

The vector preferably contains one or more (*e.g.*, several) selectable markers that permit easy selection of transformed cells. Examples of selectable markers for use in a filamentous fungal host cell include, but are not limited to, *adeA* (phosphoribosylaminoimidazole-succinocarboxamide synthase), *adeB* (phosphoribosylaminoimidazole synthase), *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae amdS* and *pyrG* genes and a *Streptomyces hygroscopicus bar* gene. Preferred for use in a *Trichoderma* cell are *adeA*, *adeB*, *amdS*, *hph*, and *pyrG* genes. Examples of bacterial selectable markers are markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin, or tetracycline resistance.

The procedures used to ligate the elements described above to construct the recombinant expression vectors are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Filamentous Fungal Host Cells

The present invention also relates to filamentous fungal strains, comprising: (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) first selectable markers, (iii)

a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

The present invention also relates to filamentous fungal strains, comprising: (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

The host cell may be any filamentous fungal cell useful in the recombinant production of polypeptides. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan,

chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

The filamentous fungal host cell may be an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* cell.

For example, the filamentous fungal host cell may be an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

In one aspect, the filamentous fungal host cell is *Aspergillus oryzae*. In another aspect, the filamentous fungal host cell is *Aspergillus niger*. In another aspect, the filamentous fungal host cell is *Fusarium venenatum*. In another aspect, the filamentous fungal host cell is *Trichoderma reesei*. In another aspect, the filamentous fungal host cell is *Trichoderma longibrachiatum*.

In another aspect, the filamentous fungal host cell is *Trichoderma reesei* RutC30. In another aspect, the filamentous fungal host cell is *Trichoderma reesei* TV10. In another aspect, the filamentous fungal host cell is a mutant of *Trichoderma reesei* RutC30. In another aspect, the filamentous fungal host cell is a mutant of *Trichoderma reesei* TV10. In another aspect, the filamentous fungal host cell is a morphological mutant of *Trichoderma*

reesei. See, for example, WO 97/26330, which is incorporated herein by reference in its entirety.

In another aspect, the filamentous fungal host cell is a *Trichoderma* strain comprising one or more (several) genes selected from the group consisting of a first subtilisin-like serine protease gene, a first aspartic protease gene, a trypsin-like serine protease gene, a second subtilisin-like serine protease gene, and a second aspartic protease gene, wherein the one or more (several) genes are modified rendering the mutant strain deficient in the production of one or more (several) enzymes selected from the group consisting of a first subtilisin-like serine protease, a first aspartic protease, a trypsin-like serine protease, a second subtilisin-like serine protease, and a second aspartic protease, respectively, compared to the parent *Trichoderma* strain when cultivated under identical conditions, as described in WO 2011/075677, which is incorporated herein by reference in its entirety.

Filamentous fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* and *Trichoderma* host cells are described in EP 238023, Yelton *et al.*, 1984, *Proc. Natl. Acad. Sci. USA* 81: 1470-1474, and Christensen *et al.*, 1988, *Bio/Technology* 6: 1419-1422. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156, and WO 1996/00787.

Methods of Production

The present invention also relates to methods for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising:

(A) cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more (*e.g.*, several) second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third

terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b); and optionally

(B) recovering the multiple recombinant polypeptides.

The present invention also relates to methods for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising:

(A) cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b); and optionally

(B) recovering the multiple recombinant polypeptides.

The filamentous fungal host cells are cultivated in a nutrient medium suitable for production of the polypeptides using methods known in the art. For example, the cells may be cultivated by shake flask cultivation, or small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors in a suitable medium and under conditions allowing the polypeptides to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptides are secreted into the nutrient medium, the polypeptides can be recovered

directly from the medium. If the polypeptides are not secreted, they can be recovered from cell lysates.

The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods include, but are not limited to, use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, enzyme assays may be used to determine the activity of the polypeptides.

The polypeptides may be recovered using methods known in the art. For example, the polypeptides may be recovered from the nutrient medium by conventional procedures including, but not limited to, collection, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. In one aspect, the whole fermentation broth is recovered.

The polypeptides may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, or extraction (see, *e.g.*, *Protein Purification*, Janson and Ryden, editors, VCH Publishers, New York, 1989) to obtain substantially pure polypeptides.

The present invention is further described by the following examples that should not be construed as limiting the scope of the invention.

Examples

Strains

Trichoderma reesei strain 981-O-8 (D4) is a mutagenized strain of *Trichoderma reesei* RutC30 (ATCC 56765; Montenecourt and Eveleigh, 1979, *Adv. Chem. Ser.* 181: 289-301).

Trichoderma reesei strain AgJg115-104-7B1 (PCT/US2010/061105, WO 2011/075677) is a ku70- derivative of *T. reesei* strain 981-O-8 (D4).

Media and Buffer Solutions

LB plates were composed of 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl, 15 g of Bacto agar, and deionized water to 1 liter.

2XYT plus ampicillin plates were composed of 16 g of tryptone, 10 g of yeast extract, 5 g of sodium chloride, 15 g of Bacto agar, and deionized water to 1 liter. One ml of a 100 mg/ml solution of ampicillin was added after the autoclaved medium was cooled to 55˚C.

SOC medium was composed of 20 g of Bacto-tryptone, 5 g of Bacto yeast extract, 0.5 g of NaCl, 2.5 ml of 1 M KCl, and deionized water to 1 liter. The pH was adjusted to 7.0

with 10 N NaOH before autoclaving. Then 20 ml of sterile 1 M glucose was added immediately before use.

COVE salt solution was composed of 26 g of KCl, 26 g of $MgSO_4 \cdot 7H_2O$, 76 g of KH_2PO_4, 50 ml of COVE trace metals solution, and deionized water to 1 liter.

COVE trace metals solution was composed of 0.04 g of $NaB_4O_7 \cdot 10H_2O$, 0.4 g of $CuSO_4 \cdot 5H_2O$, 1.2 g of $FeSO_4 \cdot 7H_2O$, 0.7 g of $MnSO_4 \cdot H_2O$, 0.8 g of $Na_2MoO_2 \cdot 2H_2O$, 10 g of $ZnSO_4 \cdot 7H_2O$, and deionized water to 1 liter.

COVE plates were composed of 342.3 g of sucrose, 20 ml of COVE salt solution, 10 ml of 1 M acetamide, 10 ml of 1.5 M CsCl, 25 g of Noble agar (Difco), and deionized water to 1 liter.

COVE2 plates were composed of 30 g of sucrose, 20 ml of COVE salt solution, 10 ml of 1 M acetamide, 25 g of Noble agar (Difco), and deionized water to 1 liter.

Trichoderma trace metals solution was composed of 216 g of $FeCl_3 \cdot 6H_2O$, 58 g of $ZnSO_4 \cdot 7H_2O$, 27 g of $MnSO_4 \cdot H_2O$, 10 g of $CuSO_4 \cdot 5H_2O$, 2.4 g of H_3BO_3, 336 g of citric acid, and deionized water to 1 liter.

CIM medium was composed of 20 g of cellulose, 10 g of corn steep solids, 1.45 g of $(NH_4)_2SO_4$, 2.08 g of KH_2PO_4, 0.28 g of $CaCl_2$, 0.42 g of $MgSO_4 \cdot 7H_2O$, 0.42 ml of *Trichoderma* trace metals solution, 1-2 drops of antifoam, and deionized water to 1 liter; pH adjusted to 6.0.

YP medium was composed of 10 g of yeast extract, 20 g of Bacto peptone, and deionized water to 1 liter.

YPG medium was composed of 4 g of yeast extract, 1 g of K_2HPO_4, 0.5 g of $MgSO_4$, 15.0 g of glucose, and deionized water to 1 liter (pH 6.0).

PEG buffer was composed of 500 g of polyethylene glycol 4000 (PEG 4000), 10 mM $CaCl_2$, 10 mM Tris-HCl pH 7.5, and deionized water to 1 liter; filter sterilized.

PDA plates were composed of 39 g of Potato Dextrose Agar (Difco) and deionized water to 1 liter.

PDA overlay medium was composed of 39 g of Potato Dextrose Agar (Difco), 2.44 g of uridine, and deionized water to 1 liter. The autoclaved medium was melted in a microwave and then cooled to 55°C before use.

STC was composed of 1 M sorbitol, 10 mM mM $CaCl_2$, and 10 mM Tris-HCl, pH 7.5; filter sterilized.

TE buffer was composed of 1 M Tris pH 8.0 and 0.5 M EDTA pH 8.0

Denaturing Solution was composed of 0.5 M NaOH and 1.5 M NaCl.

Neutralization Solution was composed of 1 M Tris pH 8.0 and 1.5 M NaCl.

20X SSC was composed of 175.3 g of NaCl, 88.2 g of sodium citrate, and deionized water to 1 liter.

TrMM-G medium was composed of 20 ml of COVE salt solution, 6 g of $(NH_4)_2SO_4$, 0.6 g of $CaCl_2$, 25 g of Nobel agar (Difco), 20 g of glucose, and deionized water to 1 liter.

NZY+ medium was composed of 5 g of NaCl, 3 g of $MgSO_4 \cdot 7H_2O$, 5 g of yeast extract, 10 g of NZ amine, 1.2 g of $MgCl_2$, 4 g of glucose, and deionized water to 1 liter.

Example 1: Cloning of an *Aspergillus fumigatus* GH61B polypeptide gene

A tblastn search (Altschul *et al.*, 1997, *Nucleic Acids Res.* 25: 3389-3402) of the *Aspergillus fumigatus* partial genome sequence (The Institute for Genomic Research, Rockville, MD, USA) was performed using as query several known GH61 polypeptides including the *Thermoascus aurantiacus* GH61A polypeptide (GeneSeqP Accession Number AEC05922). Several genes were identified as putative Family GH61 homologs based upon a high degree of similarity to the query sequences at the amino acid level. One genomic region of approximately 850 bp with greater than 70% sequence identity to the *Thermoascus aurantiacus* GH61A polypeptide sequence at the amino acid level was chosen for further study.

A. fumigatus NN051616 was grown and harvested as described in U.S. Patent No. 7,244,605. Frozen mycelia were ground, by mortar and pestle, to a fine powder and genomic DNA was isolated using a DNEASY® Plant Maxi Kit (QIAGEN Inc., Valencia, CA, USA) according to manufacturer's instructions.

Two synthetic oligonucleotide primers shown below were designed to PCR amplify the *A. fumigatus* GH61B polypeptide coding sequence from the genomic DNA. An IN-FUSION® Cloning Kit (Clontech Laboratories, Inc., Mountain View, CA, USA) was used to clone the fragment directly into the expression vector pAlLo2 (WO 2004/099228), without the need for restriction digestion and ligation.

Forward primer:

5'-ACTGGATTTACC**ATGACTTTGTCCAAGATCACTTCCA**-3' (SEQ ID NO: 31)

Reverse primer:

5'-TCACCTCTAGTTAA**TTAAGCGTTGAACAGTGCAGGACCAG**-3' (SEQ ID NO: 32)

Bold letters represent coding sequence. The remaining sequences are homologous to the insertion sites of pAlLo2.

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 204 ng of *A. fumigatus* genomic DNA, 1X *Pfx* Amplification Buffer (Invitrogen, Carlsbad, CA, USA), 1.5 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 2.5 units of PLATINUM® *Pfx* DNA Polymerase (Invitrogen Corp., Carlsbad, CA, USA), and 1 µl of 50 mM $MgSO_4$ in a final volume of 50 µl. The amplification was performed using an EPPENDORF® MASTERCYCLER® 5333 epgradient S (Eppendorf Scientific, Inc., Westbury, NY, USA) programmed for 1 cycle at 94°C for 3 minutes; and 30 cycles each at

94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minutes. The heat block was then held at 72°C for 15 minutes followed by a 4°C soak cycle. The reaction products were isolated by 1.0% agarose gel electrophoresis using 40 mM Tris base, 20 mM sodium acetate, 1 mM disodium EDTA (TAE) buffer where an approximately 850 bp product band was excised from the gel and purified using a MINELUTE® Gel Extraction Kit (QIAGEN Inc., Valencia, CA, USA) according to the manufacturer's instructions.

The 850 bp fragment was then cloned into pAlLo2 using an IN-FUSION® Cloning Kit. Plasmid pAlLo2 was digested with *Nco* I and *Pac* I. The plasmid fragment was purified by gel electrophoresis as above and a QIAQUICK® Gel Purification Kit (QIAGEN Inc., Valencia, CA, USA). The gene fragment and the digested vector were combined together in a reaction described below resulting in the expression plasmid pAG43 (Figure 1) in which transcription of the *A. fumigatus* GH61B polypeptide coding sequence was under the control of the NA2-tpi promoter. The NA2-tpi promoter is a modified promoter from the *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from the *Aspergillus nidulans* triose phosphate isomerase gene. The recombination reaction (20 µl) was composed of 1X IN-FUSION® Reaction Buffer (Clontech Laboratories, Inc., Mountain View, CA, USA), 1X BSA (Clontech Laboratories, Inc., Mountain View, CA, USA), 1 µl of IN-FUSION® Enzyme (diluted 1:10) (Clontech Laboratories, Inc., Mountain View, CA, USA), 166 ng of pAlLo2 digested with *Nco* I and *Pac* I, and 110 ng of the *A. fumigatus* GH61B polypeptide purified PCR product. The reaction was incubated at 37°C for 15 minutes followed by 15 minutes at 50°C. The reaction was diluted with 40 µl of 10 mM Tris-0.1 M EDTA buffer and 2.5 µl of the diluted reaction was used to transform *E. coli* XL10 SOLOPACK® Gold Competent Cells (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions. An *E. coli* transformant containing pAG43 (GH61B polypeptide coding sequence) was identified by restriction enzyme digestion and plasmid DNA was prepared using a BIOROBOT® 9600 (QIAGEN Inc., Valencia, CA, USA).

DNA sequencing of the 862 bp PCR fragment was performed with an Applied Biosystems Model 377 XL Automated DNA Sequencer (Applied Biosystems, Carlsbad, CA, USA) using dye-terminator chemistry (Giesecke *et al.*, 1992, *Journal of Virology Methods* 38: 47-60) and primer walking strategy. The following vector specific primers were used for sequencing:

pAllo2 5 Seq:

5'-TGTCCCTTGTCGATGCG 3' (SEQ ID NO: 33)

pAllo2 3 Seq:

5'-CACATGACTTGGCTTCC 3' (SEQ ID NO: 34)

Nucleotide sequence data were scrutinized for quality and all sequences were compared to each other with assistance of PHRED/PHRAP software (University of Washington, Seattle, WA, USA).

A gene model for the *A. fumigatus* sequence was constructed based on similarity of the encoded protein to the *Thermoascus aurantiacus* GH61A protein (GeneSeqP Accession Number AEC05922). The nucleotide sequence and deduced amino acid sequence of the *A. fumigatus* GH61B polypeptide coding sequence are shown in SEQ ID NO: 35 and SEQ ID NO: 36, respectively. The genomic fragment encodes a polypeptide of 250 amino acids, interrupted by 2 introns of 53 and 56 bp. The % G+C content of the coding sequence and the mature coding sequence are 53.9% and 57%, respectively. Using the SignalP software program (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6), a signal peptide of 21 residues was predicted. The predicted mature protein contains 221 amino acids with a predicted molecular mass of 23.39 kDa.

Example 2: Construction of pSMai214 for expression of the *Aspergillus fumigatus* GH61B polypeptide

The *Aspergillus fumigatus* GH61B polypeptide coding sequence was amplified from plasmid pAG43 (Example 1) using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*GGACTGCGCACC*ATGACTTTGTCCAAGATCACTTCCA-3' (SEQ ID NO: 37)

Reverse primer:

5'-*GCCACGGAGCTTAATTAA*TTAAGCGTTGAACAGTGCAG-3' (SEQ ID NO: 38)

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 10 ng of pAG43 DNA, 1X *Pfx* Amplification Buffer, 1.5 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 2.5 units of PLATINUM® *Pfx* DNA Polymerase, and 1 µl of 50 mM $MgSO_4$ in a final volume of 50 µl. The amplification was performed using an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 3 minutes; and 30 cycles each at 98°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minute. The heat block was then held at 72°C for 15 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 0.9 kb fragment was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit according to the manufacturer's protocol.

Plasmid pMJ09 (WO 2005/047499) was digested with *Nco* I and *Pac* I, isolated by 1.0% agarose gel electrophoresis in 1 mM disodium EDTA-50 mM Tris base-50 mM boric

acid (TBE) buffer, excised from the gel, and extracted using a QIAQUICK® Gel Extraction Kit (QIAGEN Inc., Valencia, CA, USA) according to the manufacturer's instructions.

The 0.9 kb PCR product was inserted into the gel-purified *Nco* I/*Pac* I digested pMJ09 using an IN-FUSION® PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 100 ng of the gel-purified *Nco* I/*Pac* I digested pMJ09, 37 ng of the 0.9 kb PCR product, 2 µl of 500 µg/ml BSA, and 1 µl of IN-FUSION® Enzyme in a 20 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 30 µl of TE buffer were added to the reaction. A 2.5 µl aliquot was used to transform SOLOPACK® Gold Supercompetent Cells (Agilent Technologies, Inc., Cedar Creek, TX, USA) according to the manufacturer's protocol. Transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pSMai214 (Figure 2). Plasmid pSMai214 can be digested with *Pme* I to generate an approximately 5.4 kb fragment for *T. reesei* transformation. The 5.4 kb fragment contains the expression cassette composed of the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *A. fumigatus* GH61B polypeptide coding sequence, *T. reesei* Cel7A cellobiohydrolase I gene terminator, and *Aspergillus nidulans* acetamidase (*amdS*) gene.

Example 3: Construction of a tandem construct pDM287 for expression of both *Aspergillus fumigatus* CEL3A beta-glucosidase and *Aspergillus fumigatus* GH61B polypeptide

An *A. fumigatus* GH61B polypeptide expression cassette was amplified from plasmid pSMai214 using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*CGCGGTAGTGGCGCG*GTCGACCGAATGTAGGATTGTT-3' (SEQ ID NO: 39)

Reverse primer:

5'-*TTACCAATTGGCGCG*CCACTACCGCGTTCGAGAAGA-3' (SEQ ID NO: 40)

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 25 ng of pSMai214 DNA, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer (Finnzymes Oy, Espoo, Finland), 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase (Finnzymes Oy, Espoo, Finland) in a final volume of 50 µl. The amplification was performed using an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 1 minute 30 seconds; and 1 cycle at 72°C for 10 minutes. PCR products were separated by 0.8% agarose gel electrophoresis using TAE buffer where an approximately 2.3 kb fragment was

excised from the gel and extracted using a NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Inc., Bethlehem, PA, USA) according to the manufacturer's protocol.

The approximately 2.3 kb PCR product was inserted into *Asc* I-digested pEJG107 (WO 2005/047499) using an IN-FUSION® Advantage PCR Cloning Kit (Clontech Laboratories, Inc., Mountain View, CA, USA) according to the manufacturer's protocol. Plasmid pEJG107 comprises an *Aspergillus fumigatus* CEL3A beta-glucosidase encoding sequence (SEQ ID NO: 53 [DNA sequence] and SEQ ID NO: 54 [deduced amino acid sequence]). The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 125 ng of the *Asc* I-digested pEJG107, 90 ng of the 2.33 kb PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. The transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pDM287 (Figure 3). Plasmid pDM287 can be digested with *Pme* I to generate an approximately 9.9 kb fragment for *T. reesei* transformation. The 9.9 kb fragment contains two expression cassettes composed of (1) the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *A. fumigatus* CEL3A beta-glucosidase coding sequence, and *T. reesei* Cel7A cellobiohydrolase I gene terminator; and (2) the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *A. fumigatus* GH61B polypeptide coding sequence, and *T. reesei* Cel7A cellobiohydrolase I gene terminator. The 9.9 kb fragment also contains the *Aspergillus nidulans* acetamidase (*amdS*) gene.

Example 4: *Trichoderma reesei* protoplast generation and transformation

Protoplast preparation and transformation were performed using a modified protocol by Penttila *et al.*, 1987, *Gene* 61: 155-164. Briefly, *Trichoderma reesei* strain 981-O-8 (D4) was cultivated in 25 ml of YP medium supplemented with 2% (w/v) glucose and 10 mM uridine at 27°C for 17 hours with gentle agitation at 90 rpm. Mycelia were collected by filtration using a Vacuum Driven Disposable Filtration System (Millipore, Bedford, MA, USA) and washed twice with deionized water and twice with 1.2 M sorbitol. Protoplasts were generated by suspending the washed mycelia in 20 ml of 1.2 M sorbitol containing 15 mg of GLUCANEX® 200 G (Novozymes A/S, Bagsvaerd, Denmark) per ml and 0.36 units of chitinase (Sigma Chemical Co., St. Louis, MO, USA) per ml for 15-25 minutes at 34°C with gentle shaking at 90 rpm. Protoplasts were collected by centrifuging for 7 minutes at 400 x *g* and washed twice with cold 1.2 M sorbitol. The protoplasts were counted using a

haemocytometer and resuspended to a final concentration of $1x10^8$ protoplasts/ml in STC. Excess protoplasts were stored in a Cryo 1°C Freezing Container (Nalgene, Rochester, NY, USA) at -80°C.

Approximately 100 µg of transforming plasmid (pSMai214, pDM287, or pEJG107) were digested with *Pme* I. The digestion reaction was purified by 0.8% agarose gel electrophoresis using TAE buffer. A DNA band containing the expression cassette of pSMai214, pDM287, or pEJG107, and the *Aspergillus nidulans* acetamidase (*amdS*) gene, was excised from the gel and extracted using a NUCLEOSPIN® Extract II Kit according to the manufacturer's suggested protocol.

The resulting purified DNA [1 µg of the 9.9 kb *Pme* I digested pDM287 (tandem transformation) or 1 µg of the 7.6 kb *Pme* I digested pEJG107 plus 1 µg of the 5.4 kb *Pme* I digested pSMai214 (co-transformation)] was added to 100 µl of the protoplast solution and mixed gently. PEG buffer (250 µl) was added, and the reaction was mixed and incubated at 34°C for 30 minutes. STC (3 ml) was then added, and the reaction was mixed and then spread onto COVE plates for *amdS* selection. The plates were incubated at 28°C for 6-11 days.

Example 5: Evaluation of *Trichoderma reesei* transformants expressing *Aspergillus fumigatus* CEL3A beta-glucosidase and *Aspergillus fumigatus* GH 61B polypeptide

Trichoderma reesei transformants (Example 4) were transferred from COVE transformation plates to COVE2 plates supplemented with 10 mM uridine using an inoculation loop and incubated 5-7 days at 28˚C. Spores were collected with an inoculating loop and transferred to 25 ml of CIM medium in a 125 ml plastic shake flask. The shake flask cultures were incubated for 5 days at 28°C, 200 rpm. A 1 ml aliquot of each culture was centrifuged at 13,400 x *g* in a microcentrifuge and culture supernatant was recovered. Five µl of each culture supernatant were analyzed by SDS-PAGE using a CRITERION® 8-16% Tris-HCl Gel (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer's instructions. The resulting gel was stained with BIO-SAFE™ Coomassie (Bio-Rad Laboratories, Hercules, CA, USA). SDS-PAGE profiles of 45 transformants of pDM287 (tandem construct) and 45 transformants of pEJG107 + pSMai214 (co-transformation) showed that the transformants produced major protein bands of approximately 130 kDa corresponding to the *A. fumigatus* CEL3A beta-glucosidase and approximately 24 kDa corresponding to the *A. fumigatus* GH61B polypeptide. A negative control sample, consisting of untransformed *T. reesei* strain 981-O-8 (D4) culture supernatant, showed no prominent bands at approximately 130 kDa and approximately 24 kDa.

The results shown below demonstrated that transformation with the tandem construct pDM287 yielded more positive transformants for *A. fumigatus* beta-glucosidase and *A. fumigatus* GH61B polypeptide production than co-transformation with pEJG107 and pSMai214.

Transforming DNA	Number of transformants positive for *A. fumigatus* beta-glucosidase and *A. fumigatus* GH61B polypeptide production by SDS-PAGE
pDM287 (tandem construct)	33 of 45 (73%)
pEJG107 + pSMai214 (co-transformation)	13 of 45 (29%)

Example 6: Beta-glucosidase assay of *Trichoderma reesei* transformants expressing *Aspergillus fumigatus* CEL3A beta-glucosidase and *Aspergillus fumigatus* GH61B polypeptide

The culture supernatants of Example 5 were assayed for beta-glucosidase activity using a BIOMEK® 3000, a BIOMEK® NX, and an ORCA® robotic arm (Beckman Coulter, Inc, Fullerton, CA, USA). Culture supernatants were diluted appropriately in 0.1 M succinate, 0.01% TRITON® X-100 (4-(1,1,3,3-tetramethylbutyl)phenyl-polyethylene glycol) buffer pH 5.0 (sample buffer) followed by a series of dilutions from 0-fold to 1/3-fold to 1/9-fold of the diluted sample. A total of 20 µl of each dilution was transferred to a 96-well flat bottom plate. Two hundred microliters of a *p*-nitrophenyl-beta-D-glucopyranoside substrate solution (1 mg of *p*-nitrophenyl-beta-D-glucopyranoside per ml of 0.1 M succinate pH 5.0) were added to each well and then incubated at ambient temperature for 45 minutes. Upon completion of the incubation period 50 µl of quenching buffer (1 M Tris buffer pH 9) were added to each well. An endpoint was measured at an optical density of 405 nm for the 96-well plate.

The results shown in Figure 4 confirmed the SDS-PAGE results of Example 5 that transformation with the tandem construct pDM287 yielded more positive transformants for *A. fumigatus* beta-glucosidase and *A. fumigatus* GH61B polypeptide production than co-transformation with pEJG107 and pSMai214.

