



(51) International Patent Classification:

C12N 15/80 (2006.01) C12N 9/58 (2006.01)
C12P 21/02 (2006.01) C12N 15/52 (2006.01)

(21) International Application Number:

PCT/US2017/061475

(22) International Filing Date:

14 November 2017 (14.11.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/422,413 15 November 2016 (15.11.2016) US

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

(72) Inventors: **ESTELL, David A.**; 925 Page Mill Road, Palo Alto, California 94304 (US). **MILLER, Michael C.**; 925 Page Mill Road, Palo Alto, California 94304 (US).

(74) Agent: **TODD, Stephen**; 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).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: FILAMENTOUS FUNGI WITH IMPROVED PROTEIN PRODUCTION

(57) Abstract: Described are recombinant filamentous fungal cells that produce decreased amount of proteases and/or produce increased amounts of protease inhibitors compared to comparable parental cells. Such cells are useful for producing proteins of interest, including antibodies.



FILAMENTOUS FUNGI WITH IMPROVED PROTEIN PRODUCTION

CROSS-REFERENCE TO RELATED APPLICATIONS

[01] This application claims the benefit of U.S. Provisional Application No. 62/422,413,
5 filed November 15, 2016, the disclosure of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[02] Described are recombinant filamentous fungal cells that produce decreased amount of
10 proteases and/or produce increased amounts of protease inhibitors compared to comparable parental cells. Such cells are useful for producing proteins of interest, including antibodies.

BACKGROUND

[03] Genetic engineering has allowed improvements in microorganisms used as industrial
15 bioreactors, cell factories and in food fermentations. Important enzymes and proteins produced by engineered microorganisms include glucoamylases, α -amylases, cellulases, neutral proteases, and alkaline (or serine) proteases, hormones and antibodies. However, the occurrence of protein degradation and modification in some genetically engineered systems can interfere with efficient production.

20 [04] Filamentous fungi (*e.g.*, *Aspergillus* and *Trichoderma* species) and certain bacteria (*e.g.*, *Bacillus* species) have been engineered to produce and secrete a large number of useful proteins and metabolites (see *e.g.*, *Bio/Technol.* 1987. 5: 369 - 376, 713 - 719 and 1301 -1304 and Zukowski, "Production of commercially valuable products," In: Doi and McGlouglin (eds.) *Biology of Bacilli: Applications to Industry*, © 1992, Butterworth-Heinemann,
25 Stoneham. Mass pp 311-337).

[05] As lower eukaryotic microorganisms, filamentous fungi are known for their robust
ability to secrete large quantities of proteins. Such expression can reach as high as 40 g/L
(Durand *et al.*, *Enzyme and Microbial Technology*, 1988. 10(6):341-346) with a translational
and post-translational modification process similar to that of mammalian cells except for their
30 glycosylation. They are widely used in the chemical, pharmaceutical and food industries and it is generally regarded as safe (Schuster, E., *et al.*, *Appl Microbiol Biotechnol*, 2002. 59(4-5):426-35). However, heterologous protein production in filamentous fungi is still relatively low as compared to the more common bacterial expression systems. Thus a need exists for more efficient expression systems that can produce heterologous proteins in greater quantities.

SUMMARY

[06] The present compositions and methods relate to recombinant filamentous fungal cells that produce decreased amount of proteases and/or produce increased amounts of protease inhibitors compared to comparable parental cells. Aspects and embodiments of the compositions and methods are described in the following, independently-numbered paragraphs.

1. In one aspect, a filamentous fungal cell is provided, comprising at least one mutation that decreases the amount of active protease in the cell, wherein the mutation is in a gene encoding a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24 and SEQ ID NO: 25.

2. In some embodiments of the filamentous fungal cell of paragraph 1, the mutation is located in a non-coding region of the gene.

3. In some embodiments of the filamentous fungal cell of paragraph 1, the mutation is located in a coding region of the gene.

