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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
06 February 2020 (06.02.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/028126 A1

(51) International Patent Classification:
C12N 9/42 (2006.01)

(21) International Application Number:
PCT/US2019/043348

(22) International Filing Date:
25 July 2019 (25.07.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/711,846 30 July 2018 (30.07.2018) US

(71) Applicant: **DANISCO US INC** [US/US]; 925 Page Mill Road, Palo Alto, California 94304 (US).

(72) Inventor: **PRATT II, Robert James**; 925 PAGE MILL ROAD, Palo Alto, California 94304 (US).

(74) Agent: **BRAZIL, Bill T.**; DANISCO US INC, 925 Page Mill Road, Palo Alto, California 94304 (US).

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

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

(54) Title: MUTANT AND GENETICALLY MODIFIED FILAMENTOUS FUNGAL STRAINS COMPRISING ENHANCED PROTEIN PRODUCTIVITY PHENOTYPES AND METHODS THEREOF

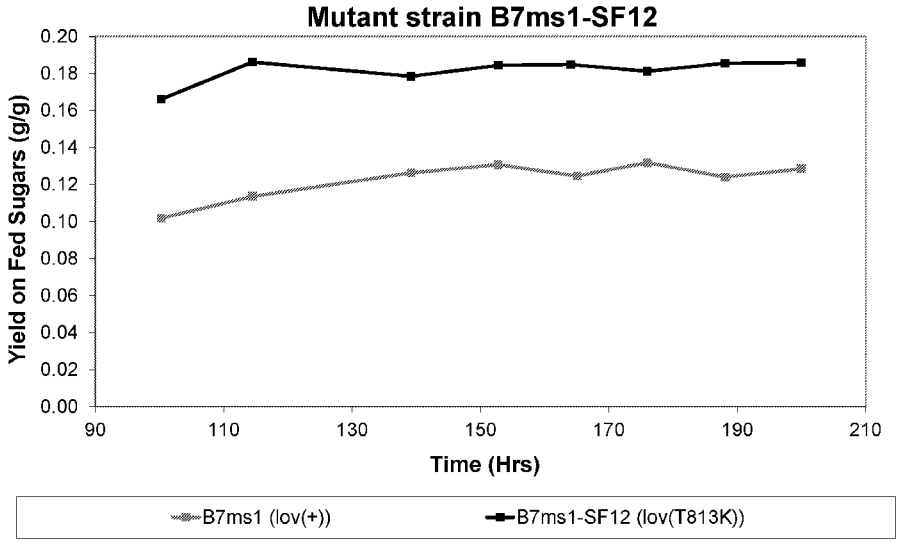

FIG. 1

(57) Abstract: The present strains and methods of the disclosure relate to genetic modifications in filamentous fungi that give rise to variant strains of filamentous fungi comprising enhanced protein productivity phenotypes. More specifically, as presented, described and exemplified herein, such variant strains of filamentous fungi comprising enhanced protein productivity phenotypes are well-suited for growth in submerged cultures, such as in large-scale production of proteins of interest for commercial applications.

[Continued on next page]

WO 2020/028126 A1

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

MUTANT AND GENETICALLY MODIFIED FILAMENTOUS FUNGAL STRAINS COMPRISING ENHANCED PROTEIN PRODUCTIVITY PHENOTYPES AND METHODS THEREOF

TECHNICAL FIELD

[0001] The present disclosure is generally related to the fields of biology, genetics, molecular biology, filamentous fungi, industrial protein production and the like. More particularly, the present strains and methods of the disclosure relate to genetic modifications in filamentous fungi that give rise to variant strains of filamentous fungi comprising enhanced protein productivity phenotypes. More specifically, as presented, described and exemplified herein, such variant strains of filamentous fungi comprising enhanced protein productivity phenotypes are well-suited for growth in submerged cultures, such as in large-scale production of proteins of interest for commercial applications.

CROSS-REFERENCE TO RELATED APPLICATIONS

[0002] The present application claims priority to U.S. Provisional Patent Application Serial No. 62/711,846, filed July 30, 2018, which is hereby incorporated by reference in its entirety.

REFERENCE TO A SEQUENCE LISTING

[0003] The contents of the electronic submission of the text file Sequence Listing, named "NB41504-WO-PCT_SequenceListing.txt" was created on July 24, 2019 and is 160 KB in size, which is hereby incorporated by reference in its entirety.

BACKGROUND

[0004] Filamentous fungi (*e.g., Aspergillus sp., Penicillium sp., Talaromyces sp., Fusarium sp., Myceliophthora sp., Neurospora sp., Trichoderma sp.* and the like) are capable of expressing native (endogenous) and heterologous proteins to high levels, making them well-suited for the large-scale production of proteins (*e.g.*, enzymes) and/or metabolites for industrial and/or commercial applications such as pharmaceutical applications, animal health applications, food applications, beverage applications and the like. Filamentous fungi are typically grown in mycelial submerged cultures in bioreactors (fermentors), which bioreactors are adapted to introduce and distribute oxygen and nutrients into the culture medium (*i.e.*, culture broth). For example, the filamentous fungus *Trichoderma reesei* (*T. reesei*; an anamorph of the fungus *Hypocrea jecorina*) is known to be an efficient producer of cellulase enzymes (*e.g.*, see PCT International Publication Nos. W01998/15619, WO2005/028636, WO2006/074005, WO 1992/06221, WO 1992/06209, WO1992/06183 and WO2002/12465).

[0005] As such, filamentous fungi have been utilized for their ability to produce proteins which are valuable in the production of commodities such as cellulosic (derived) ethanol, textile processing, grain processing, detergents, fibers, food additives, feed additives and the like. For example, recombinant gene expression in such fungal host strains is a common method for the production of proteins (*i.e.*, for industrial and commercial purposes) and as such, protein productivity improvements of a fungal host strain are an

important economic factor of protein production costs. Thus, as appreciated by one of skill in the art, such compositions and methods for enhancing protein production in filamentous fungal strains are of significant commercial interest.

[0006] As described herein, the instant disclosure, which is generally related to genetically modified filamentous fungal strains comprising enhanced/increased protein productivity phenotypes, addresses such ongoing and unmet needs in the art.

SUMMARY

[0007] Described herein are strains, cells, methods, constructs and the like relating to filamentous fungi having increased protein productivity phenotypes. Thus, certain embodiments of the disclosure are related to modified Ascomycota cells derived from parental cells, wherein the modified cells comprise a polynucleotide sequence encoding a variant LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4, wherein the modified cells comprises an enhanced protein productivity phenotype relative to the parental cells, for example, when fermented/cultivated under the same conditions. In certain embodiments, the parental cells comprise a wild-type polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0008] In related embodiments, the fungal cells further comprise a heterologous polynucleotide encoding a protein of interest (POI). In other embodiments, the cells further comprising a polynucleotide encoding a NIK1(M743T) protein of SEQ ID NO: 19. In certain other embodiments, the modified cells further comprising at least one genetic modification which deletes, disrupts or reduces the expression/production of a protein selected from the group consisting of MPG1 SFB3, SEB1, CRZ1, TSP2, SSB7 and GAS1.

[0009] In other embodiments, the Ascomycota cell is selected from a *Trichoderma sp., Aspergillus sp., Fusarium sp., Penicillium sp., a Candida sp., Chrysosporium sp., Cephalosporium sp., Talaromyces sp., Neurospora sp.* and *Myceliophthora sp.*

[0010] Thus, in certain embodiments, a parental *Aspergillus sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 11 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0011] In another embodiment a parental *Penicillium sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 12 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2. In other embodiments, a parental *Talaromyces sp.* cell comprises a polynucleotide

sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 13 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0012] In yet another embodiment, a parental *Fusarium sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 14 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0013] In certain other embodiments, a parental *Myceliophthora sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 15 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0014] In other embodiments, a parental *Neurospora sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 16 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0015] In another embodiment, a parental *Candida sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 17 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0016] In certain other embodiments, a parental *Trichoderma sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 2 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0017] In other embodiments, a protein of interest is produced by a modified cell of the disclosure and purified therefrom.

[0018] In yet other embodiments, the disclosure is related to modified Ascomycota cells derived from parental cells comprising a wild-type polynucleotide sequence which hybridizes under stringent hybridization conditions with a nucleic acid sequence encoding a native LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2, wherein the modified cells comprise a modified polynucleotide sequence which hybridizes under stringent hybridization conditions with a nucleic acid sequence encoding a LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence corresponding to position 813 of SEQ ID NO: 4. In certain embodiments, the cells further comprise a heterologous polynucleotide encoding a protein of interest. In another embodiment, the cells further comprise a polynucleotide construct encoding a NIK1(M743T) protein of SEQ ID NO: 19. In other embodiments, the

cells further comprise at least one genetic modification which deletes, disrupts or reduces the expression/production of a protein selected from the group consisting of MPG1 SFB3, SEB1, CRZ1, TSP2, SSB7 and GAS1.

[0019] In certain other embodiments, the Ascomycota cell is selected from a *Trichoderma sp., Aspergillus sp., Fusarium sp., Penicillium sp., Candida sp., Chrysosporium sp., Cephalosporium sp., Talaromyces sp., Neurospora sp.* and *Myceliophthora sp.*

[0020] In certain embodiments, the parental *Aspergillus sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 11 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0021] In other embodiments, the parental *Penicillium sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 12 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0022] In another embodiment, the parental *Talaromyces sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 13 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0023] In certain other embodiments, the parental *Fusarium sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprise SEQ ID NO: 14 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0024] In another embodiment, the parental *Myceliophthora sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 15 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0025] In yet other embodiments, the parental *Neurospora sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 16 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0026] In certain other embodiments, the parental *Candida sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 17 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

[0027] In other embodiments, the parental *Trichoderma sp.* cells comprise a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 2 or SEQ ID NO: 18 and comprise a threonine (T) residue at an amino acid sequence position corresponding to position

813 of SEQ ID NO: 2. Other embodiments are related to a protein of interest produced by a modified cell of the disclosure.

[0028] Certain other embodiments of the disclosure are related to a vector comprising a polynucleotide encoding a variant LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4.

[0029] In other embodiments, the disclosure is directed to a polynucleotide encoding a variant LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4.

[0030] In another embodiment, the disclosure is related to a polynucleotide encoding a variant LOV protein, wherein the polynucleotide encoding the variant protein hybridizes under stringent hybridization conditions with a nucleic acid sequence encoding a LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4.

[0031] In other embodiments, the disclosure is directed to a mutant Trichoderma strain derived from a parental strain, wherein the mutant strain comprises a gene encoding a LOV variant protein comprising at least 50% sequence identity to SEQ ID NO: 4 or SEQ ID NO: 18 and comprising a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 4. In certain embodiments, the mutant strain comprises an enhanced protein productivity phenotype relative to the parental strain. In other embodiments, the mutant strain further comprising a polynucleotide construct encoding a NIK1(M743T) protein of SEQ ID NO: 19. In other embodiments, the mutant strain further comprises a genetic modification which deletes, disrupts or reduces the expression/production of a gene encoding at least one protein selected from the group consisting of MPG1 SFB3, SEB1, CRZ1, TSP2, SSB7 and GAS1. Thus, certain related embodiments are directed to a protein of interest produced by the mutant strain.

[0032] Other embodiments of the disclosure are related to methods for constructing modified Ascomycota cells comprising enhanced protein productivity phenotypes, the method comprising (i) obtaining a parental cell comprising a wild-type *lov* gene which hybridizes under stringent hybridizations conditions with a nucleic acid sequence encoding a LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18, (ii) modifying the parental cell of step (i) to produce a modified cell comprising a gene encoding a LOV (variant) protein comprising a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 2, and (iii) isolating the modified cell of step (ii), wherein the modified cell comprises an enhanced protein productivity phenotype relative to the parental cell.

[0033] In other embodiments, the disclosure is directed to a method for constructing modified Ascomycota cells comprising an enhanced protein productivity phenotype, the method comprising: (i) obtaining a parental Ascomycota cell and introducing into the cell a polynucleotide construct encoding a LOV variant protein comprising a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 4 and (ii) isolating the modified cell of step (i), wherein the modified cell comprises an enhanced protein productivity phenotype relative to the parental cell.

[0034] Thus, as set forth and described herein, the various embodiments of present disclosure are generally related to variant strains of filamentous fungi comprising enhanced protein productivity phenotypes which variant strains are well-suited for growth in submerged cultures, such as in large-scale production of proteins of interest for commercial applications.

BRIEF DESCRIPTION OF THE BIOLOGICAL SEQUENCES

[0035] **SEQ ID NO: 1** is a nucleic acid sequence of a wild-type *Trichoderma sp. lov* gene encoding a (native) LOV protein of SEQ ID NO: 2.

[0036] **SEQ ID NO: 2** is the amino acid sequence of the (native) LOV protein encoded by SEQ ID NO: 1.

[0037] **SEQ ID NO: 3** is a nucleic acid sequence of a *Trichoderma sp.* mutant allele named "*lov*(T813K)" encoding a (variant) T813K (substituted) LOV protein of SEQ ID NO: 4.

[0038] **SEQ ID NO: 4** is the amino acid sequence of the (variant) T813K (substituted) LOV protein encoded by SEQ ID NO: 3.

[0039] **SEQ ID NO: 5** is the amino acid sequence of a *Trichoderma sp.* MPG1 protein.

[0040] **SEQ ID NO: 6** is the amino acid sequence of a *Trichoderma sp.* SEB1 protein.

[0041] **SEQ ID NO: 7** is the amino acid sequence of a *Trichoderma sp.* SFB3 protein.

[0042] **SEQ ID NO: 8** is the amino acid sequence of a *Trichoderma sp.* CRZ1 protein.

[0043] **SEQ ID NO: 9** is the amino acid sequence of a *Trichoderma sp.* GAS1 protein.

[0044] **SEQ ID NO: 10** is the amino acid sequence of a *Trichoderma sp.* TPS2 protein.

[0045] **SEQ ID NO: 11** is the amino acid sequence of an *Aspergillus sp.* LOV protein orthologue.

[0046] **SEQ ID NO: 12** is the amino acid sequence of a *Penicillium sp.* LOV protein orthologue.

[0047] **SEQ ID NO: 13** is the amino acid sequence of a *Talaromyces sp.* LOV protein orthologue.

[0048] **SEQ ID NO: 14** is the amino acid sequence of a *Fusarium sp.* LOV protein orthologue.

[0049] **SEQ ID NO: 15** is the amino acid sequence of a *Myceliophthora sp.* LOV protein orthologue.

[0050] **SEQ ID NO: 16** is the amino acid sequence of a *Neurospora sp.* LOV protein orthologue.

[0051] **SEQ ID NO: 17** is the amino acid sequence of a *Candida sp.* LOV protein orthologue.

[0052] **SEQ ID NO: 18** comprises the C-terminal amino acid residue positions 500-894 of the *Trichoderma sp.* (native) LOV protein (SEQ ID NO: 2).

[0053] **SEQ ID NO: 19** is the amino acid sequence of a variant *Trichoderma* histidine kinase (NIK1) comprising a methionine (M) to threonine (T) substitution at amino acid (residue) position 743 of SEQ ID

NO: 19. The gene encoding the variant NIK1 histidine kinase of SEQ ID NO: 19 has been named "*nik1*(M743T)".

[0054] SEQ ID NO: 20 is the nucleic acid sequence of a wild-type *T. reesei ssb7* gene encoding a native SSB7 protein of SEQ ID NO: 21.

[0055] SEQ ID NO: 21 is the amino acid sequence of the native *T. reesei* SSB7 protein encoded by SEQ ID NO: 20.

[0056] SEQ ID NO: 22 is the nucleic acid sequence of allele *ssb7(fs)*, comprising a deletion of G (ΔG) in exon 2, resulting in a frame-shift (*fs*) mutation, and a premature stop codon prior to the last intron of the *ssb7* gene.

[0057] SEQ ID NO: 23 is the amino acid sequence of the variant SSB7 protein encoded by allele *ssb7(fs)* of SEQ ID NO: 22.

[0058] SEQ ID NO: 23 is the amino acid sequence of a *Trichoderma harzianum* LOV protein orthologue.

BRIEF DESCRIPTION OF THE DRAWINGS

[0059] Figure 1 shows a comparison of total protein yield on fed sugars for the glucoamylase expressing B7ms1 (parental) strain (**FIG. 1**, grey data points/grey line) and mutant derived (daughter) B7ms1-SF12 strain (**FIG. 1**, black data points/black line). Thus, these glucoamylase (GA) expressing strains (comprising the viscosity reducing mutations in the *mpg1* and *seb1* genes) were evaluated for protein productivity in fermentors. The figure legend shows the strain name (see TABLE 1 for genotypes) followed by the *lov* allele in that strain in parenthesis for each line type. As shown in **FIG. 1**, the B7ms1-SF12 (daughter) strain (**FIG. 1**, black line; *i.e.*, comprising the mutant *lov*(T813K) allele)) has an enhanced protein productivity phenotype relative to the B7ms1 (parental) strain (**FIG. 1**, grey line; *i.e.*, comprising the wild-type *lov* gene). For example, as presented in **FIG. 1**, protein yield on fed sugars increased 44% when the *lov*(T813K) mutation was present.

[0060] Figure 2 is a graphical representation illustrating amino acid conservation of the LOV protein. More specifically, **FIG. 2** is a graphical representation of a Geneious multiple sequence alignment of six-hundred and ninety-one (691) Pezizomycotina homologs. At the bottom of **FIG. 2** are boxes representing the amino acid sequence of the Trichoderma LOV protein, wherein amino acid (residues) are presented in black shaded boxes if conserved in greater than 99% of the aligned sequences (or are grey otherwise). LOV residue numbers (SEQ ID NO: 2) are presented above the sequence representation. As presented in **FIG. 2**, amino acid sequence gaps present in Trichoderma LOV or otherwise in >2% of sequences in the multiple sequence alignment, are presented as grey shaded lines. Annotations in the public GenBank sequence for LOV (XP_006967324.1), are shown below the sequence representation as grey boxes. Putative active site residues, as annotated in the GenBank entry for LOV, are shown below the sequence representation as black boxes. The threonine (T) 813 residue substituted for a lysine (K) in strain B7ms1-SF12 is annotated with a light grey box. The mean hydrophobicity and isoelectric point ($_{P}$I) are also plotted and presented in **FIG. 2**. The amino acid identity in the alignment is plotted just above the sequence. The

lighter grey bars represent residues identical in at least 30% of sequences in the alignment and darker grey bars less than 30%.

[0061] **Figure 3** presents a comparison of total protein yield on fed sugars in fermentors for the whole cellulase producing T4ms strain (grey lines) and mutant derivative T4mls comprising an engineered *lov*(T813K) allele (black lines). More particularly, **FIG. 3** presents total protein yield on fed sugars for these strains under both lower cell density (LCD, dashed lines, circles) and higher cell density (HCD, solid lines, squares) fermentation conditions. The figure legend shows the fermentation condition (LCD vs. HCD) followed by strain name (see TABLE 1 for genotypes) and the *lov* allele present is shown in parenthesis. Thus, as presented in **FIG. 3**, the T4mls strain (**FIG. 3**, black lines; *i.e.*, comprising the mutant *lov*(T813K) allele)) has an enhanced protein productivity phenotype relative to the T4ms strain (**FIG. 3**, grey lines; *i.e.*, comprising the wild-type *lov* gene) under both LCD (dashed lines, circles) and HCD (solid lines, squares) fermentation conditions. For example, as presented in **FIG. 3**, protein yield on fed sugars increased 42% (LCD) and 32% (HCD) when the *lov*(T813K) mutation was present.

[0062] **Figure 4** presents a comparison of total protein yield on fed sugars in fermentors for the whole cellulase producing strains T4m (solid lines, squares) and T4s (dashed lines, circles) with the wild-type *lov* allele (grey lines) and their mutant derivatives T4ml+ (solid lines, squares) and T4sl+ (dashed lines, circles), respectively, comprising an engineered *lov*(dis) allele (black lines). The figure legend shows the strain name (see TABLE 1 for genotypes) followed by the *lov* allele in that strain shown in parenthesis for each line type. As presented in **FIG. 4**, the T4ml+ and T4sl+ (daughter) strains (**FIG. 4**, black lines; *i.e.*, comprising the mutant *lov*(dis) allele)) do not have an enhanced protein productivity phenotype relative to their T4m and T4s (parental) strains (**FIG. 4**, grey lines; *i.e.*, comprising the wild-type *lov* gene).

[0063] **Figure 5** presents a comparison of relative total protein titer in shake flask fermentations of strains with various marker insertion sites with either the *lov*(T813K) allele, the *lov*(dis) allele, or the wild-type *lov*(+) allele. More particularly, whole cellulase producing strains were evaluated in shake flasks for total protein titer relative to T4 strains and T4m strains run in parallel. Strains of various marker insertion sites were compared in the fermentations with either the wild-type *lov*(+) allele (**FIG. 5**, lightest grey bars), the mutant *lov*(T813K) allele (**FIG. 5**, medium grey bars), or the *lov*(dis) allele (**FIG. 5**, darkest grey bars). Thus, in all marker insertion sites evaluated, the total protein titers increased when the *lov*(T813K) allele was present. However, insertion of either the *pyr4* (**FIG. 5**, strain "41G") or *pyr2* marker (strains T4 and T4m) in the *lov* gene itself (the *lov*(dis) allele) showed no significant improvement in relative titers.

[0064] **Figure 6** shows a comparison of total protein yield on fed sugars for the whole cellulase producing T4mp strain (**FIG. 6**, grey data points/grey line) and *lov*(T813K) mutant T4mlp strain (**FIG. 6**, black data points/black line). Thus, these whole cellulase expressing strains were evaluated for protein productivity in fermentors. The figure legend shows the strain name (see TABLE 1 for genotypes) followed by the *lov* allele in that strain in parenthesis for each line type. As shown in **FIG. 6**, the T4mlp strain (**FIG. 6**, black line; *i.e.*, comprising the mutant *lov*(T813K) allele)) has an enhanced protein productivity phenotype relative to the T4mp (parental) strain (**FIG. 6**, grey line; *i.e.*, comprising the wild-type *lov* gene). For

example, as presented in **FIG. 6**, protein yield on fed sugars increased 28% when the *lov*(T813K) mutation was present.

[0065] **Figure 7** presents a Clustal W (1.83) multiple sequence alignment of the *Trichoderma sp.* (native) LOV protein (SEQ ID NO: 2), aligned with seven (7) different Ascomycota LOV orthologues. For example, the eight (8) protein sequences used in the Clustal alignment are shown in **FIG. 7A-7B** (SEQ ID NO: 2 and SEQ ID NOs: 11-17) and the Clustal alignment of the same is shown in **FIG. 7C-7F**. More particularly, as presented in **FIG. 7C-7F**, the *Trichoderma sp.* (native) LOV protein (SEQ ID NO: 2; labeled "2", shown in bold **CAPITAL** residues) is aligned with an *Aspergillus sp.* LOV protein orthologue (SEQ ID NO:11, labeled "11"), a *Penicillium sp.* LOV protein orthologue (SEQ ID NO:12, labeled "12"), a *Talaromyces sp.* LOV protein orthologue (SEQ ID NO:13, labeled "13"), a *Fusarium sp.* LOV protein orthologue (SEQ ID NO:14, labeled "14"), a *Myceliophthora sp.* LOV protein orthologue (SEQ ID NO:15, labeled "15"), a *Neurospora sp.* LOV protein orthologue (SEQ ID NO:16, labeled "16") and a *Candida sp.* LOV protein orthologue (SEQ ID NO:17, labeled "17"). As shown in **FIG. 7C-7F** (*i.e.*, below the aligned amino residues), an asterisk "*" indicates positions which have a single, fully conserved residue; a colon ":" indicates conservation between groups of strongly similar properties (*i.e.*, scoring > 0.5 in the Gonnet PAM 250 matrix) and a period "." indicates conservation between groups of weakly similar properties (*i.e.*, scoring < 0.5 in the Gonnet PAM 250 matrix). The highly conserved threonine (T) amino acid is indicated in **FIG. 7F**, as a bold underlined **T** residue.

[0066] **Figure 8** presents amino acid sequence alignments performed *via* BLAST protein alignment (NCBI; Blastp suite), using the C-terminal residue positions 500-894 (SEQ ID NO: 18) of the (native) *Trichoderma sp.* LOV protein sequence (SEQ ID NO: 2). For example, as presented in **FIG. 8A-8G**, SEQ ID NO: 18 comprises a C-terminal amino acid sequence of the *Trichoderma sp.* (native) LOV protein (*i.e.*, comprising 394 amino acid residue positions, which correspond to amino acid residue positions 500-894 of SEQ ID NO: 2), wherein the highly conserved threonine (T) amino acid at residue at position 813 (T_{813}) is indicated with a bold, underlined **T**. Thus, this 394 residue C-terminal amino acid sequence (SEQ ID NO: 18) was aligned with Ascomycota LOV protein orthologues (*e.g.*, SEQ ID NOs: 11-17; **FIG. 8A-8G**). As shown in **FIG. 8A**, the *Aspergillus sp.* LOV protein orthologue (SEQ ID NO: 11) comprises about 50% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ ID NO: 2 is bold and underlined (**FIG. 8A**), the *Penicillium sp.* LOV protein orthologue (SEQ ID NO: 12) comprises about 50% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ ID NO: 2 is bold and underlined (**FIG. 8B**), the *Talaromyces sp.* LOV protein orthologue (SEQ ID NO: 13) comprises about 60% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ ID NO: 2 is bold and underlined (**FIG. 8C**), the *Fusarium sp.* LOV protein orthologue (SEQ ID NO: 14) comprises about 88% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ

ID NO: 2 is bold and underlined (**FIG. 8D**), the *Myceliophthora sp.* LOV protein orthologue (SEQ ID NO: 15) comprises about 90% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ ID NO: 2 is bold and underlined (**FIG. 8E**), the *Neurospora sp.* LOV protein orthologue (SEQ ID NO: 16) comprises about 90% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ ID NO: 2 is bold and underlined (**FIG. 8F**) and the *Candida sp.* LOV protein orthologue (SEQ ID NO: 17) comprises about 50% amino acid sequence identity to SEQ ID NO: 18, wherein the highly conserved threonine (**T**) amino acid at a sequence position corresponding to position 813 of SEQ ID NO: 2 is bold and underlined (**FIG. 8G**).

DETAILED DESCRIPTION

[0067] As set forth and described herein, the present disclosure addresses certain ongoing and unmet needs in the art of filamentous fungi protein production and methods thereof, including but not limited to genetic modifications in filamentous fungi that give rise to variant strains of filamentous fungi comprising enhanced protein productivity phenotypes. More specifically, as presented, described and exemplified herein, such variant strains of filamentous fungi comprising enhanced protein productivity phenotypes are well-suited for growth in submerged cultures, such as in large-scale production of proteins of interest for commercial applications.

I. DEFINITIONS

[0068] Prior to describing the present strains and methods in detail, the following terms are defined for clarity. Terms not defined should be accorded their ordinary meanings as used in the relevant art. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the present compositions and methods apply.

[0069] Filamentous fungus cells for manipulation, construction and use as described herein are generally from the phylum Ascomycota, subphylum Pezizomycotina, particularly fungi that have a vegetative hyphae state. Such organisms include filamentous fungus cells used for the production of commercially important industrial and pharmaceutical proteins, including, but not limited to *Trichoderma sp.*, *Aspergillus sp.*, *Fusarium sp.*, *Penicillium sp.*, *Chrysosporium sp.*, *Cephalosporium sp.*, *Talaromyces sp.*, *Geosmithia sp.*, *Neurospora sp.*, *Myceliophthora sp.* and the like. For example, in certain embodiments, filamentous fungus cells and strains thereof include, but are not limited to *Trichoderma reesei* (*previously* classified as *Trichoderma longibrachiatum* and *Hypocrea jecorina*), *Aspergillus niger*, *Aspergillus fumigatus*, *Aspergillus itaconicus*, *Aspergillus oryzae*, *Aspergillus nidulans*, *Aspergillus terreus*, *Aspergillus sojae*, *Aspergillus japonicus*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium chrysogenum*, *Talaromyces (Geosmithia) emersonii*, *Fusarium venenatum*, *Myceliophthora thermophila*, *Chrysosporium lucknowense* (C1) and the like.

[0070] As used herein, terms and phrases such as "filamentous fungus strain(s)", "filamentous fungal strain(s)", "fungus strain(s)", "fungal strains(s)", "filamentous fungus cell(s)", "filamentous fungal cell(s)",

"fungus cell(s)", "fungal cell(s)" and the like may be used interchangeably for convenience of description, and are not intend to limit the scope of the disclosure.

[0071] In certain embodiments, filamentous fungus cells for manipulation, construction and use as described herein are generally from the phylum Ascomycota, subphylum Pezizomycotina, particularly fungi that have a vegetative hyphae state and comprising a *lov* gene (or *lov* gene homologue(s)).