Example 7: Construction of pDM286 expressing a *Penicillium* sp. GH61A polypeptide

The *Penicillium* sp. (*emersonii*) GH61A polypeptide coding sequence (SEQ ID NO: 43 [DNA sequence] and SEQ ID NO: 44 [deduced amino acid sequence]) was amplified from plasmid pGH61D23Y4 (WO 2011/041397) using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*CGGACTGCGCACC*ATGCTGTCTTCGACGACTCGCAC-3' (SEQ ID NO: 45)

Reverse primer:

5'-*TCGCCACGGAGC*TTATCGACTTCTTCTAGAACGTC-3' (SEQ ID NO: 46)

The amplification reaction was composed of 30 ng of pGH61D23Y4 DNA, 50 pmoles of each of the primers listed above, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 30 seconds; and 1 cycle at 72°C for 10 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 0.9 kb fragment was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's protocol.

Plasmid pMJ09 (WO 2005/047499) was digested with *Nco* I and *Pac* I, isolated by 1.0% agarose gel electrophoresis using TBE buffer, excised from the gel, and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's instructions.

The 0.9 kb PCR product was inserted into the gel-purified *Nco* I/*Pac* I digested pMJ09 using an IN-FUSION™ Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION™ reaction was composed of 1X IN-FUSION™ Reaction Buffer, 180 ng of the gel-purified *Nco* I/*Pac* I digested pMJ09, 108 ng of the 0.9 kb PCR product, and 1 µl of IN-FUSION™ Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. The transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pDM286 (Figure 5). Plasmid pDM286 can be digested with *Pme* I to generate an approximately 5.4 kb fragment for *T. reesei* transformation. The 5.4 kb fragment contains the expression cassette composed of the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *P. emersonii* GH61A polypeptide coding sequence, and *T. reesei* Cel7A cellobiohydrolase I gene terminator. The 5.4 kb fragment also contains the *Aspergillus nidulans* acetamidase (*amdS*) gene.

Example 8: Construction of a tandem construct pDM290 for expression of both *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* CEL3A beta-glucosidase

An *Aspergillus fumigatus* CEL3A beta-glucosidase expression cassette was amplified from plasmid pEJG107 using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*CGCGGTAGTGGCGCG*GTCGACCGAATGTAGGATTGTT-3' (SEQ ID NO: 47)

Reverse primer:

5'-*TTACCAATTGGCGCG*CCACTACCGCGTTCGAGAAGA-3' (SEQ ID NO: 48)

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 25 ng of pEJG107 DNA, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification was performed in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 2 minutes 30 seconds; and 1 cycle at 72°C for 10 minutes.

PCR products were separated by 0.8% agarose gel electrophoresis using TAE buffer where an approximately 4.5 kb fragment was excised from the gel and extracted using a NUCLEOSPIN® Extract II Kit according to the manufacturer's protocol.

The 4.5 kb PCR product was inserted into *Asc* I-digested pDM286 using an IN-FUSION™ Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION™ reaction was composed of 1X IN-FUSION™ Reaction Buffer, 125 ng of *Asc* I-digested pDM286, 100 ng of the 4.5 kb PCR product, and 1 µl of IN-FUSION™ Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. Transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pDM290 (Figure 6). Plasmid pDM290 can be digested with *Pme* I to generate an approximately 9.9 kb fragment for *T. reesei* transformation. The 9.9 kb fragment contains two expression cassettes: (1) the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *P. emersonii* GH61A polypeptide coding sequence, and *T. reesei* Cel7A cellobiohydrolase I gene terminator; and (2) the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *A. fumigatus* CEL3A beta-glucosidase coding sequence, and *T. reesei* Cel7A cellobiohydrolase I gene terminator. The 9.9 kb fragment also contains the *Aspergillus nidulans* acetamidase (*amdS*) gene.

Example 9: Construction of an empty *T. reesei cbhII* replacement construct pJfyS142

To generate a construct to replace the *Trichoderma reesei cbhII* gene (SEQ ID NO: 3 [DNA sequence] and SEQ ID NO: 4 [deduced amino acid sequence]) with the *Aspergillus fumigatus cbhII* coding sequence (SEQ ID NO: 49 [DNA sequence] and SEQ ID NO: 50 [deduced amino acid sequence]), the *T. reesei cbhII* promoter was first amplified from *T. reesei* RutC30 genomic DNA using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*acgaattgtttaaacgtcgac*CCAAGTATCCAGAGGTGTATGGAAATATCAGAT-3' (SEQ ID NO: 51)

Reverse primer:

5'-*cgcgtagatctgcggccat*GGTGCAATACACAGAGGGTGATCTT-3' (SEQ ID NO: 52)

Trichoderma reesei RutC30 was grown in 50 ml of YP medium supplemented with 2% glucose (w/v) in a 250 ml baffled shake flask at 28°C for 2 days with agitation at 200 rpm. Mycelia were harvested by filtration using MIRACLOTH® (Calbiochem, La Jolla, CA, USA), washed twice in deionized water, and frozen under liquid nitrogen. Frozen mycelia were ground, by mortar and pestle, to a fine powder, and genomic DNA was isolated using a DNEASY® Plant Maxi Kit with the lytic incubation extended to 2 hours.

The amplification reaction was composed of 20 ng of *T. reesei* RutC30 genomic DNA, 200 µM dNTP's, 0.4 µM primers, 1X HERCULASE® Reaction Buffer (Stratagene, La Jolla, CA, USA), and 1.875 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase (Stratagene, La Jolla, CA, USA) in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 25 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute 30 seconds; and 1 cycle at 72°C for 7 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where a 1.6 kb fragment was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit.

The 1.6 kb PCR product was inserted into *Nco* I/*Sal* I-digested pSMai155 (WO 05/074647) using an IN-FUSION® Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 125 ng of the *Nco* I/*Sal* I-digested pSMai155, 100 ng of the 1.6 kb PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. The resulting transformants were screened by restriction analysis with *Pci* I and positive clones sequenced to ensure the

absence of PCR errors. One clone containing the insert with no PCR errors was identified and designated pJfyS142-A. Plasmid pJfyS142-A was used for insertion of the *T. reesei cbhII* terminator.

The *cbhII* terminator was amplified from *T. reesei* RutC30 genomic DNA using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*atctacgcgtactagttaattaa*GGCTTTCGTGACCGGGCTTCAAACA-3' (SEQ ID NO: 53)

Reverse primer:

5'-*gcggccgttactagtggatcc*ACTCGGAGTTGTTATACGCTACTCG-3' (SEQ ID NO: 54)

The amplification reaction was composed of 150 ng of *T. reesei* RutC30 genomic DNA, 200 µM dNTP's, 0.4 µM primers, 1X HERCULASE® Reaction Buffer, and 1.875 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 25 cycles each at 95°C for 30 seconds, 54°C for 30 seconds, and 72°C for 50 seconds; and 1 cycle at 72°C for 7 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where a 0.3 kb fragment was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit.

The 0.3 kb PCR product was inserted into *Pac* I/*Bam* HI-digested pJfyS142-A using an IN-FUSION® Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 150 ng of the *PacI/Bam* HI-digested pJfyS142-A, 50 ng of the 0.3 kb PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30 seconds and 250 µl of SOC medium were added. The tubes were incubated at 37°C, 200 rpm for 1 hour and 250 µl were plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The transformants were screened by sequence analysis to identify positive clones and to ensure the absence of PCR errors. One clone containing the insert with no PCR errors was identified and designated pJfyS142-B. Plasmid pJfyS142-B was used for insertion of the *Herpes simplex tk* gene.

The *Herpes simplex tk* gene was liberated from pJfyS1579-8-6 (WO 2010/039840) by digesting the plasmid with *Bgl* II and *Bam* HI. The digestion was submitted to 1% agarose gel electrophoresis using TAE buffer where a 2.3 kb band was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The *tk* cassette was inserted into *Bam* HI-

digested, calf Intestine phoshatase-dephosphorylated pJfyS142-B using a QUICK LIGATION™ Kit according to the manufacturer's protocol. The ligation reaction was composed of 50 ng of the *Bam* HI-digested, calf Intestine phoshatase-dephosphorylated pJfyS142-B, 50 ng of the 2.3 kb *tk* gene insert, 1X QUICK LIGATION™ Buffer, and 5 units of QUICK LIGASE™ in a 20 µl ligation volume. The reaction was incubated at room temperature for 5 minutes and 2 µl of the reaction was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30 seconds and 250 µl of SOC medium were added. The tubes were incubated at 37°C, 200 rpm for 1 hour and 250 µl were plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The resulting transformants were screened by restriction digestion analysis with *Xma* I and *Bam* HI to determine the presence and orientation of the insert and a clone containing the insert was identified and designated pJfyS142-C. Plasmid pJfyS142-C was used for insertion of the *T. reesei* 3' *cbhII* gene flanking sequence.

The 3' *cbhII* gene flanking sequence was amplified from *T. reesei* RutC30 genomic DNA using the forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*atccatcacactggcggccgc*GCTTCAAACAATGATGTGCGATGGT-3' (SEQ ID NO: 55)

Reverse primer:

5'-*gatgcatgctcgagcggccgc*CTACCTTGGCAGCCCTACGAGAGAG-3' (SEQ ID NO: 56)

The amplification reaction was composed of 150 ng of *T. reesei* RutC30 genomic DNA, 200 µM dNTP's, 0.4 µM primers, 1X HERCULASE® Reaction Buffer, and 1.875 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 30 cycles each at 95°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minute 50 seconds; and 1 cycle at 72°C for 7 minutes. The PCR reaction was subjected to 1% agarose gel electrophoresis using TAE buffer where a 1.5 kb band was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The 3' *cbhII* gene flanking sequence was inserted into *Not* I-linearized pJfyS142-C using an IN-FUSION® Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 150 ng of *Not* I-linearized pJfyS142-C, 80 ng of the 1.5 kb PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30

seconds and 250 µl of SOC medium were added. The tubes were incubated at 37°C, 200 rpm for 1 hour and 250 µl were plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The resulting transformants were screened by restriction digestion analysis with *Bgl* II and positive clones were sequenced to ensure the absence of PCR errors. One clone containing the insert with no PCR errors was identified and designated pJfyS142 (Figure 7). Plasmid pJfyS142 was used for insertion of the *A. fumigatus cbhII* coding sequence.

Example 10: Construction of a *Trichoderma reesei cbhII-Aspergillus fumigatus cbhII* replacement construct pJfyS144

The *Aspergillus fumigatus cbhII* coding sequence was amplified from pAlLo33 (WO 2011/057140) using the forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction.

Forward primer:

5'-*ctctgtgtattgcacc*ATGAAGCACCTTGCATCTTCCATCG-3' (SEQ ID NO: 57)

Reverse primer:

5'-*ccggtcacgaaagcc*TTAATTAAAAGGACGGGTTAGCGTT-3' (SEQ ID NO: 58)

The amplification reaction was composed of 20 ng of pAlLo33, 200 µm dNTP's, 0.4 µM primers, 1 mM HERCULASE® Reaction Buffer, and 1.875 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2 minutes; and 1 cycle at 72°C for 7 minutes.

The PCR reaction was subjected to 1% agarose gel electrophoresis using TAE buffer where a 1.7 kb band was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The 1.7 kb PCR product was inserted into *Nco* I/*Pac* I-digested pJfyS142 (Example 9) using an IN-FUSION® Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 120 ng of the *Nco* I/*Pac* I-digested pJfyS142, 70 ng of the 1.7 kb PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30 seconds and 250 µl of SOC medium were added. The tubes were incubated at 37°C, 200 rpm for 1 hour and 250 µl were plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The resulting transformants were sequenced to

ensure the absence of PCR errors and determine the presence of the insert. One clone with error-free sequence was identified and designated pJfyS144 (Figure 8).

Example 11: Construction of a *Trichoderma reesei cbhI-Aspergillus fumigatus cbhI* replacement construct pJfyS139

The *Aspergillus fumigatus* cellobiohydrolase I (*cbhI*) coding sequence (SEQ ID NO: 59 [DNA sequence] and SEQ ID NO: 60 [deduced amino acid sequence]) was amplified from pEJG93 (WO 2011/057140) using the gene-specific forward and reverse primers shown below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction and the underlined portion is an introduced *Pac* I site.

Forward primer:

5'-*cgcggactgcgcacc*ATGCTGGCCTCCACCTTCTCCTACC-3' (SEQ ID NO: 61)

Reverse primer:

5'-ctttcgccacggag*cttaattaa*CTACAGGCACTGAGAGTAATAATCA-3' (SEQ ID NO: 62)

The amplification reaction was composed of 20 ng of pEJG93, 200 µM dNTP's, 0.4 µM primers, 1X HERCULASE® Reaction Buffer, and 1.875 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute; and 1 cycle at 72°C for 7 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where a 1.6 kb fragment was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit according to the manufacturer's protocol.

The 1.6 kb PCR product was inserted into *Nco* I/*Pac* I-digested pSMai155 (WO 05/074647) using an IN-FUSION® Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 125 ng of *Nco* I/*Pac* I-digested pSMai155, 100 ng of the 1.6 kb PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C. After the incubation period 40 µl of TE buffer were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. The resulting transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pJfyS139-A. Plasmid pJfyS139-A was used for insertion of the *Herpes simplex* virus thymidine kinase (*tk*) gene.

The *Herpes simplex* virus *tk* coding sequence (SEQ ID NO: 63 [DNA sequence] and SEQ ID NO: 64 [deduced amino acid sequence]) was liberated from pJfyS1579-8-6 (WO

2010/039840) by digesting the plasmid with *Bgl* II and *Bam* HI. The digestion was subjected to 1% agarose gel electrophoresis using TAE buffer where a 2.3 kb band was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The *tk* gene cassette was inserted into *Bam* HI-digested, calf intestine phosphatase-treated pJfyS139-A using a QUICK LIGATION™ Kit according to the manufacturer's protocol. The ligation reaction was composed of 50 ng of the *Bam* HI-digested, calf intestine phosphatase-treated pJfyS139-A, 50 ng of the 2.3 kb *tk* gene insert, 1X QUICK LIGATION™ Buffer, and 5 units of QUICK LIGASE™ in a final volume of 20 µl. The reaction was incubated at room temperature for 5 minutes and 2 µl of the reaction were used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30 seconds and 250 µl of SOC medium were added. The tubes were incubated at 37°C, 200 rpm for 1 hour and 250 µl were plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The resulting transformants were screened by restriction digestion analysis with *Xma* I to determine the presence and orientation of the insert and a clone containing the insert was identified and designated pJfyS139-B. Plasmid pJfyS139-B was used for insertion of a *T. reesei* 3' *cbhI* gene flanking sequence.

The 3' *cbhI* gene flanking sequence was amplified from *T. reesei* RutC30 genomic DNA (Example 9) using the forward and reverse primers below. The underlined portion represents an introduced *Not* I site for cloning.

Forward Primer:

5'-ttagactgcggccgcGTGGCGAAAGCCTGACGCACCGGTAGAT-3' (SEQ ID NO: 65)

Reverse Primer:

5'-agtagttagcggccgcACGGCACGGTTAAGCAGGGTCTTGC-3' (SEQ ID NO: 66)

The amplification reaction was composed of 150 ng of *T. reesei* RutC30 genomic DNA, 200 µM dNTP's, 0.4 µM primers, 1X HERCULASE® Reaction Buffer, and 1.875 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 30 cycles each at 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute 30 seconds; and 1 cycle at 72°C for 7 minutes.

The PCR reaction was subjected to a MINELUTE® Nucleotide Removal Kit (QIAGEN Inc., Valencia, CA, USA) according to the manufacturer's protocol. The resulting PCR mixture was digested with *Not* I and the digested PCR products were separated by 1% agarose gel electrophoresis using TAE buffer. A 1.3 kb fragment containing the 3' *cbhI* gene flanking sequence was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The 1.3 kb fragment was inserted into *Not* I-linearized, calf intestine phosphatase-treated pJfyS139-B using a QUICK LIGATION™ Kit. The QUICK LIGATION™

reaction was composed of 100 ng of the *Not* I-linearized, calf intestine phosphatase-treated pJfyS139-B, 20 ng of the 1.3 kb fragment, 1X QUICK LIGATION™ Buffer, and 5 units of QUICK LIGASE™ in a final volume of 20 µl. The reaction was incubated at room temperature for 5 minutes and 2 µl of the reaction was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30 seconds and 250 µl of SOC medium were added. The tubes were incubated at 37°C, 200 rpm for 1 hour and 250 µl were plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The resulting transformants were screened by restriction digestion analysis with *Xma* I to determine the presence and orientation of the insert and positive clones were sequenced. A clone containing the 3' *cbhI* gene flanking sequence with no PCR errors was designated pJfyS139 (Figure 9).

Example 12: Construction of an *Aspergillus fumigatus cbhI-Aspergillus fumigatus cbhII* tandem expression vector for replacement of the *Trichoderma reesei cbhI gene*

A tandem replacement vector, pQM21, was constructed for replacing the native *T. reesei cbhI* gene in *Trichoderma reesei* with a tandem expression cassette for expressing two recombinant proteins. Plasmid pQM21 contains the *T. reesei cbhI* 5' flanking sequence, *T. reesei* Cel7A cellobiohydrolase I gene promoter, *Aspergillus fumigatus* Cel7A cellobiohydrolase I coding sequence, *T. reesei* Cel7A cellobiohydrolase I gene terminator, *T. reesei* Cel6A cellobiohydrolase II gene promoter, *Aspergillus fumigatus* Cel6A cellobiohydrolase II coding sequence, *T. reesei* Cel6A cellobiohydrolase II gene terminator, *T. reesei* Cel7A cellobiohydrolase I gene terminator repeat, *Herpes simplex* virus thymidine kinase (*tk*) gene, *E. coli* hygromycin phosphotransferase (*hpt/hygR)* selection marker, *T. reesei cbhI* 3' flanking sequence, and ampicillin resistance marker gene.

The *Aspergillus fumigatus* cellobiohydrolase II expression cassette was amplified from pJfyS144 (Example 10) using the gene-specific forward and reverse primers shown below. The region in italics represents sequence homology to the site of insertion for an IN-FUSION® reaction and the underlined portion is an introduced *Bam* HI site and *Nhe* I site, respectively.

Forward primer:

5'-*tcaagcttggtaccgagctc*<u>ggatCC</u>AAGTATCCAGAGGTGTATGGAAAT-3' (SEQ ID NO: 67)

Reverse primer:

5'-*ctggcggccgttactagtg*<u>ctagc</u>ACTCGGAGTTGTTATACGCTAC-3' (SEQ ID NO: 68)

The amplification reaction was composed of 164 ng of pJfyS144, 1 µM primers, 1X ACCUPRIME™ *Pfx* Reaction Buffer (Invitrogen, Carlsbad, CA, USA), and 2.5 units of ACCUPRIME™ *Pfx* DNA Polymerase (Invitrogen, Carlsbad, CA, USA) in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER®

5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; and 35 cycles each at 95°C for 15 seconds, 58°C for 30 seconds, and 68°C for 5 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 3.5 kb fragment was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's protocol.

The 3.5 kb PCR product was inserted into *Bam* HI digested pJfyS139 (Example 11) using an IN-FUSION™ Advantage PCR Cloning Kit according to the manufacturer's suggested protocol. The IN-FUSION™ reaction was composed of 1X IN-FUSION™ Reaction Buffer, 103 ng of *Bam* HI digested pJfyS139, 62 ng of the 3.5 kb PCR product, and 1 µl of IN-FUSION™ Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C. After the incubation period 15 µl of TE were added to the reaction. A 2 µl aliquot was used to transform SOLOPACK® Gold Supercompetent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. The transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pQM18 (Figure 10). Plasmid pQM18 was used to insert a homologous repeat fragment from the *T. reesei* 3' *cbhI* gene flanking region after the tandem expression cassette and to modify the *T. reesei* 3' *cbhII* gene flanking region.

The homologous repeat fragment from the *T. reesei* 3' *cbhI* gene flanking region was amplified from pJfyS139 using the forward and reverse primers below. The region in italics represents sequence homology to the site of insertion for an IN-FUSION® reaction and the underlined portion represents introduced *Nhe* I site and *Xba* I sites for cloning.

Forward Primer:

5'-*gagtagcgtataacaactccgagt*gctagcTTTAAGATAACGGAATAGAAGAAAG-3' (SEQ ID NO: 69)

Reverse Primer:

5'-*ctggcggccgttactagt*ctagaCGCGCCACTACCGCGTTCG-3' (SEQ ID NO: 70)

The *T. reesei* 3' *cbhI* gene flanking sequence was amplified from pJfyS139 using the forward and reverse primers below. The region in italics represents vector homology to the site of insertion for an IN-FUSION® reaction and the underlined portion represents an introduced *Not* I site for cloning.

Forward Primer:

5'-*tctgcagatatccatcacactg*gcggccgcTTTAAGATAACGGAATAGAAGAAAG-3' (SEQ ID NO: 71)

Reverse Primer:

5'-*aaactctaggatgcatgctcga*gcggccgcACGGCACGGTTAAGCAGGGT-3' (SEQ ID NO: 72)

The amplification reaction was composed of 350 ng of pJfyS139, 1 µM primers, 1X ACCUPRIME™ *Pfx* Reaction Buffer, and 2.5 units of ACCUPRIME™ *Pfx* DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; and 35 cycles each at 95°C for 15 seconds, 58°C for 30 seconds, and 68°C for 5 minutes. The PCR products were purified by 1% agarose gel electrophoresis using TAE buffer where an approximately 1.1 kb fragment and an approximately 260 bp fragment were excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's protocol.

A DNA fragment containing the *tk* gene and *hpt* (hygromycin phosphotransferase) selection marker was liberated from pQM18 by digesting the plasmid with *Nhe* I and *Not* I. The digestion was analyzed by 1% agarose gel electrophoresis using TAE buffer where an approximately 4.4 kb band was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's protocol.

An approximately 9 kb DNA fragment from *Nhe* I and *Not* I digested pQM18 was separated by 1% agarose gel electrophoresis using TAE buffer. The 9 kb fragment was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's protocol.

An IN-FUSION™ reaction was composed of 1X IN-FUSION™ Reaction Buffer, 158 ng of the 9 kb *Nhe* I and *Not* I digested pQM18, 13 ng of the 260 bp homologous repeat fragment from *T. reesei cbhI* 3' flanking region, 39 ng of the 1.1 kb 3' *cbhI* flank, 56 ng of the 4.4 kb *tk-hpt* fragment, and 1 µl IN-FUSION™ Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2.5 µl aliquot was used to transform SOLOPACK® Gold Supercompetent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates and incubated overnight at 37°C. Transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pQM21 (Figure 11). Plasmid pQM21 was used as the vector to replace the *cbhI* gene.

Example 13: Replacement of native *Trichoderma reesei cbhI* gene with the *Aspergillus fumigatus* cellobiohydrolase I and *Aspergillus fumigatus* cellobiohydrolase II tandem expression cassette

Protoplast preparation and transformation of *Trichoderma reesei* strain AgJg115-104-7B1 were performed as described in Example 4.

In order to replace the native *cbhI* gene with the *Aspergillus fumigatus cbhI-cbhII* tandem expression cassette, approximately 137 µg of pQM21 (Example 12) was digested

with *Pme* I. The digestion reaction was purified by 1% agarose gel electrophoresis using TAE buffer where an approximately 12 kb DNA band containing the *Aspergillus fumigatus* CBHI-CBHII tandem expression cassette for targeting to the *T. reesei cbhI* locus was excised from the gel and extracted using a NUCLEOSPIN® Extract II Purification Kit. Approximately 1-3 µg of the resulting purified 12 kb DNA was added to 100 µl of the *Trichoderma reesei* ku70- strain AgJg115-104-7B1 protoplast solution and mixed gently. PEG buffer (250 µl) was added, mixed, and incubated at 34°C for 30 minutes. STC (3 ml) was then added, mixed, and spread onto each PDA plate supplemented with 1 M sucrose. After incubation at 28°C for 16 hours, 20 ml of an overlay PDA medium supplemented with 35 µg of hygromycin B per ml were added to each plate. The plates were incubated at 28°C for 4-7 days.

Seven transformants were obtained and each one was picked and transferred to a PDA plate and incubated for 7 days at 28°C. A fungal spore PCR method using the protocol described below was used to screen for transformants bearing replacement using the forward primer shown below annealing to a region upstream of the *cbhI* 5' flanking region of integration, and the reverse primer shown below annealing to a region in the *tk* region.

Forward Primer:

5'-CAAGCAAAGCGTTCCGTCGCAGTAGCAGGC-3' (SEQ ID NO: 73)

Reverse Primer:

5'-CAGTGGCGCTTATTACTCAG-3' (SEQ ID NO: 74)

An approximately 7 kb PCR product would be generated only upon the occurrence of a precise gene replacement at the *cbhII* locus. If the cassette had integrated elsewhere in the genome, no amplification would result.

A small amount of spores from each transformant was suspended in 25 µl of TE buffer and heated on high in a microwave oven for 1 minute. Each microwaved spore suspension was used as a template in the PCR reaction. The reaction was composed of 2 µl of the microwaved spore suspension, 200 µM dNTP's, 1 µM primers, 1X LONGAMP® *Taq* Reaction Buffer (New England Biolabs, Inc, Ipswich, MA, USA), and 2 units of LONGAMP® *Taq* DNA Polymerase (New England Biolabs, Inc, Ipswich, MA, USA) in a final volume of 20 µl. The reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 4 minutes; 35 cycles each at 95°C for 15 seconds, 50°C for 30 seconds, and 68°C for 7 minutes; and 1 cycle at 68°C for 15 minutes. The PCR reactions were analyzed by 1% agarose gel electrophoresis using TAE buffer. The spore PCR indicated that five of the seven transformants contained the replacement cassette at the *cbhI* locus.