4. In some embodiments of the filamentous fungal cell of any of paragraphs 1-3, the mutation results in reduced expression of a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22 and SEQ ID NO: 23.

5. In some embodiments of the filamentous fungal cell of paragraph 4, the mutation is a polynucleotide homologous to or identical to a polynucleotide selected from the group consisting of the polynucleotide of SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47 and SEQ ID NO: 48.

6. In some embodiments of the filamentous fungal cell of any of paragraphs 1-3, the mutation results in overexpression of the polypeptide homologous to a polypeptide of SEQ ID NO: 24 and/or SEQ ID NO: 25.

7. In some embodiments of the filamentous fungal cell of paragraph 6, the mutation is a polynucleotide homologous to or identical to a polynucleotide of SEQ ID NO: 49 or SEQ ID NO: 50.

8. In some embodiments of the filamentous fungal cell of any of paragraphs 1-7, the mutation comprises an insertion mutation.

9. In some embodiments of the filamentous fungal cell of paragraph 8, the insertion mutation comprises insertion of a selectable marker.

10. In some embodiments of the filamentous fungal cell of paragraph 8 or 9, the insertion mutation comprises insertion of an expression cassette for overexpressing the polypeptide of SEQ ID NO: 24 and/or SEQ ID NO: 25.

11. In some embodiments of the filamentous fungal cell of any of paragraphs 1-10, the filamentous fungal cell is an *Aspergillus* species, a *Rhizopus* species, a *Trichoderma* species or a *Mucor* species.

12. In some embodiments of the filamentous fungal cell of paragraph 11, the *Trichoderma* species is selected from the group consisting of *Trichoderma reesei*, *Trichoderma viride*, *Trichoderma koningii*, and *Trichoderma harzianum*.

13. In some embodiments of the filamentous fungal cell of any of paragraphs 1-12, the mutation results in increased production of a protein of interest compared to otherwise identical parental filamentous fungal cells that lack the deletion.

14. In some embodiments of the filamentous fungal cell of any of paragraphs 1-13, the protein of interest is an antibody or fragment, thereof.

15. In another aspect, a method for increasing expression of a protein of interest in a filamentous fungal host is provided, the method comprising: (i) introducing a mutation in a gene encoding a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24 and SEQ ID NO: 25 into filamentous fungal host cells capable of expressing a protein of interest; (ii) cultivating the filamentous fungal cell having the deletion under conditions conducive for production of the

protein of interest; and (iii) recovering the protein of interest; wherein the presence of the mutation results in increased production of the protein of interest compared to the production in otherwise identical parental filamentous fungal cells that lack the deletion.

16. In some embodiments of the method of paragraph 15, the mutation is located in a non-coding region of the gene, a coding region of the gene, or both.

17. In some embodiments of the method of paragraph 15 or 16, the mutation results in reduced expression of a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22 and SEQ ID NO: 23, and/or overexpression of a polypeptide homologous to or identical to a polypeptide of SEQ ID NO: 24 and/or SEQ ID NO: 25.

18. In some embodiments of the method of any of paragraphs 15-18, the mutation is a polynucleotide homologous to or identical to a polynucleotide selected from the group consisting of the polynucleotide of SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49 and SEQ ID NO: 50.

19. In some embodiments of the method of any of paragraphs 15-18, the mutation comprises an insertion mutation, optionally having a selectable marker.

20. In some embodiments of the method of paragraph 19, the filamentous fungal cell is an *Aspergillus* species, a *Rhizopus* species, a *Trichoderma* species or a *Mucor* species.

21. In some embodiments of the method of any of paragraphs 15-20, the mutation results in increased production of a protein of interest compared to otherwise identical parental filamentous fungal cells that lack the deletion.

22. In some embodiments of the method of any of paragraphs 15-21, the protein of interest is an antibody or fragment, thereof.