[0072] As used herein, phrases such as a "parental cell", a "parental fungal cell", a "parental strain", a "parental fungal strain", a "parental strain of filamentous fungus cells", "reference strain" and the like may be used interchangeably, and refer to "unmodified" parental filamentous fungal cells. For example, a "parental strain of filamentous fungus cells" refers to any cell or strain of filamentous fungi in which the genome of the "parental" cell is modified or modifiable (*e.g., via* only one genetic modification introduced into the parental cell) to generate a variant (daughter) strain of filamentous fungus cells such that "parental" and "daughter" cells differ by only one genetic modification.

[0073] As used herein, phrases such as a "variant cell", a "daughter cell", a "variant strain", a "daughter strain", a "variant or daughter fungal strain", a "variant or daughter strain of filamentous fungus cells" and the like may be used interchangeably, and refer to variant strains of filamentous fungus cells that are derived (*i.e.*, obtained from or obtainable from) from a parental (or reference) strain of filamentous fungus cells, wherein the variant strain comprises only one genetic modification which is not present in the parental strain, such that, by comparison, phenotypic differences between the "parental" and "variant" strains can be attributed to the one genetic modification. In other terms, parental and variant strains are otherwise isogenic except for the single genetic modification "introduced" to the variant strain. Thus, in the present disclosure, parental and variant strains can be described as having certain characteristics, such as genetic modifications, expression phenotypes, morphology phenotypes and the like; however, the skilled person will appreciate that it is technically the cells of the parental or variant strain that have such characteristics, and the "strains" are referred to for convenience.

[0074] In certain embodiments, unmodified (parental) cells may be referred to as "control cells" or "reference cells", particularly when being compared (vis-à-vis) with genetically modified (variant/daughter) cells derived therefrom.

[0075] As used herein, the terms "wild-type" and "native" are used interchangeably and refer to genes, proteins, protein mixes or strains, as found in nature.

[0076] As used herein, certain *Trichoderma reesei* strains/cells of the disclosure have been named/abbreviated as set forth below in **Table 1**.

TABLE 1

***T. reesei* Strains and Genetic Modifications**

Strain Name	Genetic Modification(s)	Marker Integration Site
B7ms1	Δ*mpg1*; Δ*seb1*; GA construct; *lov*(+)	*seb1*
B7ms1-SF12	Δ*mpg1*; Δ*seb1*; GA construct; *lov*(T813K)	*seb1*
T4	*nik1*(M743T); *lov*(+)	*pyr2*
T4_pyr2	*nik1*(M743T); *lov*(+); *pyr2*	None
T4m	*nik1*(M743T); Δ*mpg1*; *lov*(+)	*mpg1*
T4m_pyr2	*nik1*(M743T); Δ*mpg1*; *lov*(+); *pyr2*	None
T4ml+	*nik1*(M743T); Δ*mpg1; lov*(dis)	*lov*
T4ml	*nik1*(M743T); Δ*mpg1; lov*(T813K); *pyr2*	None
T4mls	*nik1*(M743T); Δ*mpg1; Δseb1; lov*(T813K)	*seb1*
T4ms	*nik1*(M743T); Δ*mpg1; Δseb1*; *lov*(+)	*seb1*
T4s	*nik1*(M743T); Δ*seb1*; *lov*(+)	*seb1*
T4s_pyr2	*nik1*(M743T); Δ*seb1*; *lov*(+); *pyr2*	None
T4sl+	*nik1*(M743T); Δ*seb1; lov*(dis)	*lov*
T4sl	*nik1*(M743T); Δ*seb1; lov*(T813K); *pyr2*	None
T4l+	*nik1*(M743T); *lov*(dis)	*lov*
41G	*nik1*(M743T); *lov*(+)	*pyr4*
41G_pyr4	*nik1*(M743T); *lov*(+); *pyr4*	None
41Gl+	*nik1*(M743T); *lov*(dis)	*lov*
T4mc	*nik1*(M743T); Δ*mpg1; lov*(+)	*site C*
T4mlc	*nik1*(M743T); Δ*mpg1; lov*(T813K)	*site C*
T4md	*nik1*(M743T); Δ*mpg1; lov*(+)	*site B*
T4mld	*nik1*(M743T); Δ*mpg1; lov*(T813K)	*site B*
T4mp	*nik1*(M743T); Δ*mpg1; lov*(+)	*site A*
T4mlp	*nik1*(M743T); Δ*mpg1; lov*(T813K)	*site A*
T4sp	*nik1*(M743T); Δ*seb1; lov*(+)	*site A*
T4slp	*nik1*(M743T); Δ*seb1; lov*(T813K)	*site A*

[0077] As used herein, a reduced viscosity Trichoderma strain "B7ms1" is a glucoamylase expressing strain referred to as "Morph TrGA 77B7Δ*mpg1*Δ*seb1*" in International PCT Publication No. WO2012/145584 (incorporated herein by reference in its entirety).

[0078] As used herein, genomic coordinates (*e.g.*, 425393 on Scaffold 16) and Protein Identification numbers (PID, *e.g.*, PID 50212) reference Version 2 of the *Trichoderma reesei* QM6a genome sequence assembly generated by the Department of Energy Joint Genome Institute (JGI). (The Genome Portal of the Department of Energy Joint Genome Institute, Grigoriev *et al.*, *Nucleic Acids Res* 2012 Jan;40(Database issue):D26-32. doi: 10.1093/nar/gkr947). The JGI assembled Scaffold sequences have also been deposited in GenBank (The National Center for Biotechnology) under the nucleotide accession numbers GL985056.1 through GL985132.1.

[0079] As used herein, a mutant (variant) Trichoderma strain named "B7ms1-SF12" (derived from the B7ms1 parental strain) comprises an increased protein productivity phenotype (*i.e.*, relative to B7ms1 (parental) strain). More particularly, the identified mutation in Trichoderma B7ms1-SF12 strain alters the coding sequence of a protein named LOV (*i.e.*, predicted protein PID 50212; SEQ ID NO: 2), wherein a highly-conserved threonine (T) amino acid at residue position 813 (T813) of the (native) LOV protein (SEQ ID NO:

2) was substituted with a lysine (T→K813) in the (mutant) B7ms1-SF12 strain (*i.e.*, a T813K substitution, *e.g.*, compare SEQ ID NO: 2 position 813 vis-à-vis SEQ ID NO: 4 position 813). For example, as described in the Examples section below, the *lov* mutant allele in strain B7ms1-SF12 comprises a single nucleotide change of G (guanine) to T (thymine) at 425393 on Scaffold 16, thereby resulting in the "T813K" substitution in the encoded LOV (variant) protein (SEQ ID NO: 4), comprising a lysine (K) at amino acid position 813 of SEQ ID NO: 4 (in contrast to the native LOV protein of SEQ ID NO: 2, comprising a threonine (T) at amino acid position 813; SEQ ID NO: 2).

[0080] As used herein, a Trichoderma strain named "T4" was derived from Trichoderma strain RL-P37 by incorporation of the *nik1*(M743T) mutation and chemical mutagenesis, as described in International PCT Publication No. WO2016/130523 (incorporated herein by reference in its entirety).

[0081] As used herein, a gene allele named "*nik1*(M743T)", comprises a mutant histidine kinase gene (*nik1*) encoding a variant histidine kinase (NIK1; SEQ ID NO: 19) comprising a methionine (M) to threonine (T) substitution at amino acid (residue) position 743 of SEQ ID NO: 19.

[0082] As used herein, a Trichoderma strain named "T4_pyr2" was derived from Trichoderma strain T4 by mutation and loss of function of the *pyr2* gene so that it may be used as a transformation selection marker.

[0083] As used herein, a Trichoderma strain named "T4m" was derived from strain T4, wherein strain T4m comprises a mutation of the *mpg1* gene (Δ*mpg1*) and a *nik1*(M743T) gene. Strain T4m therefore comprises the wild-type *lov* allele (*i.e.*, allele *lov*(+)) encoding the native lov protein (SEQ ID NO: 2).

[0084] As used herein, a Trichoderma strain named "T4m_pyr2" was derived from strain T4m, wherein strain T4m_pyr2 comprises a mutation of the *mpg1* gene (Δ*mpg1*), a *nik1*(M743T) gene and no functional *pyr2* gene.

[0085] As used herein, a Trichoderma strain named "T4ml+" was derived from strain T4m_pyr2, wherein strain T4ml+ comprises a mutation of the *mpg1* gene (Δ*mpg1*) and allele *lov*(dis).

[0086] As used herein, a Trichoderma strain named "T4ml" was derived from strain T4ml+, wherein strain T4ml comprises *nik1*(M743T), a mutation of the *mpg1* gene (Δ*mpg1*), allele *lov*(T813K), and no functional *pyr2* gene.

[0087] As used herein, a Trichoderma strain named "T4mls" was derived from strain T4ml, wherein strain T4mls comprises *nik1*(M743T), a double mutation of the *mpg1* and *seb1* genes (Δ*mpg1*; Δ*seb1*) and allele *lov*(T813K).

[0088] As used herein, a Trichoderma strain named "T4ms" was derived from strain T4m, wherein strain T4ms comprises *nik1*(M743T) and a double mutation of the *mpg1* and *seb1* genes (Δ*mpg1*; Δ*seb1*).

[0089] As used herein, a Trichoderma strain named "T4s" was derived from strain T4, wherein strain T4s comprises *nik1*(M743T) and a mutation of the *seb1* gene (Δ*seb1*).

[0090] As used herein, a Trichoderma strain named "T4s_pyr2" was derived from strain T4s, wherein strain T4s_pyr2 comprises *nik1*(M743T), a mutation of the *seb1* gene (Δ*seb1*) and no functional *pyr2* gene.

[0091] As used herein, a Trichoderma strain named "T4sl+" was derived from strain T4s_pyr2, wherein strain T4sl+ comprises *nik1*(M743T), a mutation of the *seb1* gene (Δ*seb1*) and allele *lov*(dis).

[0092] As used herein, a Trichoderma strain named "T4sl" was derived from strain T4sl+, wherein strain T4sl comprises *nik1*(M743T), a mutation of the *seb1* gene (Δ*seb1*), allele *lov*(T813K) and no functional *pyr2* gene.

[0093] As used herein, a Trichoderma strain named "T4l+" was derived from strain T4_pyr2, wherein strain T4l+ comprises *nik1*(M743T) and allele *lov*(dis).

[0094] As used herein, a Trichoderma strain named "41G" was a mutagenized derived from strain T4, wherein strain 4lG comprises *nik1*(M743T) and the (wild-type) *lov*(+) allele.

[0095] As used herein, a Trichoderma strain named "41G_pyr4" was derived from strain 41G, wherein strain 4lG_pyr4 comprises *nik1*(M743T), the (wild-type) *lov*(+) allele and deletion of the *pyr4* gene.

[0096] As used herein, a Trichoderma strain named "41Gl+" was derived from strain 4lG_pyr4, wherein strain 4lGl+ comprises *nik1*(M743T) and allele *lov*(dis).

[0097] As used herein, a Trichoderma strain named "T4mc" was derived from strain T4m_pyr2, wherein strain T4mc comprises *nik1*(M743T), Δ*mpg1, pyr2*+ insertion at site C and a wild-type *lov*(+) allele.

[0098] As used herein, a Trichoderma strain named "T4mlc" was derived from strain T4ml, wherein strain T4mlc comprises *nik1*(M743T), Δ*mpg1*, *pyr2*+ insertion at site C and allele *lov*(T813K).

[0099] As used herein, a Trichoderma strain named "T4md" was derived from strain T4m_pyr2, wherein strain T4md comprises *nik1*(M743T), Δ*mpg1*, *pyr2*+ insertion at site B and a wild-type *lov*(+) allele.

[0100] As used herein, a Trichoderma strain named "T4mld" was derived from strain T4ml, wherein strain T4mld comprises *nik1*(M743T), Δ*mpg1*, *pyr2*+ insertion at site B and allele *lov*(T813K).

[0101] As used herein, a Trichoderma strain named "T4mp" was derived from strain T4m_pyr2, wherein strain T4mp comprises *nik1*(M743T), Δ*mpg1*, *pyr2*+ insertion at site A and a wild-type *lov*(+) allele.

[0102] As used herein, a Trichoderma strain named "T4mlp" was derived from strain T4ml, wherein strain T4mlp comprises *nik1*(M743T), Δ*mpg1*, *pyr2*+ insertion at site A and allele *lov*(T813K).

[0103] As used herein, a Trichoderma strain named "T4sp" was derived from strain T4s_pyr2, wherein strain T4sp comprises *nik1*(M743T), Δ*seb1*, *pyr2*+ insertion at site A and a wild-type *lov*(+) allele.

[0104] As used herein, a Trichoderma strain named "T4slp" was derived from strain T4sl, wherein strain T4slp comprises *nik1*(M743T), Δ*seb1*, *pyr2*+ insertion at site A and allele *lov*(T813K).

[0105] As used herein, a "glucoamylase (GA) construct" or "GA construct" encodes a glucoamylase described in PCT Publication No. WO2012/145584 (specifically incorporated herein by reference in its entirety).

[0106] As used herein, "allele *lov*(+)" comprises a wild-type *lov* DNA sequence encoding a native LOV protein (*e.g.*, SEQ ID NO: 2).

[0107] As used herein, "allele *lov*(T813K)" comprises a mutated (*lov*) DNA sequence (identified in strain B7ms1-SF12, *described above*) encoding a variant LOV protein comprising the "T813K" substitution (SEQ ID NO: 4, *described above*).

[0108] As used herein, "allele *lov*(dis)" comprises a disruption of the *lov* gene in which a selectable marker, either *pyr2* or *pyr4*, was integrated into the *lov* coding sequence, as further described in Example 3 below.

[0109] As used herein, a "whole cellulase strain" refers to a Trichoderma strain where the natural secretome has not been altered by genetic engineering of the major cellulase genes.

[0110] As used herein, an "endogenous (or native) filamentous fungal gene" encoding a protein of interest includes, but is not limited to, endogenous (filamentous fungal) genes encoding glycoside hydrolase (GH) family enzymes (*e.g.*, such as EC Nos. 3.2.1.1 – 3.2.1.206), endogenous genes encoding proteases, esterases, lipases and the like, as known and understood by one skilled in the art.

[0111] Thus, as generally described herein, a substitution of a threonine (T) amino acid for a lysine (K) amino acid (T→K substitution) at an amino acid (residue) position equivalent to position 813 of the (variant) LOV protein of SEQ ID NO: 4, has been identified herein as being responsible for the observed (increased) protein productivity phenotypes. For example, filamentous fungal cells disclosed herein, comprising an introduced allele encoding the (mutant) B7ms1-SF12 strain $T_{813}\rightarrow_{813}K$ substitution (*e.g.*, allele *lov*(T813K)) demonstrate enhanced protein productivity phenotypes relative to isogenic cells lacking this mutation (*i.e.*, comprising a wild-type *lov* gene, allele *lov*(+), encoding a native LOV protein) when fermented/cultivated under the same conditions. More particularly as presented in the Examples section below, filamentous fungal cells of the disclosure (comprising such enhanced protein productivity phenotypes) are particularly well-suited for growth in (aerobic) submerged cultures (*e.g.*, such as in large-scale production of proteins of interest for commercial applications).

[0112] As used herein, the "position" of an amino acid residue in a "given amino acid sequence" is numbered herein using the amino acid residue numbering (positions) of the native *Trichoderma sp.* LOV protein of SEQ ID NO: 2. For example, phrases such as "comprises a threonine (T) residue at a sequence position corresponding to position 813 of SEQ ID NO: 2" and "comprises a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 2", the native (*Trichoderma sp.*) LOV protein's amino acid sequence (SEQ ID NO: 2) serves as a reference (parent) protein sequence.

[0113] For example, as shown in **FIG. 7C-7F**, a given amino acid sequence described herein can be aligned with the native *Trichoderma sp.* LOV protein amino acid sequence (SEQ ID NO: 2), using alignment algorithms described herein (and/or alignment algorithms known in the art,) and an amino acid residue in the given amino acid sequence that aligns (preferably, optimally aligns) with an amino acid residue in the native sequence can be conveniently numbered by reference to the corresponding amino acid residue in the LOV sequence. Thus, **FIG. 7** presents a multiple sequence alignment of the native *Trichoderma sp.* LOV protein (SEQ ID NO: 2) labeled sequence "**2**", aligned with LOV protein orthologues from various (Ascomycota) filamentous fungi, such as *Aspergillus sp.* (SEQ ID NO: 12) labeled sequence "**12**", *Penicillium sp.* (SEQ ID NO: 13) labeled sequence "**13**", *Talaromyces sp.* (SEQ ID NO: 14) labeled sequence "**14**", *Fusarium sp.* (SEQ ID NO: 15) labeled sequence "**15**", *Myceliophthora sp.* (SEQ ID NO: 16) labeled sequence "**16**", *Neurospora sp.* (SEQ ID NO: 17) labeled sequence "**17**" and *Candida sp.* (SEQ ID NO: 18) labeled sequence "**18**".

[0114] Likewise, to establish sequence homology or sequence identity to the primary (1°) sequence of the *Trichoderma sp.* LOV protein (SEQ ID NO: 2), one skilled in the art may readily compare the primary sequence of SEQ ID NO: 2 with one or more candidate LOV protein (orthologue) sequences using sequence

alignment algorithms, software and methods thereof know to one skilled in the art. Thus, after aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (*i.e.*, avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of Ascomycota LOV protein are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 98%, 95%, 90%, 85%, 80%, 75% 70%, 50% or at least 45% of conserved residues is also adequate to define equivalent residues.

[0115] Thus, as used herein, a substitution of a threonine (T) amino acid for a lysine (K) amino acid ($T_{813}\rightarrow_{813}K$ substitution) at an amino acid (residue) position corresponding (or equivalent) to position 813 of SEQ ID NO: 2, includes any $T_{813}\rightarrow_{813}K$ substitution at an amino acid (residue) position corresponding to position 813 of SEQ ID NO: 2 in any Ascomycota filamentous fungal cell.

[0116] As used herein, the term "gene" is synonymous with the term "allele" in referring to a nucleic acid that encodes and directs the expression of a protein or RNA. Vegetative forms of filamentous fungi are generally haploid, therefore a single copy of a specified gene (*i.e.*, a single allele) is sufficient to confer a specified phenotype.

[0117] As used herein, the terms "polypeptide" and "protein" (and/or their respective plural forms) are used interchangeably to refer to polymers of any length comprising amino acid residues linked by peptide bonds. The conventional one-letter or three-letter codes for amino acid residues are used herein. The polymer can be linear or branched, it can comprise modified amino acids, and it can be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), as well as other modifications known in the art.

[0118] As used herein, the term "derivative polypeptide/protein" refers to a protein which is derived or derivable from a protein by addition of one or more amino acids to either or both the N- and C-terminal end(s), substitution of one or more amino acids at one or a number of different sites in the amino acid sequence, deletion of one or more amino acids at either or both ends of the protein or at one or more sites in the amino acid sequence, and/or insertion of one or more amino acids at one or more sites in the amino acid sequence. The preparation of a protein derivative can be achieved by modifying a DNA sequence which encodes for the native protein, transformation of that DNA sequence into a suitable host, and expression of the modified DNA sequence to form the derivative protein.

[0119] Related (and derivative) proteins include "variant proteins". Variant proteins differ from a reference/parental protein (*e.g.*, a wild-type protein) by substitutions, deletions, and/or insertions at a small number of amino acid residues. The number of differing amino acid residues between the variant and parental protein can be one or more, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, or more

amino acid residues. Variant proteins can share at least about 50%, at least about 60%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or even at least about 99%, or more, amino acid sequence identity with a reference protein. A variant protein can also differ from a reference protein in selected motifs, domains, epitopes, conserved regions, and the like.

[0120] As used herein, the term "analogous sequence" refers to a sequence within a protein that provides similar function, tertiary structure, and/or conserved residues as the protein of interest (*i.e.*, typically the original protein of interest). For example, in epitope regions that contain an α-helix or a β-sheet structure, the replacement amino acids in the analogous sequence preferably maintain the same specific structure. The term also refers to nucleotide sequences, as well as amino acid sequences. In some embodiments, analogous sequences are developed such that the replacement of amino acids result in a variant enzyme showing a similar or improved function. In some embodiments, the tertiary structure and/or conserved residues of the amino acids in the protein of interest are located at or near the segment or fragment of interest. Thus, where the segment or fragment of interest contains, for example, an α-helix or a β-sheet structure, the replacement amino acids preferably maintain that specific structure.

[0121] As used herein, the term "homologous protein" refers to a protein that has similar activity and/or structure to a reference protein. It is not intended that homologues necessarily be evolutionarily related. Thus, it is intended that the term encompass the same, similar, or corresponding protein(s) (*i.e.*, in terms of structure and function) obtained from different organisms. In some embodiments, it is desirable to identify a homologue that has a quaternary, tertiary and/or primary structure similar to the reference protein.

[0122] The degree of homology between sequences can be determined using any suitable method known in the art (see, *e.g.,* Smith and Waterman,1981; Needleman and Wunsch, 1970; Pearson and Lipman, 1988; programs such as GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package (Genetics Computer Group, Madison, WI); and Devereux *et al.*, 1984).

[0123] For example, PILEUP is a useful program to determine sequence homology levels. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pair-wise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng and Doolittle (1987). The method is similar to that described by Higgins and Sharp (1989). Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps. Another example of a useful algorithm is the BLAST algorithm, described by Altschul *et al.*, 1990 and Karlin *et al.*, 1993. One particularly useful BLAST program is the WU-BLAST-2 program (see, *e.g.,* Altschul *et al.*, 1996). Parameters "W," "T," and "X" determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word-length (W) of 11, the BLOSUM62 scoring matrix (see, *e.g.,* Henikoff and Henikoff, 1989) alignments (B) of 50, expectation (E) of 10, M′5, N′-4, and a comparison of both strands.

[0124] As used herein, the phrases "substantially similar" and "substantially identical", in the context of at least two nucleic acids or polypeptides, typically means that a polynucleotide or polypeptide comprises a sequence that has at least about 70% identity, at least about 75% identity, at least about 80% identity, at least about 85% identity, at least about 90% identity, at least about 91% identity, at least about 92% identity, at least about 93% identity, at least about 94% identity, at least about 95% identity, at least about 96% identity, at least about 97% identity, at least about 98% identity, or even at least about 99% identity, or more, compared to the reference (*i.e.*, wild-type) sequence. Sequence identity can be determined using known programs such as BLAST, ALIGN, and CLUSTAL using standard parameters. (See, *e.g.,* Altschul, *et al.*,1990; Henikoff *et al.*, 1989; Karlin *et al.*, 1993; and Higgins *et al.*, 1988). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. Also, databases can be searched using FASTA (Pearson *et al.*, 1988). One indication that two polypeptides are substantially identical is that the first polypeptide is immunologically cross-reactive with the second polypeptide. Typically, polypeptides that differ by conservative amino acid substitutions are immunologically cross-reactive. Thus, a polypeptide is substantially identical to a second polypeptide, for example, where the two peptides differ only by a conservative substitution. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions (*e.g.,* within a range of medium to high stringency).

[0125] As used herein, "nucleic acid" refers to a nucleotide or polynucleotide sequence, and fragments or portions thereof, as well as to DNA, cDNA, and RNA of genomic or synthetic origin, which may be double-stranded or single-stranded, whether representing the sense or antisense strand.

[0126] As used herein, the term "expression" refers to the transcription and stable accumulation of sense (mRNA) or anti-sense RNA, derived from a nucleic acid molecule of the disclosure. Expression may also refer to translation of mRNA into a polypeptide. Thus, the term "expression" includes any step involved in the production of the polypeptide including, but not limited to transcription, post-transcriptional modification, translation, post-translational modification, secretion and the like.

[0127] As used herein, the combined term "expresses/produces", as used in phrases such as a "variant strain of filamentous fungus cells expresses/produces an 'increased' amount of a protein of interest (POI)" (*i.e.*, relative to the parental cell), the term "expresses/produces" is meant to include any steps involved in the expression and production of a protein in such variant filamentous fungus strains of the disclosure.

[0128] In certain embodiments, a gene, polynucleotide or nucleic acid sequence encoding a LOV protein comprising "sequence homology" refers to DNA or RNA (nucleic acid) sequences that have *de minimus* sequence variations from the corresponding nucleic acid sequences (to which comparison is made) and retain substantially the same biological functions as the corresponding nucleic acid sequences (to which comparison is made). For example, in certain embodiments, a nucleic acid sequence comprising substantial sequence homology to a gene, polynucleotide, or nucleic acid encoding a LOV protein is assessed by identifying the encoded gene product (LOV protein), as described herein.

[0129] In certain other embodiments, a gene, polynucleotide, or nucleic acid sequence comprising sequence homology to a gene, polynucleotide, or nucleic acid encoding a LOV protein is determined/identified using nucleic acid hybridization methods. For example, in certain embodiments, a DNA/RNA sequence comprising substantial sequence homology to a gene encoding a LOV protein (*e.g.*, SEQ ID NO: 2) is identified by the ability of such DNA/RNA sequence to hybridize with a specified nucleic acid sequence of the disclosure, under stringent conditions.

[0130] As used herein, "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Such stringent conditions are well known to those skilled in the art (see, *e.g.,* Ausubel *et al.*, 1995; Sambrook *et al.*, 1989). For example, in certain embodiments, a non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chlorine/sodium citrate (SSC), at about 65-70°C (or hybridization in 4×SSC plus 50% formamide at about 42-50°C), followed by one or more washes in 1×SSC, at about 65-70°C. Likewise, a non-limiting example of highly stringent hybridization conditions includes hybridization in 1×SSC, at about 65-70°C (or hybridization in 4×SSC plus 50% formamide at about 42-50°C), followed by one or more washes in 0.3×SSC, at about 65-70°C. Thud, highly stringent hybridization conditions include hybridization in 4×SSC, at about 50-60°C (or alternatively hybridization in 6×SSC plus 50% formamide at about 40-45°C), followed by one or more washes in 2×SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.,* at 65-70°C or at 42-50°C are also intended to be encompassed by the present disclosure. In certain embodiments, SSPE (1×SSPE is 0.15 M NaCl, 10 mM NaH2PO4, and 1.25 mM EDTA, pH 7.4) can be substituted for SSC (1×SSPE is 0.15 M NaCl and 15 mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m (°C) = 2(# of A+T bases) + 4(# of G+C bases). For hybrids between 18 and 49 base pairs in length, T_m (°C) = 81.5+16.6(log10[Na+]) + 0.41(% G+C) − (600/N), where N is the number of bases in the hybrid, and [Na+] is the concentration of sodium ions in the hybridization buffer ([Na+] for 1×SSC=0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to the hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.,* SDS) chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH2PO4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH2PO4, 1% SDS at 65°C or alternatively 0.2×SSC, 1% SDS (*see, e.g.*, Church and Gilbert, 1984).

[0131] Thus, as generally set forth above, certain embodiments of the disclosure are related to variant strains of filamentous fungus cells comprise a genetic modification of a gene encoding a LOV protein. As used

herein, the terms "modification" and "genetic modification" are used interchangeably and include, but are not limited to: (a) the introduction, substitution, or removal of one or more nucleotides in a gene, or the introduction, substitution, or removal of one or more nucleotides in a regulatory element required for the transcription or translation of the gene, (b) gene disruption, (c) gene conversion, (d) gene deletion, (e) the down-regulation of a gene (*e.g.*, *antisense* RNA, *si*RNA, *mi*RNA, and the like), (f) specific mutagenesis (including, but not limited to, CRISPR/Cas9 based mutagenesis) and/or (g) random mutagenesis of any one or more the genes disclosed herein.

[0132] As used herein, a variant strain of filamentous fungus comprising a genetic modification includes, but is not limited to a genetic modification of a gene encoding a LOV protein disclosed herein. Thus, as described in further detail below, various molecular biological methods are well known and available to one skilled in the art for generating/constructing such variant strains of filamentous fungus cells.

[0133] As used herein, "the introduction, substitution, or removal of one or more nucleotides in a gene encoding a protein", such genetic modifications include the gene's coding sequence (*i.e.*, exons) and non-coding intervening (introns) sequences.

[0134] As used herein, "disruption of a gene", "gene disruption", "inactivation of a gene" and "gene inactivation" are used interchangeably and refer broadly to any genetic modification that substantially disrupts/inactivates a target gene. Exemplary methods of gene disruptions include, but are not limited to, the complete or partial deletion of any portion of a gene, including a polypeptide coding sequence (CDS), a promoter, an enhancer, or another regulatory element, or mutagenesis of the same, where mutagenesis encompasses substitutions, insertions, deletions, inversions, and any combinations and variations thereof which disrupt/inactivate the target gene(s) and substantially reduce or prevent the expression/production of the functional gene product. In certain embodiments of the disclosure, such gene disruptions prevent a host cell from expressing/producing the encoded *lov* gene product.