Genomic DNA was isolated from four of the positive transformants according to the procedure described in Example 9 and submitted to Southern analysis to confirm the replacement cassette was in a single copy.

For Southern analysis, 2 µg of genomic DNA was digested with 10 units of *Bam* HI in a 20 µl reaction volume and subjected to 0.7% agarose electrophoresis using TAE buffer. The DNA in the gel was depurinated in 0.25 N HCl for 15 minutes, denatured in Denaturing Solution twice for 15 minutes each time, neutralized in Neutralization Solution for 10 to 30 minutes, and transferred to a NYTRAN® Supercharge membrane (Whatman, Inc., Florham Park, NJ, USA) using a TURBOBLOTTER™ System (Whatman, Inc., Florham Park, NJ, USA) according to the manufacturer's protocol. The DNA was UV crosslinked to the membrane using a STRATALINKER™ UV Crosslinker (Stratagene, La Jolla, CA, USA) and prehybridized for 1 hour at 42°C in 20 ml of DIG Easy Hyb (Roche Diagnostics Corporation, Indianapolis, IN, USA).

A probe hybridizing to the 3' flanking region of the *cbhI* gene was generated using a PCR Dig Probe Synthesis Kit (Roche Diagnostics Corporation, Indianapolis, IN, USA) according to the manufacturer's instructions with the forward and reverse primers shown below. The PCR reaction was composed of 1X EXPAND® High Fidelity PCR Buffer with $MgCl_2$ (Roche Diagnostics Corporation, Indianapolis, IN, USA), 1X PCR DIG Probe Synthesis Mix (Roche Diagnostics Corporation, Indianapolis, IN, USA), 1 µM of each primer, 100 pg of the 1.1 kb 3' *cbhI* flanking region, and 2.625 units of EXPAND® High Fidelity Enzyme Mix (Roche Diagnostics Corporation, Indianapolis, IN, USA). The PCR was performed in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 10 cycles each at 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 40 seconds; 20 cycles each at 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 40 seconds with an additional 20 seconds for each successive cycle; and 1 cycle at 72°C for 7 minutes.

Forward primer:

5'-GAGAACACAGTGAGACCATAGC-3' (SEQ ID NO: 75)

Reverse primer:

5'-TCTCAACCCAATCAGCAACATG-3' (SEQ ID NO: 76)

The probe was purified by 1% agarose gel electrophoresis using TAE buffer where a 720 bp band corresponding to the probe was excised from the gel and extracted using a NUCLEOSPIN® Extract II Purification Kit. The probe was boiled for 5 minutes, chilled on ice for 2 minutes, and added to 10 ml of DIG Easy Hyb to produce the hybridization solution. Hybridization was performed at 42°C for 15-17 hours. The membrane was then washed under low stringency conditions in 2X SSC plus 0.1% SDS for 5 minutes at room temperature followed by two high stringency washes in 0.5X SSC plus 0.1% SDS for 15

minutes each at 65°C. The probe-target hybrids were detected by chemiluminescent assay (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's instructions. Southern analysis indicated that one transformant designated *T. reesei* QMJi029-A5 contained the replacement cassette at the *cbhI* locus and was chosen for curing of the *tk* and *hpt* markers.

Spores of *T. reesei* QMJi029-A5 from a seven-day-old PDA plate grown at 28°C were collected in 5 ml of 0.01% TWEEN® 20 using a sterile spreader. The concentration of spores was determined using a hemocytometer and 10^4 and/or 10^5 spores were spread onto 150 mm plates containing TrMM-G medium supplemented with 1 µM 5-fluoro-2'-deoxyuridine (FdU).

Ten FdU-resistant spore isolates were picked and genomic DNA was extracted from three of the spore isolates as described above. The isolates were submitted to Southern analysis as described above and the results indicated that all of the spore isolates had excised the *hpt/tk* region between the homologous repeats of the replacement cassette. One strain designated *T. reesei* QMJi030-A5.6 was selected for replacing the *cbhII* gene.

Spores of *T. reesei* QMJi029-A5 and *T. reesei* QMJi030-A5.6 from seven-day-old PDA plates grown at 28°C were collected with a 10 µl inoculating loop and transferred to 25 ml of CIM medium in a 125 ml plastic shake flask. The shake flask cultures were incubated for 5 days at 28°C, 200 rpm. A 1 ml aliquot of each culture was centrifuged at 13,400 x *g* in a microcentrifuge and culture supernatant was recovered. Five µl of each culture supernatant were analyzed by SDS-PAGE using a CRITERION® 8-16% Tris-HCl Gel according to the manufacturer's instructions. The resulting gel was stained with BIO-SAFE™ Coomassie. SDS-PAGE profiles of the cultures showed that the transformants produced two major protein bands between 50 and 70 kDa, corresponding to *Aspergillus fumigatus* cellobiohydrolase I and *Aspergillus fumigatus* cellobiohydrolase II, respectively. Expression of *Aspergillus fumigatus* cellobiohydrolase I and *Aspergillus fumigatus* cellobiohydrolase II were confirmed by mass spectroscopic analysis (Example 14).

Example 14: In-gel digestion of polypeptides for peptide sequencing

A MULTIPROBE® II Liquid Handling Robot (PerkinElmer Life and Analytical Sciences, Boston, MA, USA) was used to perform in-gel digestions. A section of the SDS-PAGE gel described in Example 13 was excised between the 50 and 70 kDa MW markers containing the proteins of interest. The gel piece was reduced with 50 µl of a 10 mM dithiothreitol (DTT) in 100 mM ammonium bicarbonate pH 8.0 for 30 minutes. Following reduction, the gel piece was alkylated with 50 µl of 55 mM iodoacetamide in 100 mM ammonium bicarbonate pH 8.0 for 20 minutes. The dried gel piece was allowed to swell in 25 µl of a trypsin digestion solution containing 6 ng of sequencing grade trypsin (Promega,

Madison, WI, USA) per µl of 50 mM ammonium bicarbonate pH 8 for 30 minutes at room temperature, followed by an 8 hour digestion at 40°C. Each of the reaction steps described above was followed by numerous washes and pre-washes with the appropriate solutions following the manufacturer's standard protocol. Fifty µl of acetonitrile was used to de-hydrate the gel piece between reactions and the gel piece was air dried between steps. Peptides were extracted twice with 1% formic acid/2% acetonitrile in HPLC grade water for 30 minutes. Peptide extraction solutions were transferred to a 96 well skirted PCR type plate (ABGene, Rochester, NY, USA) that had been cooled to 10–15°C and covered with a 96-well plate lid (PerkinElmer Life and Analytical Sciences, Boston, MA, USA) to prevent evaporation. Plates were further stored at 4°C until mass spectrometry analysis could be performed.

Protein Identification. For *de novo* peptide sequencing by tandem mass spectrometry, a SYNAPT™ MS (Waters Corp., Milford, MA, USA), a hybrid orthogonal quadrupole time-of-flight mass spectrometer, was used for LC/MS/MS analysis. The SYNAPT™ MS is fully microprocessor controlled using MASSLYNX® software version 4.1 (Waters Corp., Milford, MA, USA). The SYNAPT™ MS was fitted with a NANOACQUITY UPLC® (Waters Corp, Milford, MA, USA) for concentrating and desalting samples. Samples were loaded onto a trapping column (180 µm ID X 20 mm, 5 µm SYMMETRY® C18) (Waters Corp, Milford, MA, USA) fitted in the injection loop and washed with 0.1% formic acid in water at 15 µl per minute for 1 minute using the binary solvent manager pump. Peptides were separated on a 100 µm ID x 100 mm, C18, 1.7 µm, BEH130™ C18 nanoflow fused capillary column (Waters Corp, Milford, MA, USA) at a flow rate of 400 nl/minute. A step elution gradient of 1% to 85% acetonitrile in 0.1% formic acid was applied over a 30 minute interval. The column eluent was monitored at 214 nm and introduced into the SYNAPT™ MS through an electrospray ion source fitted with the nanospray interface.

Data was acquired in survey scan mode from a mass range of m/z 250 to 1900 with switching criteria for MS to MS/MS to include an ion intensity of greater than 10.0 counts per second and charge states of +2, +3, and +4. Analysis spectra of up to 6 co-eluting species with a scan time of 2.0 seconds could be obtained. A cone voltage of 45 volts was typically used and the collision energy was programmed to vary according to the mass and charge state of the eluting peptide and in the range of 10-60 volts. The acquired spectra were combined, smoothed, and centered in an automated fashion and a peak list generated. The peak list was searched against selected databases using PROTEINLYNX GLOBAL SERVER® 2.4 software (Waters Corp, Milford, MA, USA) and MASCOT® v. 2.2 (Matrix Sciences Ltd., London, UK) Results from the PROTEINLYNX GLOBAL SERVER® and MASCOT® searches were evaluated and peptide identifications were based on peptide mass fingerprint matches to the sequence of the expected protein.

Peptide mass fingerprinting confirmed that the SDS-PAGE samples of *T. reesei* QMJi029-A5 and *T. reesei* QMJI030A5.6 contained *A. fumigatus* Cel7A cellobiohydrolase I, *A. fumigatus* Cel6A cellobiohydrolase II, *T. reesei* cellobiohydrolase II, and other minor *T. reesei* host background proteins.

Example 15: Generation of a tandem gene expression plasmid pRRAB01 encoding *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* Cel3A beta-glucosidase

The *Trichoderma reesei* cellobiohydrolase II (*cbhII*) gene promoter was amplified from plasmid pAG121 (Figure 12; SEQ ID NO: 77; from nucleotide position 6 to nucleotide position 620, see restriction map and sequence for pAG121) using the gene-specific forward and reverse primers shown below. The region in italics in the forward primer represents sequence homology to the pDM286 vector backbone and the region in italics in the reverse primer represents sequence homology to the next insert, *i.e.*, the *Aspergillus fumigatus* beta-glucosidase coding sequence for an IN-FUSION® reaction.

Forward primer:

*5'-cgaacgcggtagtgg*GAATTCTAGGCTAGGTATGC-3' (SEQ ID NO: 78)

Reverse Primer:

5'-*ccaaccgaatctcat*GGTGCAATACACAGAGGGTG-3' (SEQ ID NO: 79)

The amplification reaction was composed of 1 ng of pAG121 DNA, 100 pmoles of each of the primers listed above, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 second and 72°C for 20 seconds; and 1 cycle at 72°C for 10 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 640 bp fragment was excised from the gel and extracted using a QIAGEN® Gel Extraction Kit according to the manufacturer's protocol.

The *Aspergillus fumigatus* beta-glucosidase coding sequence was amplified from plasmid pDM290 using the gene-specific forward and reverse primers shown below. The region in italics in the forward primer represents sequence homology to the previous insert, *i.e.*, the *Trichoderma reesei* cellobiohydrolase II coding sequence, and the region in italics in the reverse primer represents sequence homology to the next insert, *i.e.*, the *Trichoderma reesei* cellobiohydrolase II gene terminator, for an IN-FUSION® reaction.

Forward primer:

5'-*tctgtgtattgcacc*ATGAGATTCGGTTGGCTCGA-3' (SEQ ID NO: 80)

Reverse Primer:

5'-*ccggtcacgaaagcc*CTAGTAGACACGGGGCAGAG-3' (SEQ ID NO: 81)

The amplification reaction was composed of 1 ng of pDM290 DNA, 100 pmoles of each of the primers listed above, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds, 65°C for 30 seconds, and 72°C for 1:35 seconds; and 1 cycle at 72°C for 10 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 3.1 kb fragment was excised from the gel and extracted using a QIAGEN® Gel Extraction Kit according to the manufacturer's protocol.

The *Trichoderma reesei* cellobiohydrolase II gene terminator was amplified from plasmid pJfyS144 using the gene-specific forward and reverse primers shown below. The region in italics in the forward primer represents sequence homology to the *Aspergillus fumigatus* beta-glucosidase coding sequence, and the region in italics in the reverse primer represents sequence homology to the pDM286 backbone for an IN-FUSION® reaction.

Forward primer:

5'-*ccccgtgtctactag*GGCTTTCGTGACCGGGCTTC-3' (SEQ ID NO: 82)

Reverse Primer:

5'-*gtcattaccaattgg*ACTCGGAGTTGTTATACGCT-3' (SEQ ID NO: 83)

The amplification reaction was composed of 1 ng of pJfyS144 DNA, 100 pmoles of each of the primers listed above, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds and 72°C for 10 seconds; and 1 cycle at 72°C for 10 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where a 330 bp fragment was excised from the gel and extracted using a QIAGEN® Gel Extraction Kit according to the manufacturer's protocol.

The *Trichoderma reesei* cellobiohydrolase II gene promoter, *Aspergillus fumigatus* beta-glucosidase coding sequence, and *Trichoderma reesei* cellobiohydrolase II gene terminator were combined in a splicing by overlap extension (SOE) PCR reaction using the gene-specific forward and reverse primers shown below. The regions in italics represent sequence homology to the site of insertion in pDM286 for an IN-FUSION® reaction.

Forward primer:

5'-*cgaacgcggtagtgg*GAATTCTAGGCTAGGTATGC-3' (SEQ ID NO: 84)

Reverse Primer:

5'-*gtcattaccaattgg*ACTCGGAGTTGTTATACGCT-3' (SEQ ID NO: 85)

The SOE PCR reaction was composed of 48 ng of the 640 bp *T. reesei cbhII* promoter amplified from pAG121, 228 ng of the 3.1 kb *A. fumigatus* beta-glucosidase gene fragment amplified from pDM290, 24 ng of the 330 bp *T. reesei cbhII* terminator amplified from pJfyS144, 100 pmoles of each of the primers listed above, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds and 72°C for 2 minutes and 5 seconds; and 1 cycle at 72°C for 10 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 4.1 kb fragment was excised from the gel and extracted using a QIAGEN® Gel Extraction Kit according to the manufacturer's protocol.

Plasmid pDM286 was digested with *Asc* I and the *Asc* I-digested pDM286 was separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 8 kb fragment was excised from the gel and extracted using a QIAGEN® Gel Extraction Kit according to the manufacturer's protocol.

The 4.1 kb PCR product was inserted into *Asc* I-digested pDM286 using an IN-FUSION™ Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION™ reaction was composed of 1X IN-FUSION™ Reaction Buffer, 200 ng of approximately 8 kb gel purified *Asc* I-digested pDM286, 203.1 ng of the 4.1 kb PCR product, and 1 µl of IN-FUSION™ Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. The transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pRRAB01 (Figure 13).

Example 16: Generation of a *Trichoderma reesei* expression vector encoding *Aspergillus fumigatus* beta-glucosidase (Cel3A) mutant gene

A variant of the *Aspergillus fumigatus* Family 3A beta-glucosidase containing the substitutions G142S, Q183R, H266Q, and D703G was constructed by performing site-directed mutagenesis on pEJG97 (WO 2005/074647) using a QUIKCHANGE® Multi Site-

Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). A summary of the oligos used for the site-directed mutagenesis are shown in Table 1.

The resulting variant plasmid pDFng128-6 was prepared using a BIOROBOT® 9600. The variant plasmid construct was sequenced using an Applied Biosystems 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) to verify the changes.

TABLE 1

Amino acid changes	Primer name	Sequences	Cloning Plasmid Name
F100D	AfBGmutF100DF	ccctttgggtatccgtGACtgtgagctatacccgcg (SEQ ID NO: 86)	pDFng128-6
S283G	AfBGmutS283GF	cgtcatgagtgactggGGCgctcaccacagcggtg (SEQ ID NO: 87)	
N456E	AfBGmutN456EF	gggtagtggtactgccGAGttcccttaccttgtcac (SEQ ID NO: 88)	
F512Y	AfBGmutF512YF	gccgactctggagagggtTACatcagtgtcgacggcaac (SEQ ID NO: 89)	

Two synthetic oligonucleotide primers shown below were designed to PCR amplify the *Aspergillus fumigatus* beta-glucosidase mutant coding sequence from plasmid pDFng128-6. An IN-FUSION™ Cloning Kit was used to clone the fragment directly into the expression vector pMJ09. Bold letters represent coding sequence. The remaining sequence is homologous to insertion sites of pMJ09.

Forward primer:

5'-CGGACTGCGCACC**ATGAGATTCGGTTGGCTCGA**-3' (SEQ ID NO: 90)

Reverse primer:

5'-TCGCCACGGAGCTTA**CTAGTAGACACGGGGCAGAG**-3' (SEQ ID NO: 91)

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 50 ng of pDFng128-6, 1X EXPAND® High Fidelity PCR Buffer with $MgCl_2$, 0.25 mM each of dATP, dTTP, dGTP, and dCTP, and 2.6 units of EXPAND® High Fidelity Enzyme Mix in a final volume of 50 µl. The amplification was performed in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 94°C for 2 minutes; 30 cycles each at 94°C for 15 seconds, 65°C for 30 seconds, and 68°C for 1 minute; and a final elongation at 68°C for 7 minutes. The heat block then went to a 4°C soak cycle. The reaction products were isolated by 0.7% agarose gel electrophoresis using TBE buffer where an

approximately 3.1 kb product band was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's instructions.

Plasmid pMJ09 was digested with *Nco* I and *Pac* I, isolated by 1.0% agarose gel electrophoresis using TBE buffer, excised from the gel, and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's instructions.

The 3.1 kb gene fragment and the digested vector were ligated together using an IN-FUSION™ Cloning Kit resulting in pDFng113-3 (Figure 14) in which transcription of the beta-glucosidase mutant coding sequence was under the control of a promoter from the *Trichoderma reesei cbhI* gene. The ligation reaction (20 µl) was composed of 1X IN-FUSION™ Reaction Buffer, 1X BSA, 1 µl of IN-FUSION™ Enzyme (diluted 1:10), 200 ng of the gel-purified *Nco* I/*Pac* I digested pMJ09, and 172.2 ng of the purified 3.1 kb PCR product. The reaction was incubated at 37°C for 15 minutes followed by 50°C for 15 minutes. Two µl of the reaction were used to transform *E. coli* XL10 SOLOPACK® Gold Supercompetent cells. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. An *E. coli* transformant containing pDFng133-3 was prepared using a BIOROBOT® 9600. The *Aspergillus fumigatus* beta-glucosidase mutant insert in pDFng133-3 was confirmed by DNA sequencing.

Example 17: Construction of a tandem expression vector pAmFs074 encoding *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* Cel3A beta-glucosidase variant

A tandem expression vector, pAmFs074, was generated by combining restriction enzyme fragments from pRRAB01 and pDFNG133-3 to generate a single vector for the expression of the *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* Cel3A beta-glucosidase variant.

One microgram of plasmid pRRAB01 (Example 15) purified using a Plasmid Midi Kit (QIAGEN Inc., Valencia, CA, USA) was combined with 20 units of *Xho* I (New England Biolabs Inc, Ipswich, MA, USA), 1X NEB Buffer 4 (New England Biolabs Inc, Ipswich, MA, USA), and 1X BSA in a final volume of 20 µl. The reaction was incubated at 37°C for 3 hours and then combined with 4 µl of 5X DNA loading dye (QIAGEN Inc., Valencia, CA, USA). The restriction digestion reaction products were separated by 1% agarose gel electrophoresis using TBE buffer where an approximately 9.4 kb fragment was excised from the gel and extracted using a NUCLEOSPIN® Extract II Kit according to the manufacturer's protocol. The 9.4 kb fragment contains the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *P. emersonii* GH61A polypeptide coding sequence, *T. reesei* Cel7A cellobiohydrolase I gene terminator, *T. reesei* Cel6A cellobiohydrolase II gene promoter, a 487 bp portion of the 3' end and a 16 bp portion of the 5' end of the *A. fumigatus* beta-glucosidase Cel3A beta-

glucosidase coding sequence, *T. reesei* Cel6A cellobiohydrolase II gene terminator, *Aspergillus nidulans* acetamidase (*amdS*) gene, and ampicillin resistance marker gene.

One microgram of plasmid pDFNG133-3 (Example 16) purified using a Plasmid Midi Kit was combined with 20 units of restriction enzyme *Xho* I, 1X NEB Buffer 4, and 1X BSA in a final volume of 20 µl. The reaction was incubated at 37°C for 3 hours and then combined with 4 µl of 5X DNA loading dye. The restriction digestion reaction products were separated by 1% agarose gel electrophoresis using TBE buffer where an approximately 2.6 kb fragment was excised from the gel and extracted using a NUCLEOSPIN® Extract II Kit according to the manufacturer's protocol. The 2.6 kb fragment contains a 1940 bp portion of the *A. fumigatus* beta-glucosidase Cel3A beta-glucosidase mutant gene.

The purified 9.4 kb pRRAB01 *Xho* I restriction fragment and the purified 2.6 kb pDFNG133-3 *Xho* I restriction fragment were ligated together using a QUICK LIGATION™ Kit according to the manufacturer's protocol. A 50 ng aliquot of the 9.4 kb pRRAB01 fragment and a 50 ng aliquot of the 2.6 kb pDFNG133-3 fragment were combined and the volume adjusted to 10 µl using sterile water. Then 10 µl of 2X QUICK LIGATION™ Buffer and 1 µl of QUICK LIGASE™ were added and mixed thoroughly. The reaction was carried out at 25°C for 5 minutes and then placed on ice. A 2 µl aliquot of the ligation reaction was used to transform ONE SHOT® TOP10 competent cells according to the manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. Transformants were screened by restriction mapping and sequencing. One clone containing the insert in the correct orientation with no sequence errors was identified and designated pAmFs074 (Figure 15).

Plasmid pAmFs074 can be digested with *Pme* I to generate an approximately 9.35 kb fragment for *T. reesei* transformation. The 9.35 kb fragment contains an expression cassette composed of (1) the *T. reesei* Cel7A cellobiohydrolase I gene promoter, *P. emersonii* GH61A polypeptide coding sequence, and *T. reesei* Cel7A cellobiohydrolase I gene terminator, and (2) *T. reesei* Cel6A cellobiohydrolase II gene promoter, *A. fumigatus* beta-glucosidase Cel3A beta-glucosidase mutant coding sequence, and *T. reesei* Cel6A cellobiohydrolase II gene terminator. The 9.35 kb fragment also contains the *Aspergillus nidulans* acetamidase (*amdS*) gene.

Example 18: Generation of a *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* Cel3A variant beta-glucosidase tandem expression vector for replacement of the *T. reesei cbhII*

The tandem replacement vector pQM22 was constructed for replacing the *T. reesei cbhII* gene in *Trichoderma reesei* with a tandem expression cassette for expressing two recombinant proteins in addition to a fungal selection marker. Vector pQM22 contains a *T.*

reesei cbhII 5' flanking region, *T. reesei* Cel7A cellobiohydrolase I promoter, *P. emersonii* GH61A polypeptide coding sequence, *T. reesei* Cel7A cellobiohydrolase I gene terminator, *T. reesei* Cel6A cellobiohydrolase II gene promoter, *A. fumigatus* beta-glucosidase Cel3A beta-glucosidase mutant coding sequence, *T. reesei* Cel6A cellobiohydrolase II gene terminator, *Herpes simplex* virus thymidine kinase (*tk*) gene, *E. coli* hygromycin phosphotransferase (*hpt/hygR)* selection marker, *T. reesei cbhII* 3' flanking region, and ampicillin resistance marker gene.

Vector pQM22 was made by inserting an approximately 6.4 kb tandem expression cassette liberated from pAmFs074 and an approximately 1.5 kb *T. reesei cbhII* 5' flanking region amplified from *T. reesei* strain AgJg115-104-7B1 genomic DNA (isolated according to Example 9) into an approximately 8.7 kb vector fragment from pJfyS142 (Example 9) digested with *Sap* I and *Pac* I.

The 6.4 kb tandem expression cassette for the *P. emersonii* GH61A polypeptide and *A. fumigatus* beta-glucosidase variant was liberated from pAmFs074 by digesting the plasmid with *Pme* I and *Eco* RV. The digestion was subjected to 1% agarose gel electrophoresis using TAE buffer where an approximately 6.4 kb band was excised from the gel and extracted using a NUCLEOSPIN® Extract II Purification Kit according to the manufacturer's protocol.

The 5' *T. reesei cbhII* flanking sequence was amplified from *Trichoderma reesei* strain QMJi030-A5.6 genomic DNA using the forward and reverse primers below. The region in italics represents sequence homology to the site of insertion for an IN-FUSION® reaction and the underlined portion represents introduced *Pac* I and *Psi* I sites for cloning.

Forward Primer:

5'-*gcgagtcagtgagcgaggaagcggaagagc*<u>ttaattaa</u>TCTTGAGTGGATGTCTGATCTAG-3' (SEQ ID NO: 92)

Reverse Primer:

5'-*gttcggataacaatcctacattcggtcgac*<u>ttataa</u>GGATGTATCAATGGGTTATACG-3' (SEQ ID NO: 93)

The amplification reaction was composed of approximately 150 ng of *Trichoderma reesei* QMJi030-A5.6 (Example 13) genomic DNA, 1 µM primers, 200 µM GeneAmp® dNTP (Appiled Biosystems, Foster City, USA), 1X PHUSION™ High-Fidelity Hot Start DNA Polymerase Buffer, and 1 unit of PHUSION™ High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 2 minutes; 35 cycles each at 98°C for 15 seconds, 60°C for 30 seconds, and 72°C for 1 minute; and 1 cycle at 72°C for 10 minutes. The PCR products were separated by 1% agarose gel

electrophoresis using TAE buffer where an approximately 1.5 kb band was excised from the gel and extracted using a NUCLEOSPIN® Extract II Purification Kit according to the manufacturer's protocol.