23. In another aspect, a method for increasing expression of a protein of interest in a filamentous fungal host is provided, the method comprising: (i) cultivating the filamentous fungal cell of any of paragraphs 1-14 under conditions conducive for production of a protein of interest; and (ii) recovering the protein of interest; wherein the presence of the mutation results

in increased production of the protein of interest compared to the production in otherwise identical parental filamentous fungal cells that lack the deletion.

[07] The and other aspects and embodiments of the compositions and methods will be apparent from the description.

5

DETAILED DESCRIPTION

[08] The present invention relates recombinant filamentous fungal cells that produce decreased amount of proteases and/or produce increased amounts of protease inhibitors compared to comparable parental cells. Preferably, the recombinant cells are capable of expressing at least one heterologous protein of interest encoded by a heterologous gene, or overexpressing a homologous protein of interest. Nucleic acids and methods for making the mutant filamentous fungal cells are provided, as well as methods for using the cells for the altered production of heterologous proteins of interest.

10

[09] The proteases and protease inhibitors relevant to the present recombinant filamentous fungal cells are listed in Table 1, below. Joint Genome Institute (JGI; U.S. Department of Energy) identifiers, a brief description, amino acid (AA) SEQ ID NOs and nucleic acid (NA) SEQ ID NOs, are indicated.

15

Table 1. Proteases and protease inhibitors

JGI Identifier	Description	SEQ ID NO:	
		AA	NA
>jgi Trire2 73897 estExt_GeneWisePlus.C_10741	chymo	1	26
>jgi Trire2 77579 estExt_GeneWisePlus.C_80475	pepsin	2	27
>jgi Trire2 22459 estExt_fgeneshl_pm.C_100158	carboxypeptidase B	3	28
>jgi Trire2 74156 estExt_GeneWisePlus.C_11223	pepsin	4	29
>jgi Trire2 119876 estExt_fgeneshl_pg.C_10816	pepsin	5	30
>jgi Trire2 111818 fgeneshl_pg.C_scaffold_29000088	pepsin	6	31
>jgi Trire2 82623 estExt_GeneWisePlus.C_450028	tripeptidyl amino peptidase	7	32
>jgi Trire2 81004 estExt_GeneWisePlus.C_220006	pepsin	8	33
>jgi Trire2 110490 fgeneshl_pg.C_scaffold_20000149	pepsin	9	34
>jgi Trire2 121133 estExt_fgeneshl_pg.C_60052	pepsin	10	35

>jgi Trire2 106661 fgenes5_pg.C_scaffold_7000462	eqolisin	11	36
>jgi Trire2 51365 estExt_Genewise1.C_220234	s8	12	37
>jgi Trire2 122076 estExt_fgenes5_pg.C_100212	pepsin	13	38
>jgi Trire2 123865 estExt_fgenes5_pg.C_270147	s8	14	39
>jgi Trire2 81517 estExt_GeneWisePlus.C_250157	s8	15	40
>jgi Trire2 123244 estExt_fgenes5_pg.C_190118		16	41
>jgi Trire2 81070 estExt_GeneWisePlus.C_220116	QM 6a	17	42
>jgi Trire2 53961 e_gwl.1.1743.1		18	43
>jgi Trire2 69555 e_gwl.29.235.1		19	44
>jgi Trire2 21668 estExt_fgenes1_pm.C_30227	QM 6a	20	45
>jgi Trire2 58698 e_gwl.5.263.1		21	46
>jgi Trire2 122703 estExt_fgenes5_pg.C_140132		22	47
>jgi Trire2 60581 e_gwl.7.573.1		23	48
>jgi Trire2 74129 estExt_GeneWisePlus.C_11150	Inhibitor	24	49
>jgi Trire2 111915 fgenes5_pg.C_scaffold_30000054	Inhibitor	25	50

[10] The amino acid sequences of the proteases and protease inhibitors are listed, below.