[0135] In certain embodiments, a gene, polynucleotide or nucleic acid sequence encoding a LOV protein is genetically modified using an established gene editing technique, such as CRISPR/Cas9 gene editing, zinc-finger nuclease (ZFN) gene editing, transcription activator-like effector nuclease editing (TALEN), homing (*mega*) nuclease editing, and the like.

[0136] In other embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by the process of gene conversion (*e.g.*, *see* Iglesias and Trautner, 1983).

[0137] In other embodiments, a protein of interest (*e.g.*, an endogenous POI or a heterologous POI) expressed/produced by the Ascomycota cells of the disclosure is detected, measured, assayed and the like, by protein quantification methods, gene transcription methods, mRNA translation methods and the like, including, but not limited to protein migration/mobility (SDS-PAGE), mass spectrometry, HPLC, size exclusion, ultracentrifugation sedimentation velocity analysis, transcriptomics, proteomics, fluorescent tags, epitope tags, fluorescent protein (GFP, RFP, *etc.*) chimeras/hybrids and the like.

[0138] As used herein, functionally and/or structurally similar proteins are considered to be "related proteins". Such related proteins can be derived from organisms of different genera and/or species, or even

different classes of organisms (*e.g.,* bacteria and fungi). Related proteins also encompass homologues and/or orthologues determined by primary sequence analysis, determined by secondary or tertiary structure analysis, or determined by immunological cross-reactivity.

[0139] The term "promoter" as used herein refers to a nucleic acid sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' (downstream) to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleic acid segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

[0140] The term "operably linked" as used herein refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence (*e.g.,* an ORF) when it is capable of affecting the expression of that coding sequence (*i.e.*, that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation. A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA encoding a secretory leader (*i.e.*, a signal peptide), is operably linked to DNA for a polypeptide if it is expressed as a pre-protein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

[0141] As defined herein, "suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, RNA processing site, effector binding site and stem-loop structure.

[0142] As defined herein, the term "introducing", as used in phrases such as "introducing into a fungal cell" at least one polynucleotide open reading frame (ORF), or a gene thereof, or a vector thereof, includes methods known in the art for introducing polynucleotides into a cell, including, but not limited to protoplast fusion, natural or artificial transformation (*e.g.,* calcium chloride, electroporation), transduction, transfection and the like.

[0143] As used herein, "transformed" or "transformation" mean a cell has been transformed by use of recombinant DNA techniques. Transformation typically occurs by insertion of one or more nucleotide sequences (*e.g.*, a polynucleotide, an ORF or gene) into a cell. The inserted nucleotide sequence may be a heterologous nucleotide sequence (*i.e.*, a sequence that is not naturally occurring in the cell that is to be transformed).

[0144] As used herein, "transformation" refers to introducing an exogenous DNA into a host cell so that the DNA is maintained as a chromosomal integrant or a self-replicating extra-chromosomal vector. As used herein, "transforming DNA", "transforming sequence", and "DNA construct" refer to DNA that is used to introduce sequences into a host cell. The DNA may be generated *in vitro* by PCR or any other suitable techniques. In some embodiments, the transforming DNA comprises an incoming sequence, while in other embodiments it further comprises an incoming sequence flanked by homology boxes. In yet a further embodiment, the transforming DNA comprises other non-homologous sequences, added to the ends (*i.e.*, stuffer sequences or flanks). The ends can be closed such that the transforming DNA forms a closed circle, such as, for example, insertion into a vector.

[0145] As used herein "an incoming sequence" refers to a DNA sequence that is introduced into the fungal cell chromosome. In some embodiments, the incoming sequence is part of a DNA construct. In other embodiments, the incoming sequence encodes one or more proteins of interest. In some embodiments, the incoming sequence comprises a sequence that may or may not already be present in the genome of the cell to be transformed (*i.e.*, it may be either a homologous or heterologous sequence). In some embodiments, the incoming sequence encodes one or more proteins of interest, a gene, and/or a mutated or modified gene. In alternative embodiments, the incoming sequence encodes a functional wild-type gene or operon, a functional mutant gene or operon, or a nonfunctional gene or operon. In some embodiments, an incoming sequence is a non-functional sequence inserted into a gene to disrupt function of the gene. In another embodiment, the incoming sequence includes a selective marker. In a further embodiment the incoming sequence includes two homology boxes.

[0146] As used herein, "homology box" refers to a nucleic acid sequence, which is homologous to a sequence in the fungal cell chromosome. More specifically, a homology box is an upstream or downstream region having between about 80 and 100% sequence identity, between about 90 and 100% sequence identity, or between about 95 and 100% sequence identity with the immediate flanking coding region of a gene or part of a gene to be deleted, disrupted, inactivated, down-regulated and the like, according to the invention. These sequences direct where in the fungal cell chromosome a DNA construct is integrated and directs what part of the fungal cell chromosome is replaced by the incoming sequence. While not meant to limit the present disclosure, a homology box may include about between 1 base pair (bp) to 200 kilobases (kb). Preferably, a homology box includes about between 1 bp and 10.0 kb; between 1 bp and 5.0 kb; between 1 bp and 2.5 kb; between 1 bp and 1.0 kb, and between 0.25 kb and 2.5 kb. A homology box may also include about 10.0 kb, 5.0 kb, 2.5 kb, 2.0 kb, 1.5 kb, 1.0 kb, 0.5 kb, 0.25 kb and 0.1 kb. In some

embodiments, the 5' and 3' ends of a selective marker are flanked by a homology box wherein the homology box comprises nucleic acid sequences immediately flanking the coding region of the gene.

[0147] As used herein, the term "selectable marker-encoding nucleotide sequence" refers to a nucleotide sequence which is capable of expression in the host cells and where expression of the selectable marker confers to cells containing the expressed gene the ability to grow in the presence of a corresponding selective agent or lack of an essential nutrient.

[0148] As used herein, the terms "selectable marker" and "selective marker" refer to a nucleic acid (*e.g.,* a gene) capable of expression in host cell which allows for ease of selection of those hosts containing the vector. Examples of such selectable markers include, but are not limited to, antimicrobials. Thus, the term "selectable marker" refers to genes that provide an indication that a host cell has taken up an incoming DNA of interest or some other reaction has occurred. Typically, selectable markers are genes that confer antimicrobial resistance or a metabolic advantage on the host cell to allow cells containing the exogenous DNA to be distinguished from cells that have not received any exogenous sequence during the transformation.

[0149] As defined herein, a host cell "genome", a fungal cell "genome", or a filamentous fungus cell "genome" includes chromosomal and extrachromosomal genes.

[0150] As used herein, the terms "plasmid", "vector" and "cassette" refer to extrachromosomal elements, often carrying genes which are typically not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA molecules. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single-stranded or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell.

[0151] As used herein, the term "vector" refers to any nucleic acid that can be replicated (propagated) in cells and can carry new genes or DNA segments (*e.g.,* an "incoming sequence") into cells. Thus, the term refers to a nucleic acid construct designed for transfer between different host cells. Vectors include viruses, bacteriophage, pro-viruses, plasmids, phagemids, transposons, and artificial chromosomes such as YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), PLACs (plant artificial chromosomes), and the like, that are "episomes" (*i.e.,* replicate autonomously) or can integrate into the chromosome of a host cell.

[0152] A used herein, a "transformation cassette" refers to a specific vector comprising a gene (or ORF thereof), and having elements in addition to the gene that facilitate transformation of a particular host cell.

[0153] An "expression vector" refers to a vector that has the ability to incorporate and express heterologous DNA in a cell. Many prokaryotic and eukaryotic expression vectors are commercially available and know to one skilled in the art. Selection of appropriate expression vectors is within the knowledge of one skilled in the art.

[0154] As used herein, the terms "expression cassette" and "expression vector" refer to a nucleic acid construct generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a target cell (*i.e.*, these are vectors or vector elements, as described above). The recombinant expression cassette can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of an expression vector includes, among other sequences, a nucleic acid sequence to be transcribed and a promoter. In some embodiments, DNA constructs also include a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a target cell. In certain embodiments, a DNA construct of the disclosure comprises a selective marker and an inactivating chromosomal or gene or DNA segment as defined herein.

[0155] As used herein, a "targeting vector" is a vector that includes polynucleotide sequences that are homologous to a region in the chromosome of a host cell into which the targeting vector is transformed and that can drive homologous recombination at that region. For example, targeting vectors find use in introducing genetic modifications into the chromosome of a host cell through homologous recombination. In some embodiments, a targeting vector comprises other non-homologous sequences, *e.g.,* added to the ends (*i.e.,* stuffer sequences or flanking sequences). The ends can be closed such that the targeting vector forms a closed circle, such as, for example, insertion into a vector.

[0156] As used herein, a variant cell (or strain) comprising an "enhanced protein productivity phenotype" includes, but is not limited to, a variant cell (or strain) comprising an enhanced/increased volumetric productivity, a variant cell (or strain) comprising an enhanced/increased carbon conversion efficiency, a variant cell (or strain) comprising an enhanced/increased protein yield, a variant cell (or strain) comprising an enhanced/increased specific protein productivity and the like. For example, in certain embodiments, a variant cell or strain comprising an enhanced protein productivity phenotype expresses/produces at least 0.1% or more total protein (g) *per* g of fed sugars (relative to parental strain), wherein fed sugars can be expressed in terms of mass of sugar added to the fermentor during production phase (*i.e.*, following feed-start).

[0157] As defined herein, the phrases "enhanced protein productivity phenotype" and "increased protein productivity phenotype", may be used interchangeably.

[0158] As used herein, when describing an "enhanced/increased protein productivity phenotype" in an unmodified (parental) cell vis-à-vis the modified (variant/daughter), it will be understood that the "parental" and "variant" cells are grown/cultivated/fermented under the same conditions (*e.g.*, the same conditions such as media, temperature, pH and the like). Similarly, when describing the "expression/production" of a protein of interest (POI) in an unmodified (parental) cell vis-à-vis the "expression/production" of the same POI in a modified (variant/daughter) cell, it will be understood that the "parental" and "variant" cells are grown/cultivated/fermented under essentially the same conditions (*e.g.*, the same conditions such as media, temperature, pH and the like).

[0159] As used herein, "aerobic fermentation" refers to growth in the presence of oxygen.

[0160] As used herein, the terms "broth", "cell broth", "fermentation broth" and/or "culture broth" are used interchangeably, and refer collectively to (i) the fermentation (culture) medium and (ii) the cells, in a liquid (submerged) culture.

[0161] As used herein, the term "cell mass" refers to the cell component (including intact and lysed cells) present in a liquid (submerged) culture. Cell mass can be expressed in dry cell weight (DCW) or wet cell weight (WCW).

[0162] As used herein, a "reduced viscosity" strain of filamentous fungus cells refers to a modified (daughter) strain that produces a cell broth that has a reduced viscosity (*i.e.*, reduced resistance to shear or tensile stress) compared to an equivalent cell broth produced by a parental strain. For example, equivalent cell broths generally have comparable cell masses. Methods for constructing reduced viscosity filamentous fungal strains, and methods for comparing the viscosities thereof, are described in detail in International PCT Publication Nos. WO2012/027580, WO2012/145596, WO2012/145596 and WO2012/145592, and International PCT Application Serial No. PCT/US2019/27590 (each specifically incorporated herein by reference in its entirety).

[0163] Thus, in certain embodiments, a variant strain of the disclosure (*e.g.*, a variant strain comprising a genetic modification encoding allele *lov*(T813K) comprising an enhanced/increased protein productivity phenotype, the variant strain further comprises a genetic modification of a gene encoding a MPG1, SFB3, SEB1, CRZ1, GAS1, TPS2 and/or SSB7 protein.

[0164] As used herein, a *Trichoderma sp.* "MPG1 protein" comprises an amino acid sequence of SEQ ID NO: 5 (described in International PCT Publication No. WO2012/145584, incorporated herein by reference in its entirety).

[0165] As used herein, a *Trichoderma sp.* "SEB1 protein" comprises an amino acid sequence of SEQ ID NO: 6 (described in International PCT Publication No. WO2012/145595, incorporated herein by reference in its entirety).

[0166] As used herein, a *Trichoderma sp.* "SFB3 protein" comprises an amino acid sequence of SEQ ID NO: 7, (described in International PCT Publication No. WO2012/027580, incorporated herein by reference in its entirety).

[0167] As used herein, a *Trichoderma sp.* "CRZ1 protein" comprises an amino acid sequence of SEQ ID NO: 8 (as described in International PCT Publication No. WO2012/145596, incorporated herein by reference in its entirety).

[0168] As used herein, a *Trichoderma sp.* "GAS1 protein" comprises an amino acid sequence of SEQ ID NO: 9 (as described in International PCT Publication Nos. WO2012/145596 and WO2012/145592, each incorporated herein by reference in their entirety).

[0169] As used herein, a *Trichoderma sp.* "TSP2 protein" comprises an amino acid sequence of SEQ ID NO: 10 (as described in International PCT Publication No. WO2012/145598, incorporated herein by reference in its entirety).

[0170] As used herein, a wild-type "*ssb7* gene" encodes a native "SSB7 protein", described in International PCT Application Serial No. PCT/US2019/27590, filed April 16, 2019. For example, a wild-type *T. reesei* "*ssb7* gene" (SEQ ID NO: 20) encodes a native "SSB7 protein" of SEQ ID NO: 21.

[0171] As used herein, "allele *ssb7(fs)*" (SEQ ID NO: 22) encodes a variant SSB7 protein of SEQ ID NO: 23, as described in International PCT Application Serial No. PCT/US2019/27590 (specifically incorporated herein by reference in its entirety).

II. FILAMENTOUS FUNGAL STRAINS COMPRISING ENHANCED PROTEIN PRODUCTIVITY PHENOTYPES

[0172] As generally set forth above, and further described in the Examples section below, certain embodiments of the disclosure are related to mutant and genetically modified (variant) strains of filamentous fungus derived from parental strains. More particularly, certain embodiments are related to mutant and genetically modified (variant) strains of filamentous fungus (and methods thereof), wherein such strains comprise enhanced protein productivity phenotypes, such as improved volumetric efficiencies, higher specific productivities, improved yield on carbon sources, reduced bioreactor (fermentor) operating costs and the like.

[0173] More particularly, as further described below in Example 1, a mutant (daughter) Trichoderma strain was identified, isolated and named "B7ms1-SF12", which (mutant) strain had a 40% higher protein yield on fed sugars in fermentors relative to the parental strain (B7ms1) from which it was derived (*e.g.*, see, **FIG. 1**). Genome sequence analysis and genetic analysis were subsequently performed on the mutant B7ms1-SF12 strain to identify one or more mutation(s) in a gene (or genes) therein, as being responsible for the observed (increased productivity) phenotype of the B7ms1-SF12 (mutant) strain. More specifically, the identified mutation was determined to alter the coding sequence of the (native) LOV protein of SEQ ID NO: 2, wherein a highly-conserved threonine (T) at amino acid (residue) position 813 (T813) of the (native) LOV protein (SEQ ID NO: 2) was mutated (substituted) to a lysine at position 813 (813K) of the (variant) LOV protein of SEQ ID NO: 4 (*i.e.*, a T813K substitution; *e.g.*, see **FIG. 2**).

[0174] A review of the scientific literature and related art indicate that the functional characterization of the LOV protein (SEQ ID NO: 2) and/or related LOV protein orthologues (*e.g.*, SEQ ID NOs: 11-17) have not been described in the literature for any organism. However, as generally set forth in Example 1 of the disclosure, a conserved domain analysis (NCBI) identified a (conserved) region distantly related to a glycosyl transferase family group 2 (pfam13632, E-value of 3.25x10-44, e.g., see "Glyco trans 2 3" in FIG. 2). For example, members of this (glycosyl transferase) family of prokaryotic proteins include putative glucosyltransferases, which are involved in bacterial capsule biosynthesis (PFAM). More particularly, a fungal protein with the (glycosyl transferase) family group 2 domain, "ZtGT2", orthologous to Trichoderma PID 79396 (*i.e.*, which is not LOV, PID 50212), is important for hyphal growth on solid surfaces (King *et al.*, 2017).

[0175] Most surprisingly, the LOV protein is widely conserved among filamentous fungi of Basidiomycetes and Ascomycete phyla. For example, the threonine (T) at position 813 of the (native) LOV protein (*i.e.*, which mutation to lysine (K) is beneficial as disclosed herein) is highly conserved (691/691; 100%) among the top BLASTp search results of the NCBI non-redundant database, within the Pezizomycotina subphylum, to which Trichoderma and most industrially relevant filamentous fungi belong. For example, a graphical representation from a Geneious multiple sequence alignment (www.geneious.com, Geneious 11.0, Biomatters Ltd.) of these 691 Pezizomycotina homologs is presented in **FIG. 2**. Likewise, among the top 1,000 BLASTp hits, amino acid residue position 813 was identical in 975/1000 (97.5%) hits (*i.e.*, residue T813) in a MUSCLE multiple sequence alignment (Geneious software package), wherein residue position 813 never occurs as a lysine (K) (*i.e.*, the substitution described herein resulting in enhanced protein productivity was not found in the top 1000 orthologs identified by BLASTp).

[0176] Thus, to further validate that the *lov*(T813K) allele was causative for the observed enhanced protein productivity phenotypes, the *lov*(T813K) mutation was introduced into a different *T. reesei* strain lineage named "T4" (*e.g.*, see TABLE 1), wherein the B7ms1 and T4 parental lineages are both mutagenized derivatives from different strain improvement programs of *Trichoderma* strain RL-P37 (Sheir-Neiss *et al.*, 1984; Montenecourt, 1987). For example, the T4 strain notably differs from the B7ms1 strain in that the T4 strain expresses its (endogenous) native cocktail of cellulases and comprises a *nik1*(M743T) mutation which increases total protein production (*e.g.*, see International PCT Publication No. WO2016/130523). Likewise, similar to the observed results of B7ms1 lineage set forth in Example 1, the T4 lineage (strains) comprising the *lov*(T813K) allele (*e.g.*, see Example 2 and Example 4), had on average a 34% higher total protein yield on fed sugars relative to otherwise isogenic strains comprising the wild-type *lov*(+) allele.

[0177] In addition, to further demonstrate utility of the *lov*(T813K) allele for improvement of fungal strain protein productivity phenotypes, Applicant integrated a *pyr2* marker at three other convenient genomic locations (*i.e.*, named sites A-C; also see Example 4, TABLE 1 and TABLE 2). For example, protein production by these different strains was evaluated in shake flasks, where in all such cases, the presence of the *lov*(T813K) mutation improved total protein production relative to the otherwise isogenic (parental) strains comprising the wild-type *lov*(+) allele.

III. MOLECULAR BIOLOGY

[0178] As generally described above, certain embodiments of the disclosure are related to mutant and genetically modified (variant) strains of filamentous fungus derived from parental strains. More particularly, certain embodiments are related to mutant and genetically modified (variant) strains of filamentous fungus (and methods thereof), wherein such strains comprise enhanced protein productivity phenotypes, such as improved volumetric efficiencies, higher specific productivities, improved yield on carbon sources, reduced bioreactor (fermentor) operating costs and the like.

[0179] Thus, certain embodiments of the disclosure are related to mutant (daughter) Trichoderma strains (*e.g.*, mutant B7ms1-SF12) derived from parental Trichoderma strains (*e.g.*, parent B7ms1), wherein the

mutant (daughter) Trichoderma strains comprise a mutant *lov* gene encoding a variant LOV protein comprising a lysine (K) residue at an amino acid position corresponding to amino acid residue position 813 of SEQ ID NO: 4.

[0180] Certain other embodiments of the disclosure are therefore related to genetically modified strains (hereinafter, "variant" strains) of filamentous fungus derived from parental strains described herein. For example, in certain embodiments, variant strains of filamentous fungus comprise a modified *lov* gene (or a modified polynucleotide sequence thereof) encoding a variant LOV protein comprising sequence homology to SEQ ID NO: 2 (*i.e.*, the native LOV protein sequence) and comprise a lysine (K) residue at an amino acid sequence position corresponding to amino acid position 813 of SEQ ID NO: 4.

[0181] Thus, in certain embodiments, variants strains of the disclosure comprise a modified *lov* gene (or a modified polynucleotide sequence thereof) encoding a LOV protein comprising at least about 50% to about 100% sequence homology to a LOV protein of SEQ ID NO: 2 (or SEQ ID NO: 4), and comprising a lysine (K) residue at an amino acid sequence position corresponding to amino acid position 813 of SEQ ID NO: 4. In certain embodiments, variants strains of the disclosure comprise a modified *lov* gene (or a modified polynucleotide sequence thereof) encoding a LOV protein comprising at least about 50% to about 100% sequence identity to a LOV protein of SEQ ID NO: 2 (or SEQ ID NO: 4), and comprising a lysine (K) residue at an amino acid sequence position corresponding to amino acid position 813 of SEQ ID NO: 4.

[0182] Thus, in other embodiments, variants strains of the disclosure comprise a modified *lov* gene (or a modified polynucleotide sequence thereof) encoding a LOV protein comprising at least about 50% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 55% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 60% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 65% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 70% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 75% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 80% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 85% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 90% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, at least 95% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, or up to and about 100% sequence homology or identity to SEQ ID NO: 2 or SEQ ID NO: 4, and comprising a lysine (K) residue at an amino acid sequence position corresponding to amino acid position 813 of SEQ ID NO: 4.

[0183] Thus, in certain embodiments, variant strains of filamentous fungus comprise at least a genetic modification which introduces allele *lov*(T813K) into the strain. For example, in certain embodiments, variant strains of filamentous fungus may comprise genetic modifications including, but is not limited to: (a) the introduction, substitution, or removal of one or more nucleotides in a gene (or ORF or polynucleotide thereof), or the introduction, substitution, or removal of one or more nucleotides in a regulatory element required for the transcription or translation of the gene, (b) a gene disruption, (c) a gene conversion, (d) a gene deletion, (e) a gene down-regulation, (f) specific mutagenesis and/or (g) random mutagenesis. In other embodiments, variant strains of filamentous fungus comprising allele *lov*(T813K) further comprises one or more genetic

modifications of a gene encoding NIK1 protein, an SSB7 protein, a MPG1 protein, SFB3 protein, SEB1 protein, CRZ1 protein, a TSP2 protein and/or GAS1 protein, as described herein.

[0184] Thus, in certain embodiments, variant strains of filamentous fungus comprising a genetic modification may be constructed by gene deletion to eliminate the expression/production of given gene product (*e.g.,* LOV(+), NIK1, SSB7, MPG1 SFB3, SEB1, CRZ1, TSP2, GAS1, endogenous (background) proteases, cellulases and the like). In other embodiments, variant strains of filamentous fungus comprising a genetic modification may be constructed by partial gene deletion to eliminate (or reduce) the expression/production of a given gene product. For example, in certain embodiments, modified filamentous fungal strains may comprise a partial deletion of a gene, wherein a partial deletion includes the partial deletion of any portion of the gene's coding sequence. For example, in certain embodiments, such variant strains do not express/produce the encoded protein, or such variant strains express/produce a reduced amount of the encoded protein (relative to the parental strain), wherein a "reduced" amount of the encoded protein can be measured, detected, assayed and the like as described herein.

[0185] Thus, as generally set forth and described in the Examples section, one skilled in the art may perform the following genetic modifications (and molecular biology methods thereof described in this section) and construct such (variant) filamentous fungus strains thereof, by reference to one or more amino acid sequences (SEQ ID NOs: 2, 4, 11-19) and/or nucleic acid sequences (SEQ ID NOs: 1 and 2) of the instant disclosure.

[0186] For example, gene deletion techniques enable the partial or complete removal of the gene, thereby eliminating or reducing expression/production of the protein, and/or thereby eliminating or reducing expression/production the encoded protein. In such methods, the deletion of the gene may be accomplished by homologous recombination using an integration plasmid/vector that has been constructed to contiguously contain the 5′ and 3′ regions flanking the gene. The contiguous 5′ and 3′ regions may be introduced into a filamentous fungal cell, for example, on an integrative plasmid/vector in association with a selectable marker to allow the plasmid to become integrated in the cell.

[0187] In other embodiments, a variant strain of filamentous fungus comprises genetic modification which disrupts or inactivates a gene encoding a protein of the disclosure. Exemplary methods of gene disruption/inactivation include disrupting any portion of a gene, including the polypeptide coding sequence (CDS), promoter, enhancer, or another regulatory element, which disruption includes substitutions, insertions, deletions, inversions, and combinations thereof and variations thereof. Thus, in certain embodiments, a variant strain of filamentous fungus is constructed by a gene disruption technique. A non-limiting example of a gene disruption technique includes inserting (integrating) into one or more of the genes of the disclosure an integrative plasmid containing a nucleic acid fragment homologous to the gene, which will create a duplication of the region of homology and incorporate (insert) vector DNA between the duplicated regions.

[0188] Thus, in certain other non-limiting examples, a gene disruption technique includes inserting into a gene an integrative plasmid containing a nucleic acid fragment homologous to the gene, which will create

a duplication of the region of homology and incorporate (insert) vector DNA between the duplicated regions, wherein the vector DNA inserted separates, *e.g.*, the promoter of the gene from the protein coding region, or interrupts (disrupts) the coding, or non-coding, sequence of the gene. Thus, a disrupting construct may be a selectable marker gene (*e.g., pyr2*; see Examples) accompanied by 5′ and 3′ regions homologous to the gene. The selectable marker enables identification of transformants containing the disrupted gene. Thus, in certain embodiments, gene disruption includes modification of control elements of the gene, such as the promoter, ribosomal binding site (RBS), untranslated regions (UTRs), codon changes, and the like.

[0189] In other embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by introducing, substituting, or removing one or more nucleotides in the gene, or a regulatory element required for the transcription or translation thereof. For example, nucleotides may be inserted or removed so as to result in the introduction of a stop codon, the removal of the start codon, or a frame-shift of the open reading frame (ORF). Such a modification may be accomplished by site-directed mutagenesis or PCR generated mutagenesis in accordance with methods known in the art (*e.g., see*, Botstein and Shortle, 1985; Lo *et al.,* 1985; Higuchi *et al.*, 1988; Shimada, 1996; Ho *et al.*, 1989; Horton *et al.*, 1989 and Sarkar and Sommer, 1990). Likewise, allele *lov*(T813K) comprising the T813K substitution described herein may be constructed by substituting nucleotides encoding the position 813 threonine of the wild-type *lov*(+) gene for nucleotides encoding the position 813 lysine of allele *lov*(T813K).

[0190] In other embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by the process of gene conversion (*e.g., see* Iglesias and Trautner, 1983). For example, in the gene conversion method, a nucleic acid sequence corresponding to the target gene is mutagenized *in vitro* to produce a defective nucleic acid sequence, which is then transformed into the parental cell to produce a variant cell comprising a defective gene. By homologous recombination, the defective nucleic acid sequence replaces the endogenous gene. It may be desirable that the defective gene or gene fragment also encodes a marker which may be used for selection of transformants containing the defective gene. For example, the defective gene may be introduced on a non-replicating or temperature-sensitive plasmid in association with a selectable marker. Selection for integration of the plasmid is effected by selection for the marker under conditions not permitting plasmid replication. Selection for a second recombination event leading to gene replacement is effected by examination of colonies for loss of the selectable marker and acquisition of the mutated gene (Perego, 1993).

[0191] In other embodiments, a variant strain of filamentous fungus is constructed by established anti-sense (gene-silencing) techniques, using a nucleotide sequence complementary to the nucleic acid sequence of the gene (Parish and Stoker, 1997). More specifically, expression of a gene by a filamentous fungus strain may be reduced (down-regulated) or eliminated by introducing a nucleotide sequence complementary to the nucleic acid sequence of the gene, which is transcribed in the cell and is capable of hybridizing to the mRNA produced in the cell. Under conditions allowing the complementary anti-sense nucleotide sequence to hybridize to the mRNA, the amount of protein translated is thus reduced or

eliminated. Such anti-sense methods include, but are not limited to RNA interference (RNAi), small interfering RNA (siRNA), microRNA (miRNA), antisense oligonucleotides, and the like, all of which are well known to the skilled artisan.

[0192] In other embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by random or specific mutagenesis using methods well known in the art, including, but not limited to, chemical mutagenesis (*see, e.g.,* Hopwood, 1970) and transposition (see, *e.g.,* Youngman *et al.*, 1983). Modification of the gene may be performed by subjecting the parental cell to mutagenesis and screening for mutant cells in which expression of the gene has been reduced or eliminated. For example, one of skill in the art may readily adapt and/or modify the screening methods set forth in the Example section herewith to identify such (mutagenized) variant strains of filamentous fungus cells comprising a reduced viscosity phenotype.