Approximately 36 µg of pJfyS142 were digested with *Sap* I and *Pac* I. The digestion reaction was separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 8.7 kb band was excised from the gel and extracted using a NUCLEOSPIN® Extract II Purification Kit according to the manufacturer's protocol.

An IN-FUSION™ reaction was composed of 1X IN-FUSION™ Reaction Buffer, 138 ng of the 8.7 kb pJfyS142 fragment digested with *Sap* I and *Pac* I, 205 ng of the 6.4 kb fragment from pAmFs074, 49 ng of the 1.5 kb fragment of the *T. reesei cbhII* 5' flanking region, and 1 µl of IN-FUSION™ Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C and 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 2.5 µl aliquot was used to transform ONE SHOT® TOP10 competent cells according to manufacturer's protocol. The *E. coli* transformation reactions were spread onto 2XYT plus ampicillin plates. Transformants were screened by sequencing and one clone containing the insert with no PCR errors was identified and designated pQM22 (Figure 16). Plasmid pQM22 was used as the vector to replace the *T. reesei cbhII* gene.

Example 19: Replacement of the native *Trichoderma reesei cbhII* gene with *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* beta-glucosidase mutant tandem expression cassette

To replace the native *T. reesei cbhII* gene with the *P. emersonii* GH61 polypeptide-*Aspergillus fumigatus* beta-glucosidase mutant tandem expression cassette, *Trichoderma reesei* QMJi030-A5.6 (Example 13) was transformed with 6.3 µg of *Pme* I-linearized pQM22 (Example 18). Thirty-one transformants were obtained and each one was picked and transferred to a PDA plate and incubated for 7 days at 28°C. After screening out transformants containing an intact *cbhII* gene locus using the fungal spore PCR method described in Example 13 with primer set A shown below, genomic DNA was isolated from eight transformants according to the procedure of Example 9 and analyzed by PCR using primer set B shown below, a forward primer annealing to a region upstream of the *cbhII* 5' flanking region of integration, a reverse primer-1 in the *tk* region after the *A. fumigatus* beta-glucosidase mutant region, and another reverse primer-2 downstream of the *T. reesei cbhII* 3' flanking region.

Primer set A Forward primer:

5'-TCAACCAGCTTCTTTATTGG-3' (SEQ ID NO: 94)

Primer set A Reverse primer:

5'-GATCGCCATAGGCTCATGCTCCGCA-3' (SEQ ID NO: 95)

Primer set B Forward primer:

5'-GCGGCATCAAACACGAACCTG-3' (SEQ ID NO: 96)

Primer set B Reverse-1 primer:

5'-CAGTGGCGCTTATTACTCAG-3' (SEQ ID NO: 97)

Primer set B Reverse-2 primer:

5'-GATCGCCATAGGCTCATGCTCCGCA-3' (SEQ ID NO: 98)

The reaction was composed of 2 µl of the spore suspension, 200 µM dNTP's, 1 µM primers, 1X LONGAMP® *Taq* Reaction Buffer, and 2 units of LONGAMP® *Taq* DNA Polymerase in a final volume of 20 µl. The reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 4 minutes; 35 cycles each at 95°C for 15 seconds, 55°C for 30 seconds, 68°C for 11 minutes; and 1 cycle at 68°C for 15 minutes. The PCR reactions were analyzed by 1% agarose gel electrophoresis using TAE buffer. The PCR results indicated one transformant contained the replacement cassette at the targeted *cbhII* locus. Southern analysis was performed as described below to confirm the replacement cassette as a single copy for both the *P. emersonii* GH61A polypeptide sequence and the *A. fumigatus* beta-glucosidase mutant coding sequence.

Genomic DNA was isolated from the transformants according to the procedure described In Example 9 and each transformant submitted to Southern analysis. For Southern analysis, 6 µg of genomic DNA was digested with 10 units of *Cla* I and/or a restriction enzyme mixture containing 5 units of *Stu* I and 5 units of *Sex* AI in a 70 µl reaction volume. The digested DNA reaction was mixed with 14 µl of 5X DNA loading dye and 25 µl of each mixture was subjected to 1% agarose electrophoresis using TAE buffer. The DNA in the gel was depurinated in 0.25 N HCl for 15 minutes, denatured in Denaturing Solution twice for 15 minutes each time, neutralized in Neutralization Solution for 10 to 30 minutes, and transferred to a NYTRAN® Supercharge membrane. The DNA was UV crosslinked to the membrane using a UV STRATALINKER™ UV crosslinker and prehybridized for 1 hour at 42°C in 20 ml of DIG Easy Hyb.

Approximately 1 µg of pQM22 was digested with *Stu* I and *Xba* I. The digestion reaction was separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 800 bp band from the *P. emersonii* GH61A polypeptide coding sequence was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit according to the manufacturer's protocol.

Approximately 1 µg of pQM22 was digested with *Xho* I and *Nru* I. The digestion reaction was separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 1 kb band from the *A. fumigatus* Cel3A beta-glucosidase mutant coding

sequence was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit according to the manufacturer's protocol.

Approximately 1 µg of pQM22 was digested with *Not* I and *Pvu* I. The digestion reaction was separated by 1% agarose gel electrophoresis using TAE buffer where an approximately 1.1 kb band from the *hpt* gene was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit according to the manufacturer's protocol.

Amounts of 266 ng of the 800 bp fragment, 168 ng of the 1 kb fragment, and 240 ng of the 1.1 kb fragment were each combined in a final volume of 16 µl to generate a probe using a DIG-High Prime DNA Labeling Kit (Roche Diagnostics Corporation, Indianapolis, IN, USA) according to the manufacturer's instructions. The DNA mixture was boiled for ten minutes and then quickly chilled on ice before adding 4 µl of a DIG-High Prime Mix (Roche Diagnostics Corporation, Indianapolis, IN, USA). The reaction was incubated at 37°C for approximately 20 hours before adding 2 µl of 0.2 M EDTA and then heating at 65°C for 10 minutes to stop the reaction.

The probe was purified using a MINELUTE® Gel Extraction Kit according to the manufacturer's protocol. The probe was boiled for 5 minutes, chilled on ice for 2 minutes, and added to 10 ml of DIG Easy Hyb to produce the hybridization solution. Hybridization was performed at 42°C for approximately 17 hours. The membrane was then washed under low stringency conditions in 2X SSC plus 0.1% SDS for 5 minutes at room temperature followed by two high stringency washes in 0.5X SSC plus 0.1% SDS for 15 minutes each at 65°C. The probe-target hybrids were detected by chemiluminescent assay (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's instructions. Southern analysis indicated that the transformant contained a single copy of the *P. emersonii* GH61A polypeptide coding sequence and *A. fumigatus* beta-glucosidase (Cel3A) mutant coding sequence at the targeted *cbhII* locus. The transformant was designated *T. reesei* QMJi033.

Example 20: *Trichoderma reesei* protoplast generation and transformation

Protoplast preparation and transformation were performed using a modified protocol by Penttila *et al.*, 1987, *Gene* 61: 155-164. Briefly, *Trichoderma reesei* strain AgJg115-104-7B1 (PCT/US2010/061105; WO 2011/075677) was cultivated in 25 ml of YP medium supplemented with 2% (w/v) glucose and 10 mM uridine at 27°C for 17 hours with gentle agitation at 90 rpm. Mycelia were collected by filtration using a Vacuum Driven Disposable Filtration System and washed twice with deionized water and twice with 1.2 M sorbitol. Protoplasts were generated by suspending the washed mycelia in 20 ml of 1.2 M sorbitol containing 15 mg of GLUCANEX® 200 G per ml and 0.36 units of chitinase per ml for 15-25 minutes at 34°C with gentle shaking at 90 rpm. Protoplasts were collected by centrifuging for

7 minutes at 400 x *g* and washed twice with cold 1.2 M sorbitol. The protoplasts were counted using a haemocytometer and re-suspended to a final concentration of $1x10^8$ protoplasts per ml in STC. Excess protoplasts were stored in a Cryo 1°C Freezing Container at -80°C.

Approximately 100 µg of a transforming plasmid described in the following examples were digested with *Pme* I. The digestion reaction was purified by 1% agarose gel electrophoresis using TAE buffer. A DNA band was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit. The resulting purified DNA was added to 100 µl of the protoplast solution and mixed gently. PEG buffer (250 µl) was added, mixed, and incubated at 34°C for 30 minutes. STC (3 ml) was then added, mixed, and spread onto PDA plates supplemented with 1 M sucrose. After incubation at 28°C for 16 hours, 20 ml of an overlay PDA medium supplemented with 35 µg of hygromycin B per ml was added to each plate. The plates were incubated at 28°C for 4-7 days.

Example 21: Replacement of native *Trichoderma reesei cbhI* gene with the *Aspergillus fumigatus cbhI* gene

In order to replace the *Trichoderma reesei* native *cbhI* gene (SEQ ID NO: 1 [DNA sequence] and SEQ ID NO: 2 [deduced amino acid sequence]) with the *Aspergillus fumigatus cbhI* coding sequence (SEQ ID NO: 57 [DNA sequence] and SEQ ID NO: 58 [deduced amino acid sequence]), *Trichoderma reesei* ku70- strain AgJg115-104-7B1 (PCT/US2010/061105; WO 2011/075677) was transformed with 4 x 2 µg of *Pme* I-linearized pJfyS139 (Example 11) according to the procedure described in Example 20. Seven transformants were obtained and each one was picked and transferred to a PDA plate and incubated for 7 days at 28°C. Genomic DNA was isolated from the transformants according to the procedure described in Example 9 and each transformant submitted to Southern analysis.

For Southern analysis, 2 µg of genomic DNA was digested with 33 units of *Bgl* II in a 50 µl reaction volume and subjected to 1% agarose electrophoresis using TAE buffer. The DNA in the gel was depurinated with one 10 minute wash in 0.25 N HCl, denatured with two 15 minute washes in 0.5 N NaOH-1.5 M NaCl, neutralized with one 30 minute wash in 1 M Tris pH 8-1.5 M NaCl, and incubated in 20X SSC for 5 minutes. The DNA was transferred to a NYTRAN® Supercharge membrane using a TURBOBLOTTER™ System according to the manufacturer's protocol. The DNA was UV crosslinked to the membrane using a STRATALINKER™ UV Crosslinker and prehybridized for 1 hour at 42°C in 20 ml of DIG Easy Hyb.

A probe hybridizing to the 3' *cbhI* gene flanking sequence was generated using a PCR Dig Probe Synthesis Kit according to the manufacturer's instructions with the forward and reverse primers shown below. The PCR reaction was composed of 1X HERCULASE® Reaction Buffer, 400 nM of each primer, 200 µM DIG-labeled dUTP-containing dNTPs, 20 ng of pJfyS139, and 1.5 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 25 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 40 seconds; and 1 cycle at 72°C for 7 minutes.

Forward primer:

5'-AAAAAACAAACATCCCGTTCATAAC-3' (SEQ ID NO: 99)

Reverse primer:

5'-AACAAGGTTTACCGGTTTCGAAAAG-3' (SEQ ID NO: 100)

The probe was purified by 1% agarose gel electrophoresis using TAE buffer where a 0.5 kb band corresponding to the probe was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The probe was boiled for 5 minutes, chilled on ice for 2 minutes, and added to 10 ml of DIG Easy Hyb to produce the hybridization solution. Hybridization was performed at 42°C for 15-17 hours. The membrane was then washed under low stringency conditions in 2X SSC plus 0.1% SDS for 5 minutes at room temperature followed by two high stringency washes in 0.5X SSC plus 0.1% SDS for 15 minutes each at 65°C. The probe-target hybrids were detected by chemiluminescent assay (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's instructions. Southern analysis indicated that 3 of the 7 transformants contained the replacement cassette at the *cbhI* locus and one transformant, *T. reesei* JfyS139-8, was chosen for curing the *hpt* and *tk* markers.

A fresh plate of spores was generated by transferring spores of a 7 day old PDA plate grown at 28°C to a PDA plate and incubating for 7 days at 28°C. Spores were collected in 10 ml of 0.01% TWEEN® 20 using a sterile spreader. The concentration of spores was determined using a hemocytometer and 10^5 spores were spread onto 150 mm plates containing TrMM-G medium supplemented with 1 µM 5-fluoro-2'-deoxyuridine (FdU).

Three hundred FdU-resistant spore isolates were obtained and DNA was extracted from 2 of the spore isolates as described above. The isolates were submitted to Southern analysis as described above and the results indicated that both spore isolates had excised the *hpt/tk* region between the homologous repeats of the replacement cassette. One strain designated *T. reesei* JfyS139-8A was chosen for replacing the *cbhII* gene.

Example 22: Replacement of the native *Trichoderma reesei cbhII* gene with the *Aspergillus fumigatus cbhII* gene

In order to replace the native *T. reesei cbhII* gene with the *Aspergillus fumigatus cbhII* coding sequence, *Trichoderma reesei* JfyS139-8A (Example 21) was transformed according to the procedure described in Example 20 with 2 µg of *Pme* I-linearized and gel purified pJfyS144 (Example 10). Seven transformants were obtained and each one was picked and transferred to a PDA plate and incubated for 7 days at 28°C. A fungal spore PCR method described below was used to screen for transformants bearing gene replacement using the forward primer shown below annealing to a region upstream of the 5' *cbhII* gene flanking sequence beyond the region of integration, and the reverse primer shown below for the *A. fumigatus cbhII* coding sequence.

Forward primer:

5'-AGCCACATGCCGCATATTGACAAAG-3' (SEQ ID NO: 101)

Reverse primer:

5'-AGGGATTCAGTGTGCTACAGGCTGC-3' (SEQ ID NO: 102)

A 1.8 kb PCR product would be generated only upon the occurrence of a precise gene replacement at the *cbhII* locus. If the cassette had integrated elsewhere in the genome, no amplification would result.

A small amount of spores from each transformant was suspended in 25 µl of TE buffer and heated on high in a microwave oven for 1 minute. Each microwaved spore suspension was used as a template in the PCR reaction. The reaction was composed of 1 µl of the microwaved spore suspension, 1 µl of a 10 mM dNTPs, 12.5 µl of 2X ADVANTAGE® GC-Melt LA Buffer (Clontech, Mountain View, CA, USA), 25 pmol of forward primer, 25 pmol of reverse primer, 1.25 units of ADVANTAGE® GC Genomic LA Polymerase Mix (Clontech, Mountain View, CA, USA), and 9.25 µl of water. The reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 10 minutes; 35 cycles each at 95°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minute 40 seconds; 1 cycle at 72°C for 7 minutes; and a 4°C hold. The PCR reactions were subjected to 1% agarose gel electrophoresis using TAE buffer. The spore PCR indicated that four of the seven transformants contained the replacement cassette at the targeted locus and three of them were submitted to Southern analysis to confirm the replacement cassette was in a single copy.

Genomic DNA was isolated from the three transformants according to the procedure described in Example 9 and each transformant submitted to Southern analysis. For Southern analysis, 2 µg of genomic DNA was digested with 50 units of *Dra* I in a 50 µl reaction volume and subjected to 1% agarose electrophoresis using TAE buffer. The DNA in the gel was depurinated with one 10 minute wash in 0.25 N HCl, denatured with two 15 minute washes in 0.5 N NaOH-1.5 M NaCl, neutralized with one 30 minute wash in 1 M Tris pH 8-1.5 M NaCl, and incubated in 20X SSC for 5 minutes. The DNA was transferred to a NYTRAN®

Supercharge membrane. The DNA was UV crosslinked to the membrane using a STRATALINKER™ UV crosslinker and prehybridized for 1 hour at 42°C in 20 ml of DIG Easy Hyb.

A probe hybridizing to the 3' *cbhII* gene flanking sequence was generated using a PCR Dig Probe Synthesis Kit according to the manufacturer's instructions with the forward and reverse primers indicated below. The PCR reaction was composed of 1X HERCULASE® Reaction Buffer, 400 nM each primer, 200 µM DIG-labeled dUTP-containing dNTPs, 150 ng of *T. reesei* RutC30 genomic DNA, and 1.5 units of HERCULASE® Hot Start High-Fidelity DNA Polymerase. The reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 2 minutes; 30 cycles each at 95°C for 30 seconds, 51°C for 30 seconds, and 72°C for 40 seconds; and 1 cycle at 72°C for 7 minutes.

Forward primer:

5'-AAAAAACAAACATCCCGTTCATAAC-3' (SEQ ID NO: 103)

Reverse primer:

5'-AACAAGGTTTACCGGTTTCGAAAAG-3' (SEQ ID NO: 104)

The probe was purified by 1% agarose gel electrophoresis using TAE buffer where a 0.5 kb band corresponding to the probe was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit. The probe was boiled for 5 minutes, chilled on ice for 2 minutes, and added to 10 ml of DIG Easy Hyb to produce the hybridization solution. Hybridization was performed at 42°C for approximately 17 hours. The membrane was then washed under low stringency conditions in 2X SSC plus 0.1% SDS for 5 minutes at room temperature followed by two high stringency washes in 0.5X SSC plus 0.1% SDS for 15 minutes each at 65°C. The probe-target hybrids were detected by chemiluminescent assay (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's instructions. Southern analysis indicated that the three transformants contained the replacement cassette at the *cbhII* locus and all three (designated JfyS139/144-5, -6, and -10) were chosen for curing the *hpt* and *tk* markers.

A fresh plate of spores was generated by transferring a plug of a 7 day old culture grown on a PDA plate at 28°C to a new PDA plate and incubating for 7 days at 28°C. Spores were collected in 10 ml of 0.01% TWEEN® 20 using a sterile spreader. The concentration of spores was determined using a hemacytometer and 10^5 and 10^4 spores were spread onto 150 mm plates containing TrMM-G medium supplemented with 1 µM FdU.

Approximately 500 FdU-resistant spore isolates for each transformant were obtained from the plate containing 10^5 spores and approximately 100 FdU-resistant spore isolates for each transformant from the plate containing 10^4 spores. Eight spore isolates were picked for strains JfyS139/144-5 and -6 and four were picked for strain JfyS139/144-10. Each isolate 1

to 8 from primary transformant 5 was designated JfyS139/144-5A to -5H. Isolates 1 to 8 from primary transformant 6 were designated JfyS139/144-6A to 6H. Isolates from primary transformant 10 were designated JfyS139/144-10A to 10D for isolates 1 to 4. Spore PCR was conducted as described above, using the forward and reverse primers shown below, to confirm the *hpt* and *tk* markers had been correctly excised.

Forward primer:

5'-GTTAAGCATACAATTGAACGAGAATGG-3' (SEQ ID NO: 105)

Reverse primer:

5'-GATGATATAATGGAGCAAATAAGGG-3' (SEQ ID NO: 106)

The PCR reactions were performed as described above with the following cycling parameters: 1 cycle at 95°C for 2 minutes; 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 6 minutes seconds; and 1 cycle at 72°C for 7 minutes.

The primers annealed to the 5' (forward) and 3' (reverse) flanking sequences used for the *cbhII* gene replacement. Strains from which the *hpt/tk* cassette had been correctly excised would display a 3.5 kb fragment while those with the markers intact would display an 8 kb fragment. The PCR screen indicated that all of the spore isolates had correctly excised the *hpt/tk* cassette.

DNA was extracted from the A and B spore isolates from each primary transformant and submitted to Southern analysis as described above. The Southern analysis confirmed that each spore isolate had correctly excised the *hpt/tk* cassette. Spore isolate *T. reesei* JfyS139/144-10B was chosen to represent the strain containing both the *T. reesei cbhI* and *cbhII* genes replaced with the respective homologs from *Aspergillus fumigatus.*

Example 23: Generation of *Trichoderma reesei ku70* gene repair plasmid pTH239

Four DNA segments were combined using an IN-FUSION® Advantage PCR Cloning Kit to generate a construct to replace the disrupted *Trichoderma reesei ku70* coding sequence with the native *Trichoderma reesei ku70* coding sequence (SEQ ID NO: 107 [DNA sequence] and SEQ ID NO: 108 [deduced amino acid sequence]). The ampicillin resistance marker region including the prokaryotic origin of replication was amplified from pJfyS139-B (Example 11) using the sequence-specific forward and reverse primers shown below (SEQ ID NOs: 109 and 110). The *T. reesei ku70* gene upstream sequence (consisting of 989 bp from upstream of the *ku70* coding sequence and the first 1010 bp of the *ku70* coding sequence) was amplified from *T. reesei* 981-O-8 genomic DNA using the sequence-specific forward and reverse primers shown below (SEQ ID NOs: 111 and 112). The *T. reesei ku70* gene downstream sequence (consisting of a 500 bp segment repeated from the 3' end of the 1010 bp segment of the *ku70* coding sequence amplified in the upstream PCR product, and a 1067 bp segment containing the remainder of the *ku70* coding sequence, and 461 bp from

downstream of the *ku70* coding sequence) was amplified from *T. reesei* 981-O-8 genomic DNA using the sequence-specific forward and reverse primers shown below (SEQ ID NOs: 113 and 114). *T. reesei* 981-O-8 genomic DNA was prepared according to the procedure described in Example 9.

Forward primer:
5'-GTGTGCGGCCGCTCGAGCATGCATGTTTAAACAGCTTGGCACTGGCCGTCGTTTT-3' (SEQ ID NO: 109)
Reverse primer:
5'-ATCAGCCCCGAGACGGCGCCGCGTTTAAACAATTCGTAATCATGGTCATAGCTGT-3' (SEQ ID NO: 110)
Forward primer:
5'-CATGATTACGAATTGTTTAAACGCGGCGCCGTCTCGGGGCTGATCTTGTCGAGGA-3' (SEQ ID NO: 111)
Reverse primer:
5'-GGCGGCCGTTACTAGTGGATCCAGCCCTTGACAGTGATCTTGAGTCCAGGTGCAA-3' (SEQ ID NO: 112)
Forward primer:
5'-TGCAGATATCCATCACACTGGCGGCCGCAGTTTCCATGTCCAACGTGTTGTTTTGCGC-3' (SEQ ID NO: 113)
Reverse primer:
5'-GCCAGTGCCAAGCTGTTTAAACATGCATGCTCGAGCGGCCGCACACGCCCTCTCCTCG-3' (SEQ ID NO: 114)

For amplification of the ampicillin resistance marker and prokaryotic origin of replication region, the reaction was composed of 100 ng of *T. reesei* 981-O-8 genomic DNA, 200 µM dNTPs, 1 µM of each primer (SEQ ID NOs: 109 and 110), 1X PHUSION® High-Fidelity Hot Start DNA Polymerase Buffer, and 1.0 unit of PHUSION® High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 30 cycles each at 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 1 minute 30 seconds; and 1 cycle at 72°C for 7 minutes. The PCR product was separated by 1% agarose gel electrophoresis using TAE buffer where a 2.692 kb fragment was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit.

For amplification of the *ku70* gene upstream sequence and downstream sequence, the reactions were composed of 100 ng of pJfyS139-B, 200 µM dNTPs, 1 µM of each primer (SEQ ID NOs: 111 and 112 or 113 and 114), 1X PHUSION® High-Fidelity Hot Start DNA Polymerase Buffer, and 1.0 unit of PHUSION® High-Fidelity Hot Start DNA Polymerase in a final volume of 50 µl. The amplification reactions were incubated in an EPPENDORF®

MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 30 cycles each at 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 1 minute 30 seconds; and 1 cycle at 72°C for 7 minutes. The PCR products were separated by 1% agarose gel electrophoresis using TAE buffer where 1.999 kb and 2.028 kb fragments were separately excised from the gels and extracted using a MINELUTE® Gel Extraction Kit.

The fourth DNA segment was generated from a restriction enzyme digestion of pJfyS139-B with *Not* I and *Bam* HI. The reaction was composed of 5 µg of pJfyS139-B, 10 units of *Not* I, 20 units of *Bam* HI, and 20 µl of Restriction Enzyme Buffer 2 (New England Biolabs, Inc., Ipswich, MA, USA) in a total volume of 50 µl. The reaction was incubated for 1 hour at 37°C and then separated by 1% agarose gel electrophoresis using TAE buffer where a 4.400 kb fragment was excised from the gel and extracted using a MINELUTE® Gel Extraction Kit.

The three PCR products of 2,028 bp, 1,999 bp, and 2,692 bp were inserted into *Not* I and *Bam* HI-digested pJfyS139-B using an IN-FUSION® Advantage PCR Cloning Kit according to the manufacturer's protocol. The IN-FUSION® reaction was composed of 1X IN-FUSION® Reaction Buffer, 50 ng of the *Not* I/*Bam* HI-digested pJfyS139-B, 50 ng of the 1.999 kb *ku70* gene upstream PCR product, 50 ng of the 2.028 kb *ku70* gene downstream PCR product, 50 ng of the 2.692 kb ampicillin resistance marker and prokaryotic origin of replication PCR product, and 1 µl of IN-FUSION® Enzyme in a 10 µl reaction volume. The reaction was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C. After the incubation period 40 µl of TE were added to the reaction. A 3 µl aliquot was used to transform *E. coli* XL10 GOLD® competent cells (Stratagene, La Jolla, CA, USA) according to the manufacturer's protocol. The cells were heat shocked at 42°C for 30 seconds and then 500 µl of NZY+ medium, pre-heated to 42°C, were added. The tubes were incubated at 37°C with shaking at 200 rpm for 40 minutes and then plated onto 150 mm diameter 2XYT plus ampicillin plates and incubated at 37°C overnight. The resulting transformants were screened by restriction digestion analysis with *Hind* III and *Xba* I and positive clones sequenced to ensure the absence of PCR errors. One clone containing the insert with no PCR errors was identified and designated pTH239.