[11] >jgi|Trire2|73897|estExt_GeneWisePlus.C_10741 polypeptide (SEQ ID NO: 1)

MAPASQVVSALMLPALALGAAIQPRGADIVGGTAASLGEFPYIVSLQNPNQGGHFCG
 5 GVLVNANTVVTAHCSVVYPASQIRVRAGTLTWNSGGTLVGVSQIIVNPSYNDRTTD
 FDVAVWHLSSPIRESSTIGYATLPAQGS DPVAGSTVTTAGWGTTSENSNSIPSRLNKVS
 VPVVARSTCQADYRSQGLSVTNNMFCAGLTQGGKDSCSGDSGGPIVDANGVLQGVV
 SWGIGCAEAGFPGVYTRIGNFVNYINQNLA

[12] >jgi|Trire2|77579|estExt_GeneWisePlus.C_80475 polypeptide (SEQ ID NO: 2)

10 MKSALLAAAALVGSAQAGIHKMKLQKVSLEQQLEGSSIEAHVQQLGQKYMGVRPTS
 RAEVMFNDKPPKVQGGHPVPVTNFMNAQYFSEITIGTPPQSFKVVLDTGSSNLWVPSQ
 SCNSIACFLHSTYDSSSSSTYKPNGSDFEIHYGSGSLTGFI SNDVVTIGDLKIKGQDFAEA
 TSEPLAFAFGRFDGILGLGYDTISVNGIVPPFYQMVNQKLIDEPVFAFYLGSSDEGSEA
 VFGGVDDAHYEGKIEYIPLRRKAYWEVDLDSIAFGDEVAELENTGAILDTGTSLNVLP
 15 SGLAELLNAEIGAKKGFGGQYTVDCSKRDSL PDITFSLAGSKYSLPASDYIEMSGNCIS
 SFQGMDFPEPVGPLVILGDAFLRRYYSVYDLGRDAVGLAKAK

[13] >jgi|Trire2|22459|estExt_fgenes1_pm.C_100158 polypeptide (SEQ ID NO: 3)

MKTFIPVALLALSQATSACLLPHERDEASGLKPV VRRQSSNGTPIGTGDRFSGGSTAPR
 GLGTQSSSTSFSSTILNVKEVTSGLQGLANTYGVQTFNTPYKTAQGATVVGAKVGGTG

GNCTDAYRVFFNGNIHARERGSADSVLYFISDLLYANAHTGLTYGSKTYSNAQVKT
 ALAAGIVFIPLSNPDGVAYDQSTNSCWKRNRNPNSGQSPGVDLNRNFDLWDFRNLF
 ASSAQSSVGSTSPSSEYHGHASAFSEPETKNIKWVFDITYSKVRWFVDLHSYAGDVLW
 NWGSDENQVDYPTMNFLNGTYNKVRGILTDTSPGRGYGEYVPQADLDVKEAAAKR
 5 VASALTAGGGRSYTAFQSAQLYPTSGASDDYAYSRLFSDPTKNLIHSYTIEFGFANNA
 ASCPFYPSVSQYNSNLRATSAGFMELLLAATDYGLGDATT

[14] >|Trire2|74156|estExt_GeneWisePlus.C_11223 polypeptide (SEQ ID NO: 4)

MQTFGAFLVSFLAASGLAAALPTEGQKTASVEVQYNKNYVPHGPTALFKAKRKYGA
 PISDNLKSLVAARQAKQALAKRQTGSAPNHPSDSADSEYITSVSIGTPAQVLPLDFDTG
 10 SSDLWVFSSETPKSSATGHAIYTPSKSSTSKKVSGASWSISYGDGSSSSGDVYTDKVTI
 GGFSVNTQGVESATRVSTEFVQDTVISGLVGLAFDSGNQVRPHQKTFWFSNAASSLAE
 PLFTADLRHNGSYNFGYIDTSVAKGPVAYTPVDNSQGFWEFTASGYSVGGGKLNK
 NSIDGIADTGTLLLLDDNVVDAYYANVQSAQYDNQQEGVVFDCDEDLPFSFGVGS
 STITIPGDLLNLTPLEEGSSTCFGGLQSSSGIGINIFGDVALKAALVVFDLGNERLGWAQ
 15 K