[0193] The mutagenesis, which may be specific or random, may be performed, for example, by use of a suitable physical or chemical mutagenizing agent, use of a suitable oligonucleotide, or subjecting the DNA sequence to PCR generated mutagenesis. Furthermore, the mutagenesis may be performed by use of any combination of these mutagenizing methods. Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), N-methyl-N'-nitrosoguanidine (NTG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulphite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically performed by incubating the parental cell to be mutagenized in the presence of the mutagenizing agent of choice under suitable conditions, and selecting for mutant cells exhibiting reduced or no expression of the gene.

[0194] For example, such genetic modifications in the one or more of the genes disclosed herein can reduce the efficiency of the gene's promoter, reduce the efficiency of an enhancer, interfere with the splicing or editing of the gene's mRNA, interfere with the translation of the gene's mRNA, introduce a stop codon into the gene's-coding sequence to prevent the translation of full-length protein, change the coding sequence of the protein to produce a less active or inactive protein, reduce the protein interaction with other nuclear protein components, change the coding sequence of the protein to produce a less stable protein, or target the protein for destruction, or cause the protein to misfold or be incorrectly modified (*e.g.,* by glycosylation), or interfere with cellular trafficking of the protein.

[0195] In certain other embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by means of site specific gene editing techniques. For example, in certain embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by use of transcriptional activator like endonucleases (TALENs), zinc-finger endonucleases (ZFNs), homing (mega) endonuclease and the like. More particularly, the portion of the gene to be modified (*e.g.,* a coding region, a non-coding region, a leader sequence, a pro-peptide sequence, a signal sequence, a transcription terminator, a transcriptional activator, or other regulatory elements required for expression of the coding region) is subjected genetic modification by means of ZFN gene editing, TALEN gene editing, homing (mega)

endonuclease and the like, which modification methods are well known and available to one skilled in the art.

[0196] Thus, in certain embodiments, a variant strain of filamentous fungus is constructed (*i.e.*, genetically modified) by means of CRISPR/Cas9 editing. More specifically, compositions and methods for fungal genome modification by CRISPR/Cas9 systems are described and well known in the art (*e.g., see*, International PCT Publication Nos: WO2016/100571, WO2016/100568, WO2016/100272, WO2016/100562 and the like). For example, a gene encoding a native LOV protein can be genetically modified by means of nucleic acid guided endonucleases, that find their target DNA by binding either a guide RNA (*e.g.,* Cas9) or a guide DNA (*e.g.,* NgAgo), which recruits the endonuclease to the target sequence on the DNA, wherein the endonuclease can generate a single or double stranded break in the DNA. This targeted DNA break becomes a substrate for DNA repair, and can recombine with a provided editing template to disrupt or delete the gene. For example, the gene encoding the nucleic acid guided endonuclease (*e.g.*, a Cas9 from *S. pyogenes*, or a codon optimized gene encoding the Cas9 nuclease) is operably linked to a promoter active in the filamentous fungal cell and a terminator active in filamentous fungal cell, thereby creating a filamentous fungal Cas9 expression cassette. Likewise, one or more target sites unique to the gene of interest are readily identified by a person skilled in the art.

[0197] For example, to build a DNA construct encoding a gRNA-directed to a target site within the gene of interest, the variable targeting domain (VT) will comprise nucleotides of the target site which are 5′ of the (PAM) proto-spacer adjacent motif (TGG), which nucleotides are fused to DNA encoding the Cas9 endonuclease recognition domain for *S. pyogenes* Cas9 (CER). The combination of the DNA encoding a VT domain and the DNA encoding the CER domain thereby generate a DNA encoding a gRNA. Thus, a filamentous fungal expression cassette for the gRNA is created by operably linking the DNA encoding the gRNA to a promoter active in filamentous fungal cells and a terminator active in filamentous fungal cells.

[0198] In certain embodiments, the DNA break induced by the endonuclease is repaired/replaced with an incoming sequence. For example, to precisely repair the DNA break generated by the Cas9 expression cassette and the gRNA expression cassette described above, a nucleotide editing template is provided, such that the DNA repair machinery of the cell can utilize the editing template. For example, about 500bp 5′ of targeted gene can be fused to about 500bp 3′ of the targeted gene to generate an editing template, which template is used by the filamentous fungal host's machinery to repair the DNA break generated by the RGEN (RNA-guided endonuclease).

[0199] The Cas9 expression cassette, the gRNA expression cassette and the editing template can be co-delivered to filamentous fungal cells using many different methods (*e.g.,* protoplast fusion, electroporation, natural competence, or induced competence). The transformed cells are screened by PCR, by amplifying the target locus with a forward and reverse primer. These primers can amplify the wild-type locus or the modified locus that has been edited by the RGEN. These fragments are then sequenced using a sequencing primer to identify edited colonies.

IV. PROTEINS OF INTEREST

[0200] As briefly stated in the preceding section, the present strains and methods find use in the production of commercially important proteins in submerged cultures of filamentous fungi. A protein of interest (POI) of the instant disclosure can be any endogenous or heterologous protein, and it may be a variant of such a POI. The protein can contain one or more disulfide bridges or is a protein whose functional form is a monomer or a multimer, *i.e.*, the protein has a quaternary structure and is composed of a plurality of identical (homologous) or non-identical (heterologous) subunits, wherein the POI or a variant POI thereof is preferably one with properties of interest.

[0201] In certain embodiments, a variant strain of filamentous fungus exhibits an increased protein titer relative to the (unmodified) parental strain, wherein protein titer is defined as the amount of protein *per* volume (g/L). For example, titers can be measured by methods known in the art (*e.g.,* ELISA, HPLC, Bradford assay, LC/MS and the like). Thus, in certain embodiments, a variant strain of filamentous fungus comprises a protein titer increase of at least about 0.1 %, at least about 1%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10% or more as compared to the unmodified (parental) cell.

[0202] In certain embodiments, a variant strain of filamentous fungus exhibits an increased volumetric productivity relative to the (unmodified) parental strain, wherein volumetric productivity is defined as the amount of protein produced (g) during the fermentation *per* nominal volume (L) of the bioreactor *per* total fermentation time (h). For example, volumetric productivities can be measured by methods know in the art (*e.g.,* ELISA, HPLC, Bradford assay, LC/MS and the like). Thus, in certain embodiments, a variant strain of filamentous fungus comprises a volumetric productivity increase of at least about 0.1 %, at least about 1%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10% or more as compared to the unmodified (parental) cell.

[0203] In certain other embodiments, a variant strain of filamentous fungus exhibits an increased total protein yield, wherein total protein yield is defined as the amount of protein produced (g) *per* gram of carbohydrate fed, relative to the (unmodified) parental strain. Thus, as used herein, total protein yield (g/g) may be calculated using the following equation:

$$Y_f = T_p/T_c$$

[0204] wherein "Y_f" is total protein yield (g/g), "T_p" is the total protein produced during the fermentation (g) and "T_c" is the total carbohydrate (g) fed during the fermentation (bioreactor) run. In certain embodiments, the increase in total protein yield of the modified strain (*i.e.*, relative to the parental strain) is an increase of at least about 0.1 %, at least about 1%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10% or more as compared to the unmodified (parental) cell.

[0205] Total protein yield may also be described as carbon conversion efficiency/carbon yield, for example, as in the percentage (%) of carbon fed that is incorporated into total protein. Thus, in certain embodiments, a variant strain of filamentous fungus comprises an increased carbon conversion efficiency (*e.g.*, an increase in the percentage (%) of carbon fed that is incorporated into total protein), relative to the (unmodified) parental

strain. In certain embodiments, the increase in carbon conversion efficiency of the modified strain (*i.e.*, relative to the parental strain) is an increase of at least about 0.1 %, at least about 1%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10% or more as compared to the unmodified (parental) cell.

[0206] In certain embodiments, a variant strain of filamentous fungus exhibits an increased specific productivity (Qp) of a POI relative the (unmodified) parental strain. For example, the detection of specific productivity (Qp) is a suitable method for evaluating rate of protein production. The specific productivity (Qp) can be determined using the following equation:

"Qp = gP/gDCW•hr"

wherein, "gP" is grams of protein produced in the tank; "gDCW" is grams of dry cell weight (DCW) in the tank and "hr" is fermentation time in hours from the time of inoculation, which includes the time of production as well as growth time. Thus, in certain embodiments, a variant strain of filamentous fungus comprises a specific productivity (Qp) increase of at least about 0.1 %, at least about 1%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10% or more as compared to the unmodified (parental) cell.

[0207] In certain embodiments, a POI or a variant POI thereof is selected from the group consisting of acetyl esterases, aminopeptidases, amylases, arabinases, arabinofuranosidases, carbonic anhydrases, carboxypeptidases, catalases, cellulases, chitinases, chymosins, cutinases, deoxyribonucleases, epimerases, esterases, α-galactosidases, β-galactosidases, α-glucanases, glucan lyases, endo-β-glucanases, glucoamylases, glucose oxidases, α-glucosidases, β-glucosidases, glucuronidases, glycosyl hydrolases, hemicellulases, hexose oxidases, hydrolases, invertases, isomerases, laccases, ligases, lipases, lyases, mannanases, mannosidases, oxidases, oxidoreductases, pectate lyases, pectin acetyl esterases, pectin depolymerases, pectin methyl esterases, pectinolytic enzymes, perhydrolases, polyol oxidases, peroxidases, phenoloxidases, phytases, polygalacturonases, proteases, peptidases, rhamno-galacturonases, ribonucleases, transferases, transport proteins, transglutaminases, xylanases, hexose oxidases, and combinations thereof.

[0208] In certain embodiments, a POI or a variant POI thereof is selected from an Enzyme Commission (EC) Number selected from the group consisting of EC 1, EC 2, EC 3, EC 4, EC 5 or EC 6.

[0209] For example, in certain embodiments a POI is an oxidoreductase enzyme, including, but not limited to, an EC1 (oxidoreductase) enzyme selected from EC 1.10.3.2 (*e.g.,* a laccase), EC 1.10.3.3 (*e.g.,* L-ascorbate oxidase), EC 1.1.1.1 (*e.g.,* alcohol dehydrogenase), EC 1.11.1.10 (*e.g.,* chloride peroxidase), EC 1.11.1.17 (*e.g.,* peroxidase), EC 1.1.1.27 (*e.g.,* L-lactate dehydrogenase), EC 1.1.1.47 (*e.g.,* glucose 1-dehydrogenase), EC 1.1.3.X (*e.g.,* glucose oxidase), EC 1.1.3.10 (*e.g.,* pyranose oxidase), EC 1.13.11.X (*e.g.,* dioxygenase), EC 1.13.11.12 (*e.g.,* lineolate 13S-lipozygenase), EC 1.1.3.13 (*e.g.,* alcohol oxidase), EC 1.14.14.1 (*e.g.,* monooxygenase), EC 1.14.18.1 (*e.g.,* monophenol monooxigenase), EC 1.15.1.1 (*e.g.,* superoxide dismutase), EC 1.1.5.9 (formerly EC 1.1.99.10, *e.g.,* glucose dehydrogenase), EC 1.1.99.18 (*e.g.,* cellobiose dehydrogenase), EC 1.1.99.29 (*e.g.,* pyranose dehydrogenase), EC 1.2.1.X (*e.g.,* fatty acid

reductase), EC 1.2.1.10 (*e.g.*, acetaldehyde dehydrogenase), EC 1.5.3.X (*e.g.*, fructosyl amine reductase), EC 1.8.1.X (*e.g.*, disulfide reductase) and EC 1.8.3.2 (*e.g.*, thiol oxidase).

[0210] In certain embodiments a POI is a transferase enzyme, including, but not limited to, an EC 2 (transferase) enzyme selected from EC 2.3.2.13 (*e.g.*, transglutaminase), EC 2.4.1.X (*e.g.*, hexosyltransferase), EC 2.4.1.40 (*e.g.*, alternasucrase), EC 2.4.1.18 (*e.g.*, 1,4 alpha-glucan branching enzyme), EC 2.4.1.19 (*e.g.*, cyclomaltodextrin glucanotransferase), EC 2.4.1.2 (*e.g.*, dextrin dextranase), EC 2.4.1.20 (*e.g.*, cellobiose phosphorylase), EC 2.4.1.25 (*e.g.*, 4-alpha-glucanotransferase), EC 2.4.1.333 (*e.g.*, 1,2-beta-oligoglucan phosphor transferase), EC 2.4.1.4 (*e.g.*, amylosucrase), EC 2.4.1.5 (*e.g.*, dextransucrase), EC 2.4.1.69 (*e.g.*, galactoside 2-alpha-L-fucosyl transferase), EC 2.4.1.9 (*e.g.*, inulosucrase), EC 2.7.1.17 (*e.g.*, xylulokinase), EC 2.7.7.89 (formerly EC 3.1.4.15, *e.g.*, [glutamine synthetase]-adenylyl-L-tyrosine phosphorylase), EC 2.7.9.4 (*e.g.*, alpha glucan kinase) and EC 2.7.9.5 (*e.g.*, phosphoglucan kinase).

[0211] In other embodiments a POI is a hydrolase enzyme, including, but not limited to, an EC 3 (hydrolase) enzyme selected from EC 3.1.X.X (*e.g.*, an esterase), EC 3.1.1.1 (*e.g.*, pectinase), EC 3.1.1.14 (*e.g.*, chlorophyllase), EC 3.1.1.20 (*e.g.*, tannase), EC 3.1.1.23 (*e.g.*, glycerol-ester acylhydrolase), EC 3.1.1.26 (*e.g.*, galactolipase), EC 3.1.1.32 (*e.g.*, phospholipase A1), EC 3.1.1.4 (*e.g.*, phospholipase A2), EC 3.1.1.6 (*e.g.*, acetylesterase), EC 3.1.1.72 (*e.g.*, acetylxylan esterase), EC 3.1.1.73 (*e.g.*, feruloyl esterase), EC 3.1.1.74 (*e.g.*, cutinase), EC 3.1.1.86 (*e.g.*, rhamnogalacturonan acetylesterase), EC 3.1.1.87 (*e.g.*, fumosin B1 esterase), EC 3.1.26.5 (*e.g.*, ribonuclease P), EC 3.1.3.X (*e.g.*, phosphoric monoester hydrolase), EC 3.1.30.1 (*e.g.*, Aspergillus nuclease S1), EC 3.1.30.2 (*e.g.*, *Serratia marcescens* nuclease), EC 3.1.3.1 (*e.g.*, alkaline phosphatase), EC 3.1.3.2 (*e.g.*, acid phosphatase), EC 3.1.3.8 (*e.g.*, 3-phytase), EC 3.1.4.1 (*e.g.*, phosphodiesterase I), EC 3.1.4.11 (*e.g.*, phosphoinositide phospholipase C), EC 3.1.4.3 (*e.g.*, phospholipase C), EC 3.1.4.4 (*e.g.*, phospholipase D), EC 3.1.6.1 (*e.g.*, arylsufatase), EC 3.1.8.2 (*e.g.*, diisopropyl-fluorophosphatase), EC 3.2.1.10 (*e.g.*, oligo-1,6-glucosidase), EC 3.2.1.101 (*e.g.*, mannan endo-1,6-alpha-mannosidase), EC 3.2.1.11 (*e.g.*, alpha-1,6-glucan-6-glucanohydrolase), EC 3.2.1.131 (*e.g.*, xylan alpha-1,2-glucuronosidase), EC 3.2.1.132 (*e.g.*, chitosan N-acetylglucosaminohydrolase), EC 3.2.1.139 (*e.g.*, alpha-glucuronidase), EC 3.2.1.14 (*e.g.*, chitinase), EC 3.2.1.151 (*e.g.*, xyloglucan-specific endo-beta-1,4-glucanase), EC 3.2.1.155 (*e.g.*, xyloglucan-specific exo-beta-1,4-glucanase), EC 3.2.1.164 (*e.g.*, galactan endo-1,6-beta-galactosidase), EC 3.2.1.17 (*e.g.*, lysozyme), EC 3.2.1.171 (*e.g.*, rhamnogalacturonan hydrolase), EC 3.2.1.174 (*e.g.*, rhamnogalacturonan rhamnohydrolase), EC 3.2.1.2 (*e.g.*, beta-amylase), EC 3.2.1.20 (*e.g.*, alpha-glucosidase), EC 3.2.1.22 (*e.g.*, alpha-galactosidase), EC 3.2.1.25 (*e.g.*, beta-mannosidase), EC 3.2.1.26 (*e.g.*, beta-fructofuranosidase), EC 3.2.1.37 (*e.g.*, xylan 1,4-beta-xylosidase), EC 3.2.1.39 (*e.g.*, glucan endo-1,3-beta-D-glucosidase), EC 3.2.1.40 (*e.g.*, alpha-L-rhamnosidase), EC 3.2.1.51 (*e.g.*, alpha-L-fucosidase), EC 3.2.1.52 (*e.g.*, beta-N-Acetylhexosaminidase), EC 3.2.1.55 (*e.g.*, alpha-N-arabinofuranosidase), EC 3.2.1.58 (*e.g.*, glucan 1,3-beta-glucosidase), EC 3.2.1.59 (*e.g.*, glucan endo-1,3-alpha-glucosidase), EC 3.2.1.67 (*e.g.*, galacturan 1,4-alpha-galacturonidase), EC 3.2.1.68 (*e.g.*, isoamylase), EC 3.2.1.7 (*e.g.*, 1-beta-D-fructan fructanohydrolase), EC

3.2.1.74 (*e.g.*, glucan 1,4-β-glucosidase), EC 3.2.1.75 (*e.g.*, glucan endo-1,6-beta-glucosidase), EC 3.2.1.77 (*e.g.*, mannan 1,2-(1,3)-alpha-mannosidase), EC 3.2.1.80 (*e.g.*, fructan beta-fructosidase), EC 3.2.1.82 (*e.g.*, exo-poly-alpha-galacturonosidase), EC 3.2.1.83 (*e.g.*, kappa-carrageenase), EC 3.2.1.89 (*e.g.*, arabinogalactan endo-1,4-beta-galactosidase), EC 3.2.1.91 (*e.g.*, cellulose 1,4-beta-cellobiosidase), EC 3.2.1.96 (*e.g.*, mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase), EC 3.2.1.99 (*e.g.*, arabinan endo-1,5-alpha-L-arabinanase), EC 3.4.X.X (*e.g.*, peptidase), EC 3.4.11.X (*e.g.*, aminopeptidase), EC 3.4.11.1 (*e.g.*, leucyl aminopeptidase), EC 3.4.11.18 (*e.g.*, methionyl aminopeptidase), EC 3.4.13.9 (*e.g.*, Xaa-Pro dipeptidase), EC 3.4.14.5 (*e.g.*, dipeptidyl-peptidase IV), EC 3.4.16.X (*e.g.*, serine-type carboxypeptidase), EC 3.4.16.5 (*e.g.*, carboxypeptidase C), EC 3.4.19.3 (*e.g.*, pyroglutamyl-peptidase I), EC 3.4.21.X (*e.g.*, serine endopeptidase), EC 3.4.21.1 (*e.g.*, chymotrypsin), EC 3.4.21.19 (*e.g.*, glutamyl endopeptidase), EC 3.4.21.26 (*e.g.*, prolyl oligopeptidase), EC 3.4.21.4 (*e.g.*, trypsin), EC 3.4.21.5 (*e.g.*, thrombin), EC 3.4.21.63 (*e.g.*, oryzin), EC 3.4.21.65 (*e.g.*, thermomycolin), EC 3.4.21.80 (*e.g.*, streptogrisin A), EC 3.4.22.X (*e.g.*, cysteine endopeptidase), EC 3.4.22.14 (*e.g.*, actinidain), EC 3.4.22.2 (*e.g.*, papain), EC 3.4.22.3 (*e.g.*, ficain), EC 3.4.22.32 (*e.g.*, stem bromelain), EC 3.4.22.33 (*e.g.*, fruit bromelain), EC 3.4.22.6 (*e.g.*, chymopapain), EC 3.4.23.1 (*e.g.*, pepsin A), EC 3.4.23.2 (*e.g.*, pepsin B), EC 3.4.23.22 (*e.g.*, endothiapepsin), EC 3.4.23.23 (*e.g.*, mucorpepsin), EC 3.4.23.3 (*e.g.*, gastricsin), EC 3.4.24.X (*e.g.*, metalloendopeptidase), EC 3.4.24.39 (*e.g.*, deuterolysin), EC 3.4.24.40 (*e.g.*, serralysin), EC 3.5.1.1 (*e.g.*, asparaginase), EC 3.5.1.11 (*e.g.*, penicillin amidase), EC 3.5.1.14 (*e.g.*, N-acyl-aliphatic-L-amino acid amidohydrolase), EC 3.5.1.2 (*e.g.*, L-glutamine amidohydrolase), EC 3.5.1.28 (*e.g.*, N-acetylmuramoyl-L-alanine amidase), EC 3.5.1.4 (*e.g.*, amidase), EC 3.5.1.44 (*e.g.*, protein-L-glutamine amidohydrolase), EC 3.5.1.5 (*e.g.*, urease), EC 3.5.1.52 (*e.g.*, peptide-N(4)-(N-acetyl-beta-glucosaminyl)asparagine amidase), EC 3.5.1.81 (*e.g.*, N-Acyl-D-amino-acid deacylase), EC 3.5.4.6 (*e.g.*, AMP deaminase) and EC 3.5.5.1 (*e.g.*, nitrilase).

[0212] In other embodiments a POI is a lyase enzyme, including, but not limited to, an EC 4 (lyase) enzyme selected from EC 4.1.2.10 (*e.g.*, mandelonitrile lyase), EC 4.1.3.3 (*e.g.*, N-acetylneuraminate lyase), EC 4.2.1.1 (*e.g.*, carbonate dehydratase), EC 4.2.2.- (*e.g.*, rhamnogalacturonan lyase), EC 4.2.2.10 (*e.g.*, pectin lyase), EC 4.2.2.22 (*e.g.*, pectate trisaccharide-lyase), EC 4.2.2.23 (*e.g.*, rhamnogalacturonan endolyase) and EC 4.2.2.3 (*e.g.*, mannuronate-specific alginate lyase).

[0213] In certain other embodiments a POI is an isomerase enzyme, including, but not limited to, an EC 5 (isomerase) enzyme selected from EC 5.1.3.3 (*e.g.*, aldose 1-epimerase), EC 5.1.3.30 (*e.g.*, D-psicose 3-epimerase), EC 5.4.99.11 (*e.g.*, isomaltulose synthase) and EC 5.4.99.15 (*e.g.*, (1→4)-α-D-glucan 1-α-D-glucosylmutase).

[0214] In yet other embodiments, a POI is a ligase enzyme, including, but not limited to, an EC 6 (ligase) enzyme selected from EC 6.2.1.12 (*e.g.*, 4-coumarate: coenzyme A ligase) and EC 6.3.2.28 (*e.g.*, L-amino-acid alpha-ligase).

[0215] These and other aspects and embodiments of the present strains and methods will be apparent to the skilled person in view of the present description and the following Examples.

EXAMPLES

[0216] Certain aspects of the present disclosure may be further understood in light of the following examples, which should not be construed as limiting. Modifications to materials and methods will be apparent to those skilled in the art.

EXAMPLE 1

IDENTIFICATION OF THE *lov* GENE AS BEING RESPONSIBLE FOR PROTEIN PRODUCTION INCREASES IN FILAMENTOUS FUNGI

A. OVERVIEW

[0217] In the present example, a reduced viscosity Trichoderma strain named "B7ms1", expressing a glucoamylase (GA) construct, was specifically evolved towards reducing the propensity to pellet under shake flask conditions. More particularly, a mutant (strain) of B7ms1 was identified, isolated and named "B7ms1-SF12". The (mutant) B7ms1-SF12 strain had a 40% higher protein yield on fed sugars in fermentors relative to the B7ms1 (parental) strain (*e.g.*, see, **FIG. 1**).

[0218] Thus, as described herein, genome sequence analysis and genetic analysis were performed on the B7ms1-SF12 (mutant) strain, to identify one or more mutation(s) in a gene (or genes) therein, as being responsible for the observed (increased productivity) phenotype of the B7ms1-SF12 (mutant) strain. As disclosed herein, the identified mutation was determined to alter the coding sequence of the (native) LOV protein, wherein a highly-conserved threonine (T) at amino acid (residue) position 813 (T813) of the (native) LOV protein (SEQ ID NO: 2) was changed (substituted) to a lysine at position 813 (813K) of the (variant) LOV protein (SEQ ID NO: 4, *i.e.*, a T813K substitution) (see **FIG. 2**).

B. EVOLUTION STRATEGY

[0219] Filamentous fungal species (*e.g., Trichoderma sp., Aspergillus sp., Fusarium sp., Penicillium sp., Chrysosporium sp., Cephalosporium sp., Talaromyces sp., Geosmithia sp., Neurospora sp., Myceliophthora sp.* and the like) are aerobic fungi that generally produce a thick, viscous fermentation broth when used in commercial/industrial fermentations. The high fermentation (broth) viscosity typically reduces dissolved oxygen (DO) transfer, thereby limiting the amount of cell mass and reducing the volumetric productivity that can be achieved in such aerobic filamentous fungus fermentations. For example, isolation of reduced viscosity filamentous fungal mutants has resulted in mutant strains/cells that produce lower viscosity fermentation broths, wherein fermentations using such reduced viscosity strains/cells can utilize more cell mass leading to increases in protein productivity (*e.g.*, see Applicant's International PCT Publication Nos. WO2012/145584, WO2012/027580, WO2012/145595, WO2012/145596 and WO2012/145592). More particularly, as Applicant began to combine certain reduced viscosity mutations to generate further reduced viscosity strains thereof, it was surprisingly observed that certain viscosity (reducing) combinations (*i.e,* mutations) thereof seemed to have a greater propensity to form mycelial pellets in shake flask assays.

[0220] More particularly, as contemplated herein, in the event that the observed increased propensity to form mycelial pellets in shake flasks may scale to fermentors, Applicant sought to mitigate the aforementioned pelting phenotype by means of directed evolution experimentation. For example, without wishing to be bound by any particular theory, mechanism or mode of action, it is contemplated herein that such propensity to form mycelial pellets may negatively interfere with protein production and/or the downstream processing of such proteins. Applicant therefore rationalized that "evolved" mutants of these strains could be generated, screened and isolated, wherein such evolved strains comprise a reduced propensity for mycelial pelleting in shake flasks, fermentors, bioreactors and the like.

[0221] For example, cultures of Trichoderma strain B7ms1 were grown for 12-24 hours in shake flasks containing complete complex media, then passed through a 70 micrometer (μm) sieve. The flow through, depending on the cell density, was used to inoculate fresh culture media, or the whole of the flow through was transferred to a fresh shake flask. After incubation, the process was repeated serially for weeks, occasionally taking samples that were then plated to isolate possible mutants. Isolates were (individually) visually screened for gross changes in shake flask morphology (*e.g.*, changes in shake flask morphology including, but not limited to, pelleting or mycelia chunkiness). Those isolates with more homogeneous growth in shake flasks were additionally tested for protein production in shake flasks, and then in fermentors.

C. Isolation and Characterization of *Trichoderma* Mutant Strain B7ms1-SF12

[0222] In such an evolution scheme set forth above, Applicant isolated a mutant strain named "B7ms1-SF12", wherein the (mutant) B7ms1-SF12 strain demonstrated higher protein production in shake flasks, relative to the (parental) B7ms1 strain (*data not shown*). Thus, independent of any influence on the aforementioned pelleting phenotype, any mutation(s) enhancing the protein production phenotype of the host strain is of particular value in its potential to reduce fermentation/protein production costs.

[0223] Thus, the (parental) B7ms1 strain and the spontaneous mutant (daughter) strain B7ms1-SF12 were assayed for total protein production in fermentors. These strains were grown under identical conditions in submerged (liquid culture), and their total protein yield on fed sugars compared in 14 L fermentors. As presented in **FIG. 1**, the mutant B7ms1-SF12 strain (**FIG. 1**; black line/black squares) showed a 44% increased improvement in yield on fed sugars vis-à-vis the (parental) B7ms1 strain (**FIG. 1**; grey line/grey squares).

[0224] Briefly, spores of each strain were added separately to 500 mL of medium in a 3 L flask with both side and bottom baffles. The cultures were grown in a minimal medium for 48 hours at 34°C in a shaking incubator. After 48 hours, the contents of each flask were added separately to 14 L fermentors containing 9.5 L of medium containing 4.7 g/L KHPO, 1.0 g/L $MgSO_7HO$, 4.3 g/L (NH)SO and 2.5 mL/L of 400X *T. reesei* trace elements solution (citric Acid (anhydrous), 175 g/L; $FeSO_4 \cdot 7\ H_2O$, 200 g/L, $ZnSO_4 \cdot 7\ H_2O$, 16 g/L, $CuSO_4 \cdot 5\ H_2O$, 3.2 g/L; $MnSO_4H_2O$, 1.4 g/L; H_3BO_3, 0.8 g/L.). These components were heat sterilized together at 121°C for 30 minutes. A solution of 60% glucose and 0.48% $CaCl_2 \cdot 2\ H_2O$ was

separately autoclaved, cooled, and added to the fermentor to a final concentration of 75 g/L glucose and 0.6 g/L $CaCl_2 \cdot 2\ H_2O$. The medium was adjusted to pH 3.5 with 28% NH and the temperature was maintained at 34°C during the growth period. Once glucose was exhausted, the temperature was dropped to 28°C, and the cultures were fed glucose-sophorose. The dry cell weight (DCW), total protein concentrations and other parameters were measured, and specific total protein production rates and yield on fed sugars were calculated.