Example 24: Repair of the *ku70* gene in the *A. fumigatus cbh1* and *cbh2* replacement strain JfyS139/144-10B

The native *Trichoderma reesei ku70* gene was repaired in strain *T. reesei* JfyS139/144-10B (Example 22) in order to facilitate strain manipulation steps requiring the function of the *ku70* gene in non-homologous end-joining. *T. reesei* JfyS129/144-10B was transformed with 23 x 2 µg of *Pme* I-linearized pTH239 (Example 23) according to the

procedure described in Example 4. Nineteen transformants were obtained and each one was separately transferred to a PDA plate and incubated for 7 days at 28°C.

All nineteen transformants were screened by PCR to confirm homologous integration of the pTH239 *Pme* I fragment at the disrupted *ku70* gene locus. For each of the transformants a sterile inoculation loop was used to collect spores from a 7 day old PDA plate. The spores were transferred to a tube containing 25 µl of 1 mM EDTA-10 mM Tris buffer and microwaved on high for 1 minute. A 1 µl aliquot of the microwaved spore mixture was added directly to the PCR reaction as template DNA. A set of PCR primers shown below were designed to amplify across the disrupted region of the *ku70* coding sequence to distinguish between the host genome with the disruption in the *ku70* coding sequence (848 bp) and the pTH239 targeted strain of interest (606 bp). The PCR reaction was composed of 1X ADVANTAGE® Genomic LA Polymerase Reaction Buffer (Clontech, Mountain View, CA, USA), 400 nM of each primer, 200 µM dNTPs, 1 µl of microwaved TE-spore mixture (described above), and 1.0 unit of ADVANTAGE® Genomic LA Polymerase (Clontech, Mountain View, CA, USA). The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 10 minutes; 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 60 seconds; and 1 cycle at 72°C for 7 minutes.

Forward primer:

5'-CAATGACGATCCGCACGCGT-3' (SEQ ID NO: 115)

Reverse primer:

5'-CAATGACGATCCGCACGCGT-3' (SEQ ID NO: 116)

Only one of the nineteen transformants (#19) was positive for the 606 bp PCR product and negative for the 848 bp PCR product indicative of a strain containing the pTH239 *Pme*I fragment homologously integrated at the *ku70* locus.

Spores from the 7 day old PDA plate of transformant #19 were collected in 10 ml of 0.01% TWEEN® 20 using a sterile spreader. The concentration of spores was determined using a hemocytometer and 10^6 spores were spread onto 150 mm plates containing TrMM-G medium supplemented with 1 µM 5-fluoro-2'-deoxyuridine (FdU) and cultured for 5 days at 28°C. Twenty-two FdU-resistant spore isolates were obtained and transferred to PDA plates and cultivated at 28°C for five days.

All twenty-two spore isolates (#19A-V) were screened by PCR for excision of the *hpt/tk* marker region present between the homologous repeats of the *ku70* coding sequence within the repair cassette. For each of the spore isolates a sterile inoculating loop was used to collect spores from a 7 day old PDA plate. The spores were transferred to a tube containing 25 µl of 1 mM EDTA-10 mM Tris buffer and microwaved on high for 1 minute. A 1 µl aliquot of the spore mixture was added directly to the PCR reaction as template genomic

DNA. A set of PCR primers shown below were designed to amplify across the *hpt/tk* region to distinguish between the presence (6 kb) or absence (1.1 kb) of the *hpt/tk* region. The PCR reaction was composed of 1X ADVANTAGE® Genomic LA Polymerase Reaction Buffer, 400 nM of each primer (below), 200 µM dNTPs, 1 µl of microwaved TE-spore mixture (described above), and 1.0 unit of ADVANTAGE® Genomic LA Polymerase. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 95°C for 10 minutes; 30 cycles each at 95°C for 30 seconds, 50°C for 30 seconds, and 72°C for 6 minutes; and 1 cycle at 72°C for 7 minutes.

Forward primer:

5'-GACACTCTTTTCTCCCATCT-3' (SEQ ID NO: 117)

Reverse primer:

5'-GAGGAGCAGAAGAAGCTCCG-3' (SEQ ID NO: 118)

All twenty-two spore isolates were negative for the 6 kb PCR product corresponding to the *hpt/tk* marker region.

Spores from the 7 day old PDA plates of isolates #19A and #19L were collected in 10 ml of 0.01% TWEEN® 20 using a sterile spreader. The concentration of spores was determined using a hemocytometer and 10^3, 10^2, and 10^1 spores were spread onto 150 mm PDA plates containing 1 M sucrose and cultured for 3 days at 28°C. Ten spore isolates were selected from the PDA plates for both strains #19A and #19L and transferred to fresh PDA plates and placed at 28°C.

Genomic DNA was extracted from 6 spore isolates of both #19L and #19A, according to the procedure described in Example 9 and submitted to Southern analysis.

For Southern analysis, 2 µg of genomic DNA was digested with (1) 5 units and 10 units, respectively, of *Asc* I *and Xho* I or (2) 5 units and 25 units, respectively, of *Asc* I *and Apa* I in a 50 µl reaction volume and subjected to 1% agarose electrophoresis using TAE buffer. The DNA in the gel was depurinated with one 10 minute wash in 0.25 N HCl, denatured with two 15 minute washes in 0.5 N NaOH-1.5 M NaCl, neutralized with one 30 minute wash in 1 M Tris pH 8-1.5 M NaCl, and incubated in 20X SSC for 5 minutes. The DNA was transferred to a NYTRAN® Supercharge membrane using a TURBOBLOTTER™ System according to the manufacturer's protocol. The DNA was UV crosslinked to the membrane using a STRATALINKER™ UV Crosslinker and prehybridized for 1 hour at 42°C in 20 ml of DIG Easy Hyb.

A probe hybridizing to the 3' end of the *ku70* coding sequence was generated using a PCR Dig Probe Synthesis Kit according to the manufacturer's instructions with the forward and reverse primers shown below. In order to generate a pure template for the probe PCR reaction, the 3' end of the *ku70* coding sequence was amplified from *T. reesei* 981-O-8 genomic DNA. The PCR reaction was composed of 1X PHUSION® High-Fidelity Hot Start

DNA Polymerase Buffer, 1 µM of each primer, 200 µM dNTPs, 165 ng of *T. reesei* 981-O-8 genomic DNA, and 1.0 unit of PHUSION® High-Fidelity Hot Start DNA Polymerase. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 epgradient S programmed for 1 cycle at 98°C for 30 seconds; 35 cycles each at 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 15 seconds; and 1 cycle at 72°C for 10 minutes.

Forward primer:

5'-gcatatataacccactcaagta-3' (SEQ ID NO: 119)

Reverse primer:

5'-attatcttggaccggccgcagg-3' (SEQ ID NO: 120)

The 0.5 kb probe template was purified by 1% agarose gel electrophoresis using TAE buffer and excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The purified PCR product was used to generate a DIG-labeled probe as specified by the manufacturer's instructions using the primers and amplification conditions specified above. The 0.5 kb DIG-labeled probe was purified by 1% agarose gel electrophoreseis using TAE buffer and excised from the gel and extracted using a MINELUTE® Gel Extraction Kit. The probe was boiled for 5 minutes, chilled on ice for 2 minutes, and added to 10 ml of DIG Easy Hyb to produce a hybridization solution. Hybridization was performed at 42°C for 15-17 hours. The membrane was then washed under low stringency conditions in 2X SSC plus 0.1% SDS for 5 minutes at room temperature followed by two high stringency washes in 0.5X SSC plus 0.1% SDS for 15 minutes each at 65°C. The probe-target hybrids were detected by chemiluminescent assay (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's instructions. Southern analysis indicated that all spore isolates contained the repair/replacement cassette at the *ku70* locus and were cured of the *hpt* and *tk* markers. One strain designated *T. reesei* 981-O-8.5#10B+Ku70#19L3 was chosen for further transformations.

Example 25: *Trichoderma reesei* 981-O-8.5#10B+Ku70#19L3 protoplast generation and transformation

Protoplast preparation and transformation were performed using a modified protocol by Penttila *et al.*, 1987, *Gene* 61: 155-164. Briefly, *Trichoderma reesei* strain 981-O-8.5#10B+Ku70#19L3 was cultivated in 25 ml of YP medium supplemented with 2% (w/v) glucose and 10 mM uridine at 27°C for 17 hours with gentle agitation at 90 rpm. Mycelia were collected by filtration using a Vacuum Driven Disposable Filtration System and washed twice with deionized water and twice with 1.2 M sorbitol. Protoplasts were generated by suspending the washed mycelia in 20 ml of 1.2 M sorbitol containing 15 mg of GLUCANEX®

200 G per ml and 0.36 units of chitinase per ml for 15-25 minutes at 34°C with gentle shaking at 90 rpm. Protoplasts were collected by centrifuging for 7 minutes at 400 x *g* and washed twice with cold 1.2 M sorbitol. The protoplasts were counted using a haemocytometer and resuspended to a final concentration of $1x10^8$ protoplasts per ml in STC. Excess protoplasts were stored in a Cryo 1°C Freezing Container at -80°C.

Approximately 100 µg of plasmid pAmFs074 were digested with *Pme* I. The digestion reaction was purified by 0.8% agarose gel electrophoresis using TAE buffer. An approximately 9.4 kb DNA band containing the tandem expression cassette comprising the *Penicillium emersonii* GH61A polypeptide coding sequence, *Aspergillus fumigatus* beta-glucosidase mutant coding sequence, and *amdS* marker was excised from the gel and extracted using a QIAQUICK® Gel Extraction Kit according to the manufacturer's suggested protocol.

The resulting 9.4 kb fragment (5 x 2 µg or 2 x 3.5 µg of the 9.4 kb *Pme* I digested pAmFs074) was added to 100 µl of the protoplast solution and mixed gently. PEG buffer (250 µl) was added, and the reaction was mixed and incubated at 34°C for 30 minutes. STC (3 ml) was then added, and the reaction was mixed and then spread onto COVE plates for *amdS* selection. The plates were incubated at 28°C for 6-11 days.

Example 26: Evaluation of *Trichoderma reesei* transformants in shake flasks for expression of *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* beta-glucosidase variant

One hundred and twelve *Trichoderma reesei* 981-O-8.5#10B+Ku70#19L3 transformants were transferred from COVE transformation plates to COVE2 plates supplemented with 10 mM uridine using an inoculation loop and incubated 5-7 days at 28°C. Spores were collected with an inoculating loop and transferred to 25 ml of CIM medium in 125 ml plastic shake flasks. The shake flask cultures were incubated for 5 days at 28°C, 200 rpm. Approximately 14-15 ml of each of the cultures were poured into 15 ml tubes. The tubes were centrifuged at 863 x *g* for 20 minutes. The supernatants were decanted into fresh tubes. The supernatants were assayed for beta-glucosidase activity using a BIOMEK® 3000, a BIOMEK® NX, and an ORCA® robotic arm (Beckman Coulter, Inc, Fullerton, CA, USA). The supernatants were diluted appropriately in 0.1 M succinate, 0.01% TRITON® X-100 (4-(1,1,3,3-tetramethylbutyl)phenyl-polyethylene glycol) buffer pH 5.0 (sample buffer) followed by a series of dilutions from 0-fold to 1/3-fold to 1/9-fold of the diluted sample. A total of 20 µl of each dilution was transferred to a 96-well flat bottom plate. Two hundred microliters of a 1 mg/ml *p*-nitrophenyl-beta-D-glucopyranoside substrate in 0.1 M succinate pH 5.0 buffer were added to each well and then incubated at ambient temperature for 45 minutes. Upon

completion of the incubation period 50 µl of quenching buffer (1 M Tris pH 9) was added to each well. An endpoint was measured at an optical density of 405 nm for the 96-well plate. Thirty-three samples showing activity greater than 16,000 µM/min/ml were analyzed by SDS-PAGE to determine expression of the GH61A polypeptide.

Ten µl of each culture supernatant were analyzed by SDS-PAGE using a CRITERION® 8-16% Tris-HCl Gel according to the manufacturer's instructions. The resulting gel was stained with BIO-SAFE™ Coomassie. SDS-PAGE profiles of the cultures showed that the transformants produced major protein bands of approximately 130 kDa corresponding to the *Aspergillus fumigatus* beta-glucosidase variant and approximately 27 kDa corresponding to the *Penicillium emersonii* GH61A polypeptide. Based on expression of the GH61A polypeptide on SDS-PAGE profiles, twenty-one transformants were selected for further spore purification and analysis.

Spore purification was performed by touching a plate of spores with a 1 µl inoculating loop before dipping the loop in 1 ml of 0.01% TWEEN® 20 followed by vortexing. One microliter and ten microliters of this spore mixture were mixed with 100 µl of 0.01% TWEEN® 20 and plated onto a large COVE plate. Plates were incubated at 28°C until single spore isolates appeared on the plates.

Single spore isolates for all twenty-one candidates were transferred to COVE2 + 10 mM uridine plates and incubated at 28°C for about a week before inoculating 125 ml plastic shake flasks containing 25 ml of CIM medium. The resulting shake flask cultures were analyzed by SDS-PAGE as described above and compared with the shake flask culture of primary transformants. Six spore purified transformants, namely 597A, 676D, 679C, 680A, 683A and 686C, were chosen for 2 liter fermentations to determine yield.

Example 27: Fermentation

The top six spore purified *T. reesei* 981-O-8.5#10B+Ku70#19L3 transformants (597A, 676D, 679C, 680A, 683A and 686C) and *T. reesei* QMJi033 were cultivated on PDA plates at 28°C for about a week.

Shake flask medium was composed of 20 g of dextrose, 10 g of corn steep solids, 1.45 g of $(NH_4)_2SO_4$, 2.08 g of KH_2PO_4, 0.36 g of $CaCl_2$, 0.42 g of $MgSO_4 \cdot 7H_2O$, and 0.42 ml of trace metals solution. The trace metals solution was composed of 216 g of $FeCl_3 \cdot 6H_2O$, 58 g of $ZnSO_4 \cdot 7H_2O$, 27 g of $MnSO_4 \cdot H_2O$, 10 g of $CuSO_4 \cdot 5H_2O$, 2.4 g of H_3BO_3, 336 g of citric acid, and deionized water to 1 liter.

One hundred ml of shake flask medium were added to a 500 ml shake flask. The shake flask was inoculated with two plugs from a solid plate culture and incubated at 28°C

on an orbital shaker at 200 rpm for 48 hours. Fifty ml of the shake flask broth were used to inoculate a 2 liter fermentation vessel.

Fermentation batch medium was composed per liter of 30 g of cellulose, 4 g of dextrose, 10 g of corn steep solids, 3.8 g of $(NH_4)_2SO_4$, 2.8 g of KH_2PO_4, 2.64 g of $CaCl_2$, 1.63 g of $MgSO_4.7H_2O$, 1.8 ml of anti-foam, and 0.66 ml of trace metals solution. The trace metals solution was composed per liter of 216 g of $FeCl_3 \cdot 6H_2O$, 58 g of $ZnSO_4 \cdot 7H_2O$, 27 g of $MnSO_4 \cdot H_2O$, 10 g of $CuSO_4 \cdot 5H_2O$, 2.4 g of H_3BO_3, and 336 g of citric acid. Fermentation feed medium was composed of dextrose.

A total of 1.8 liters of the fermentation batch medium was added to an APPLIKON® three liter glass jacketed fermentor (Applikon Biotechnology Inc., Foster City CA USA). Fermentation feed medium was dosed at a rate of 0 to 4 g/l/hr for a period of 185 hours. The fermentation vessel was maintained at a temperature of 28°C and pH was controlled using an APPLIKON® 1030 control system (Applikon Biotechnology Inc., Foster City CA USA) to a set-point of 4.5 +/- 0.1. Air was added to the vessel at a rate of 1 vvm and the broth was agitated by Rushton impeller rotating at 1100 to 1300 rpm. At the end of the fermentation, whole broth was harvested from the vessel and centrifuged at 3000 *x g* to remove the biomass. The supernatant was sterile filtered and stored at 5 to 10°C.

Example 28: Evaluation of *Trichoderma reesei* transformants in 2 liter fermentation broth for expression of *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* beta-glucosidase variant in 2 liter fermentation broth

Fermentation broth of the top six spore purified *T. reesei* 981-O-8.5#10B+Ku70#19L3 transformants (597A, 676D, 679C, 680A, 683A and 686C) and *T. reesei* QMJi033 were analyzed by SDS-PAGE to evaluate the expression level for the *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* beta-glucosidase variant. A 0.5 µl volume of fermentation broth supernatant of each strain was mixed with 25 µl of deionized water and 25 µl of Laemmli dye (Bio-Rad Laboratories, Hercules, CA, USA) containing 5% 2-mercaptoethanol. The mixture was heated at 95°C for 5 minutes and 10 µl (equivalent to 0.1 µl fermentation broth) were loaded onto a stain-free CRITERION® 8-16% Tris-HCl Gel (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer's instructions. The gel was run at 200 volts for about 55 minutes and then analyzed by a Gel Doc™ EZ Imager (Bio-Rad Laboratories, Hercules, CA, USA). SDS-PAGE profiles of the fermentation broth supernatants showed that expression of the *Penicillium emersonii* GH61A polypeptide and *Aspergillus fumigatus* beta-glucosidase variant in *T. reesei* QMJi033 were comparable to the top six *T. reesei* 981-O-8.5#10B+Ku70#19L3 transformants (597A, 676D, 679C, 680A, 683A and 686C).

The present invention is further described by the following numbered paragraphs:

[1] A method for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising: (a) replacing an endogenous first gene by targeted integration by introducing into the filamentous fungal strain a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) replacing an endogenous second gene by targeted integration by introducing into the filamentous fungal strain a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

[2] The method of paragraph 1, wherein the first gene is a cellobiohydrolase I gene.

[3] The method of paragraph 2, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[4] The method of paragraph 1, wherein the second gene is a cellobiohydrolase II gene.

[5] The method of paragraph 4, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[6] The method of any of paragraphs 1-5, wherein each of the tandem constructs integrates by homologous recombination into the chromosome of the filamentous fungal strain.

[7] The method of any of paragraphs 1-6, wherein the homologous 5' region of the first gene, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[8] The method of any of paragraphs 1-7, wherein the homologous 3' region of the first gene, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[9] The method of any of paragraphs 1-8, wherein the homologous 5' region of the second gene, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[10] The method of any of paragraphs 1-9, wherein the homologous 3' region of the second gene, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[11] The method of any of paragraphs 1-10, which further comprises replacing one or more additional endogenous genes each by targeted integration by introducing into the filamentous fungal strain a corresponding tandem construct for each gene comprising (i) a homologous 5' region of the gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

[12] The method of any of paragraphs 1-11, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

[13] The method of paragraph 12, wherein the first and second homologous repeats are identical or have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other.

[14] The method of paragraph 12 or 13, wherein the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[15] The method of any of paragraphs 12-14, wherein upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for replacing the one or more additional endogenous genes each by targeted integration with the corresponding tandem construct for each gene.

[16] The method of any of paragraphs 1-15, further comprising transforming the filamentous fungal host cell with a tandem construct comprising (i) one or more selectable markers, (ii) a fifth polynucleotide encoding a fifth polypeptide having biological activity operably linked to a fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

[17] The method of any of paragraphs 1-16, wherein the polypeptides having biological activity are different polypeptides.

[18] The method of any of paragraphs 1-16, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[19] The method of any of paragraphs 1-18, wherein the promoters are different promoters.

[20] The method of any of paragraphs 1-18, wherein two or more of the promoters are the same promoter.

[21] The method of any of paragraphs 1-20, wherein the terminators are different terminators.

[22] The method of any of paragraphs 1-20, wherein two or more of the terminators are the same terminator.

[23] The method of any of paragraphs 1-22, wherein one or more of the tandem constructs are contained in an expression vector.

[24] The method of any of paragraphs 1-23, wherein the filamentous fungal strain is an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* strain.

[25] The method of paragraph 24, wherein the *Trichoderma* strain is selected from the group consisting of *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[26] The method of paragraph 24, wherein the *Trichoderma* strain is *Trichoderma reesei*.

[27] A filamentous fungal strain obtained according to the methods of any of paragraphs 1-26.

[28] A filamentous fungal strain, comprising: (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3'

region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

[29] The filamentous fungal strain of paragraph 28, wherein the first gene is a cellobiohydrolase I gene.

[30] The filamentous fungal strain of paragraph 29, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[31] The filamentous fungal strain of paragraph 28, wherein the second gene is a cellobiohydrolase II gene.

[32] The filamentous fungal strain of paragraph 31, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least

high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[33] The filamentous fungal strain of any of paragraphs 28-32, wherein each of the tandem constructs integrated by homologous recombination into the chromosome of the filamentous fungal strain.

[34] The filamentous fungal strain of any of paragraphs 28-33, wherein the homologous 5' region of the first gene, the homologous flanking region thereof, or the combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[35] The filamentous fungal strain of any of paragraphs 28-34, wherein the homologous 3' region of the first gene, the homologous flanking region thereof, or the combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[36] The filamentous fungal strain of any of paragraphs 28-35, wherein the homologous 5' region of the second gene, the homologous flanking region thereof, or the combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[37] The filamentous fungal strain of any of paragraphs 28-36, wherein the homologous 3' region of the second gene, the homologous flanking region thereof, or the combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[38] The filamentous fungal strain of any of paragraphs 28-37, which further comprises one or more additional endogenous genes each replaced by targeted integration with a corresponding tandem construct for each gene comprising (i) a homologous 5' region of the gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

[39] The filamentous fungal strain of any of paragraphs 28-38, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more

selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

[40] The filamentous fungal strain of paragraph 39, wherein the first and second homologous repeats are identical or have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other.

[41] The filamentous fungal strain of paragraph 39 or 40, wherein the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[42] The filamentous fungal strain of any of paragraphs 39-41, wherein upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for replacing the one or more additional endogenous genes each by targeted integration with a corresponding tandem construct for each gene.

[43] The filamentous fungal strain of any of paragraphs 28-42, further comprising a tandem construct comprising (i) one or more selectable markers, (ii) a fifth polynucleotide encoding a fifth polypeptide having biological activity operably linked to a fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

[44] The filamentous fungal strain of any of paragraphs 28-43, wherein the polypeptides having biological activity are different polypeptides.

[45] The filamentous fungal strain of any of paragraphs 28-43, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[46] The filamentous fungal strain of any of paragraphs 28-45, wherein the promoters are different promoters.

[47] The filamentous fungal strain of any of paragraphs 28-45, wherein two or more of the promoters are the same promoter.

[48] The filamentous fungal strain of any of paragraphs 28-47, wherein the terminators are different terminators.

[49] The filamentous fungal strain of any of paragraphs 28-47, wherein two or more of the terminators are the same terminator.

[50] The filamentous fungal strain of any of paragraphs 28-49, wherein one or more of the tandem constructs are contained in an expression vector.

[51] The filamentous fungal strain of any of paragraphs 28-50, wherein the filamentous fungal strain is an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*,

Ceriporiopsis, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* strain.

[52] The filamentous fungal strain of paragraph 51, wherein the *Trichoderma* strain is selected from the group consisting of *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[53] The filamentous fungal strain of paragraph 51, wherein the *Trichoderma* strain is *Trichoderma reesei*.

[54] A method for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising: cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

[55] The method of claim 54, further comprising recovering the multiple recombinant polypeptides.

[56] The method of paragraph 54 or 55, wherein the first gene is a cellobiohydrolase I gene.

[57] The method of paragraph 56, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least

89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[58] The method of paragraph 54 or 55, wherein the second gene is a cellobiohydrolase II gene.

[59] The method of paragraph 58, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[60] The method of any of paragraphs 54-59, wherein the homologous 5' region of the first gene, the homologous flanking region thereof, or the combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[61] The method of any of paragraphs 54-60, wherein the homologous 3' region of the first gene, the homologous flanking region thereof, or the combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[62] The method of any of paragraphs 54-61, wherein the homologous 5' region of the second gene, the homologous flanking region thereof, or the combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[63] The method of any of paragraphs 54-62, wherein the homologous 3' region of the second gene, the homologous flanking region thereof, or the combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[64] The method of any of paragraphs 54-63, wherein the filamentous fungal host cell further comprises one or more additional endogenous genes each replaced by targeted integration with a corresponding tandem construct for each gene comprising (i) a homologous 5' region of the gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

[65] The method of any of paragraphs 54-64, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

[66] The method of paragraph 65, wherein the first and second homologous repeats are identical or have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other.

[67] The method of paragraph 65 or 66, wherein the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[68] The method of any of paragraphs 65-67, wherein upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for replacing the one or more additional endogenous genes each by targeted integration with the corresponding tandem construct for each gene.

[69] The method of any of paragraphs 54-68, wherein the filamentous fungal host cell further comprises a tandem construct comprising (i) one or more selectable markers, (ii) a fifth polynucleotide encoding a fifth polypeptide having biological activity operably linked to a

fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

[70] The method of any of paragraphs 54-69, wherein the polypeptides having biological activity are different polypeptides.

[71] The method of any of paragraphs 54-69, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[72] The method of any of paragraphs 54-71, wherein the promoters are different promoters.

[73] The method of any of paragraphs 54-71, wherein two or more of the promoters are the same promoter.

[74] The method of any of paragraphs 54-73, wherein the terminators are different terminators.

[75] The method of any of paragraphs 54-73, wherein two or more of the terminators are the same terminator.

[76] The method of any of paragraphs 54-75, wherein one or more of the tandem constructs are contained in an expression vector.

[77] The method of any of paragraphs 54-76, wherein the filamentous fungal strain is an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* strain.

[78] The method of paragraph 77, wherein the *Trichoderma* strain is selected from the group consisting of *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[79] The method of paragraph 77, wherein the *Trichoderma* strain is *Trichoderma reesei*.

[80] A tandem construct comprising (i) a homologous 5' region of a gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

[81] The tandem construct of paragraph 80, wherein the gene is a cellobiohydrolase I gene.