D. IDENTIFICATION OF THE CAUSATIVE *lov*(T813K) MUTATION IN TRICHODERMA MUTANT STRAIN B7MS1-SF12

[0225] Applicant sequenced the genomes of the (parental) B7ms1 strain and the (mutant) B7ms1-SF12 (daughter) strain, leading to the identification of two (2) new mutations predicted to alter a coding sequence in the B7ms1-SF12 genome, either of which mutations alone (or in combination), could have been necessary for the observed protein productivity (improvement) phenotypes. More particularly, to determine/identify which of these mutations were of importance, complementation analysis was used by transforming the (mutant) B7ms1-SF12 strain with DNA encoding each of the wild-type loci. For example, only one locus when transformed complemented the mutant phenotype, which locus was called *lov*. The *lov* gene encodes the predicted protein PID 50212 (SEQ ID NO: 2), wherein in the (mutant) B7ms1-SF12 strain the *lov* mutant (allele) comprises a single nucleotide change of G (guanine) to T (thymine) at 425393 on Scaffold 16, resulting in amino acid substitution T813K in the encoded LOV (variant) protein (SEQ ID NO: 4), which mutant allele is referred to herein as allele "*lov*(T813K)" and the native (wild-type) allele is referred to herein as "*lov*(+)".

[0226] Functional characterization of the LOV protein or LOV protein orthologues, has not been described in the literature for any organism. As described herein, a conserved domain analysis (NCBI) identified a region distantly related to a glycosyl transferase family group 2 (pfam13632, E-value of 3.25x10-44, e.g., see "Glyco trans 2 3" in FIG. 2). For example, members of this (glycosyl transferase) family of prokaryotic proteins include putative glucosyltransferases, which are involved in bacterial capsule biosynthesis (PFAM). More particularly, a fungal protein with the (glycosyl transferase) family group 2 domain, "ZtGT2", orthologous to Trichoderma PID 79396 (*i.e.*, which is not LOV, PID 50212), is important for hyphal growth on solid surfaces (King *et al.*, 2017).

[0227] Most surprisingly, the LOV protein is widely conserved among filamentous fungi of Basidiomycetes and Ascomycete phyla. For example, the threonine (T) at position 813 of the (native) LOV protein (*i.e.*, which mutation to lysine (K) is beneficial as disclosed herein) is highly conserved (691/691; 100%) among the top BLASTp search results of the NCBI non-redundant database, within the Pezizomycotina subphylum, to which Trichoderma and most industrially relevant filamentous fungi belong. A graphical representation from a Geneious multiple sequence alignment (www.geneious.com, Geneious 11.0, Biomatters Ltd.) of these 691 Pezizomycotina homologs is presented in **FIG. 2**. Likewise, among the top 1,000 BLASTp hits, residue position 813 was identical in 975/1000 (97.5%) hits (*i.e.*, T813)

in a MUSCLE multiple sequence alignment (Geneious software package), wherein residue 813 never occurs as a lysine (K) (i.e., the substitution described herein resulting in enhanced protein productivity was not found in the top 1000 orthologs identified by BLASTp). While a conserved position and mutation is useful for improving fungal protein production, the T813 residue is not within any of the regions previously annotated in Genebank (**FIG. 2**). The LOV protein is the only predicted member of this class of proteins in the Trichoderma genome, whereas many genomes have more than one.

EXAMPLE 2

TARGETED INTRODUCTION OF THE *lov*(T813K) ALLELE IN REDUCED VISCOSITY STRAINS INCREASES PROTEIN PRODUCTIVITY IN FERMENTORS

A. OVERVIEW

[0228] To further validate that the *lov*(T813K) allele was causative for the observed protein productivity improvements, the *lov*(T813K) mutation was introduced into a different *T. reesei* strain lineage (herein named "T4"). B7ms1 (Example 1) and T4 lineages are both mutagenized derivatives from different strain improvement programs of *Trichoderma* strain RL-P37 (Sheir-Neiss *et al.*, 1984; Montenecourt, 1987). The T4 strain notably differs from B7ms1 in that the T4 strain expresses the native cocktail of cellulases and contains a *nik1*(M743T) mutation that increases total protein production (US20180037919). As before with the B7ms1 linage (Example 1), reduced viscosity double mutants (Δ*mpg1*; Δ*seb1*) were developed in the T4 lineage both with or without the presence of the *lov*(T813K) allele. For example, in fermentors, the T4 Δ*mpg1*; Δ*seb1* strains comprising the *lov*(T813K) allele had on average, a 37% higher total protein yield on fed sugars.

B. CONSTRUCTION OF PLASMIDS COMPRISING *lov* DISRUPTION CASSETTES WITH EITHER *pyr4* OR *pyr2* SELECTION MARKERS

[0229] The Trichoderma *lov* disruption cassette plasmids were prepared using standard molecular biology procedures, wherein one of skill in the relevant art may readily recreate this plasmid from the information disclosed herein. The plasmid included a DNA sequence having a 1.6 Kb homology box identical to the DNA sequence corresponding to Scaffold 16, 426947 to 425393 (Left Flank). The last nucleotide of the Left Flank introduced a single nucleotide G to T mutation, corresponding to the mutation identified in the (mutant) B7ms1-SF12 strain (Example 1). Also included within the plasmid was a DNA sequence having a 1.5 Kb homology box corresponding to the DNA sequence identical to Scaffold 16, 425392 to 423880 (Right Flank). These sequences were designed to target the *lov* gene and replace the nucleotide of the genome between the Left and Right Flanks (Scaffold 16, 425393) with the intervening cassette sequences.

[0230] These intervening (cassette) sequences included either a *pyr4* selection marker from *Trichoderma reesei* (pRATT308) or a *pyr2* selection marker from *Trichoderma atroviride* (pRATT312). Immediately downstream of the selection marker was a DNA sequence having a 0.5 Kb region homologous to the 3′-most 0.5 Kb region of the Left Flank (Repeat). These repeated sequences were intended to facilitate the

subsequent loss of the selection marker, enabling subsequent use of this marker in future strain development, and leaving the single nucleotide G to T mutation (Scaffold 16, 425393) in the *lov* gene, encoding the LOV(T813K) protein (SEQ ID NO: 4), as the only new targeted genome modification.

C. DEVELOPMENT OF T4 DERIVED STRAINS WITH THE *lov*(T813K) ALLELE

[0231] Derivatives of the T4 strain, herein named strain "T4m" and "T4m_pyr2", comprising a viscosity reducing disruption of gene *mpg1* (Δ*mpg1*), were developed essentially as described in U.S. Patent No. 9,725,727 (incorporated herein by referenced in its entirety). Thus, the T4m_pyr2 strain (Δ*mpg1; pyr2*) was transformed with a *lov* disruption cassette from pRATT312 (*pyr2* selection marker), using PEG-mediated transformation and plated on Vogel's minimal medium (Vogel, *1956*) containing 1.2 M sorbitol to select for candidates based on uridine prototrophy acquired by the *pyr2* marker. Trichoderma transformations are well known and described in the art (*e.g.*, see U.S. Patent No. 5,246,853). Individual transformants were isolated and propagated by transfer to Vogel's minimal medium. PCR analysis was used to identify transformants in which the *lov* disruption cassette integrated at the *lov* locus by homologous recombination, using methods known to one skilled in the art per guidance below.

[0232] Only a subset of recombinant cells may successfully utilize the homologous flanks to correctly target the disruption of the gene of interest, so many transformants may need to be screened to identify one with the desired event. PCR can be used to test which recombinant cells have the desired targeted disruption. Primers must be designed that amplify across each of the homology box regions, where one primer primes at a location within the selectable marker greater than 100 bp from the closest end and the other primes at a location greater than 100 bp beyond the end of a homology box region within the adjacent genomic sequence. Cells likely containing the correct targeted disruption will successfully create PCR products spanning the Left Flank and Right Flank of the disruption cassette, whereas unsuccessful transformation events will not generate a product of the expected size. At this stage the culture may be a mix of transformed and untransformed cells, so a step of purification may be needed. Purification of the culture can be tested by PCR for loss of a short PCR product spanning the disruption site.

[0233] More particularly, one such T4m derived strain comprising a *pyr2* disruption of the *lov* gene was identified, isolated and named "T4ml+" (*i.e.*, comprising allele *lov*(dis) in contrast to allele *lov*(T813K) encoding the specific T813K substitution). Thus, to generate strains comprising the specific *lov*(T813K) allele, spore suspensions of strain T4ml+ (comprising the *lov*(dis) allele) were plated on media containing 5-fluoro-orotic acid (FOA) to select for derivative strains in which there was a spontaneous recombination between the repeated regions flanking the inserted selection marker and concomitant loss of the selection marker from the genome. The uridine auxotrophs were isolated and analyzed by PCR to test for loss of the selection marker and sequencing of PCR products to confirm presence of the *lov*(T813K) allele. One strain comprising the *lov*(T813K) allele was generated in this manner and is referred to herein as "T4ml". To restore the pyrimidine prototrophy for fermentor evaluation, both the T4ml and T4m_pyr2 strains had the *pyr2* marker targeted to the viscosity reducing locus *seb1*, as essentially described in U.S. Patent No.

9,725,727, resulting in strains herein referred to as "T4mls" (comprising the *lov*(T813K) allele) and "T4ms" (comprising the *lov*(+) allele).

D. FERMENTOR EVALUATION OF T4 DERIVED STRAINS COMPRISING THE *lov*(T813K) ALLELE

[0234] Thus, strains T4mls (comprising the *lov*(T813K) allele) and T4ms (comprising the *lov*(+) allele) were evaluated in fermentors as described in Example 1. In addition, the T4mls and T4ms strains were evaluated at a lower cell density fermentation relative to Example 1, by adjusting the amount of glucose provided before feed start, to ensure that the protein production phenotypes observed with the *lov*(T813K) allele were not limited to higher cell density fermentations. As shown in **FIG. 3**, under both lower cell density (LCD; **FIG. 3**, T4mls, black circles/black dashed lines; T4ms, grey circles/grey dashed lines) fermentation conditions and higher cell density (HCD; **FIG. 3**, T4mls, black squares/black solid lines; T4ms, grey squares/grey solid lines) fermentation conditions, the protein yield on fed sugars increased 42% and 32%, respectively, when the *lov*(T813K) allele was present.

EXAMPLE 3
TARGETED DISRUPTION OF THE *lov* GENE BY INSERTION OF *pyr2* SELECTION MARKER FAILED TO IMPROVE PRODUCTIVITY IN FERMENTORS

A. OVERVIEW

[0235] Given the high conservation of the threonine (T) amino acid at residue 813 position of (native) LOV protein (SEQ ID NO: 2), it seemed probable that disruption of the locus in any way that would block the expression/production of the LOV protein, would also improve strain productivity as well. To test this hypothesis, fungal strains with and without the *lov*(dis) allele were compared in fermentors, wherein a *pyr2* selection marker was inserted into the *lov* gene to disrupt LOV protein expression. Surprisingly, the *lov*(dis) allele did not improve production in fermentors.

B. FERMENTOR EVALUATION OF STRAINS WITH THE *lov*(dis) ALLELE

[0236] Development of the T4m and T4ml+ strains are described above in Example 2. Strain T4ml+ is a derivative of T4m in which the *pyr2* marker was inserted into the *lov* coding sequence of stain T4m_pyr2. The T4m and T4ml+ strains were evaluated under identical conditions in 14L fermentors that were essentially the same as described in Example 1, except at a lower cell density fermentation, by adjusting the amount of glucose provided before feed start.

[0237] Derivatives of T4, named "T4s" and "T4s_pyr2" (comprising a viscosity reducing disruption of gene *seb1*), were developed essentially as described in U.S. Patent No. 9,725,727. The *lov*(dis) allele was introduced into T4s_pyr2 analogously as described for strain T4m_pyr2 in Example 2, to generate strain T4sl+. Thus, the T4s and T4sl+ strains were evaluated under identical conditions in 2 L bioreactors. Specifically, for 2 L bioreactors, to create a seed culture, the spores of each strain were added separately to 50 mL of citrate minimal medium in a 250 mL flask. The cultures were grown for 48 h at 30°C and 170

rpm in a shaking incubator. After 48 hours, 145.6 mL of 50% glucose, and 0.6 g/kg of $CaCl_2$, adjusted to pH 3.5, was inoculated with the seed culture. Thereafter, the temperature was maintained at 34°C, and pH at 3.5. Following exhaustion of batched glucose, a glucose-sophorose feed was thereafter introduced, and the temperature was dropped to 25°C, and pH increased to 4.8. The dry cell weight (DCW), total protein concentrations, and other parameters were measured, and specific total protein production rates and yield on fed sugars were calculated. As shown in **FIG. 4**, protein yield on fed sugars decreased 17% when the *lov*(dis) mutation was present in the T4m background and was unchanged when present in the T4s background.

EXAMPLE 4
TARGETED INTRODUCTION OF THE *lov*(T813K) ALLELE IN OTHER FUNGAL STRAINS IMPROVES PROTEIN PRODUCTIVITY

A. OVERVIEW

[0238] In Example 2, the *pyr2* auxotrophy generated concomitant with allele *lov*(T813K) generation and marker loss had been restored by integration of the *pyr2* marker at the *seb1* locus. Therefore, the improvement in protein production observed with the *lov*(T813K) allele has thus far been exemplified, in Example 1 and Example 2, in strains always containing both the *mpg1* and *seb1* viscosity mutations. To demonstrate the more general application of the *lov*(T813K) allele for improvement of strain productivity, the *pyr2* marker was integrated at three other convenient genomic locations: site A, site B and site C. Protein production by these different strains was evaluated in shake flasks, where in all cases, presence of the *lov*(T813K) mutation improved total protein production relative to the isogenic strains with a wild-type *lov*(+) allele. On the contrary, strains comprising a *lov*(dis) mutation did not show significant improvement in protein production in shake flasks.

B. RESTORED PYRIMIDINE AUXOTROPHY IN T4ml AND T4sl BY TARGETED INTEGRATION OF A *pyr2* SELECTION MARKER AT DIFFERENT GENOMIC LOCATIONS

[0239] Strains either with or without the *lov*(T318K) allele were transformed independently with cassettes targeting integration of *pyr2* to three convenient genomic locations named sites A-C, using methods known to one skilled in the art. Briefly, following PEG-mediated transformation with the *pyr2* cassettes, protoplasts were plated on Vogel's minimal medium containing 1.2 M sorbitol to select for candidates based on uridine prototrophy acquired by the *pyr2* marker. Individual transformants were isolated and propagated by transfer to Vogel's minimal medium. PCR analysis was used to identify transformants in which the *pyr2* cassette integrated at the intended genomic location. Following spore purification, further PCR analysis was done to ensure integration occurred correctly and that the transformants were homokaryotic.

[0240] Thus, the *pyr2* marker was integrated at site A, site B and site C in *pyr2* mutant strains T4m_pyr2 (*lov*(+)) and T4ml (Δ*mpg*1; *lov*(T318K)) which were described in Example 2 (see, **Table 1**). Likewise, the *pyr2* marker was integrated at site A in mutants strain T4s_pyr2 (Δ*seb1; lov*(+)) and T4sl (Δ*seb1*; *lov*(T318K)) described in Example 3. In addition, disruption allele *lov*(dis) was added to whole cellulase strains T4_pyr2 (*pyr2*; *lov*(+)) and 41G_pyr4 (*pyr4*; *lov*(+)) using plasmids pRATT312 and pRATT308 respectively as described in Example 2.

C. Shake Flask Evaluation of Strains with and without Either the *lov*(T813K) or *lov*(dis) Alleles

[0241] Strains were evaluated in shake flask fermentations for their protein production. The strains evaluated are listed in TABLE 1. In all genetic backgrounds tested, the total protein titers increased when the *lov*(T813K) allele was present. However, disruption of the *lov* gene, the lov(dis) allele, with either the *pyr4* (41G) or *pyr2* marker (T4 and T4m) showed no significant improvement in relative titers.

[0242] Liquid defined (LD) culture medium (*e.g.*, see, US Patent No. 8,455,631), contained the following components. Casamino acids, 9 g/L; $(NH_4)_2SO_4$, 5 g/L; $MgSO_4 \cdot 7H_2O$, 1 g/L; KH_2PO_4, 4.5 g/L; $CaCl_2 \cdot 2H2O$, 1 g/L, PIPPS, 33 g/L, 400× *T. reesei* trace elements, 2.5 ml/L; pH adjusted to 5.5 with NaOH. After sterilization, lactose or a glucose/sophorose mixture was added to a final concentration of 1.6% w/v.

[0243] To create a seed culture, the spores of each strain were added separately to 50 mL of YEG (5 g/L yeast extract, 22 g/L glucose, H_2O) in a 250 mL flask. The cultures were grown for 36-48 hours at 28°C and 200 rpm in a shaking incubator. After incubation, 0.3 mL of seed culture were added to 50 mL of LD medium in a baffled shake flask. This production culture was grown for 5 days at 28°C and 180 rpm. Secreted protein was harvested by centrifugation to pellet cells, and then collecting the supernatant. Proteins were precipitated from the supernatant with an equal volume of trichloroacetic acid (TCA), followed by dissolution in 0.1 N sodium hydroxide (NaOH). Total protein was then measured with a BCA protein assay (ThermoFisher Scientific, Grand Island, N.Y., USA) *per* manufacturer protocol. In all shake flask experiments, prototrophic T4 and T4m strains were included in duplicate. BCA assay numbers were normalized to the average of these controls which were run in parallel to minimize the influence of any week-to-week variation in total protein production efficiency.

[0244] As shown in **FIG. 5**, protein titer increased when the *lov*(T813K) mutation was present in the T4mc (site C), T4md (site B), T4mp (site A) and T4sp (site A) backgrounds in addition to the T4ms background which was exemplified above in bioreactors (Example 2). On the contrary, genetic backgrounds 41G, T4 and T4m comprising an insertion of the *pyr2* marker or *pyr4* marker disrupting the *lov* gene, *i.e.*, allele *lov*(dis), did not show significant improvement in protein production in shake flasks.

D. Fermentor Evaluation of Strain T4mp with and without the *lov*(T813K) Allele

[0245] One pair of these strains, T4mp and T4mlp (insertion of *pyr2* at site A), was further evaluated in 14L fermentors as generally described in Example 1, but at lower cell density. As shown in **FIG. 6**, in the

presence of the *lov*(T813K) allele, this genetic background showed a 28% increase in total protein yield on fed sugars.

CLAIMS

1. A modified Ascomycota cell derived from a parental cell, the modified cell comprising a polynucleotide sequence encoding a variant LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4, wherein the modified cell comprises an enhanced protein productivity phenotype relative to the parental cell when fermented under the same conditions.

2. The modified cell of claim 1 comprising a polynucleotide encoding a protein of interest (POI).

3. The modified cell of claim 1, comprising a polynucleotide construct encoding a NIK1(M743T) protein of SEQ ID NO: 19.

4. The modified cell of any one of claims 1-3, comprising at least one genetic modification which deletes, disrupts or reduces the expression/production of a protein selected from the group consisting of SSB7, MPG1 SFB3, SEB1, CRZ1, TSP2 and/or GAS1.

5. The modified cell of any one of claims 1-4, wherein the Ascomycota cell is selected from a *Trichoderma sp.*, *Aspergillus sp.*, *Fusarium sp.*, *Penicillium sp.*, a *Candida sp.*, *Chrysosporium sp.*, *Cephalosporium sp.*, *Talaromyces sp.*, *Neurospora sp.* and *Myceliophthora sp.*

6. The modified cell of claim 5, wherein the parental *Aspergillus sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 11 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

7. The modified cell of claim 5, wherein the parental *Penicillium sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 12 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

8. The modified cell of claim 5, wherein the parental *Talaromyces sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 13 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

9. The modified cell of claim 5, wherein the parental *Fusarium sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ

ID NO: 14 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

10. The modified cell of claim 5, wherein the parental *Myceliophthora sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 15 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

11. The modified cell of claim 5, wherein the parental *Neurospora sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 16 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

12. The modified cell of claim 5, wherein the parental *Candida sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 17 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

13. The modified cell of claim 5, wherein the parental *Trichoderma sp.* cell comprises a polynucleotide sequence encoding a native LOV protein comprising at least 50% sequence identity comprises SEQ ID NO: 2 or SEQ ID NO: 18 and comprises a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2.

14. A POI produced by the modified cell of any one of claims 1-13.

15. A modified Ascomycota cell derived from a parental cell comprising a wild-type polynucleotide sequence which hybridizes under stringent hybridization conditions with a nucleic acid sequence encoding a native LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a threonine (T) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 2, wherein the modified cell comprises a modified polynucleotide sequence which hybridizes under stringent hybridization conditions with a nucleic acid sequence encoding a LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence corresponding to position 813 of SEQ ID NO: 4.

16. A vector comprising a polynucleotide encoding a variant LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a

lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4.

16. A polynucleotide encoding a variant LOV protein comprising at least 50% sequence identity to SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4.

17. A polynucleotide encoding a variant LOV protein, wherein the polynucleotide encoding the variant protein hybridizes under stringent hybridization conditions with a nucleic acid sequence encoding a LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and comprising a lysine (K) residue at an amino acid sequence position corresponding to position 813 of SEQ ID NO: 4.

18. A mutant Trichoderma strain derived from a parental strain, wherein the mutant strain comprises a gene encoding a LOV variant protein comprising at least 75% sequence identity to SEQ ID NO: 4 or SEQ ID NO: 18 and comprises a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 4.

19. The mutant strain of claim 18, wherein the mutant strain comprises an enhanced protein productivity phenotype relative to the parental strain.

20. The mutant strain of claim 18 or claim 19, comprising a polynucleotide construct encoding a NIK1(M743T) protein of SEQ ID NO: 19.

21. The mutant strain of any one of claims 18-20, comprising a genetic modification which deletes, disrupts or reduced the expression/production of a gene encoding at least one protein selected from the group consisting of SSB7, MPG1 SFB3, SEB1, CRZ1, TSP2 and GAS1.

22. A POI produced by the mutant strain of any one of claims 18-21.

23. A method for constructing a modified Ascomycota cell comprising an enhanced protein productivity phenotype, the method comprising:

(a) obtaining a parental Ascomycota cell and introducing into the cell a polynucleotide construct encoding a LOV variant protein comprising a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 4, and

(b) isolating the modified cell of step (a),

wherein the modified cell comprises an enhanced protein productivity phenotype relative to the parental cell.

24. A method for constructing a modified Ascomycota cell comprising an enhanced protein productivity phenotype, the method comprising:

(a) obtaining a parental cell comprising a wild-type *lov* gene which hybridizes under stringent hybridizations conditions with a nucleic acid sequence encoding a LOV protein of SEQ ID NO: 2, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18,

(b) modifying the parental cell of step (a) to produce a modified cell comprising a gene encoding a LOV (variant) protein comprising a lysine (K) residue at a sequence position corresponding to position 813 of SEQ ID NO: 2, and

(c) isolating the modified cell of step (b),

wherein the modified cell comprises an enhanced protein productivity phenotype relative to the parental cell.

REFERENCES

PCT Application Serial No. PCT/US2019/27590

PCT International Publication No. WO1992/06183

PCT International Publication No. WO1992/06209

PCT International Publication No. WO1992/06221

PCT International Publication No. WO1998/15619

PCT International Publication No. WO2002/12465

PCT International Publication No. WO2005/028636

PCT International Publication No. WO2006/074005

PCT International Publication No. WO2012/027580

PCT International Publication No. WO2012/145584

PCT International Publication No. WO2012/145592

PCT International Publication No. WO2012/145595

PCT International Publication No. WO2012/145596

PCT International Publication No. WO2012/145598

PCT International Publication No. WO2016/100272

PCT International Publication No. WO2016/100562

PCT International Publication No. WO2016/100568

PCT International Publication No. WO2016/100571

PCT International Publication No. WO2016/130523

U.S. Patent No. 5,246,853

U.S. Patent No. 8,455,631

U.S. Patent No. 9,725,727

U.S. Patent Publication No. US2018/0037919

Altschul *et al.*, *J. Mol. Biol. 215:403-10, 1990.*

Altschul *et al.*, *Meth. Enzymol. 266:460-80, 1996.*

Ausubel *et al.*, "Current Protocols in Molecular Biology", *Eds., John Wiley & Sons, Inc., 1995, sections 2, 4 and 6.*

Botstein and Shortle, *Science 229: 4719, 1985.*

Church and Gilbert, *"Genomic Sequencing", PNAS, U.S.A., 81, 1991-1995, 1984.*

Devereux *et al.*, *Nucleic Acids Res. 12:387-95, 1984.*

Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA 89:10915, 1989.*

Higgins and Sharp, *CABIOS 5:151-53, 1989.*

Higgins *et al.*, *Gene 73:237-244, 1998.*

Higuchi *et al., Nucleic Acids Research 16: 7351, 1988.*

Ho *et al., Gene 77: 61, 1989.*

Hopwood, The Isolation of Mutants in Methods in Microbiology *(J. R. Norris and D. W. Ribbons, eds.) pp 363-433, Academic Press, New York, 1970.*

Horton *et al., Gene 77: 61, 1989.*

Iglesias and Trautner, *Molecular General Genetics 189: 73-76, 1983.*

Karlin *et al.*, *Proc. Natl. Acad. Sci. USA 90:5873-87, 1993.*

King *et al., "*A conserved fungal glycosyltransferase facilitates pathogenesis of plants by enabling hyphal growth on solid surfaces", PLoS Pathology, 13(10), 1-26, 2017.

Lo *et al., Proceedings of the National Academy of Sciences USA 81: 2285, 1985.*

Montenecourt B. S., *Can., 1-20, 1987.*

Needleman and Wunsch, *J. Mol. Biol., 48:443, 1970.*

Parish and Stoker, *FEMS Microbiology Letters 154: 151-157, 1997.*

Perego, *1993, In A. L. Sonneshein, J. A. Hoch, and R. Losick, editors, Bacillus subtilis and Other Gram-Positive Bacteria, Chapter 42, American Society of Microbiology, Washington, D.C.*

Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual", *Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1989), chapters 7, 9 and 11.*

Sarkar and Sommer, *BioTechniques 8: 404, 1990.*

Sheir-Neiss *et al.*, *Appl. Microbiol. Biotechnology, 20: 46-53, 1984.*

Shimada, *Meth. Mol. Biol. 57: 157; 1996.*

Smith and Waterman, *Adv. Appl. Math. 2:482, 1981.*

The Genome Portal of the Department of Energy Joint Genome Institute, I. V. Grigoriev, H. Nordberg, I. Shabalov, A. Aerts, M. Cantor, D. Goodstein, A. Kuo, S. Minovitsky, R. Nikitin, R. A. Ohm, R. Otillar,

A. Poliakov, I. Ratnere, R. Riley, T. Smirnova, D. Rokhsar, and I. Dubchak, *Nucleic Acids Res., 2011 0: gkr947v1-gkr947.*

Vogel, *Microbiol. Genet. Bull. 13:42-43, 1956.*

Wisconsin Genetics Software Package (*Genetics Computer Group, Madison, WI*)

Youngman *et al., Proc. Natl. Acad. Sci. USA 80: 2305-2309, 1983.*

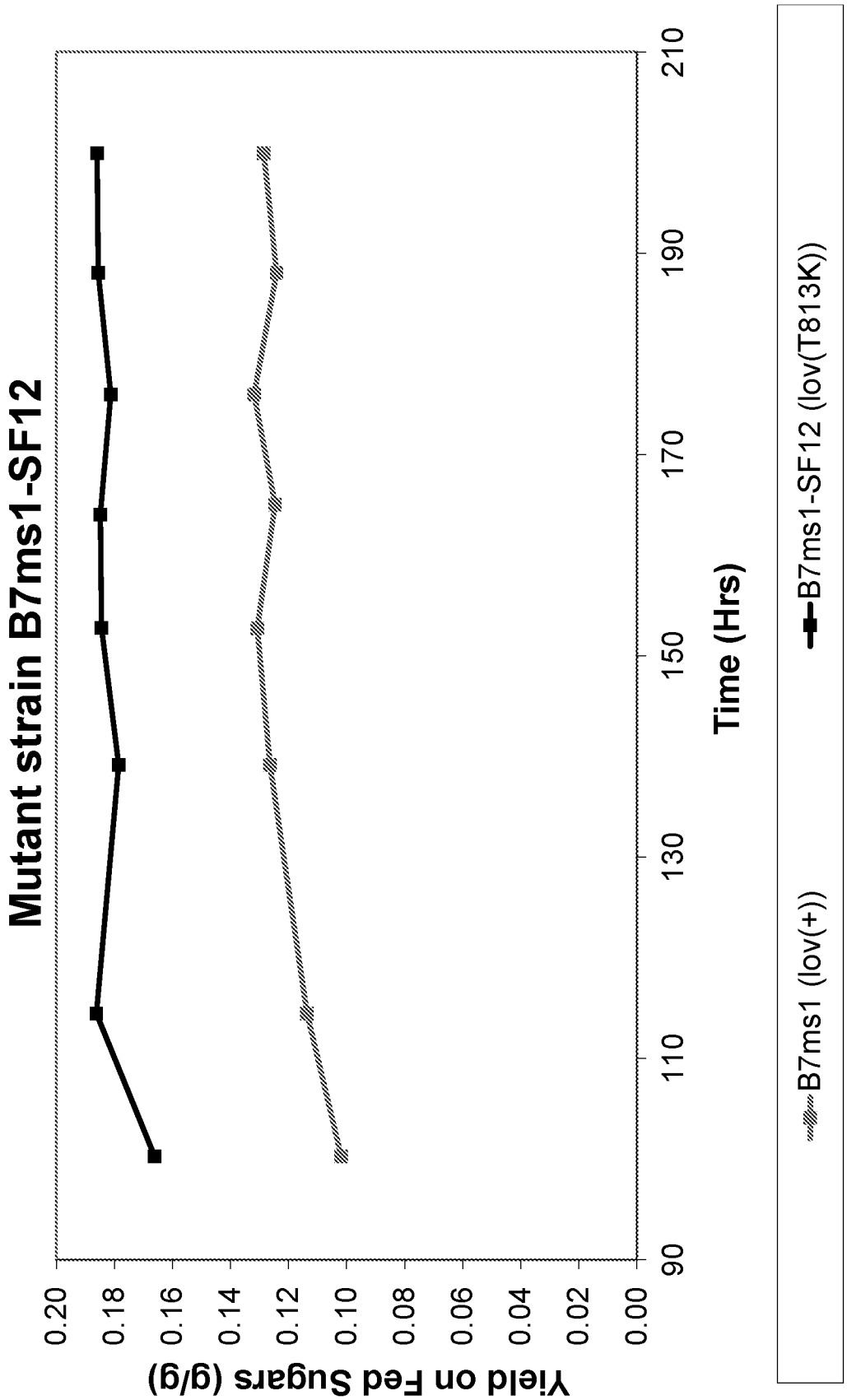
Mutant strain B7ms1-SF12
Yield on Fed Sugars (g/g)
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
90
110
130
150
170
190
210
Time (Hrs)
B7ms1 (lov(+))
B7ms1-SF12 (lov(T813K))

FIG. 1

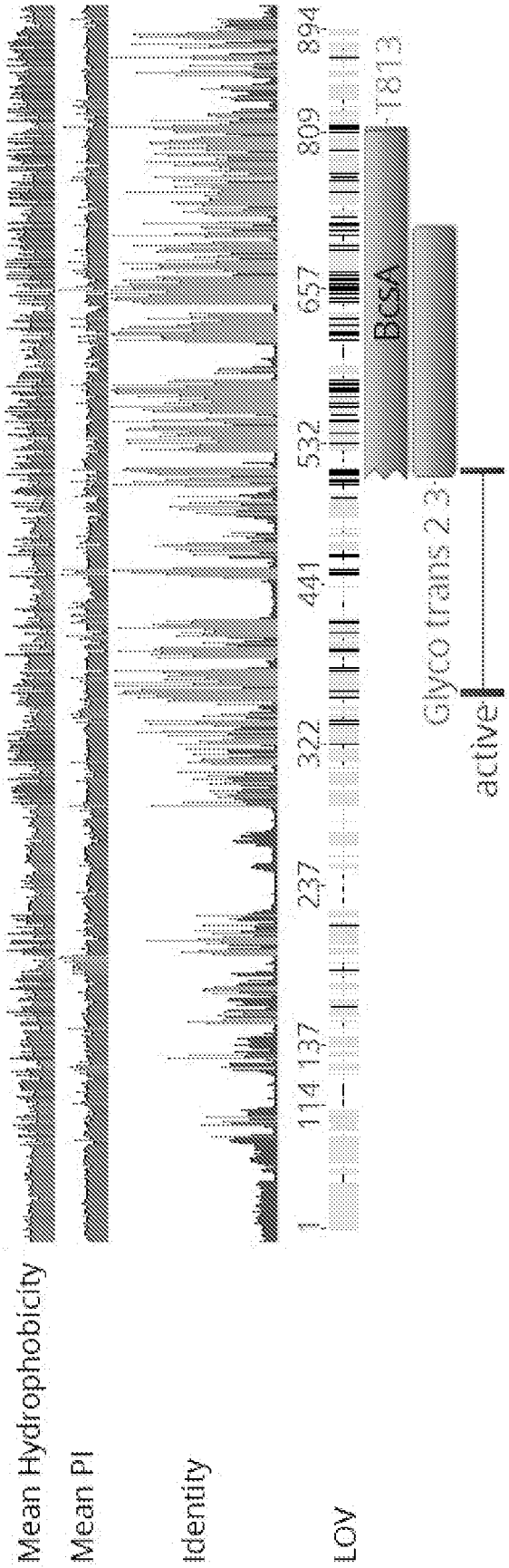
Mean Hydrophobicity
Mean PI
Identity
LOV
1
114
137
237
322
441
532
657
809
894
BCSA
T813
Glyco trans 2 3
active

FIG. 2

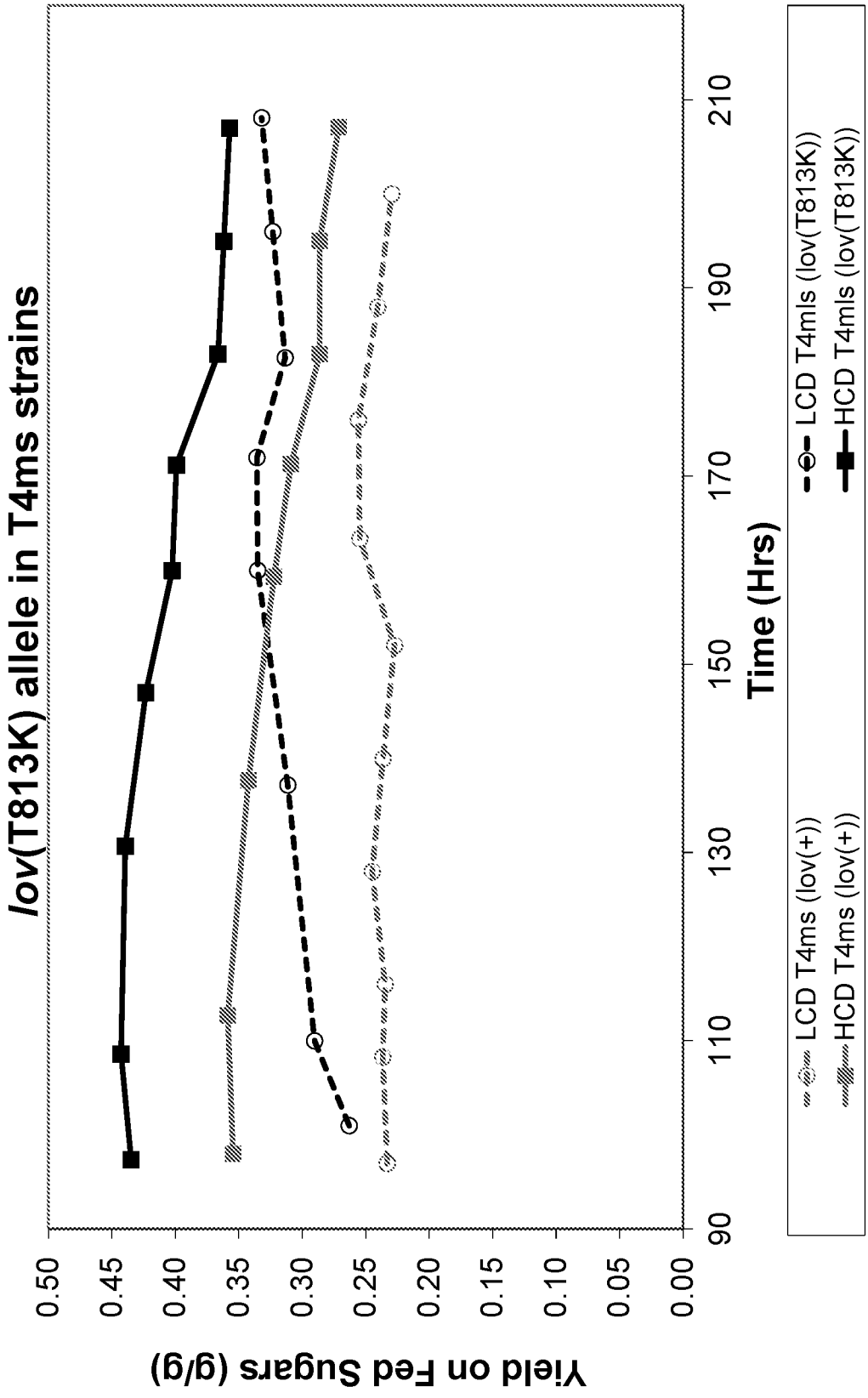
lov(T813K) allele in T4ms strains
Yield on Fed Sugars (g/g)
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
90
110
130
150
170
190
210
Time (Hrs)
LCD T4ms (lov(+))
HCD T4ms (lov(+))
LCD T4mls (lov(T813K))
HCD T4mls (lov(T813K))

FIG. 3

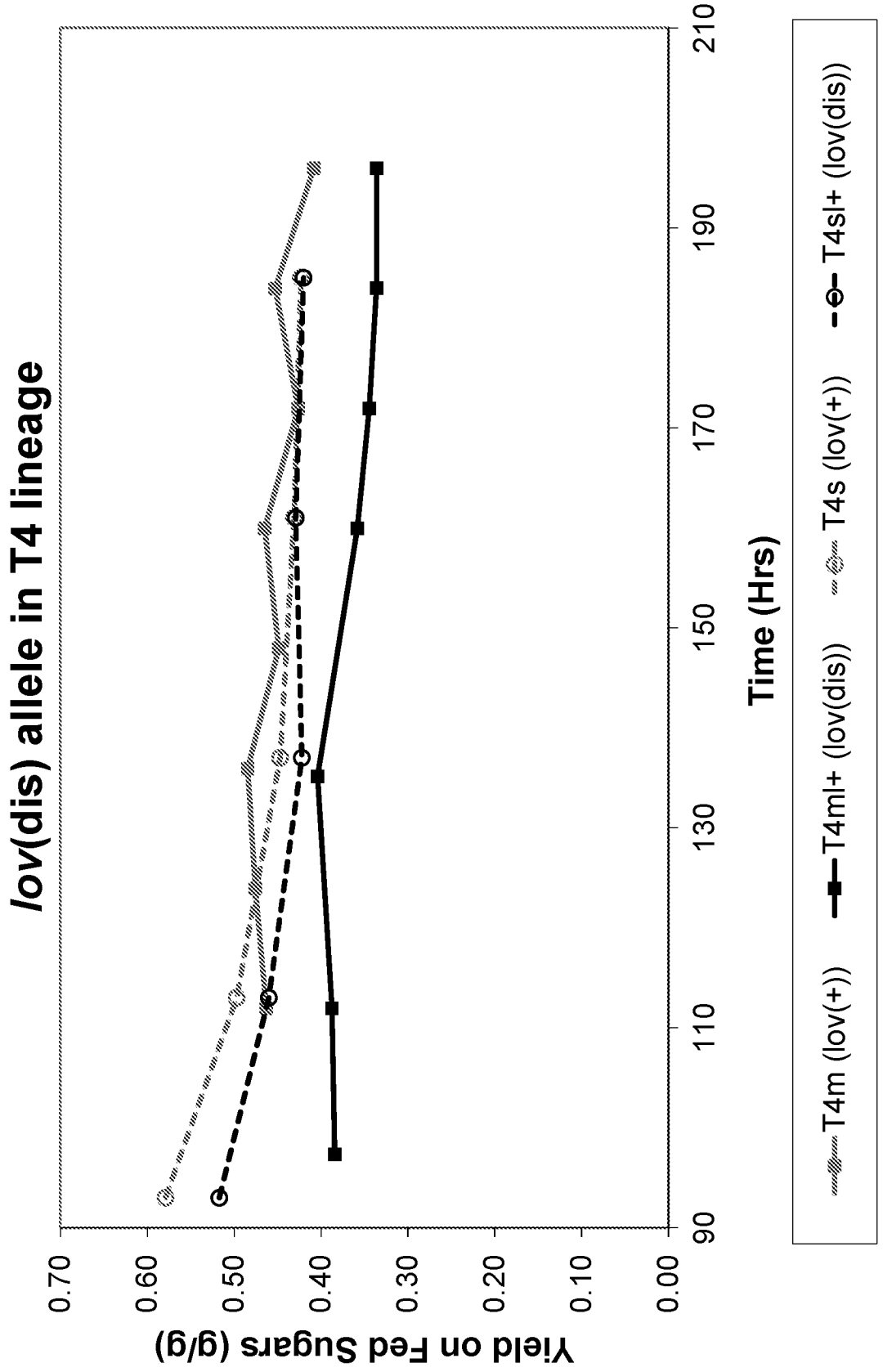
lov(dis) allele in T4 lineage
Yield on Fed Sugars (g/g)
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
90
110
130
150
170
190
210
Time (Hrs)
T4m (lov(+))
T4ml+ (lov(dis))
T4s (lov(+))
T4sl+ (lov(dis))

FIG. 4

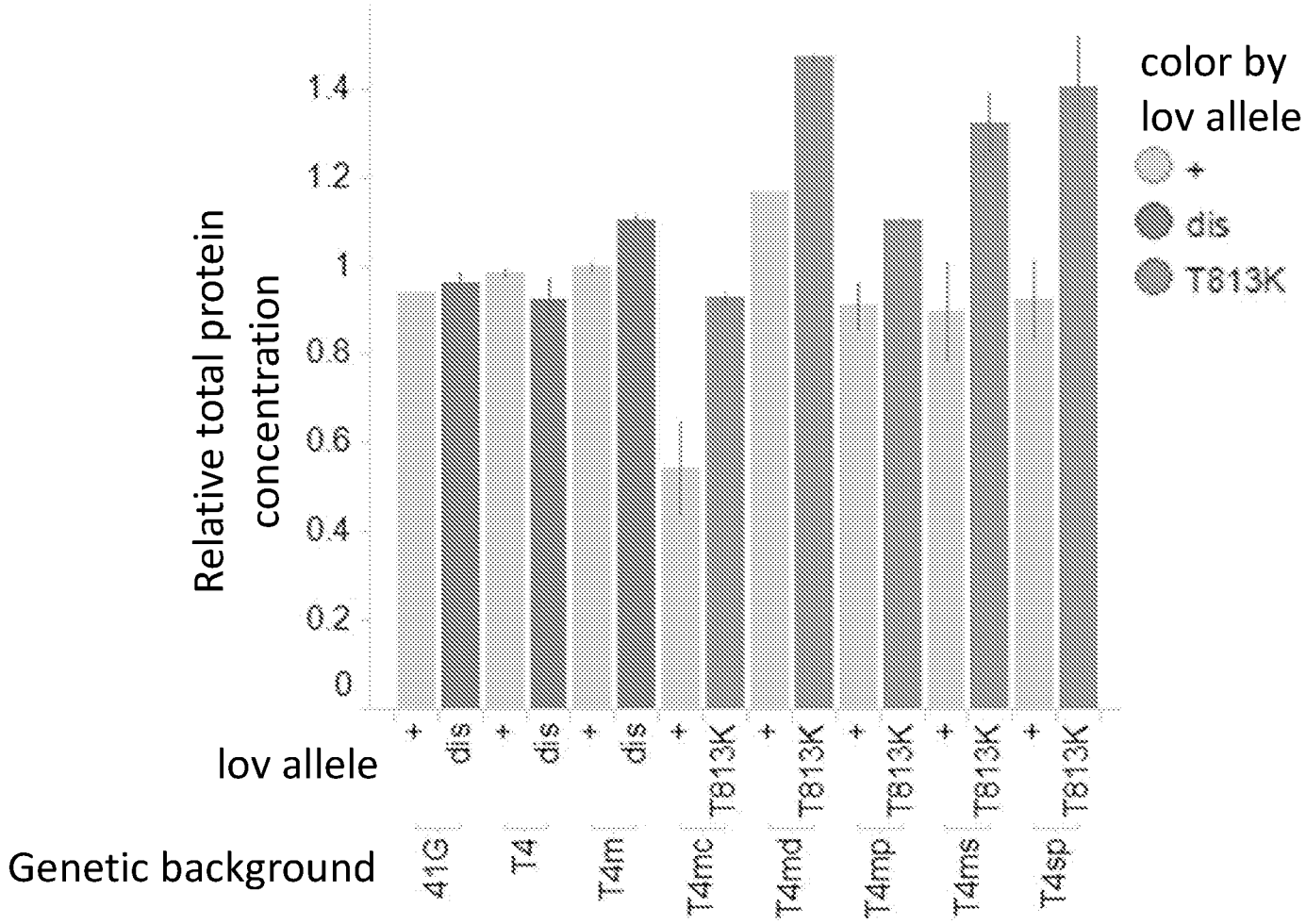
Relative total protein concentration
0
0.2
0.4
0.6
0.8
1
1.2
1.4
color by lov allele
+
dis
T813K
lov allele
Genetic background
41G
T4
T4m
T4mc
T4md
T4mp
T4ms
T4sp

FIG. 5

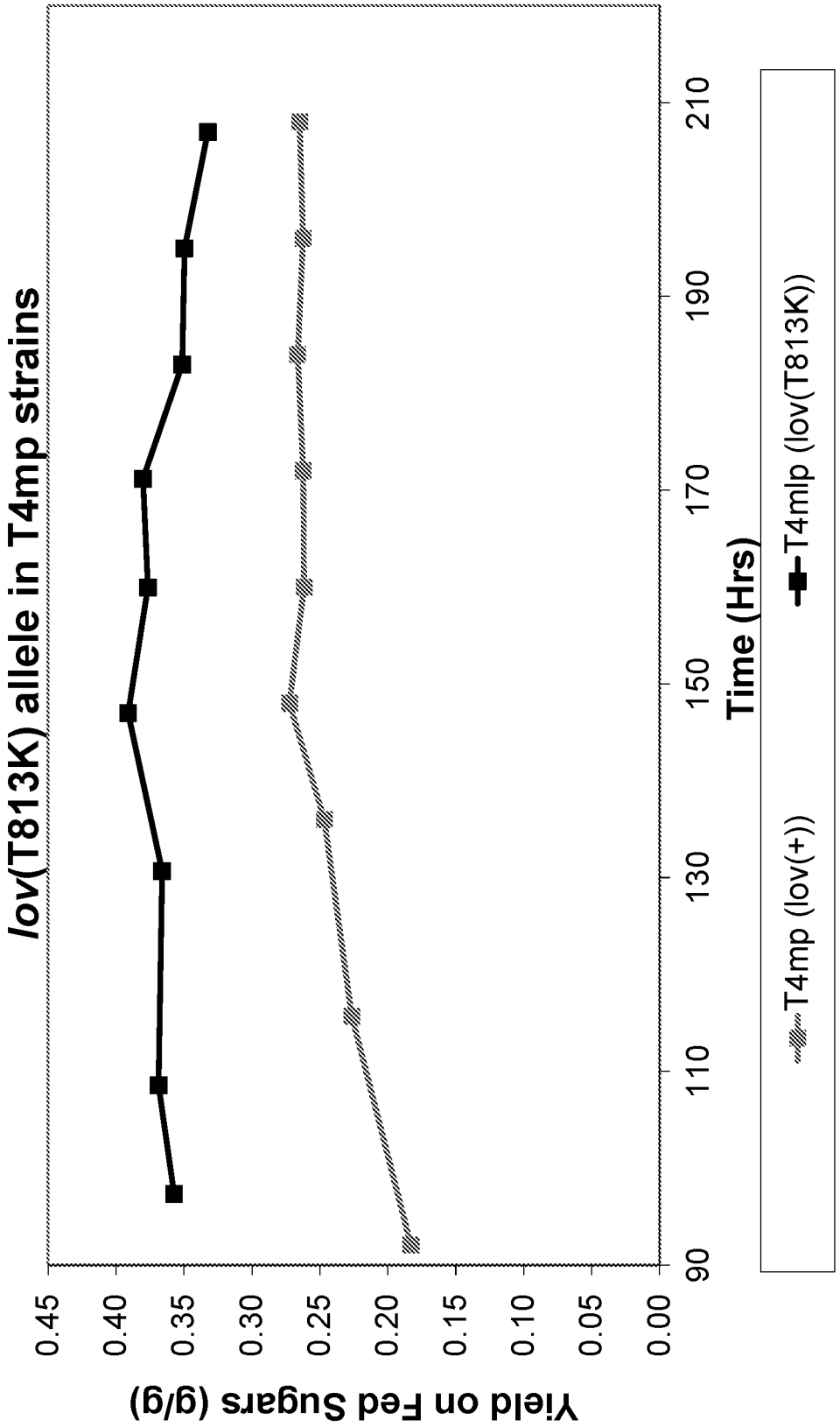
lov(T813K) allele in T4mp strains
Yield on Fed Sugars (g/g)
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
90
110
130
150
170
190
210
Time (Hrs)
T4mp (lov(+))
T4mlp (lov(T813K))

FIG. 6

SEQ ID NO: 2 (*Trichoderma sp.*)
MGIGSYFKASTKNAEQAPPAPPPAKLAPPPRLQLEGRQSVMSEKPPSSTSGNDVDLQPPTPRFHSRPQSSSGRSTPSTQSSMFLDD
IKHEVMVNYLYQQQCSQLWVSDGSGEIEGVLLRKARGHYMACPPQLASSPFALACAALNVQCAMTVNSRVIKTFLQWSPDAVDVPL
MNGLRVQILPTIDDLPRARKYQFAAFVASEGLLVVWDDDALHLVARAKAIESELMELVWKAGNVDEEGGDEKGGQPVTEVEIDEES
GEIKPEKRPIHLQNTVLVSLTLALVTVSLGAAWRQLAIEVSVDSNYIRLALVALAPVQVFFTLFFAQVIIGCLAQIFGPIRQLTIN
SKFYSARPPPRLQSAVLPHVTIQCPVYKEGLQGVIMPTVKSIKQAMSTYELQGGSANMFINDDGLQLISEEDRLARIEFYADNSIG
WVARPKHGENGFTRKGKFKKASNMNFALMISCKVEEKLQAIERHPGWSQNDEALAYEEALKEVLEADGRAWADGNIRMGDYILLID
SDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQD
EDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRF
LFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYAL
FENFMWTLMLAIFLGGLSLHVSQALLAHMFEINMTWGA**X**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDW
RIKDFVAILPMATVAASHFLLPLALNPALMTFSW

SEQ ID NO: 11(*Aspergillus sp.*)
MTMLSPTSDHPQGASMKDIENPGNDSPALEGSTQAEKTVDNPSRSPTPQLQTAELAQVLSPLAVRQTQSISTRSQHNRLEYLCLPN
ATNRGKLSRSLARRATLLGWFDQAATQSGQESQSSVCVAVKLPDDGYSFSPSNITPALSKAITRLNEMAVVALSSRVVDKGLSGIL
PGQKSFLVESTSTRIPIVATLDDVEPTLAHYSRACIVMEQKIVLVWSHDTAGILNVAYDVETQLGGQSPKISNATTPRISGRSTPM
DPYRTDDLQKPANVLQPVRGALDEKQAIYSKAVALEEGEDNEVPDLERNAAPRPVLLVHTVKISLAIMLVIVTQSLGVARLLNEFQ
WDGQYTRFALVVTIPPLALFSLFFFIVLVTSVFQLFLPASFCLKNSKFHSGDTEDSSAVKPNPKAHGEYELPHITIQMPVYKEGLK
GVIVPTMISVLAAVQYYEEQGGTASVFVNDDGMQVIQSDLAEARKQYYRENGIGFTARPPNKKSPVQKGGWGSWFRKSKPATTEPA
DSEEPEGPYTPQALANKIGFERKGKFKKASNMNYGLAFSLRVEDELARLTQIETERRGCTVDDLTAEDDDRLYQQALDNMLAADEE
RTWAEGNIRIGELILLIDCDTRVPVDCLLYGALEMHESPEVAILQHGSGVMQVVHNVFENGITYFTNVVYTAIKYGVGSGDVSPFV
GHNAFLRWRALQSIEFVDPSDGQTKWWSDTHVSEDFDISLRLQMQGMVVRLATYHNGEFKEGVSLTLYDELTRWEKYAYGCNELVF
HPFYQWVYKGPVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIGSGMLLATVNYVILGLFDSDIDHLYLPSWGIWCSLVVVFNGFSS
IAFSMVRHQLKEETFWRALLDAIKWLPFLIIYFGGISLNCAKAILCHAFSINLEWASTAKEMGPTGFYIGMDKMVRRFKWTWAICL
VLAGVMIYFAVGAPWGWTITPGPYSTANVAIAPLAIQICSASFLPFFLGLN

SEQ ID NO: 12 (*Penicillium sp.*)
MEEWPRNADEPYGSQRYQEDVQDLVSLPSRSSTPHLDRAEIGRVLSPSVQPGRNSRSRRSNPSIAHSFTSPSISESGEGEDPSSTA
RYLARRATLLGWFDEVDVEDGRNSWSSTCISLRRADNEYTFFPYNVNPSLIDAITRLGETAAMAMSSRFTSMLIDSILPGQKSLVV
ESTSARIPVVHSLNDVSSNLVHFARACIVVQEKLVLIWSHNAGTILHVAHDVERQLGKKSPRISTATTPRASGRSSPFDVDADVAP
STQISIKAKIDSFPVRGALDEKHEVYSKAVALEEGTVEDDTPVDLEGNLPPRPAHRIHAVKISLAIMLVILTQSLGVSRLVNEFAW
DGSFTRFALVVTIPPLALFSLFFFIVLVTSAFQLFLPASFCLKNSKFHSAIKPNPRFHRDYELPHITVQMPVYKEGLKGVIVPTMM
SVLAAVQYYEEQGGTASVFINDDGMQCIQPELAEARKQYYRENGIGYTARLPNRKTASKKKRGFGWFRKAKSAEGDAETEAEEDTS
SPQAIANKIGFERKGKFKKASNMNYGLAFSNRVEDELARLADLECQQRGCSNDDLTFEDDDRLYQQALSNMLAEDEGRTWAEGNVR
IGEIILIIDCDTRVPVDCLLYGALEMHESPEVAIVQHGSGVMQVVHNMFENGITYFTDVVYTAIKYGVGSGDVSPFVGHNAFLRWK
ALQSIQFVDPADGQTKWWSDAHVSEDFDISLRLQMQGMTVRLATYHNGGFKEGVSLTLYDELTRWEKYAYGCNELVFNPFYQWPYK
GPVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIASGMMLSVVNYVIVGLFNSEVDHIYLRSWGIWISLVVVFNGVASVAFSMARHQ
LKEMVFWKALLKSALWLPFLVVFFGGISLNCAKAILCHAFSINIEWASTAKEPGPSGFFIGLDKMVKQFKYTWAICLFLAAVMIFM
ALGTPWGWQIKPGEYSTASIAIGPLAIQICNAAILPLILGLN