[82] The tandem construct of paragraph 81, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[83] The tandem construct of paragraph 80, wherein the gene is a cellobiohydrolase II gene.

[84] The tandem construct of paragraph 83, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[85] The tandem construct of any of paragraphs 80-84, wherein the homologous 5' region of the gene, the homologous flanking region thereof, or the combination thereof is at

least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[86] The tandem construct of any of paragraphs 80-85, wherein the homologous 3' region of the gene, the homologous flanking region thereof, or the combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[87] The tandem construct of any of paragraphs 80-86, wherein the polypeptides having biological activity are different polypeptides.

[88] The tandem construct of any of paragraphs 80-86, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[89] The tandem construct of any of paragraphs 80-88, wherein the promoters are different promoters.

[90] The tandem construct of any of paragraphs 80-88, wherein two or more of the promoters are the same promoter.

[91] The tandem construct of any of paragraphs 80-90, wherein the terminators are different terminators.

[92] The tandem construct of any of paragraphs 80-90, wherein two or more of the terminators are the same terminator.

[93] An expression vector comprising the tandem construct of any of paragraphs 80-92.

[94] A method for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising: (a) inserting into an endogenous first locus by targeted integration a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) inserting into an endogenous second locus by targeted integration a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

[95] The method of paragraph 94, wherein the first locus is a cellobiohydrolase I gene.

[96] The method of paragraph 95, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[97] The method of paragraph 94, wherein the second locus is a cellobiohydrolase II gene.

[98] The method of paragraph 97, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[99] The method of any of paragraphs 94-98, wherein each of the tandem constructs integrates by homologous recombination into the chromosome of the filamentous fungal strain.

[100] The method of any of paragraphs 94-99, wherein the homologous 5' region of the first locus, the homologous flanking region thereof, or a combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[101] The method of any of paragraphs 94-100, wherein the homologous 3' region of the first locus, the homologous flanking region thereof, or a combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[102] The method of any of paragraphs 94-101, wherein the homologous 5' region of the second locus, the homologous flanking region thereof, or a combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[103] The method of any of paragraphs 94-102, wherein the homologous 3' region of the second locus, the homologous flanking region thereof, or a combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[104] The method of any of paragraphs 94-103, which further comprises inserting into one or more additional endogenous loci each by targeted integration a corresponding tandem construct for each locus comprising (i) a homologous 5' region of the locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

[105] The method of any of paragraphs 94-104, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

[106] The method of paragraph 105, wherein the first and second homologous repeats are identical or have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%,

at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other.

[107] The method of paragraph 105 or 106, wherein the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[108] The method of any of paragraphs 105-107, wherein upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for modifying by insertion the one or more additional endogenous loci each by targeted integration with the corresponding tandem construct for each locus.

[109] The method of any of paragraphs 94-108, further comprising transforming the filamentous fungal host cell with a tandem construct comprising (i) one or more selectable markers, (ii) a fifth polynucleotide encoding a fifth polypeptide having biological activity operably linked to a fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

[110] The method of any of paragraphs 94-109, wherein the polypeptides having biological activity are different polypeptides.

[111] The method of any of paragraphs 94-109, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[112] The method of any of paragraphs 94-111, wherein the promoters are different promoters.

[113] The method of any of paragraphs 94-111, wherein two or more of the promoters are the same promoter.

[114] The method of any of paragraphs 94-113, wherein the terminators are different terminators.

[115] The method of any of paragraphs 94-113, wherein two or more of the terminators are the same terminator.

[116] The method of any of paragraphs 94-115, wherein one or more of the tandem constructs are contained in an expression vector.

[117] The method of any of paragraphs 94-116, wherein the filamentous fungal strain is an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* strain.

[118] The method of paragraph 117, wherein the *Trichoderma* strain is selected from the group consisting of *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[119] The method of paragraph 117, wherein the *Trichoderma* strain is *Trichoderma reesei*.

[120] A filamentous fungal strain obtained according to the methods of any of paragraphs 94-119.

[121] A filamentous fungal strain, comprising: (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

[122] The filamentous fungal strain of paragraph 121, wherein the first locus is a cellobiohydrolase I gene.

[123] The filamentous fungal strain of paragraph 122, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at

least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[124] The filamentous fungal strain of paragraph 121, wherein the second locus is a cellobiohydrolase II gene.

[125] The filamentous fungal strain of paragraph 124, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[126] The filamentous fungal strain of any of paragraphs 121-125, wherein each of the tandem constructs integrated by homologous recombination into the chromosome of the filamentous fungal strain.

[127] The filamentous fungal strain of any of paragraphs 121-126, wherein the homologous 5' region of the first locus, the homologous flanking region thereof, or a combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[128] The filamentous fungal strain of any of paragraphs 121-127, wherein the homologous 3' region of the first locus, the homologous flanking region thereof, or a combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[129] The filamentous fungal strain of any of paragraphs 121-128, wherein the homologous 5' region of the second locus, the homologous flanking region thereof, or a

combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[130] The filamentous fungal strain of any of paragraphs 121-129, wherein the homologous 3' region of the second locus, the homologous flanking region thereof, or a combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[131] The filamentous fungal strain of any of paragraphs 121-130, which further comprises one or more additional endogenous loci each modified by insertion by targeted integration with a corresponding tandem construct for each locus comprising (i) a homologous 5' region of the locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

[132] The filamentous fungal strain of any of paragraphs 121-131, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

[133] The filamentous fungal strain of paragraph 132, wherein the first and second homologous repeats are identical or have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other.

[134] The filamentous fungal strain of paragraph 132 or 133, wherein the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[135] The filamentous fungal strain of any of paragraphs 132-134, wherein upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for inserting into the one or more additional endogenous loci each by targeted integration a corresponding tandem construct for each locus.

[136] The filamentous fungal strain of any of paragraphs 121-135, further comprising a tandem construct comprising (i) one or more selectable markers, (ii) a fifth polynucleotide

encoding a fifth polypeptide having biological activity operably linked to a fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

[137] The filamentous fungal strain of any of paragraphs 121-136, wherein the polypeptides having biological activity are different polypeptides.

[138] The filamentous fungal strain of any of paragraphs 121-136, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[139] The filamentous fungal strain of any of paragraphs 121-138, wherein the promoters are different promoters.

[140] The filamentous fungal strain of any of paragraphs 121-138, wherein two or more of the promoters are the same promoter.

[141] The filamentous fungal strain of any of paragraphs 121-140, wherein the the terminators are different terminators.

[142] The filamentous fungal strain of any of paragraphs 121-140, wherein two or more of the terminators are the same terminator.

[143] The filamentous fungal strain of any of paragraphs 121-142, wherein one or more of the tandem constructs are contained in an expression vector.

[144] The filamentous fungal strain of any of paragraphs 121-143, wherein the filamentous fungal strain is an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* strain.

[145] The filamentous fungal strain of paragraph 144, wherein the *Trichoderma* strain is selected from the group consisting of *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[146] The filamentous fungal strain of paragraph 144, wherein the *Trichoderma* strain is *Trichoderma reesei*.

[147] A method for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising: cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide

encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

[148] The method of claim 147, further comprising recovering the multiple recombinant polypeptides.

[149] The method of paragraph 147 or 148, wherein the first locus is a cellobiohydrolase I gene.

[150] The method of paragraph 149, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[151] The method of paragraph 147 or 148, wherein the second locus is a cellobiohydrolase II gene.

[152] The method of paragraph 151, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at

least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[153] The method of any of paragraphs 147-152, wherein the homologous 5' region of the first locus, the homologous flanking region thereof, or a combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[154] The method of any of paragraphs 147-153, wherein the homologous 3' region of the first locus, the homologous flanking region thereof, or a combination thereof of the first tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[155] The method of any of paragraphs 147-154, wherein the homologous 5' region of the second locus, the homologous flanking region thereof, or a combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[156] The method of any of paragraphs 147-155, wherein the homologous 3' region of the second locus, the homologous flanking region thereof, or a combination thereof of the second tandem construct is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[157] The method of any of paragraphs 147-156, wherein the filamentous fungal host cell further comprises one or more additional endogenous loci each modified by insertion by targeted integration with a corresponding tandem construct for each locus comprising (i) a homologous 5' region of the locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a polynucleotide encoding a polypeptide having biological activity operably linked to a promoter and a terminator, (iv) another polynucleotide encoding another polypeptide having biological activity operably linked to another promoter and another terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

[158] The method of any of paragraphs 147-157, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

[159] The method of paragraph 158, wherein the first and second homologous repeats are identical or have a sequence identity of at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% to each other.

[160] The method of paragraph 158 or 159, wherein the first and second homologous repeats are each at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[161] The method of any of paragraphs 158-160, wherein upon the excision of the one or more selectable markers, the one or more selectable markers can be reused for modifying by insertion the one or more additional endogenous loci each by targeted integration with the corresponding tandem construct for each locus.

[162] The method of any of paragraphs 147-161, wherein the filamentous fungal host cell further comprises a tandem construct comprising (i) one or more selectable markers, (ii) a fifth polynucleotide encoding a fifth polypeptide having biological activity operably linked to a fifth promoter and a fifth terminator, and (iii) a sixth polynucleotide encoding a sixth polypeptide having biological activity operably linked to a sixth promoter and a sixth terminator, wherein the tandem construct integrates by ectopic integration.

[163] The method of any of paragraphs 147-162, wherein the polypeptides having biological activity are different polypeptides.

[164] The method of any of paragraphs 147-162, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[165] The method of any of paragraphs 147-164, wherein the promoters are different promoters.

[166] The method of any of paragraphs 147-164, wherein two or more of the promoters are the same promoter.

[167] The method of any of paragraphs 147-166, wherein the terminators are different terminators.

[168] The method of any of paragraphs 147-166, wherein two or more of the terminators are the same terminator.

[169] The method of any of paragraphs 147-168, wherein one or more of the tandem constructs are contained in an expression vector.

[170] The method of any of paragraphs 147-169, wherein the filamentous fungal strain is an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* strain.

[171] The method of paragraph 170, wherein the *Trichoderma* strain is selected from the group consisting of *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[172] The method of paragraph 170, wherein the *Trichoderma* strain is *Trichoderma reesei*.

[173] A tandem construct comprising (i) a homologous 5' region of a locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

[174] The tandem construct of paragraph 173, wherein the locus is a cellobiohydrolase I gene.

[175] The tandem construct of paragraph 174, wherein the cellobiohydrolase I gene encodes a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high

stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof.

[176] The tandem construct of paragraph 173, wherein the locus is a cellobiohydrolase II gene.

[177] The tandem construct of paragraph 176, wherein the cellobiohydrolase II gene encodes a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

[178] The tandem construct of any of paragraphs 173-177, wherein the homologous 5' region of the locus, the homologous flanking region thereof, or a combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[179] The tandem construct of any of paragraphs 173-178, wherein the homologous 3' region of the locus, the homologous flanking region thereof, or a combination thereof is at least 50 bp, *e.g.*, at least 100 bp, at least 200 bp, at least 400 bp, at least 800 bp, at least 1000 bp, at least 1500 bp, or at least 2000 bp.

[180] The tandem construct of any of paragraphs 173-179, wherein the polypeptides having biological activity are different polypeptides.

[181] The tandem construct of any of paragraphs 173-179, wherein two or more of the polypeptides having biological activity are the same polypeptide.

[182] The tandem construct of any of paragraphs 173-181, wherein the promoters are different promoters.

[183] The tandem construct of any of paragraphs 173-181, wherein two or more of the promoters are the same promoter.

[184] The tandem construct of any of paragraphs 173-183, wherein the terminators are different terminators.

[185] The tandem construct of any of paragraphs 173-183, wherein two or more of the terminators are the same terminator.

[186] An expression vector comprising the tandem construct of any of paragraphs 173-185.

The invention described and claimed herein is not to be limited in scope by the specific aspects herein disclosed, since these aspects are intended as illustrations of several aspects of the invention. Any equivalent aspects are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. In the case of conflict, the present disclosure including definitions will control.

Claims

What is claimed is

1. A method for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising:

(a) replacing an endogenous first gene by targeted integration by introducing into the filamentous fungal strain a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof;

(b) replacing an endogenous second gene by targeted integration by introducing into the filamentous fungal strain a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or

(c) a combination of (a) and (b).

2. The method of claim 1, wherein the first gene is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence

identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; and wherein the second gene is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

3. The method of claim 1 or 2, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

4. A filamentous fungal strain, comprising: (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a

combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

5. The filamentous fungal strain of claim 4, wherein the first gene is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; and wherein the second gene is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions,

with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

6. The filamentous fungal strain of claim 4 or 5, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

7. A method for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising: cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first gene replaced by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first gene, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second gene replaced by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second gene, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

8. The method of claim 7, further comprising recovering the multiple recombinant polypeptides.

9. The method of claim 7 or 8, wherein the first gene is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%,

at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; and wherein the second gene is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

10. The method of any of claims 7-9, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

11. A tandem construct comprising (i) a homologous 5' region of a gene, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to

a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the gene, a homologous flanking region thereof, or a combination thereof.

12. The tandem construct of claim 11, wherein the gene is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; or wherein the gene is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

13. A method for constructing a filamentous fungal strain for production of multiple recombinant polypeptides having biological activity, comprising:

(a) inserting into an endogenous first locus by targeted integration a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof;

(b) inserting into an endogenous second locus by targeted integration a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more second selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or

(c) a combination of (a) and (b).

14. The method of claim 13, wherein the first locus is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; and wherein the

second locus is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

15. The method of claim 13 or 14, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

16. A filamentous fungal strain, comprising: (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth

terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

17. The filamentous fungal strain of claim 16, wherein the first locus is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; and wherein the second locus is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

18. The filamentous fungal strain of claim 16 or 17, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

19. A method for producing multiple recombinant polypeptides having biological activity in a filamentous fungal strain, comprising: cultivating a filamentous fungal host cell under conditions conducive for production of the polypeptides, wherein the filamentous fungal host cell comprises (a) an endogenous first locus modified by insertion by targeted integration with a first tandem construct comprising (i) a homologous 5' region of the first locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second terminator, and (v) a homologous 3' region of the first locus, a homologous flanking region thereof, or a combination thereof; (b) an endogenous second locus modified by insertion by targeted integration with a second tandem construct comprising (i) a homologous 5' region of the second locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more selectable markers, (iii) a third polynucleotide encoding a third polypeptide having biological activity operably linked to a third promoter and a third terminator, (iv) a fourth polynucleotide encoding a fourth polypeptide having biological activity operably linked to a fourth promoter and a fourth terminator, and (v) a homologous 3' region of the second locus, a homologous flanking region thereof, or a combination thereof; or (c) a combination of (a) and (b).

20. The method of claim 19, further comprising recovering the multiple recombinant polypeptides.

21. The method of claim 19 or 20, wherein the first locus is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide

comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; and wherein the second locus is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

22. The method of any of claims 19-21, wherein one or more of the tandem constructs further comprise a first homologous repeat flanking 5' of the one or more selectable markers and a second homologous repeat flanking 3' of the one or more selectable markers, wherein the first homologous repeat and the second homologous repeat undergo homologous recombination to excise the one or more selectable markers.

23. A tandem construct comprising (i) a homologous 5' region of a locus, a homologous flanking region thereof, or a combination thereof, (ii) one or more first selectable markers, (iii) a first polynucleotide encoding a first polypeptide having biological activity operably linked to a first promoter and a first terminator, (iv) a second polynucleotide encoding a second polypeptide having biological activity operably linked to a second promoter and a second

terminator, and (v) a homologous 3' region of the locus, a homologous flanking region thereof, or a combination thereof.

24. The tandem construct of claim 23, wherein the locus is a cellobiohydrolase I gene encoding a cellobiohydrolase I selected from the group consisting of: (i) a cellobiohydrolase I comprising the mature polypeptide of SEQ ID NO: 2; (ii) a cellobiohydrolase I comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 2; (iii) a cellobiohydrolase I encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1; and (iv) a cellobiohydrolase I encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 1 or the full-length complement thereof; or wherein the locus is a cellobiohydrolase II gene encoding a cellobiohydrolase II selected from the group consisting of: (i) a cellobiohydrolase II comprising the mature polypeptide of SEQ ID NO: 4; (ii) a cellobiohydrolase II comprising an amino acid sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide of SEQ ID NO: 4; (iii) a cellobiohydrolase II encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, *e.g.*, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3; and (iv) a cellobiohydrolase II encoded by a polynucleotide that hybridizes under at least high stringency conditions, *e.g.*, very high stringency conditions, with the mature polypeptide coding sequence of SEQ ID NO: 3 or the full-length complement thereof.

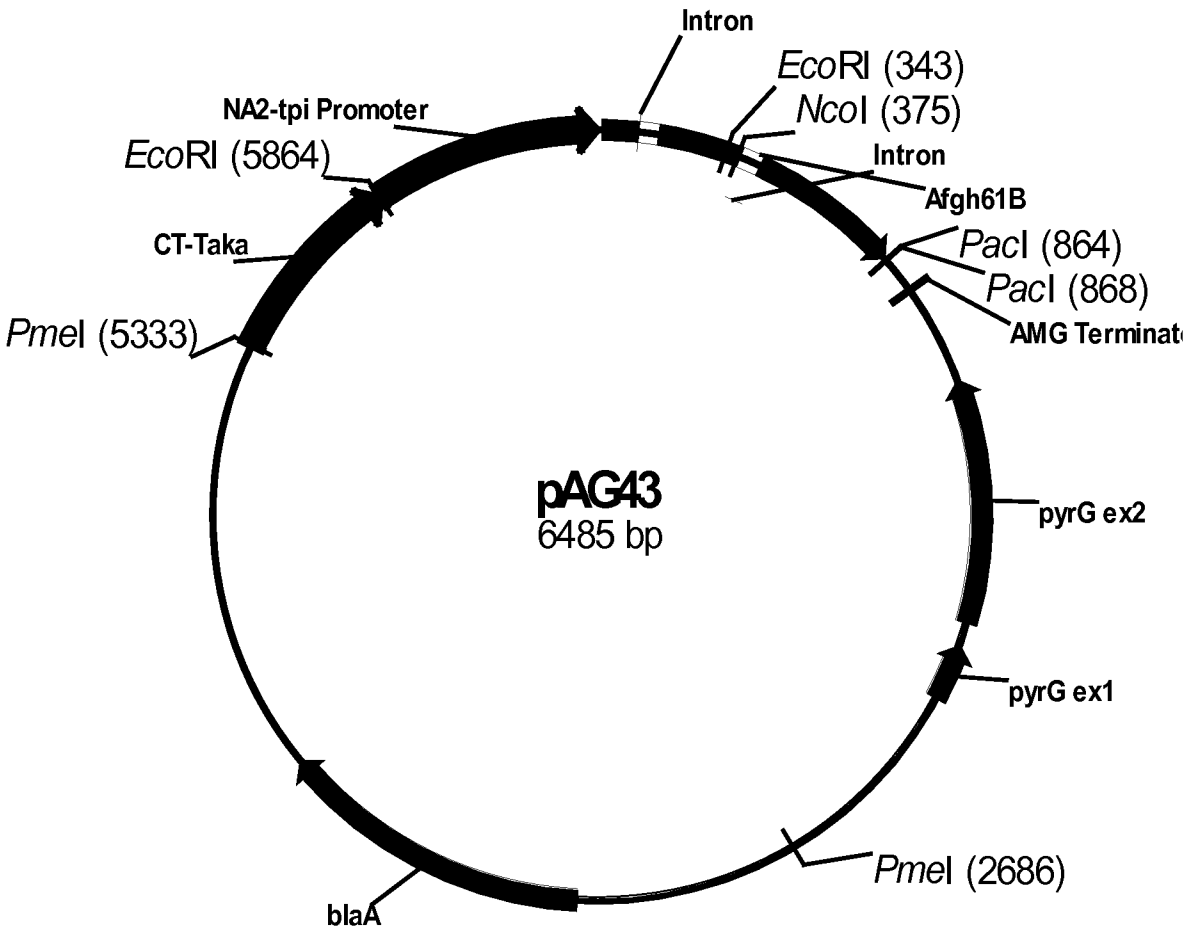
Intron
EcoRI (343)
NcoI (375)
Intron
Afgh61B
PacI (864)
PacI (868)
AMG Terminat
pyrG ex2
pyrG ex1
PmeI (2686)
blaA
PmeI (5333)
CT-Taka
EcoRI (5864)
NA2-tpi Promoter
pAG43
6485 bp

Fig. 1

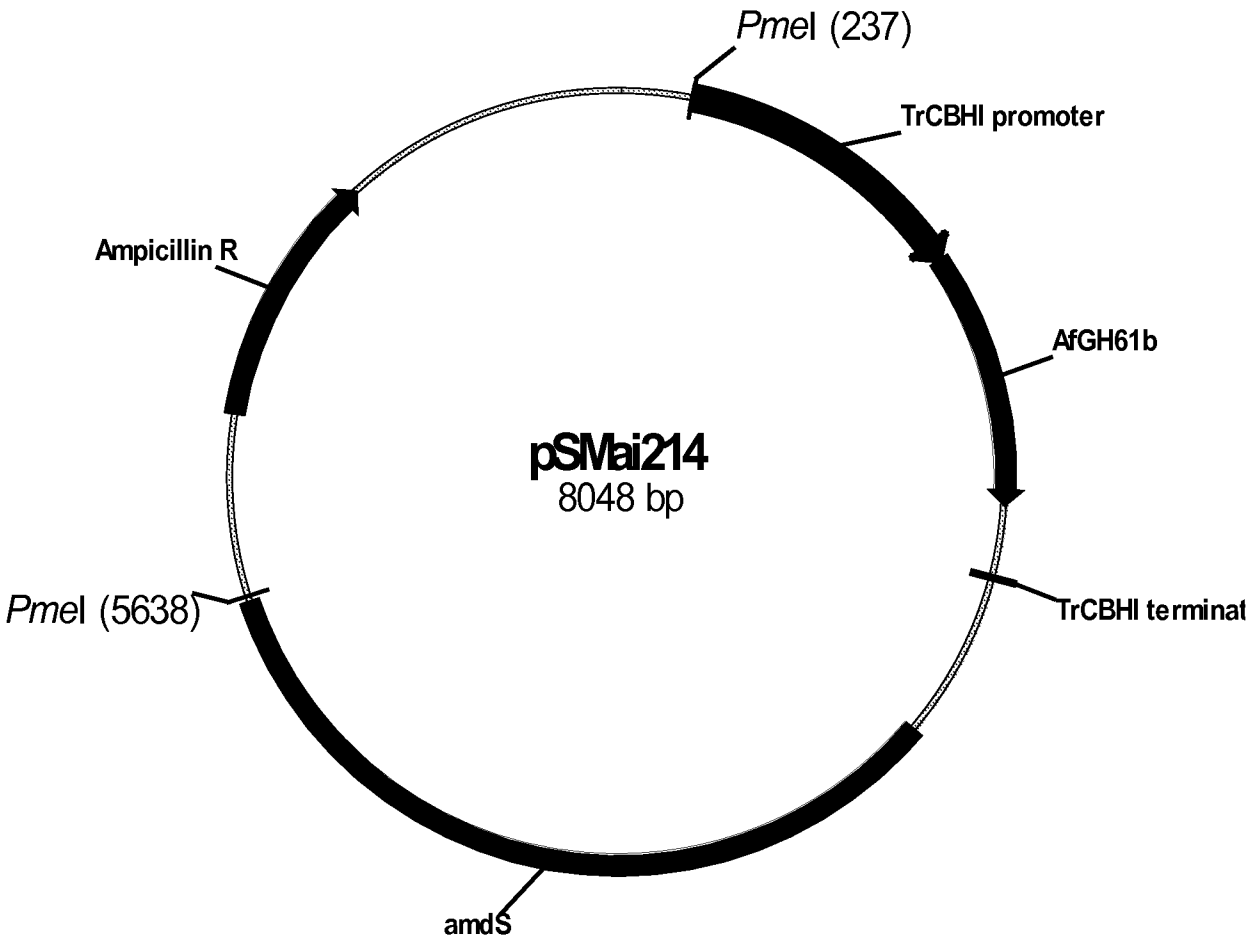
PmeI (237)
TrCBHI promoter
AfGH61b
TrCBHI terminat
amdS
PmeI (5638)
Ampicillin R
pSMai214
8048 bp

Fig. 2

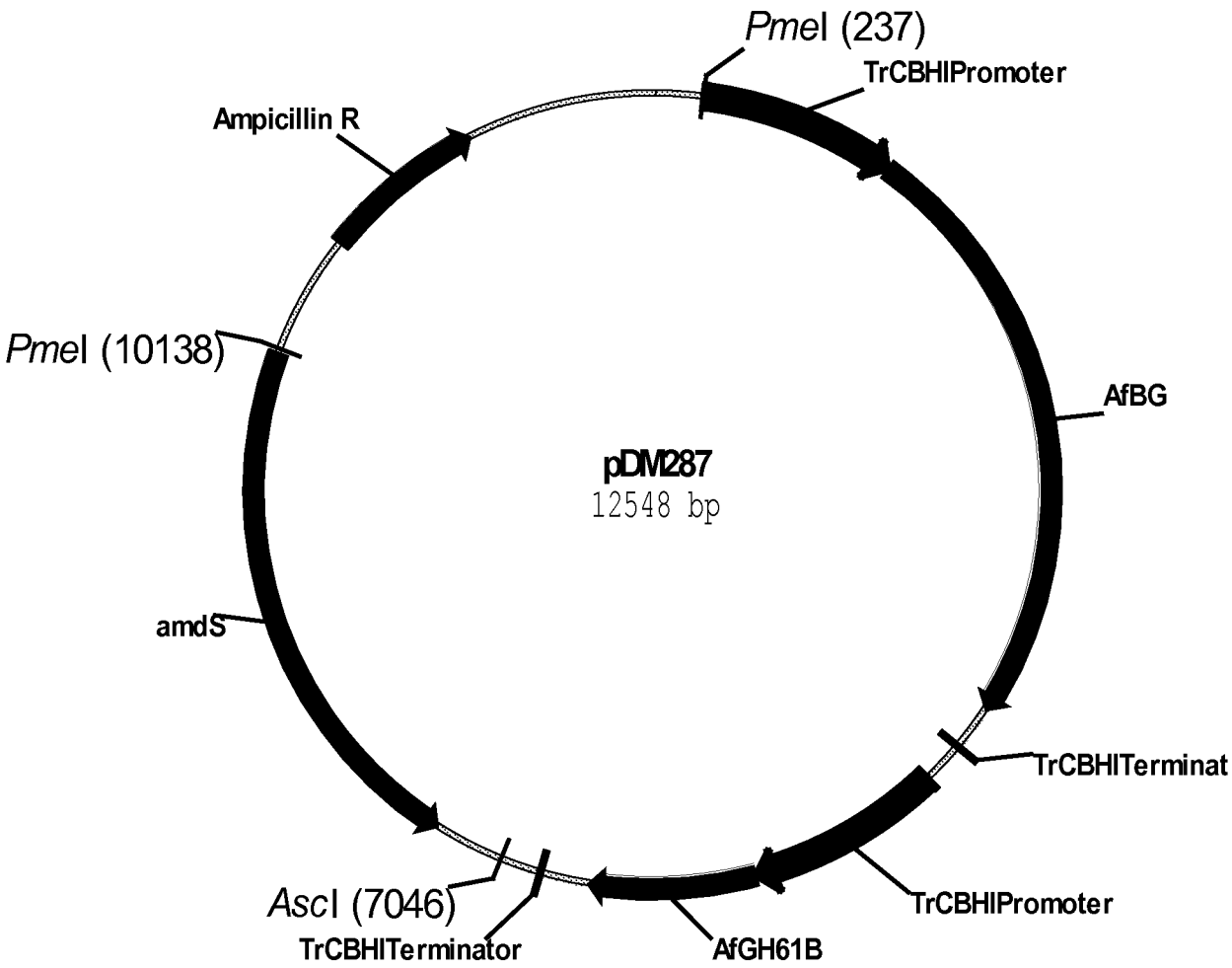
PmeI (237)
TrCBHIPromoter
Ampicillin R
PmeI (10138)
AfBG
pDM287
12548 bp
amdS
TrCBHITerminat
AscI (7046)
TrCBHITerminator
AfGH61B
TrCBHIPromoter

Fig. 3

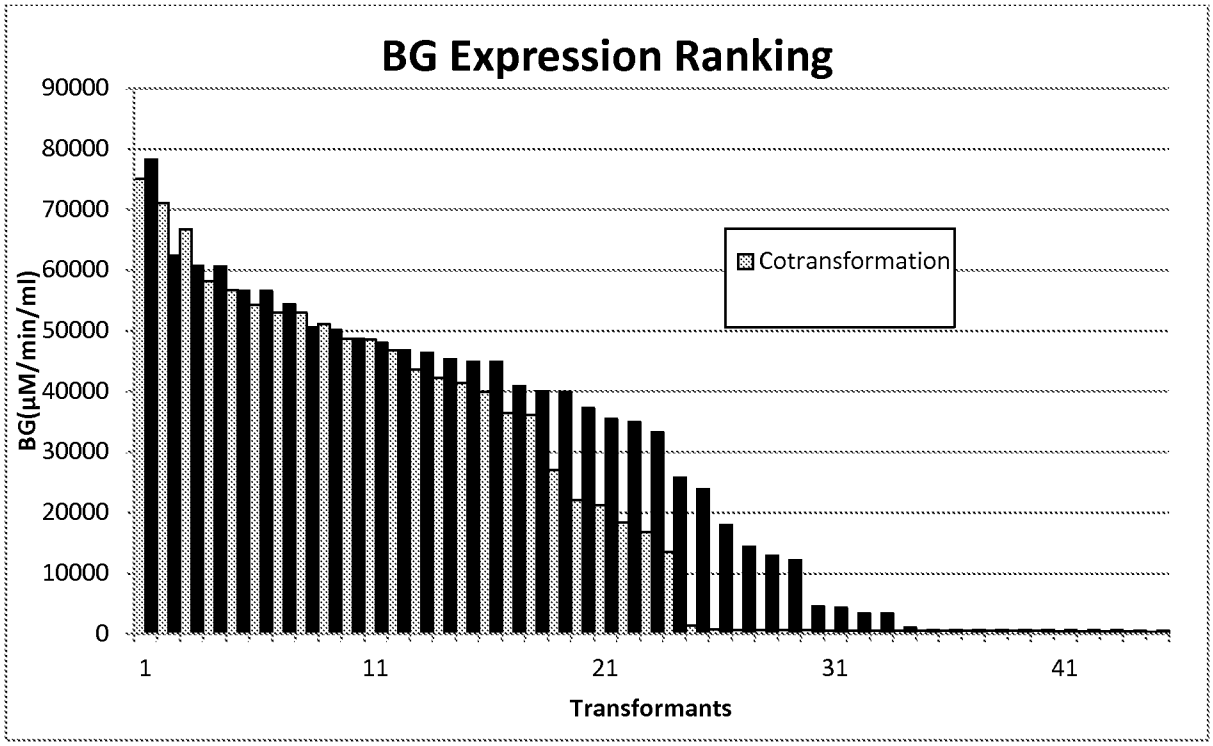
BG Expression Ranking
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
BG(μM/min/ml)
Cotransformation
1
11
21
31
41
Transformants

Fig. 4

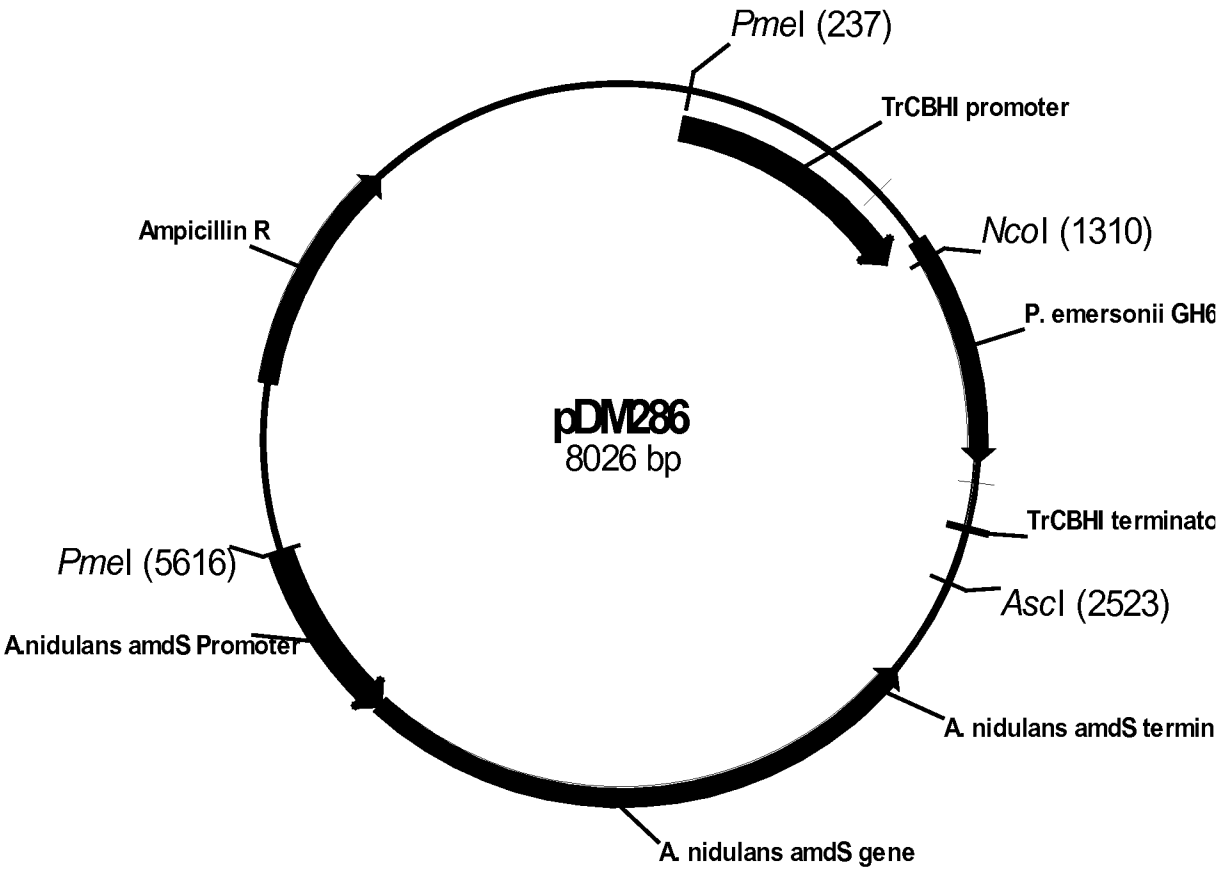
PmeI (237)
TrCBHI promoter
NcoI (1310)
P. emersonii GH6
TrCBHI terminato
AscI (2523)
A. nidulans amdS termin
A. nidulans amdS gene
A.nidulans amdS Promoter
PmeI (5616)
Ampicillin R
pDM286
8026 bp

Fig. 5

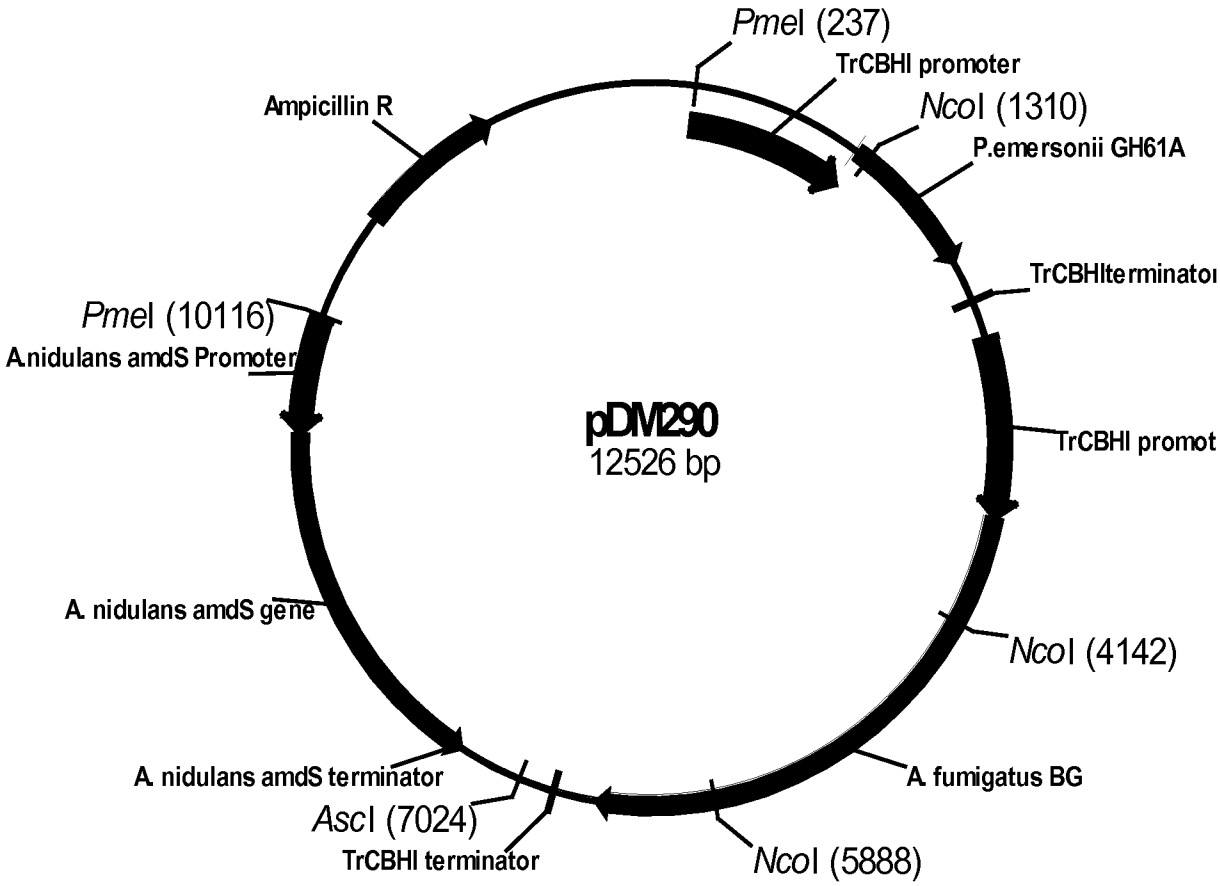
PmeI (237)
TrCBHI promoter
NcoI (1310)
P.emersonii GH61A
TrCBHIterminator
TrCBHI promot
NcoI (4142)
A. fumigatus BG
NcoI (5888)
TrCBHI terminator
AscI (7024)
A. nidulans amdS terminator
A. nidulans amdS gene
PmeI (10116)
A.nidulans amdS Promoter
Ampicillin R
pDM290
12526 bp

Fig. 6

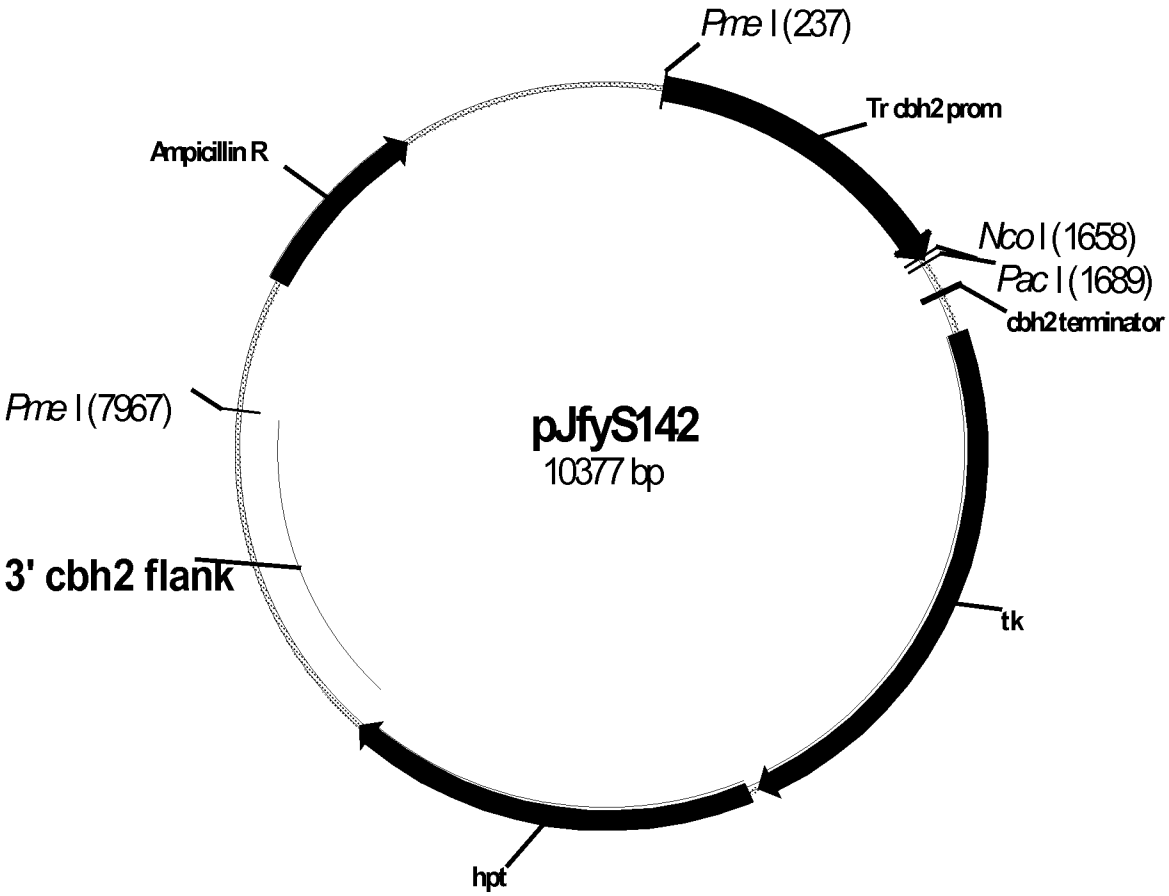
Pme I (237)
Tr cbh2 prom
Ampicillin R
Nco I (1658)
Pac I (1689)
cbh2 terminator
Pme I (7967)
pJfyS142
10377 bp
3' cbh2 flank
tk
hpt

Fig. 7

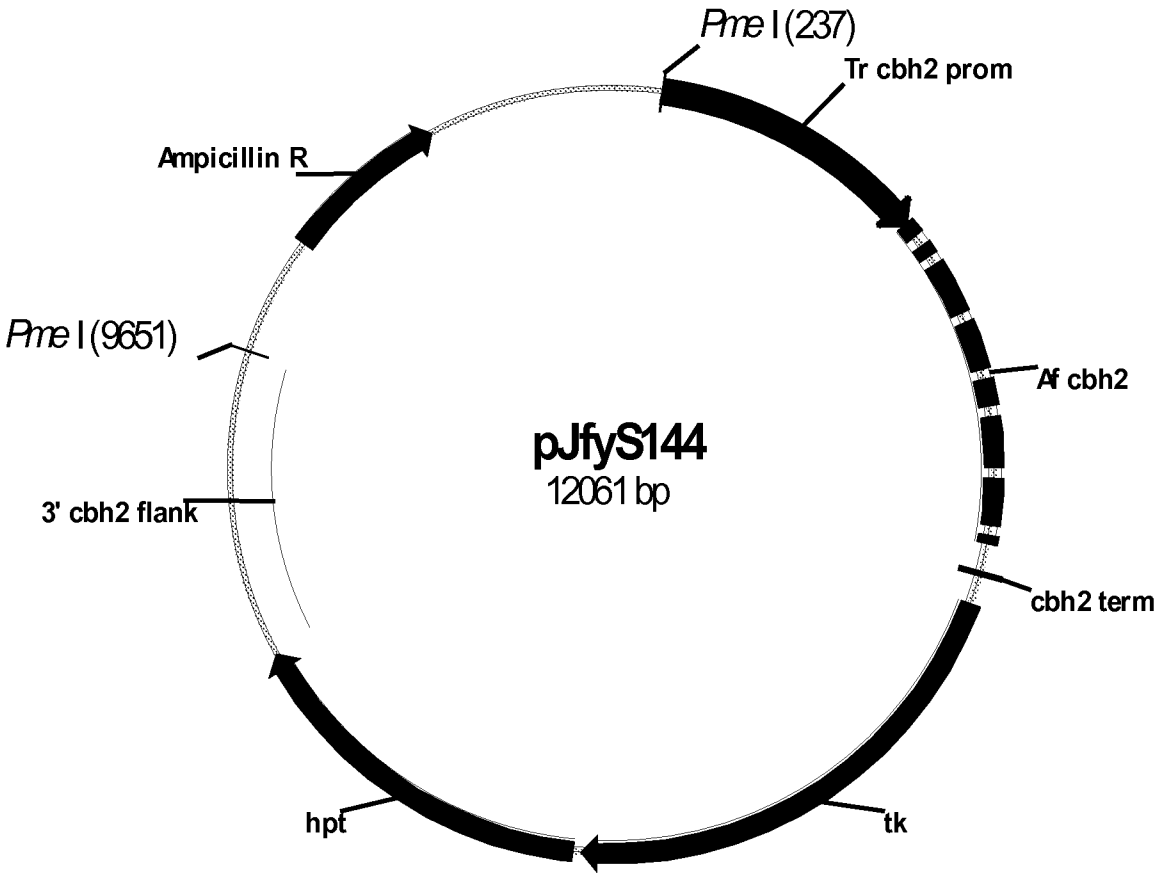
Pme I (237)
Tr cbh2 prom
Af cbh2
cbh2 term
tk
hpt
3' cbh2 flank
Pme I (9651)
Ampicillin R
pJfyS144
12061 bp

Fig. 8

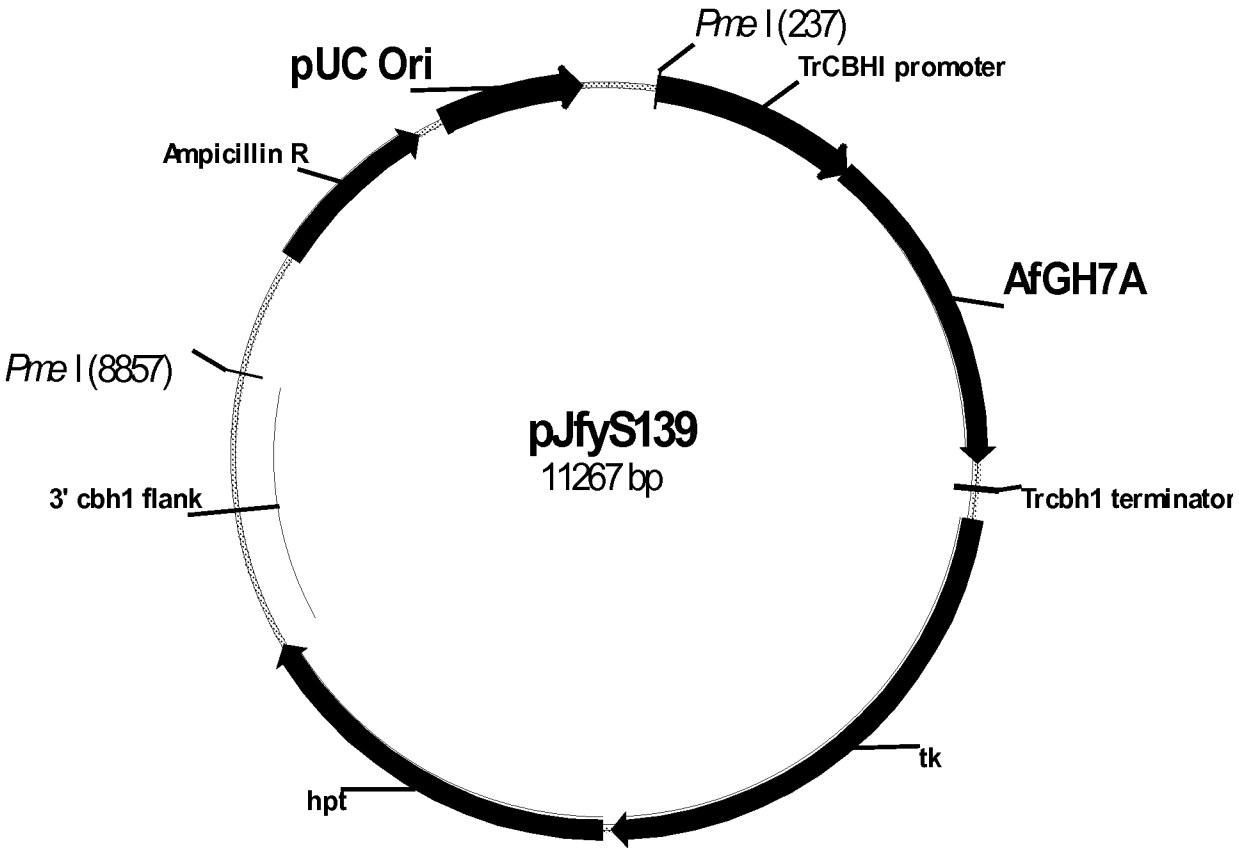
pUC Ori
Pme I (237)
TrCBHI promoter
Ampicillin R
AfGH7A
Pme I (8857)
pJfyS139
11267 bp
3' cbh1 flank
Trcbh1 terminator
tk
hpt