SEQ ID NO: 13(*Talaromyces sp.*)
MSSLFPTRWSFRGHKTPKGDGTPTETASQRDSLPVSESGRSSRGQGYFDRRHVNNSIDDASKYKAMIKFFHVRLTAYQWLPPPTHP
QHSSTGVFLRRSRGIYMSEPEDINPLLLAAIQRINATIAFTMMTETTSIITSQLAPGQTELILPNGYQVQIIESYADIVGSHSNMV
KKYQYAALIREEQLLLVWNDDLNAILNHAADVEGKLLSLIWGSPIPTFNLQAVPMMTPGESVVASPNDSLYHLALEPRESPAAAED
SGTSRDASPRRMINEEVKRPKESLERPLAVTSAIFVGMAGMLLVILLLGFGISNLLLEYSVDGGAMRFALTATIPFFLLFSIFFMI
VIFTDIFQAVGPVKTLKSNSRFYSPIAPDLKTAYSLGFTPPRVTIQMPIYTESLEGVIKPTISSLKTAISHYESHGGTANIFINDD
GFALLSEEQQHERINFYHDNNIGWVARPKNNEDGYIRKGKFKKASNMNFALNVSNKVEMELIQRMAPTLEKSDMVDPMEEELVYRE
AFDHVIQSDPRIRGAGGDIRVGEFILIVDSDTRIPADCLLYGAAEMFLSPEVAIIQHSTSVMQVSQDYFENGITYFTNLIYSAIRF
AVGSGETAPFVGHNAFLRWQAVQSVGRPDDGYVSFWSESHVSEDFDIALRLQIQGNIIRLASYHNDEFKEGVSLTIYDELSRWEKY
AYGCNELVFNPVHTWFYRGPLTKLFMTFLWSNLQLSSKITILGYISSYYALASGFPLTVLNYFLVGWFEGYLDKFYMESWKVFLSL
LVVFSAAGNVCLAIIRYRLGEKPLLASLVENFMWMPMFAIFFGGLSFHLSLAILSHMFGINMSWGTTAKEKDDSNFFKEIPKIFKS
FKWMYAVVLPFFPAMIYLACFAPNGWTITEVGAIVPMSVTLASHALLPLLLNPSLMVFNY

FIG. 7A

SEQ ID NO: 14 (*Fusarium sp.*)
MGIGSYFKADIPPQTPIGPPPRRPSHRPSMSLHAPIIEEKIPPAAVVELHSAAGPKFTAGPPSIAPSSKSGNSNLLDDIKHEVMVN
YLYQQQCSQLWVGDGSGEIEGVLLRKSRGQYMACPPQLGQSPFAIACTALNVQCAMTVNSRVIKTFLQWSPDAVDVPLMNGLRVQI
LPTVNDLPRARKHQFAAFIASDGLLVVWDDDALNLMARAKIIESELMELVWNSGQSVDEDERDSTIAAEYEIDEESGEIKPEARPV
HLQNAVLVSLTLLLVMASLGAAWRQLAVEIAIDGDYKRLGLVALFPIQIFFSLFFAQVIVGCLAQIFGPIRQLTINSKFYSARPPR
RLQGATLPHITIQCPVYKEGLNAVILPTVKSIKQAMSTYELQGGSANMFINDDGLQLLSEEERDARIDFYADNSIGWVARPKHGED
GFIRKGKFKKASNMNFGLMISCKVEERLQLIKRPADWSQSDEALAYEQALKDVLEENGRAWADGNIRVGDYILLIDSDTRVPTDCL
LDSVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTDLIYTAIRFTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSE
SHVSEDFDMSLRLQCNGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAFGCNELLFNPIRTWLWRGPFTPLFRRFCFSNIRFTSK
ITVVSYIGTYYAIGAAWIMTTANYFLMGWFNGYLDKYYLDSWQVWFSIILVFNGLGNLALAIMRYRSGERTLGYAIYENFKWTLML
AVFLGGLSLHVSQALLAHMFEIDMSWGSTSKEAEFSNFFIEVPKVLKKFRVSMTLSLLAIVALIIMATADFIPHYWRINDFVAILP
MATVAGSHLLLPLALNPALMTFSW

SEQ ID NO: 15 (*Myceliophthora sp.*)
MGIGDYFKAEKLGSKPTPTPASPPREHGRHQQQPSASEDHPAPSVQPASELQPPTPRFSSRPQSISGRSVRSTGSSVLDEIKHEVM
VNYLYQQQCSHLWISDGSGEIEGVLLRKARGQYMACPPQLVNSPLAAACTALNVQCAMTVNSRVIKTFLQWSPDAVDVPLMNGMRV
QILATIDDLPRARKHQFAAFVASEGLLIVWDDDALHLVQRAKAIESELMELVWKVGAEDNEDEKGVAAVEEPEVDEESGELKPEKR
PVHLLNAYLVSLSLILVTVSLGAAFRQLAIEVSVDGNYVRLALVALFPVQMFFTLFFAQVIVGCLAQIFGPIRQLTVNSKFYSARP
PPRLRSSVLPHVTVQCPVYKEGLNAVIAPTVKSIKQAMSTYELQGGSANMFINDDGLQLISEEDRRARIEFYADNSIGWVARPKHG
ENGFQRRGKFKKASNMNFALMISCKVEDKLAAIQRTPDWTQHDEALAYERALKEVLEEDGRAWADGNIRIGDYILLVDSDTRVPAD
CLLDAVSEMELSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVSYEDEDGYEKFW
SESHVSEDFDMSLRLQCAGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRKWIYKGPFTPLFRRFLFSNIRFT
SKVTVISYIGTYYAIAAAWIMTSINYFIMGWFNGYLDKYYVDSWQVWFSIILVFNGLGNIALAVMRYRVGERSILYALYENFKWTF
LLAVFLGGLSLHLSQALLAHMFEIDMTWGATAKEAEFSNFFIEVPKVLKKFKISMLFATIFIAGMIILAVAPFIPYSWHIKDFVAI
LPMATVAASHLLLPLVLNPALMTFS

SEQ ID NO: 16(*Neurospora sp.*)
MGIGSYFKAKKPEPAGQQHAASTPSRGRQPSMGNTKAGQDDGDILAPPQIRYGSRSRSATRSMMSSTSSVILEDIKHEVMVNYLYQ
QQCSYLWVANGSGEIEGVLLRKSRGQYMACPPALGNSPFAMACAALNVQCAMTVNSRVIKTFLQWSPDAVDVPLLNGLRVQILPTI
EDLPRARKHQFAAFIASEGLLVVWDDDALHLIPRAKEIESELMQLVWKTGEPGEMDEKANPIVGATEIDEESGEPRPEARPVHLLN
TYLVSITMAVVTVSLGAAWRQLAIEVMVDGDYVRLALVALAPVQIFFTLFFAQVIIGCLAQIFGPIKQLSVNSRFYSAKPPPRLQT
AVLPHVTVQCPVYKEGLSGVIAPTVKSIKHAMSTYELQGGSANMFINDDGLQLLSEEDRQARIDFYADHSIGWVARPRHGENGFQR
RGKFKKASNMNYALMISCKVEEKLAQVPRHSEWSQHDEAQAYERALKDVLEENGRAWADGNIRMGDYILLIDSDTRVPSDCLLDAV
SEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYTAIRYTVSNGDVAPFVGHNAILRWSAIQQVSYEDEDGYEKFWSESHVS
EDFDMSLRLQCNDYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFYPIRKWIWKGPFTPLFRRFLFSNIRFTSKITII
SYIGTYYAIGAAWILTSVNYFLMGWYNGFLDKYYVDSWQVWFSIILVFNGLGNVALAVMRYRVGERSILGSILENFKWTLMLAIFL
GGLSLHVSQALLAHMFEIDMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSIGFIIGMVILATAPFIPHSWHITDFVAILPMATV
AASHLLLPLALNPALMTFSW

SEQ ID NO: 17 (*Candida sp.*)
MGITDYFVSKGATEAKQGKNSTDIQQNETLTFPEMLAGPTGASPRAQFTIGETHLSSDIRTLQGHNEMISAGKSDFKDPQFIVVNY
LHDICLGNGWLKLVDPLEPCVVKMNSKKGNEIGYRYLPACDEIAPYSFVDCAARFLRSDVCVRVSFPVIHAILDILSSKGTVSLDA
DHNIQIIETVADLQWVRKSQNCAFIRNEKSLVCWADSVQEVTGFVRNLENKMVDYVWKKGNAVDVKGEDYVPRVTFASVFQSSSES
DVGSEGAVEIIGQNAVSQVSISEKSSDSSTHSDGNLNEKKNLDLEQQSSERPVIYIHATVSAFAITLVLAWAGLQFAQVTKEIRAE
GNYLILLSLLMVLPYFLFTSFFASSVMSTLLYVFGPISQMNKNSYSYSVHKAPRLKAAHGSLPHVTIQCPVYKEKLESVIKPTIKS
LQAAIRTYELQGGSANIFINDDGLQLIDRKEALERIEYYEECGLGYVARPGHGVNGFIRKGRFKKASNMNYCLHISKLVDTRFHER
LELIENPTPKEESGLYLKVLEEVVREEGKCWAGGDILLGDIILIIDSDTRVPEDCFVDSVSEMEQSPEVAIIQHASGVMMVVGNYW
EKMIAWFTNMIYFSISCVSGNGLTMAAFVGHNAFLRWSAIQELAYIDEDDGRTKYWSESHVSEDFEMTLKLASLGYTIRIATYHDG
GFKEGVSLTVYDEITRWSKYAFGCAEIMFSPFKDWWKGKIFARLFFVFLNSHISLPCKFSILGYMGTYYAIATSLIMLVANYFIVG
YYDWGYSRVYIDAMKVFVSVMVVFGCATQVAYIIGRYRIYKHSIYTMVLEFRYSILFSVFLGGLSWHMIVSIGSYFFSLNLQWGAT
AKDIDDSNFFKELPKAIKNYKFMYILCIFLIAGMIVLAFFVPYAFQIRLLTCALPLGWSVASHFLSPIVLNPQLMTFAW

FIG. 7B

CLUSTAL W (1.83) Multiple Sequence Alignment

SEQ ID NO: 2, residues 1-894 (bold single letter **AA** residues)

```
11                ------------------------------------------------------------
12                MEEWPRNADEPYGSQRYQEDVQDLVSLPSRSSTPHLDRAEIGRVLSPSVQPGRNSRSRRS
13                ------------------------------------------------------------
15                ------------------------------------------------------------
16                MGIGSYFKA---------KKPEPAGQQHAASTPSRGRQPSMGNTKAGQDDGD--ILAPPQ
2           (1)MGIGSYFKASTKNAEQAPPAPPPAKLAPPPRLQLEGRQSVMSEKPPSSTSGNDVDLQPPT
14                MGIGSYFKADIP--PQTPIGPPPRRPSHRPSMSLH--APIIEEKIPPAAVVELHSAAGPK
17                MGITDYFVSKGATEAKQGKNSTDIQQNETLTFPEMLAGPTGASPRAQFTIGETHLSSDIR

11                ------------------------------------------------------------
12                NPSIAHSFTSPSISESGEGEDPSSTARYLARRAT--LLGWFDEVDVEDGRNSWSSTCISL
13                ------------------------------------------------------------
15                ------------------------------------------------------------
16                IRYGSRSRSAT-RSMMSSTSSVILEDIKHEVMVN--YLYQQQCSYLWVANGSGEIEGVLL
2                 PRFHSRPQSSSGRSTPSTQSSMFLDDIKHEVMVN--YLYQQQCSQLWVSDGSGEIEGVLL
14                FTAGPPSIAPSSKSGNSN----LLDDIKHEVMVN--YLYQQQCSQLWVGDGSGEIEGVLL
17                TLQGHNEMISAGKSDFKDPQFIVVNYLHDICLGNGWLKLVDPLEPCVVKMNSKKGNEIGY

11                ------------------------------------------------------------
12                RRADNEYTFFPYNVNPS-LIDAITRLGETAAMAMSSRFTSMLIDSILPGQKSLVVESTSA
13                ------------------------------------------------------------
15                ------------------------------------------------------------
16                RKSRGQYMACPPALGNSPFAMACAALNVQCAMTVNSRVIKTFLQWSPDAVDVPLLN--GL
2                 RKARGHYMACPPQLASSPFALACAALNVQCAMTVNSRVIKTFLQWSPDAVDVPLMN--GL
14                RKSRGQYMACPPQLGQSPFAIACTALNVQCAMTVNSRVIKTFLQWSPDAVDVPLMN--GL
17                RYLPACDEIAPYSFVDCAARFLRSDVCVRVSFPVIHAILDILSSKGTVSLDA------DH

11                ------------------------------------------------------------
12                RIPVVHSLNDVSSNLVHFARACIVVQEKLVLIWSHNAGTILHVAHDVERQLGKKSPRIST
13                ------------------------------------------------------------
15                ------------------------------------------------------------
16                RVQILPTIEDLPR-ARKHQFAAFIASEGLLVVWDDDALHLIPRAKEIESELMQLVWKTGE
2                 RVQILPTIDDLPR-ARKYQFAAFVASEGLLVVWDDDALHLVARAKAIESELMELVWKAGN
14                RVQILPTVNDLPR-ARKHQFAAFIASDGLLVVWDDDALNLMARAKIIESELMELVWNSGQ
17                NIQIIETVADLQW-VRKSQNCAFIRNEKSLVCWADSVQEVTGFVRNLENKMVDYVWKKGN

11                ------------------------------------------------------------
12                ATTPRASGRSSPFDVDADVAPSTQISIKAKIDSFPVRG---ALDEKHEVYSKAVALEEGT
13                ------------------------------------------------------------
15                ------------------------------------------------------------
16                PGEM--DEKANPIVGATEIDE----------------------------------------
2                 VDEEGGDEKGGQPVTEVEIDE----------------------------------------
14                SVDE--DERDSTIAAEYEIDE----------------------------------------
17                AVDVKGEDYVPRVTFASVFQSSSSESDVGSEGAVEIIGQNAVSQVSISEKSSDSSTHSDGN
```

FIG. 7C

CLUSTAL W (1.83) Multiple Sequence Alignment

SEQ ID NO: 2, residues 1-894 (bold single letter **AA** residues)—*Continued*

```
11              ------------------------------------------------------------
12              VEDDTPVDLEGNLPPRPAHRIHAVKISLAIMLVILTQSLGVSRLVNEFAWDGSFTRFALV
13              ------------------------------------------------------------
15              ------------------------------------------------------------
16              ------ESGEPRPEARPVHLLNTYLVSITMAVVTVSLGAAWRQLAIEVMVDGDYVRLALV
2               ------ESGEIKPEKRPIHLQNTVLVSLTLALVTVSLGAAWRQLAIEVSVDSNYIRLALV
14              ------ESGEIKPEARPVHLQNAVLVSLTLLLVMASLGAAWRQLAVEIAIDGDYKRLGLV
17              LNEKKNLDLEQQSSERPVIYIHATVSAFAITLVLAWAGLQFAQVTKEIRAEGNYLILLSL

11              ------------------------------------------------------------
12              VTIPPLALFSLFFFIVLVTSAFQLFLPASFCLKNSKFHSAIKPNPRFHRDYELPHITVQM
13              ------------------------------------------------------------
15              ------------------------------------------------------------
16              ALAPVQIFFTLFFAQVIIGCLAQIFGPIKQLSVNSRFYSAKPPPRLQT--AVLPHVTVQC
2               ALAPVQVFFTLFFAQVIIGCLAQIFGPIRQLTINSKFYSARPPPRLQS--AVLPHVTIQC
14              ALFPIQIFFSLFFAQVIVGCLAQIFGPIRQLTINSKFYSARPPRRLQG--ATLPHITIQC
17              LMVLPYFLFTSFFASSVMSTLLYVFGPISQMNKNSYSYSVHKAPRLKAAHGSLPHVTIQC

11              ------------------------------------------------------------
12              PVYKEGLKGVIVPTMMSVLAAVQYYEEQGGTASVFINDDGMQCIQPELAEARKQYYRENG
13              ------------------------------------------------------------
14              ------------------------------------------------------------
15              PVYKEGLSGVIAPTVKSIKHAMSTYELQGGSANMFINDDGLQLLSEEDRQARIDFYADHS
2               PVYKEGLQGVIMPTVKSIKQAMSTYELQGGSANMFINDDGLQLISEEDRLARIEFYADNS
16              PVYKEGLNAVILPTVKSIKQAMSTYELQGGSANMFINDDGLQLLSEEERDARIDFYADNS
17              PVYKEKLESVIKPTIKSLQAAIRTYELQGGSANIFINDDGLQLIDRKEALERIEYYEECG

11              ----------KPATTEP---------------ADSEEPEGPYTPQALANKIGFERKGKFK
12              IGYTARLPNRKTASKKKRGFGWFRKAKSAEGDAETEAEEDTSSPQAIANKIGFERKGKFK
13              ------------------------------------------------------------
15              ------------------------------------------------------------
16              IG--------------------------------------WVARPRHGENGFQRRGKFK
2               IG--------------------------------------WVARPKHGENGFTRKGKFK
14              IG--------------------------------------WVARPKHGEDGFIRKGKFK
17              LG--------------------------------------YVARPGHGVNGFIRKGRFK

11              KASNMNYGLAFSLRVEDELARLTQIETERRGCTVDDLTAEDDDRLYQQALDNMLAADEER
12              KASNMNYGLAFSNRVEDELARLADLECQQRGCSNDDLTFEDDDRLYQQALSNMLAEDEGR
13              ----------------------------------------EEELVYREAFDHVIQSDPRI
15              ------------------------------------------------------------
16              KASNMNYALMISCKVEEKLAQVPRHS---------EWSQHDEAQAYERALKDVLEEN-GR
2               KASNMNFALMISCKVEEKLQAIERHP---------GWSQNDEALAYEEALKEVLEAD-GR
14              KASNMNFGLMISCKVEERLQLIKRPA---------DWSQSDEALAYEQALKDVLEEN-GR
17              KASNMNYCLHISKLVDTRFHERLELIEN--------PTPKEESGLYLKVLEEVVREE-GK
```

FIG. 7D

CLUSTAL W (1.83) Multiple Sequence Alignment

SEQ ID NO: 2, residues 1-894 (bold single letter **AA** residues)—*Continued*

```
11              TWAEGNIRIGELILLIDCDTRVPVDCLLYGALEMHESPEVAILQHGSGVMQVVHNVFENG
12              TWAEGNVRIGEIILIIDCDTRVPVDCLLYGALEMHESPEVAIVQHGSGVMQVVHNMFENG
13              RGAGGDIRVGEFILIVDSDTRIPADCLLYGAAEMFLSPEVAIIQHSTSVMQVSQDYFENG
15              -----------------SDTRVPADCLLDAVSEMELSPDVGIMQFSSGVMQVVHTYFENG
16              AWADGNIRMGDYILLIDSDTRVPSDCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG
2               AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG
14              AWADGNIRVGDYILLIDSDTRVPTDCLLDSVSEMEQSPDVGIMQFSSGVMQVVHTYFENG
17              CWAGGDILLGDIILIIDSDTRVPEDCFVDSVSEMEQSPEVAIIQHASGVMMVVGNYWEKM
                             .***:* **:: .. **  **:*.*:*..:.** *    :*:

11              ITYFTNVVYTAIKYGVGSG-DVSPFVGHNAFLRWRALQSIEFVDPSDGQTKWWSDTHVSE
12              ITYFTDVVYTAIKYGVGSG-DVSPFVGHNAFLRWKALQSIQFVDPADGQTKWWSDAHVSE
13              ITYFTNLIYSAIRFAVGSG-ETAPFVGHNAFLRWQAVQSVG--RPDDGYVSFWSESHVSE
15              ITFFTNLIYSAIRYTVSNG-DVAPFVGHNAILRWSAIQQVSY-EDEDGYEKFWSESHVSE
16              ITFFTNLIYTAIRYTVSNG-DVAPFVGHNAILRWSAIQQVSY-EDEDGYEKFWSESHVSE
2               ITFFTNLIYSAIRYTVSNG-DVAPFVGHNAILRWSAIQQVAY-QDEDGYDKFWSESHVSE
14              ITFFTDLIYTAIRFTVSNG-DVAPFVGHNAILRWSAIQQVAY-QDEDGYDKFWSESHVSE
17              IAWFTNMIYFSISCVSGNGLTMAAFVGHNAFLRWSAIQELAYIDEDDGRTKYWSESHVSE
                *::**:::* :*    ..*   :.******:*** *:*.:      **  .:**::****

11              DFDISLRLQMQGMVVRLATYHNGEFKEGVSLTLYDELTRWEKYAYGCNELVFHPFYQWVY
12              DFDISLRLQMQGMTVRLATYHNGGFKEGVSLTLYDELTRWEKYAYGCNELVFNPFYQWPY
13              DFDIALRLQIQGNIIRLASYHNDEFKEGVSLTIYDELSRWEKYAYGCNELVFNPVHTWFY
15              DFDMSLRLQCAGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRKWIY
16              DFDMSLRLQCNDYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFYPIRKWIW
2               DFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLW
14              DFDMSLRLQCNGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAFGCNELLFNPIRTWLW
17              DFEMTLKLASLGYTIRIATYHDGGFKEGVSLTVYDEITRWSKYAFGCAEIMFSPFKDWWK
                **:::*:*   .  :*:*:: .  ********:***::**.***:** *::* *.  *

11              KGPVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIGSGMLLATVNYVILGLFDSDIDHLYL
12              KGPVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIASGMMLSVVNYVIVGLFNSEVDHIYL
13              RGPLTKLFMTFLWSNLQLSSKITILGYISSYYALASGFPLTVLNYFLVGWFEGYLDKFYM
15              KGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIAAAWIMTSINYFIMGWFNGYLDKYYV
16              KGPFTPLFRRFLFSNIRFTSKITIISYIGTYYAIGAAWILTSVNYFLMGWYNGFLDKYYV
2               RGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYV
14              RGPFTPLFRRFCFSNIRFTSKITVVSYIGTYYAIGAAWIMTTANYFLMGWFNGYLDKYYL
17              GKIFARLFFVFLNSHISLPCKFSILGYMGTYYAIATSLIMLVANYFIVGYYDWGYSRVYI
                   .: **  *  *:: ...*.::::.*: :***:.:.  :   **.::* ::   .: *:

11              PSWGIWCSLVVVFNGFSSIAFSMVRHQLKEETFWRALLDAIKWLPFLIIYFGGISLNCAK
12              RSWGIWISLVVVFNGVASVAFSMARHQLKEMVFWKALLKSALWLPFLVVFFGGISLNCAK
13              ESWKVFLSLLVVFSAAGNVCLAIIRYRLGEKPLLASLVENFMWMPMFAIFFGGLSFHLSL
15              DSWQVWFSIILVFNGLGNIALAVMRYRVGERSILYALYENFKWTFLLAVFLGGLSLHLSQ
16              DSWQVWFSIILVFNGLGNVALAVMRYRVGERSILGSILENFKWTLMLAIFLGGLSLHVSQ
2               DSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQ
14              DSWQVWFSIILVFNGLGNLALAIMRYRSGERTLGYAIYENFKWTLMLAVFLGGLSLHVSQ
17              DAMKVFVSVMVVFGCATQVAYIIGRYRIYKHSIYTMVLE-FRYSILFSVFLGGLSWHMIV
                 :  :: *:::**.   .:.  : *::  :  :   : .   :  :: :::**:* :
```

FIG. 7E

CLUSTAL W (1.83) Multiple Sequence Alignment

SEQ ID NO: 2, residues 1-894 (bold single letter **AA** residues)—*Continued*

```
11              AILCHAFSINLEWASTAKEMGPTGFYIGMDKMVRRFKWTWAICLVLAGVMIYFAVG--AP
12              AILCHAFSINIEWASTAKEPGPSGFFIGLDKMVKQFKYTWAICLFLAAVMIFMALG--TP
13              AILSHMFGINMSWGTTAKEKDDSNFFKEIPKIFKSFKWMYAVVLPFFPAMIYLACF--AP
15              ALLAHMFEIDMTWGATAKEAEFSNFFIEVPKVLKKFKISMLFATIFIAGMIILAVAPFIP
16              ALLAHMFEIDMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSIGFIIGMVILATAPFIP
2               ALLAHMFEINMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVP
14              ALLAHMFEIDMSWGSTSKEAEFSNFFIEVPKVLKKFRVSMTLSLLAIVALIIMATADFIP
17              SIGSYFFSLNLQWGATAKDIDDSNFFKELPKAIKNYKFMYILCIFLIAGMIVLAFF--VP
                :: .: * ::: *.:*:*:   :.*:  : * .: ::    .       :: :*     *

11              WGWTITPGPYSTANVAIAPLAIQICSASFLPFFLGLN------
12              WGWQIKPGEYSTASIAIGPLAIQICNAAILPLILGLN------
13              NGWTIT------EVGAIVPMSVTLASHALLPLLLNPSLMVFNY
15              YSWHIK------DFVAILPMATVAASHLLLPLVLNPALMTFS-
16              HSWHIT------DFVAILPMATVAASHLLLPLALNPALMTFSW
2               FDWRIK------DFVAILPMATVAASHFLLPLALNPALMTFSW(894)
14              HYWRIN------DFVAILPMATVAGSHLLLPLALNPALMTFSW
17              YAFQIR------LLTCALPLGWSVASHFLSPIVLNPQLMTFAW
                  :*      . *:.  . :*: *.
```

FIG. 7F

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 11 (*Aspergillus* sp. LOV protein orthologue)

MTMLSPTSDHPQGASMKDIENPGNDSPALEGSTQAEKTVDNPSRSPTPQLQTAELAQVLSPLAVRQTQSISTRSQHN
RLEYLCLPNATNRGKLSRSLARRATLLGWFDQAATQSGQESQSSVCVAVKLPDDGYSFSPSNITPALSKAITRLNEM
AVVALSSRVVDKGLSGILPGQKSFLVESTSTRIPIVATLDDVEPTLAHYSRACIVMEQKIVLVWSHDTAGILNVAYD
VETQLGGQSPKISNATTPRISGRSTPMDPYRTDDLQKPANVLQPVRGALDEKQAIYSKAVALEEGEDNEVPDLERNA
APRPVLLVHTVKISLAIMLVIVTQSLGVARLLNEFQWDGQYTRFALVVTIPPLALFSLFFFIVLVTSVFQLFLPASF
CLKNSKFHSGDTEDSSAVKPNPKAHGEYELPHITIQMPVYKEGLKGVIVPTMISVLAAVQYYEEQGGTASVFVNDDG
MQVIQSDLAEARKQYYRENGIGFTARPPNKKSPVQKGGWGSWFRKSKPATTEPADSEEPEGPYTPQALANKIGFERK
GKFKKASNMNYGLAFSLRVEDELARLTQIETERRGCTVDDLTAEDDDRLYQQALDNMLAADEERTWAEGNIRIGELI
LLIDCDTRVPVDCLLYGALEMHESPEVAILQHGSGVMQVVHNVFENGITYFTNVVYTAIKYGVGSGDVSPFVGHNAF
LRWRALQSIEFVDPSDGQTKWWSDTHVSEDFDISLRLQMQGMVVRLATYHNGEFKEGVSLTLYDELTRWEKYAYGCN
ELVFHPFYQWVYKGPVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIGSGMLLATVNYVILGLFDSDIDHLYLPSWGI
WCSLVVVFNGFSSIAFSMVRHQLKEETFWRALLDAIKWLPFLIIYFGGISLNCAKAILCHAFSINLEWASTAKEMGP
TGFYIGMDKMVRRFKWTWAICLVLAGVMIYFAVGAPWGWTITPGPYSTANVAIAPLAIQICSASFLPFFLGLN

Score	Expect	Method	Identities	Positives	Gaps
409 bits(1052)	9e-139	Compositional matrix	201/395(51%)	273/395(69%)	11/395(2%)

```
Query  501  WADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGI  560
            WA+GNIR+G+ ILLID DTRVP DCLL    EM +SP+V I+Q  SGVMQVVH  FENGI
Sbjct  605  WAEGNIRIGELILLIDCDTRVPVDCLLYGALEMHESPEVAILQHGSGVMQVVHNVFENGI  664

Query  561  TFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQD-EDGYDKFWSESHVSEDF  619
            T+FTN++Y+AI+Y V +GDV+PFVGHNA LRW A+Q + + D  DG  K+WS++HVSEDF
Sbjct  665  TYFTNVVYTAIKYGVGSGDVSPFVGHNAFLRWRALQSIEFVDPSDGQTKWWSDTHVSEDF  724

Query  620  DMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRG  679
            D+SLRLQ  G ++RLA +    FKEGVSLT+YDEL RWEKYAYGCNEL+FHP   W+++G
Sbjct  725  DISLRLQMQGMVVRLATYHNGEFKEGVSLTLYDELTRWEKYAYGCNELVFHPFYQWVYKG  784

Query  680  PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDS  739
            P T LF RFL+SN+  TSKVT+I+YI TYYAIG+  +L  VNY ++G F+  +D  Y+ S
Sbjct  785  PVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIGSGMLLATVNYVILGLFDSDIDHLYLPS  844

Query  740  WQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQAL  799
            W +W S+++VFNG  +IA +++R+++ E     AL +   W   L I+ GG+SL+ ++A+
Sbjct  845  WGIWCSLVVVFNGFSSIAFSMVRHQLKEETFWRALLDAIKWLPFLIIYFGGISLNCAKAI  904

Query  800  LAHMFEINMTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFD  859
            L H F IN+ W +T+KE   + F+I + K++++FK++    L+    MI  A     P+
Sbjct  905  LCHAFSINLEWAS**T**AKEMGPTGFYIGMDKMVRRFKWTWAICLVLAGVMIYFAVG--APWG  962

Query  860  WRIKD------FVAILPMATVAASHFLLP--LALN  886
            W I         VAI P+A    S   LP  L LN
Sbjct  963  WTITPGPYSTANVAIAPLAIQICSASFLPFFLGLN  997
```