Fig. 9

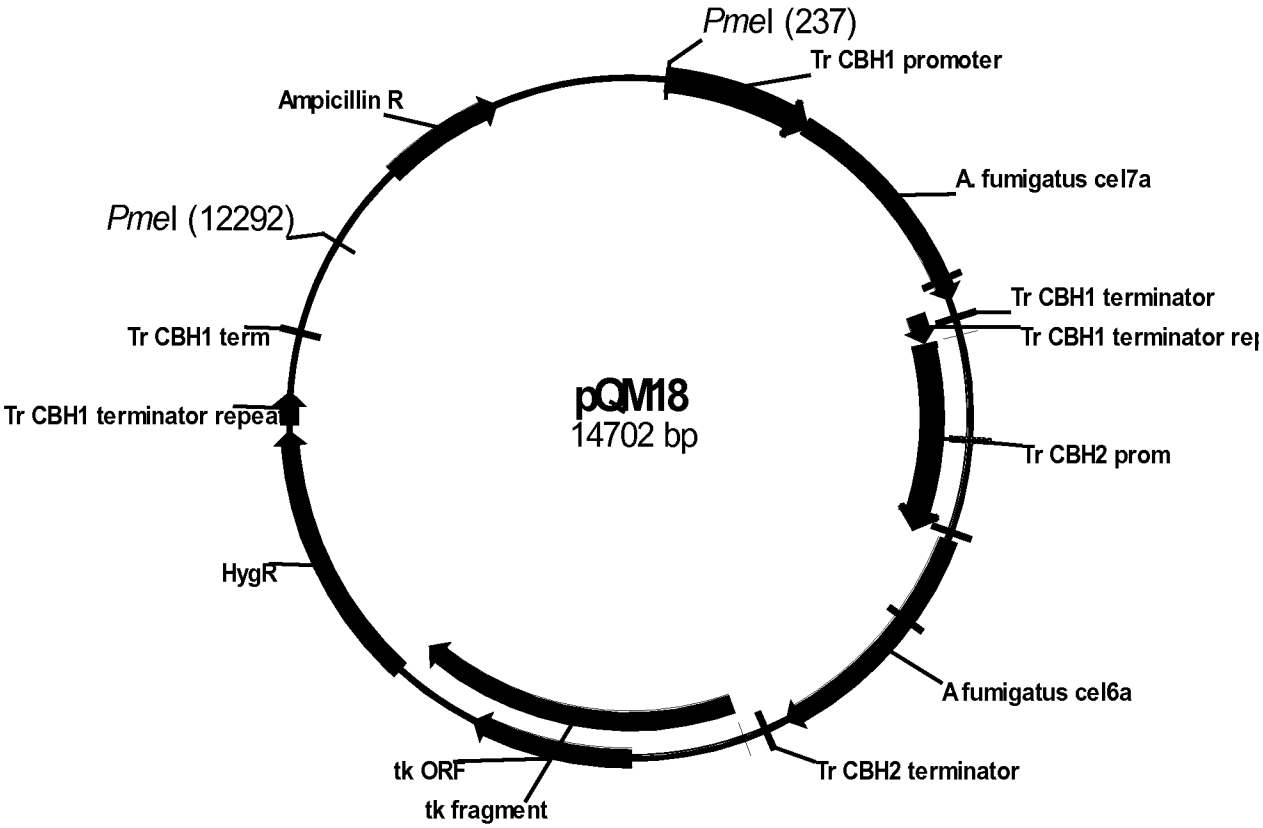
PmeI (237)
Tr CBH1 promoter
Ampicillin R
A. fumigatus cel7a
PmeI (12292)
Tr CBH1 terminator
Tr CBH1 terminator re
Tr CBH1 term
pQM18
14702 bp
Tr CBH1 terminator repeat
Tr CBH2 prom
HygR
A fumigatus cel6a
tk ORF
Tr CBH2 terminator
tk fragment

Fig. 10

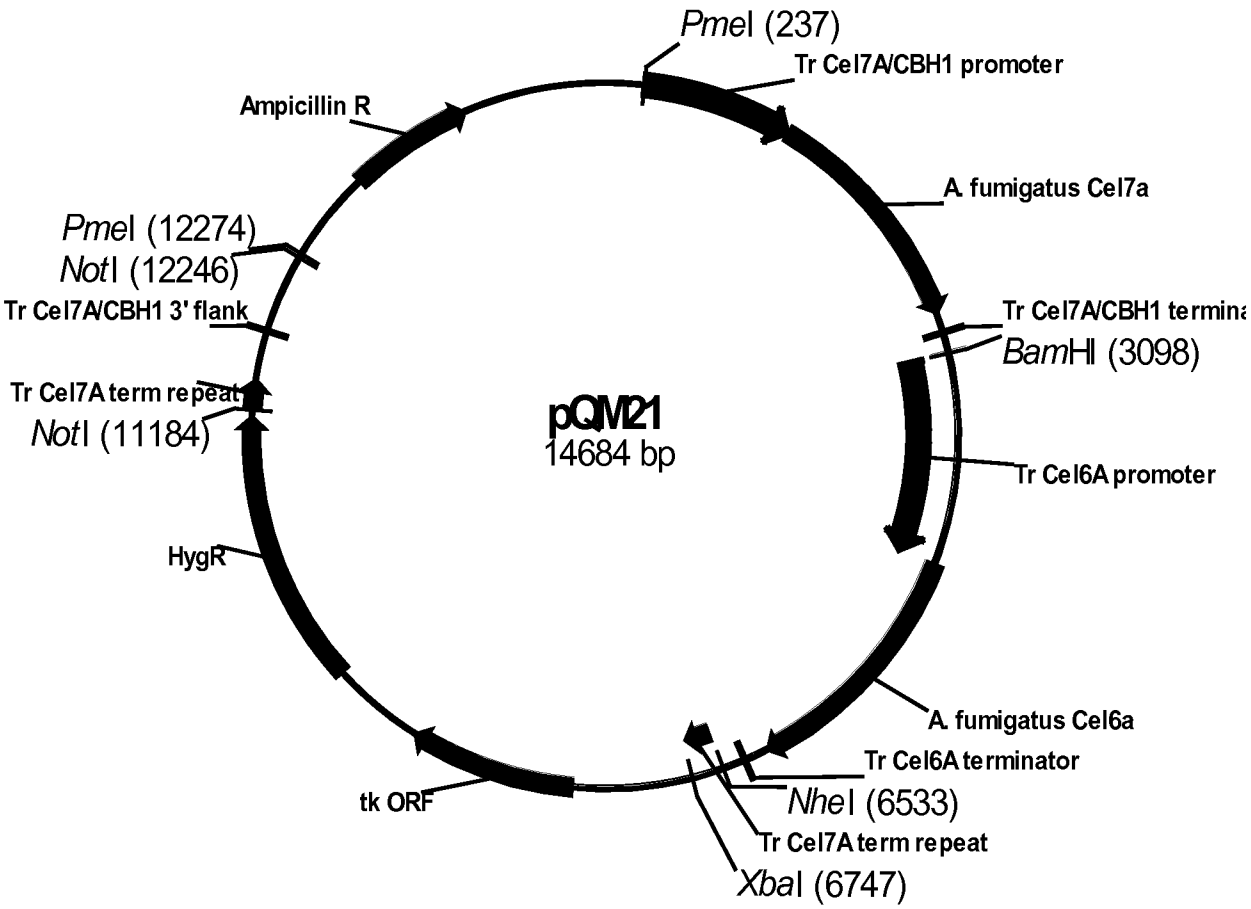
PmeI (237)
Tr Cel7A/CBH1 promoter
Ampicillin R
A. fumigatus Cel7a
PmeI (12274)
NotI (12246)
Tr Cel7A/CBH1 3' flank
Tr Cel7A/CBH1 termina
BamHI (3098)
Tr Cel7A term repeat
NotI (11184)
pQM21
14684 bp
Tr Cel6A promoter
HygR
A. fumigatus Cel6a
Tr Cel6A terminator
NheI (6533)
tk ORF
Tr Cel7A term repeat
XbaI (6747)

Fig. 11

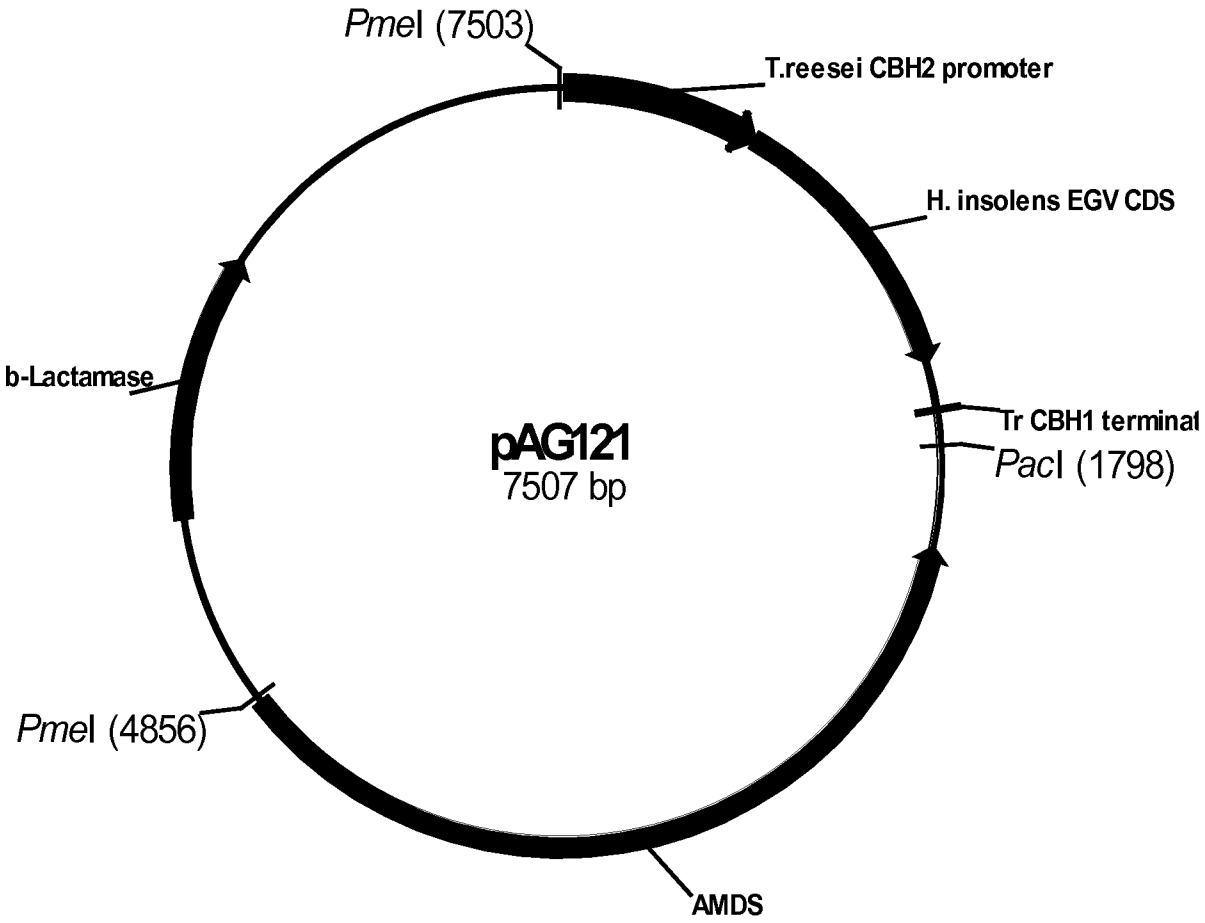
PmeI (7503)
T.reesei CBH2 promoter
H. insolens EGV CDS
Tr CBH1 terminat
PacI (1798)
AMDS
PmeI (4856)
b-Lactamase
pAG121
7507 bp

Fig. 12

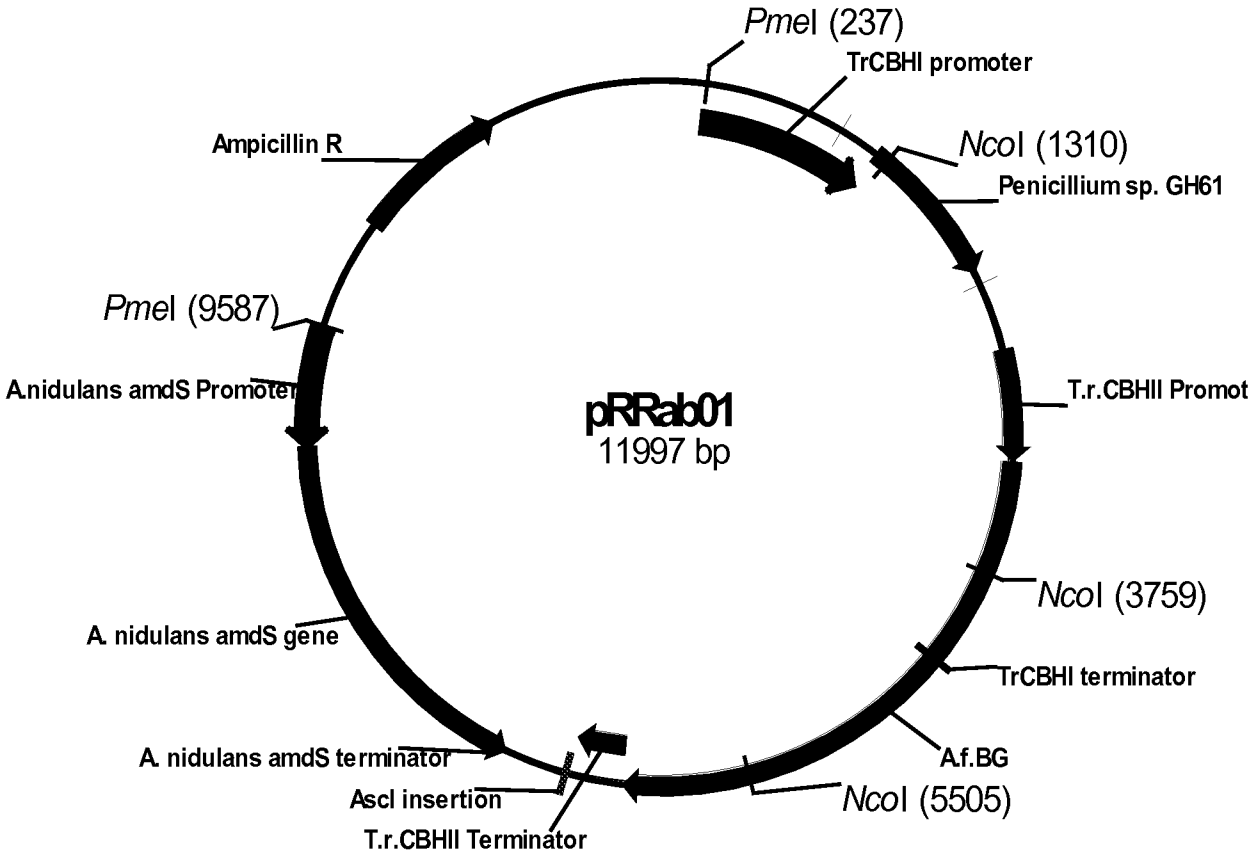
PmeI (237)
TrCBHI promoter
NcoI (1310)
Penicillium sp. GH61
T.r.CBHII Promot
NcoI (3759)
TrCBHI terminator
A.f.BG
NcoI (5505)
T.r.CBHII Terminator
AscI insertion
A. nidulans amdS terminator
A. nidulans amdS gene
A.nidulans amdS Promoter
PmeI (9587)
Ampicillin R
pRRab01
11997 bp

Fig. 13

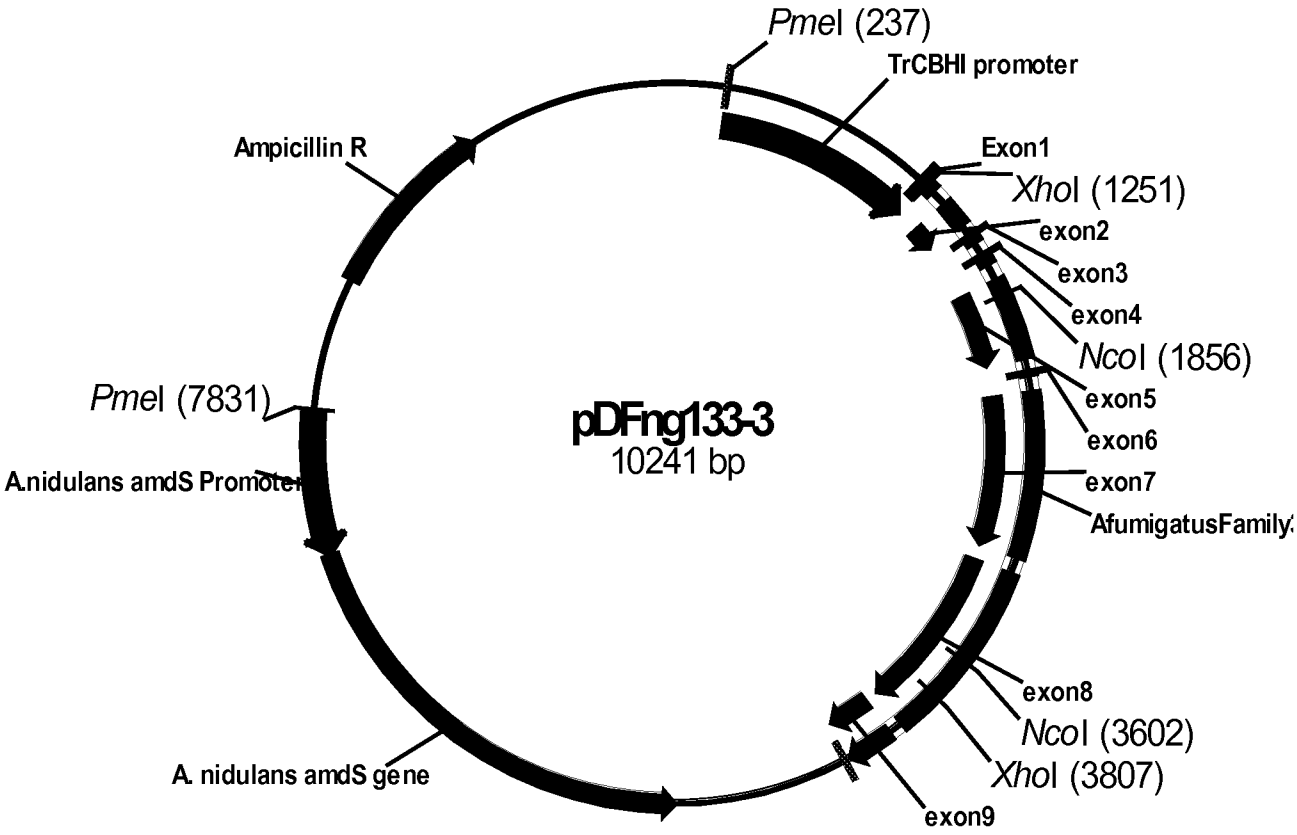
PmeI (237)
TrCBHI promoter
Exon1
XhoI (1251)
exon2
exon3
exon4
NcoI (1856)
exon5
exon6
exon7
AfumigatusFamily
exon8
NcoI (3602)
XhoI (3807)
exon9
A. nidulans amdS gene
PmeI (7831)
A.nidulans amdS Promoter
Ampicillin R
pDFng133-3
10241 bp

Fig. 14

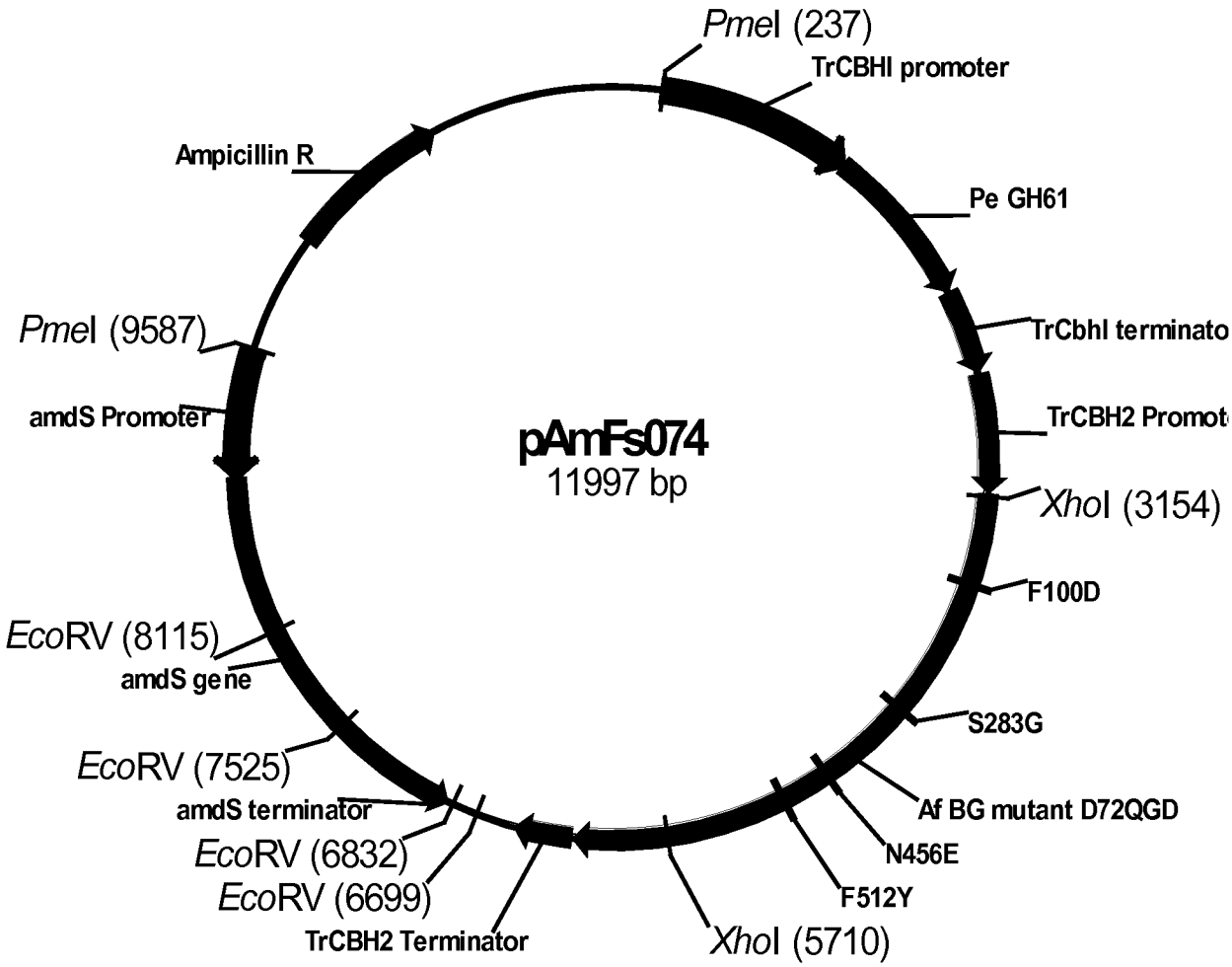
PmeI (237)
TrCBHI promoter
Pe GH61
TrCbhI terminato
TrCBH2 Promot
XhoI (3154)
F100D
S283G
Af BG mutant D72QGD
N456E
F512Y
XhoI (5710)
TrCBH2 Terminator
EcoRV (6699)
EcoRV (6832)
amdS terminator
EcoRV (7525)
amdS gene
EcoRV (8115)
amdS Promoter
PmeI (9587)
Ampicillin R
pAmFs074
11997 bp

Fig. 15

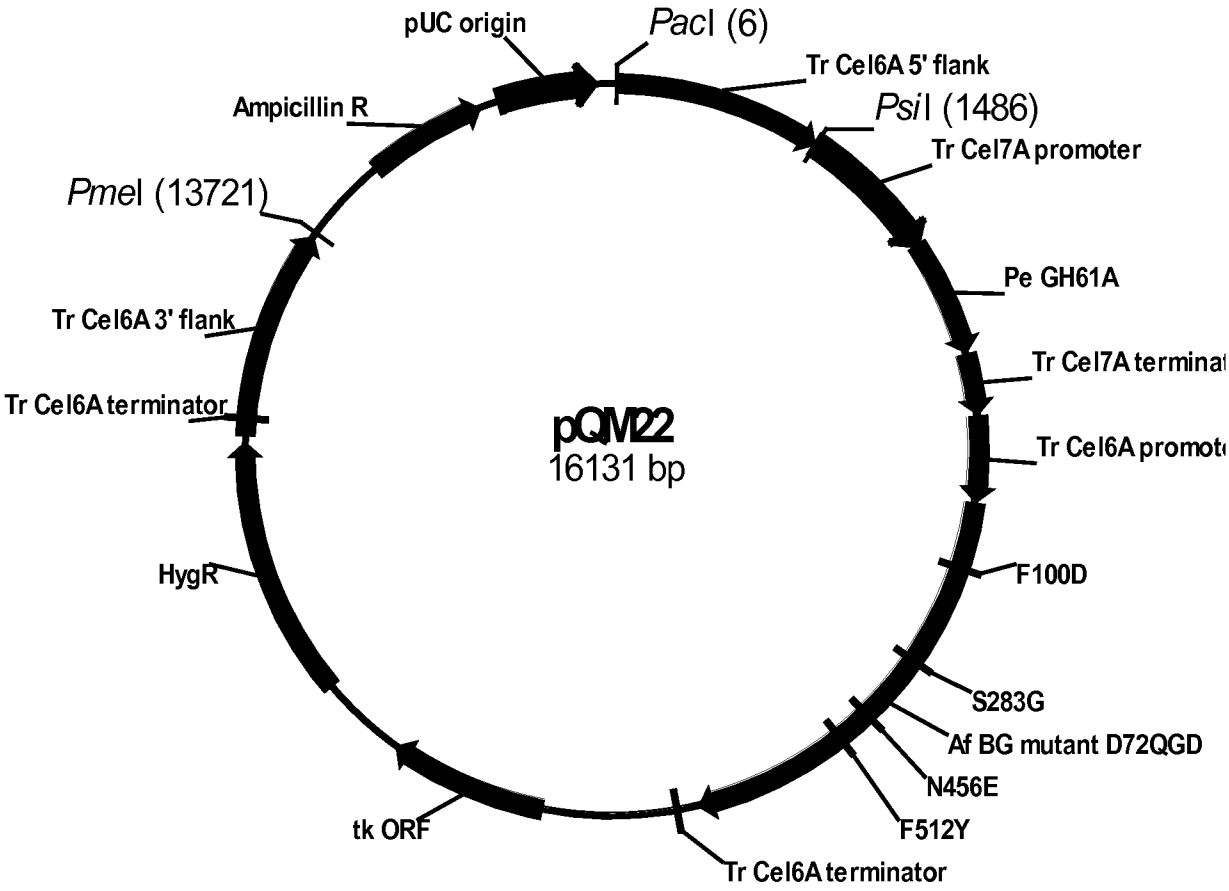
pUC origin
PacI (6)
Tr Cel6A 5' flank
Ampicillin R
PsiI (1486)
Tr Cel7A promoter
PmeI (13721)
Pe GH61A
Tr Cel6A 3' flank
Tr Cel7A terminat
Tr Cel6A terminator
pQM22
16131 bp
Tr Cel6A promot
HygR
F100D
S283G
Af BG mutant D72QGD
N456E
F512Y
tk ORF
Tr Cel6A terminator

Fig. 16