FIG. 8A

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 12 (*Penicillium sp.* LOV protein orthologue)

MEEWPRNADEPYGSQRYQEDVQDLVSLPSRSSTPHLDRAEIGRVLSPSVQPGRNSRSRRSNPSIAHSFTSPSISESG
EGEDPSSTARYLARRATLLGWFDEVDVEDGRNSWSSTCISLRRADNEYTFFPYNVNPSLIDAITRLGETAAMAMSSR
FTSMLIDSILPGQKSLVVESTSARIPVVHSLNDVSSNLVHFARACIVVQEKLVLIWSHNAGTILHVAHDVERQLGKK
SPRISTATTPRASGRSSPFDVDADVAPSTQISIKAKIDSFPVRGALDEKHEVYSKAVALEEGTVEDDTPVDLEGNLP
PRPAHRIHAVKISLAIMLVILTQSLGVSRLVNEFAWDGSFTRFALVVTIPPLALFSLFFFIVLVTSAFQLFLPASFC
LKNSKFHSAIKPNPRFHRDYELPHITVQMPVYKEGLKGVIVPTMMSVLAAVQYYEEQGGTASVFINDDGMQCIQPEL
AEARKQYYRENGIGYTARLPNRKTASKKKRGFGWFRKAKSAEGDAETEAEEDTSSPQAIANKIGFERKGKFKKASNM
NYGLAFSNRVEDELARLADLECQQRGCSNDDLTFEDDDRLYQQALSNMLAEDEGRTWAEGNVRIGEIILIIDCDTRV
PVDCLLYGALEMHESPEVAIVQHGSGVMQVVHNMFENGITYFTDVVYTAIKYGVGSGDVSPFVGHNAFLRWKALQSI
QFVDPADGQTKWWSDAHVSEDFDISLRLQMQGMTVRLATYHNGGFKEGVSLTLYDELTRWEKYAYGCNELVFNPFYQ
WPYKGPVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIASGMMLSVVNYVIVGLFNSEVDHIYLRSWGIWISLVVVFN
GVASVAFSMARHQLKEMVFWKALLKSALWLPFLVVFFGGISLNCAKAILCHAFSINIEWASTAKEPGPSGFFIGLDK
MVKQFKYTWAICLFLAAVMIFMALGTPWGWQIKPGEYSTASIAIGPLAIQICNAAILPLILGLN

Score	Expect	Method	Identities	Positives	Gaps
411 bits(1057)	1e-139	Compositional	199/392(51%)	277/392(70%)	9/392(2%)

```
Query  501  WADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGI  560
            WA+GN+R+G+ IL+ID DTRVP DCLL    EM +SP+V I+Q  SGVMQVVH  FENGI
Sbjct  596  WAEGNVRIGEIILIIDCDTRVPVDCLLYGALEMHESPEVAIVQHGSGVMQVVHNMFENGI  655

Query  561  TFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQDE-DGYDKFWSESHVSEDF  619
            T+FT+++Y+AI+Y V +GDV+PFVGHNA LRW A+Q + + D  DG  K+WS++HVSEDF
Sbjct  656  TYFTDVVYTAIKYGVGSGDVSPFVGHNAFLRWKALQSIQFVDPADGQTKWWSDAHVSEDF  715

Query  620  DMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRG  679
            D+SLRLQ  G  +RLA +   GFKEGVSLT+YDEL RWEKYAYGCNEL+F+P   W ++G
Sbjct  716  DISLRLQMQGMTVRLATYHNGGFKEGVSLTLYDELTRWEKYAYGCNELVFNPFYQWPYKG  775

Query  680  PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDS  739
            P T LF RFL+SN+  TSKVT+I+YI TYYAI +  +L+ VNY ++G FN  +D  Y+ S
Sbjct  776  PVTRLFLRFLWSNMPVTSKVTIIAYIFTYYAIASGMMLSVVNYVIVGLFNSEVDHIYLRS  835

Query  740  WQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQAL  799
            W +W S+++VFNG+ ++A ++ R+++ E     AL ++ +W   L +F GG+SL+ ++A+
Sbjct  836  WGIWISLVVVFNGVASVAFSMARHQLKEMVFWKALLKSALWLPFLVVFFGGISLNCAKAI  895

Query  800  LAHMFEINMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFD  859
            L H F IN+ W +T+KE   S FFI + K++K+FK++    L   A MI +A     P+
Sbjct  896  LCHAFSINIEWASTAKEPGPSGFFIGLDKMVKQFKYTWAICLFLAAVMIFMALG--TPWG  953

Query  860  WRIKD------FVAILPMATVAASHFLLPLAL  885
            W+IK        +AI P+A    +  +LPL L
Sbjct  954  WQIKPGEYSTASIAIGPLAIQICNAAILPLIL  985
```

FIG. 8B

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 13 (*Talaromyces sp.* LOV protein orthologue)

MSSLFPTRWSFRGHKTPKGDGTPTETASQRDSLPVSESGRSSRGQGYFDRRHVNNSIDDASKYKAMIKFFHVRLTAY
QWLPPPTHPQHSSTGVFLRRSRGIYMSEPEDINPLLLAAIQRINATIAFTMMTETTSIITSQLAPGQTELILPNGYQ
VQIIESYADIVGSHSNMVKKYQYAALIREEQLLLVWNDDLNAILNHAADVEGKLLSLIWGSPIPTFNLQAVPMMTPG
ESVVASPNDSLYHLALEPRESPAAAEDSGTSRDASPRRMINEEVKRPKESLERPLAVTSAIFVGMAGMLLVILLLGF
GISNLLLEYSVDGGAMRFALTATIPFFLLFSIFFMIVIFTDIFQAVGPVKTLKSNSRFYSPIAPDLKTAYSLGFTPP
RVTIQMPIYTESLEGVIKPTISSLKTAISHYESHGGTANIFINDDGFALLSEEQQHERINFYHDNNIGWVARPKNNE
DGYIRKGKFKKASNMNFALNVSNKVEMELIQRMAPTLEKSDMVDPMEEELVYREAFDHVIQSDPRIRGAGGDIRVGE
FILIVDSDTRIPADCLLYGAAEMFLSPEVAIIQHSTSVMQVSQDYFENGITYFTNLIYSAIRFAVGSGETAPFVGHN
AFLRWQAVQSVGRPDDGYVSFWSESHVSEDFDIALRLQIQGNIIRLASYHNDEFKEGVSLTIYDELSRWEKYAYGCN
ELVFNPVHTWFYRGPLTKLFMTFLWSNLQLSSKITILGYISSYYALASGFPLTVLNYFLVGWFEGYLDKFYMESWKV
FLSLLVVFSAAGNVCLAIIRYRLGEKPLLASLVENFMWMPMFAIFFGGLSFHLSLAILSHMFGINMSWGTTAKEKDD
SNFFKEIPKIFKSFKWMYAVVLPFFPAMIYLACFAPNGWTITEVGAIVPMSVTLASHALLPLLLNPSLMVFNY

Score	Expect	Method	Identities	Positives	Gaps
477 bits(1227)	9e-166	Compositional	222/377(59%)	290/377(76%)	3/377(0%)

```
Query  502  ADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGIT  561
            A G+IR+G++IL++DSDTR+PADCLL   +EM  SP+V I+Q S+ VMQV   YFENGIT
Sbjct  531  AGGDIRVGEFILIVDSDTRIPADCLLYGAAEMFLSPEVAIIQHSTSVMQVSQDYFENGIT  590

Query  562  FFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDM  621
            +FTNLIYSAIR+ V +G+ APFVGHNA LRW A+Q V   D DGY  FWSESHVSEDFD+
Sbjct  591  YFTNLIYSAIRFAVGSGETAPFVGHNAFLRWQAVQSVGRPD-DGYVSFWSESHVSEDFDI  649

Query  622  SLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRGPF  681
            +LRLQ  G IIRLA++  + FKEGVSLT+YDEL+RWEKYAYGCNEL+F+P+ TW +RGP
Sbjct  650  ALRLQIQGNIIRLASYHNDEFKEGVSLTIYDELSRWEKYAYGCNELVFNPVHTWFYRGPL  709

Query  682  TPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDSWQ  741
            T LF  FL+SN++ +SK+T++ YI +YYA+ + + LT +NYF++GWF GYLDK+Y++SW+
Sbjct  710  TKLFMTFLWSNLQLSSKITILGYISSYYALASGFPLTVLNYFLVGWFEGYLDKFYMESWK  769

Query  742  VWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLA  801
            V+ S+++VF+  GN+ LA++RYR+GE+ LL +L ENFMW  M AIF GGLS H+S A+L+
Sbjct  770  VFLSLLVVFSAAGNVCLAIIRYRLGEKPLLASLVENFMWMPMFAIFFGGLSFHLSLAILS  829

Query  802  HMFEINMTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWR  861
            HMF INM+WG T+KE + SNFF E+PK+ K FK+     L F   MI LA   F P  W
Sbjct  830  HMFGINMSWGT**T**AKEKDDSNFFKEIPKIFKSFKWMYAVVLPFFPAMIYLAC--FAPNGWT  887

Query  862  IKDFVAILPMATVAASH  878
            I +  AI+PM+   ASH
Sbjct  888  ITEVGAIVPMSVTLASH  904
```

FIG. 8C

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 14 (*Fusarium sp.* LOV protein orthologue)

MGIGSYFKADIPPQTPIGPPPRRPSHRPSMSLHAPIIEEKIPPAAVVELHSAAGPKFTAGPPSIAPSSKSGNSNLLD
DIKHEVMVNYLYQQQCSQLWVGDGSGEIEGVLLRKSRGQYMACPPQLGQSPFAIACTALNVQCAMTVNSRVIKTFLQ
WSPDAVDVPLMNGLRVQILPTVNDLPRARKHQFAAFIASDGLLVVWDDDALNLMARAKIIESELMELVWNSGQSVDE
DERDSTIAAEYEIDEESGEIKPEARPVHLQNAVLVSLTLLLVMASLGAAWRQLAVEIAIDGDYKRLGLVALFPIQIF
FSLFFAQVIVGCLAQIFGPIRQLTINSKFYSARPPRRLQGATLPHITIQCPVYKEGLNAVILPTVKSIKQAMSTYEL
QGGSANMFINDDGLQLLSEEERDARIDFYADNSIGWVARPKHGEDGFIRKGKFKKASNMNFGLMISCKVEERLQLIK
RPADWSQSDEALAYEQALKDVLEENGRAWADGNIRVGDYILLIDSDTRVPTDCLLDSVSEMEQSPDVGIMQFSSGVM
QVVHTYFENGITFFTDLIYTAIRFTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQ
CNGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAFGCNELLFNPIRTWLWRGPFTPLFRRFCFSNIRFTSKITVVS
YIGTYYAIGAAWIMTTANYFLMGWFNGYLDKYYLDSWQVWFSIILVFNGLGNLALAIMRYRSGERTLGYAIYENFKW
TLMLAVFLGGLSLHVSQALLAHMFEIDMSWGSTSKEAEFSNFFIEVPKVLKKFRVSMTLSLLAIVALIIMATADFIP
HYWRINDFVAILPMATVAGSHLLLPLALNPALMTFSW

Score	**Expect**	**Method**	**Identities**	**Positives**	**Gaps**
717 bits(1852)	0.0	Compositional	347/395(88%)	374/395(94%)	0/395(0%)

```
Query  500  AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG  559
            AWADGNIR+GDYILLIDSDTRVP DCLLD+VSEMEQSPDVGIMQFSSGVMQVVHTYFENG
Sbjct  490  AWADGNIRVGDYILLIDSDTRVPTDCLLDSVSEMEQSPDVGIMQFSSGVMQVVHTYFENG  549

Query  560  ITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDF  619
            ITFFT+LIY+AIR+TVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDF
Sbjct  550  ITFFTDLIYTAIRFTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDF  609

Query  620  DMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRG  679
            DMSLRLQCNGYIIRLAAWAG+GFKEGVSLTVYDELARWEKYA+GCNELLF+PIRTWLWRG
Sbjct  610  DMSLRLQCNGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAFGCNELLFNPIRTWLWRG  669

Query  680  PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDS  739
            PFTPLFRRF FSNIRFTSK+TV+SYIGTYYAIGAAWI+T  NYF+MGWFNGYLDKYY+DS
Sbjct  670  PFTPLFRRFCFSNIRFTSKITVVSYIGTYYAIGAAWIMTTANYFLMGWFNGYLDKYYLDS  729

Query  740  WQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQAL  799
            WQVWFSII+VFNGLGN+ALA+MRYR GER L YA++ENF WTLMLA+FLGGLSLHVSQAL
Sbjct  730  WQVWFSIILVFNGLGNLALAIMRYRSGERTLGYAIYENFKWTLMLAVFLGGLSLHVSQAL  789

Query  800  LAHMFEINMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFD  859
            LAHMFEI+M+WG+TSKEAEFSNFFIEVPKVLKKF+ SM  SL+ I  +II+A A F+P
Sbjct  790  LAHMFEIDMSWGSTSKEAEFSNFFIEVPKVLKKFRVSMTLSLLAIVALIIMATADFIPHY  849

Query  860  WRIKDFVAILPMATVAASHFLLPLALNPALMTFSW  894
            WRI DFVAILPMATVA SH LLPLALNPALMTFSW
Sbjct  850  WRINDFVAILPMATVAGSHLLLPLALNPALMTFSW  884
```

FIG. 8D

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 15 (*Myceliophthora sp.* LOV protein orthologue)

MGIGDYFKAEKLGSKPTPTPASPPREHGRHQQQPSASEDHPAPSVQPASELQPPTPRFSSRPQSISGRSVRSTGSSV
LDEIKHEVMVNYLYQQQCSHLWISDGSGEIEGVLLRKARGQYMACPPQLVNSPLAAACTALNVQCAMTVNSRVIKTF
LQWSPDAVDVPLMNGMRVQILATIDDLPRARKHQFAAFVASEGLLIVWDDDALHLVQRAKAIESELMELVWKVGAED
NEDEKGVAAVEEPEVDEESGELKPEKRPVHLLNAYLVSLSLILVTVSLGAAFRQLAIEVSVDGNYVRLALVALFPVQ
MFFTLFFAQVIVGCLAQIFGPIRQLTVNSKFYSARPPPRLRSSVLPHVTVQCPVYKEGLNAVIAPTVKSIKQAMSTY
ELQGGSANMFINDDGLQLISEEDRRARIEFYADNSIGWVARPKHGENGFQRRGKFKKASNMNFALMISCKVEDKLAA
IQRTPDWTQHDEALAYERALKEVLEEDGRAWADGNIRIGDYILLVDSDTRVPADCLLDAVSEMELSPDVGIMQFSSG
VMQVVHTYFENGITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVSYEDEDGYEKFWSESHVSEDFDMSLR
LQCAGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRKWIYKGPFTPLFRRFLFSNIRFTSKVTV
ISYIGTYYAIAAAWIMTSINYFIMGWFNGYLDKYYVDSWQVWFSIILVFNGLGNIALAVMRYRVGERSILYALYENF
KWTFLLAVFLGGLSLHLSQALLAHMFEIDMTWGATAKEAEFSNFFIEVPKVLKKFKISMLFATIFIAGMIILAVAPF
IPYSWHIKDFVAILPMATVAASHLLLPLVLNPALMTFS

Score	Expect	Method	Identities	Positives	Gaps
759 bits(1961)	0.0	Compositional	356/394(90%)	380/394(96%)	0/394(0%)

```
Query  500  AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG  559
            AWADGNIR+GDYILL+DSDTRVPADCLLDAVSEME SPDVGIMQFSSGVMQVVHTYFENG
Sbjct  492  AWADGNIRIGDYILLVDSDTRVPADCLLDAVSEMELSPDVGIMQFSSGVMQVVHTYFENG  551

Query  560  ITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDF  619
            ITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQV+Y+DEDGY+KFWSESHVSEDF
Sbjct  552  ITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVSYEDEDGYEKFWSESHVSEDF  611

Query  620  DMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRG  679
            DMSLRLQC GYIIRLAAWAG+GFKEGVSLTVYDELARWEKYAYGCNELLFHPIR W+++G
Sbjct  612  DMSLRLQCAGYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRKWIYKG  671

Query  680  PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDS  739
            PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAI AAWI+T++NYF+MGWFNGYLDKYYVDS
Sbjct  672  PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIAAAWIMTSINYFIMGWFNGYLDKYYVDS  731

Query  740  WQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQAL  799
            WQVWFSII+VFNGLGNIALAVMRYRVGER +LYAL+ENF WT +LA+FLGGLSLH+SQAL
Sbjct  732  WQVWFSIILVFNGLGNIALAVMRYRVGERSILYALYENFKWTFLLAVFLGGLSLHLSQAL  791

Query  800  LAHMFEINMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFD  859
            LAHMFEI+MTWGAT+KEAEFSNFFIEVPKVLKKFK SMLF+ IFIAGMIILA APF+P+
Sbjct  792  LAHMFEIDMTWGATAKEAEFSNFFIEVPKVLKKFKISMLFATIFIAGMIILAVAPFIPYS  851

Query  860  WRIKDFVAILPMATVAASHFLLPLALNPALMTFS  893
            W IKDFVAILPMATVAASH LLPL LNPALMTFS
Sbjct  852  WHIKDFVAILPMATVAASHLLLPLVLNPALMTFS  885
```

FIG. 8E

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 16 (*Neurospora sp.* LOV protein orthologue)

MGIGSYFKAKKPEPAGQQHAASTPSRGRQPSMGNTKAGQDDGDILAPPQIRYGSRSRSATRSMMSSTSSVILEDIKH
EVMVNYLYQQQCSYLWVANGSGEIEGVLLRKSRGQYMACPPALGNSPFAMACAALNVQCAMTVNSRVIKTFLQWSPD
AVDVPLLNGLRVQILPTIEDLPRARKHQFAAFIASEGLLVVWDDDALHLIPRAKEIESELMQLVWKTGEPGEMDEKA
NPIVGATEIDEESGEPRPEARPVHLLNTYLVSITMAVVTVSLGAAWRQLAIEVMVDGDYVRLALVALAPVQIFFTLF
FAQVIIGCLAQIFGPIKQLSVNSRFYSAKPPPRLQTAVLPHVTVQCPVYKEGLSGVIAPTVKSIKHAMSTYELQGGS
ANMFINDDGLQLLSEEDRQARIDFYADHSIGWVARPRHGENGFQRRGKFKKASNMNYALMISCKVEEKLAQVPRHSE
WSQHDEAQAYERALKDVLEENGRAWADGNIRMGDYILLIDSDTRVPSDCLLDAVSEMEQSPDVGIMQFSSGVMQVVH
TYFENGITFFTNLIYTAIRYTVSNGDVAPFVGHNAILRWSAIQQVSYEDEDGYEKFWSESHVSEDFDMSLRLQCNDY
IIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFYPIRKWIWKGPFTPLFRRFLFSNIRFTSKITIISYIGT
YYAIGAAWILTSVNYFLMGWYNGFLDKYYVDSWQVWFSIILVFNGLGNVALAVMRYRVGERSILGSILENFKWTLML
AIFLGGLSLHVSQALLAHMFEIDMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSIGFIIGMVILATAPFIPHSWH
ITDFVAILPMATVAASHLLLPLALNPALMTFSW

Score	Expect	Method	Identities	Positives	Gaps
735 bits(1898)	0.0	Compositional	357/395(90%)	381/395(96%)	0/395(0%)

```
Query  500  AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG  559
            AWADGNIRMGDYILLIDSDTRVP+DCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG
Sbjct  486  AWADGNIRMGDYILLIDSDTRVPSDCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENG  545

Query  560  ITFFTNLIYSAIRYTVSNGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDF  619
            ITFFTNLIY+AIRYTVSNGDVAPFVGHNAILRWSAIQQV+Y+DEDGY+KFWSESHVSEDF
Sbjct  546  ITFFTNLIYTAIRYTVSNGDVAPFVGHNAILRWSAIQQVSYEDEDGYEKFWSESHVSEDF  605

Query  620  DMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRG  679
            DMSLRLQCN YIIRLAAWAG+GFKEGVSLTVYDELARWEKYAYGCNELLF+PIR W+W+G
Sbjct  606  DMSLRLQCNDYIIRLAAWAGDGFKEGVSLTVYDELARWEKYAYGCNELLFYPIRKWIWKG  665

Query  680  PFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDS  739
            PFTPLFRRFLFSNIRFTSK+T+ISYIGTYYAIGAAWILT+VNYF+MGW+NG+LDKYYVDS
Sbjct  666  PFTPLFRRFLFSNIRFTSKITIISYIGTYYAIGAAWILTSVNYFLMGWYNGFLDKYYVDS  725

Query  740  WQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQAL  799
            WQVWFSII+VFNGLGN+ALAVMRYRVGER +L ++ ENF WTLMLAIFLGGLSLHVSQAL
Sbjct  726  WQVWFSIILVFNGLGNVALAVMRYRVGERSILGSILENFKWTLMLAIFLGGLSLHVSQAL  785

Query  800  LAHMFEINMTWGAtSKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFD  859
            LAHMFEI+MTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFS+ FI GM+ILA APF+P
Sbjct  786  LAHMFEIDMTWGATSKEAEFSNFFIEVPKVLKKFKFSMLFSIGFIIGMVILATAPFIPHS  845

Query  860  WRIKDFVAILPMATVAASHFLLPLALNPALMTFSW  894
            W I DFVAILPMATVAASH LLPLALNPALMTFSW
Sbjct  846  WHITDFVAILPMATVAASHLLLPLALNPALMTFSW  880
```

FIG. 8F

SEQ ID NO: 18 (*Trichoderma sp.* LOV protein C-terminal residue positions 500-894)

AWADGNIRMGDYILLIDSDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS
NGDVAPFVGHNAILRWSAIQQVAYQDEDGYDKFWSESHVSEDFDMSLRLQCNGYIIRLAAWAGEGFKEGVSLTVYDE
LARWEKYAYGCNELLFHPIRTWLWRGPFTPLFRRFLFSNIRFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNG
YLDKYYVDSWQVWFSIIIVFNGLGNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEIN
MTWGA**T**SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMATVAASHFLLPLA
LNPALMTFSW

SEQ ID NO: 17 (*Candida sp.* LOV protein orthologue)

MGITDYFVSKGATEAKQGKNSTDIQQNETLTFPEMLAGPTGASPRAQFTIGETHLSSDIRTLQGHNEMISAGKSDFKDPQFIVVNY
LHDICLGNGWLKLVDPLEPCVVKMNSKKGNEIGYRYLPACDEIAPYSFVDCAARFLRSDVCVRVSFPVIHAILDILSSKGTVSLDA
DHNIQIIETVADLQWVRKSQNCAFIRNEKSLVCWADSVQEVTGFVRNLENKMVDYVWKKGNAVDVKGEDYVPRVTFASVFQSSSES
DVGSEGAVEIIGQNAVSQVSISEKSSDSSTHSDGNLNEKKNLDLEQQSSERPVIYIHATVSAFAITLVLAWAGLQFAQVTKEIRAE
GNYLILLSLLMVLPYFLFTSFFASSVMSTLLYVFGPISQMNKNSYSYSVHKAPRLKAAHGSLPHVTIQCPVYKEKLESVIKPTIKS
LQAAIRTYELQGGSANIFINDDGLQLIDRKEALERIEYYEECGLGYVARPGHGVNGFIRKGRFKKASNMNYCLHISKLVDTRFHER
LELIENPTPKEESGLYLKVLEEVVREEGKCWAGGDILLGDIILIIDSDTRVPEDCFVDSVSEMEQSPEVAIIQHASGVMMVVGNYW
EKMIAWFTNMIYFSISCVSGNGLTMAAFVGHNAFLRWSAIQELAYIDEDDGRTKYWSESHVSEDFEMTLKLASLGYTIRIATYHDG
GFKEGVSLTVYDEITRWSKYAFGCAEIMFSPFKDWWKGKIFARLFFVFLNSHISLPCKFSILGYMGTYYAIATSLIMLVANYFIVG
YYDWGYSRVYIDAMKVFVSVMVVFGCATQVAYIIGRYRIYKHSIYTMVLEFRYSILFSVFLGGLSWHMIVSIGSYFFSLNLQWGAT
AKDIDDSNFFKELPKAIKNYKFMYILCIFLIAGMIVLAFFVPYAFQIRLLTCALPLGWSVASHFLSPIVLNPQLMTFAW

Score	Expect	Method	Identities	Positives	Gaps
369 bits(947)	6e-124	Compositional	185/381(49%)	264/381(69%)	7/381(1%)

```
Query  517  SDTRVPADCLLDAVSEMEQSPDVGIMQFSSGVMQVVHTYFENGITFFTNLIYSAIRYTVS  576
            SDTRVP DC +D+VSEMEQSP+V I+Q +SGVM VV  Y+E  I +FTN+IY +I
Sbjct  563  SDTRVPEDCFVDSVSEMEQSPEVAIIQHASGVMMVVGNYWEKMIAWFTNMIYFSISCVSG  622

Query  577  NG-DVAPFVGHNAILRWSAIQQVAYQDED-GYDKFWSESHVSEDFDMSLRLQCNGYIIRL  634
            NG  +A FVGHNA LRWSAIQ++AY DED G  K+WSESHVSEDF+M+L+L   GY IR+
Sbjct  623  NGLTMAAFVGHNAFLRWSAIQELAYIDEDDGRTKYWSESHVSEDFEMTLKLASLGYTIRI  682

Query  635  AAWAGEGFKEGVSLTVYDELARWEKYAYGCNELLFHPIRTWLWRGP-FTPLFRRFLFSNI  693
            A +   GFKEGVSLTVYDE+ RW KYA+GC E++F P + W W+G  F  LF  FL S+I
Sbjct  683  ATYHDGGFKEGVSLTVYDEITRWSKYAFGCAEIMFSPFKDW-WKGKIFARLFFVFLNSHI  741

Query  694  RFTSKVTVISYIGTYYAIGAAWILTAVNYFVMGWFNGYLDKYYVDSWQVWFSIIIVFNGL  753
                K +++ Y+GTYYAI  + I+   NYF++G+++    + Y+D+ +V+ S+++VF
Sbjct  742  SLPCKFSILGYMGTYYAIATSLIMLVANYFIVGYYDWGYSRVYIDAMKVFVSVMVVFGCA  801

Query  754  GNIALAVMRYRVGERGLLYALFENFMWTLMLAIFLGGLSLHVSQALLAHMFEINMTWGA**T**  813
              +A  + RYR+ +   +Y +   F ++++ ++FLGGLS H+  ++ ++ F +N+ WGAT
Sbjct  802  TQVAYIIGRYRIYKHS-IYTMVLEFRYSILFSVFLGGLSWHMIVSIGSYFFSLNLQWGA**T**  860

Query  814  SKEAEFSNFFIEVPKVLKKFKFSMLFSLIFIAGMIILAQAPFVPFDWRIKDFVAILPMAT  873
            +K+ + SNFF E+PK +K +KF  +  +  IAGMI+LA   FVP+ ++I+     LP+
Sbjct  861  AKDIDDSNFFKELPKAIKNYKFMYILCIFLIAGMIVLAF--FVPYAFQIRLLTCALPLGW  918

Query  874  VAASHFLLPLALNPALMTFSW  894
              ASHFL P+ LNP LMTF+W
Sbjct  919  SVASHFLSPIVLNPQLMTFAW  939
```

FIG. 8G

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/043348

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/42
ADD.

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

B. FIELDS SEARCHED

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

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ZHANG FEI ET AL: "Improvement of cellulase production inTrichoderma reeseiRut-C30 by overexpression of a novel regulatory geneTrvib-1", BIORESOURCE TECHNOLOGY, ELSEVIER, AMSTERDAM, NL, vol. 247, 20 September 2017 (2017-09-20), pages 676-683, XP085298937, ISSN: 0960-8524, DOI: 10.1016/J.BIORTECH.2017.09.126	14-19,22
A	abstract figure 3	1-13,20, 21,23,24

-/--

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 August 2019	06/09/2019

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Sprinks, Matthew

3

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/043348

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MONIKA SCHMOLL ET AL: "Envoy, a PAS/LOV Domain Protein of Hypocrea jecorina (Anamorph Trichoderma reesei), Modulates Cellulase Gene Transcription in Response to Light", EUKARYOTIC CELL, vol. 4, no. 12, 1 December 2005 (2005-12-01), pages 1998-2007, XP055615560, US ISSN: 1535-9778, DOI: 10.1128/EC.4.12.1998-2007.2005	14-19,22
A	the whole document	1-13,20, 21,23,24

3