[15] >|Trire2|119876|estExt_fgenes5_pg.C_10816 polypeptide (SEQ ID NO: 5)

MRVQYVSLAGLFAAATVSAGVVTVPFEKRNLPDFAPSLRRDGSVSLDAINNLTG
 GGYAQAQFSVGTPPQKLSFLLDTGSSDTWVNSVTADLCTDEFTQQTVGEYCFRQFNPR
 RSSYKASTEVDITYLDGRRIRGNFTDTVTINQANITGQKIGLALQSVRGTGILGLGF
 20 RENEAADTKYPTVIDNLVSQKVIPVPAFSLYLNDLQTSQGILLFGGVDTDKFHGGLAT
 LPLQSLPPSIAETQDIVMYSVNLDGFSASDVDPDVSAKAVLDSGSTITLLPDAVVQEL
 FDEYDVLNIQGLPVPFIDCAKANIKDATFNFKFDGKTIKVPIDEMVLNNLAAASDEIMS
 DPSLSKFFKGWSGVCTFGMGSTKTFGIQSDEFVLLGDTFLRSAYVVYDLQNKQIGIAQ
 ATLNSTSTIVEFKAGSKTIPGPASTGDDSDSDSDSDSAGAALHPTFSIALAGTLFT
 25 AVSMMMSVL

[16] >|Trire2|111818|fgenes5_pg.C_scaffold_29000088 polypeptide (SEQ ID NO: 6)

MHAALVSLSIASLAAAALPEGVVEVPLQRIKNQTAYGVEFQVGNPPQKAVISADTGSP
 TYAFESPRNTVCQQGLCSAYGTYDNTTSTTSKWLSDGYSMDLIDHGFGSFINDTLRIG
 GVTLNDMMFQVVEQNFASFPINTQQTAFGLGAFCQTPACDTYSTFLNQLYEHGAIR
 30 RAFSVYLGPNPDATGSLLFGGIDLAKRQGPVHKLKVLDPVTAAANLQPNWVSLSSIE
 LQLSNGTTITSTYDNGTYALWDTGSPGWYVQQDMFDALTSYWGLTNPNNNDIIVDC
 KFREPSDDSLAVNIAPGVNTINPLSSLPIDNGDGTCTARVSPWGRVMGDTFLRNVYFTF
 DYEDLTVEFALVKYTDENIVKIE

[17] >|Trire2|82623|estExt_GeneWisePlus.C_450028 polypeptide (SEQ ID NO: 7)

MAKLSTLRLASLLSLVSVQVSASVHLLSLEKLPHGWKAAETPSPSSQIVLQVALTQQ
 NIDQLESRLAAVSTPTSSTYGYLDVDEINSIFAPSDASSSAVESWLQSHGVTSYTKQG
 SSIWFQTNISTANAMLSTNFHTYSDLTGAKKVRTLKYSIPESLIGHVDLISPTTYFGTTK
 AMRKLKSSGVSPAADALAAARQEPSSCKGTLVFEGETFNVFQPDCLRTEYSVDGYTPSV
 5 KSGSRIGFGSFLNESASFADQALFEKHFNIPSQNFSVVLINGGTDLPQPPSDANDGEANL
 DAQTILTIAHPLPITEFITAGSPPYFPDPVEPAGTPNENEPYLQYYEFLLSKSNAEIPQVIT
 NSYGDEEQTVPRSYAVRVCNLIGLLGLRGISVLHSSGDEGVGASCVATNSTTPQFNPIF
 PATCPYVTSVGGTVSFNPEVAWAGSSGGFSYYFSRPWYQQEAVGTYLEKYVSAETKK
 YYGPYVDFSGRGFPDVAHHSVSPDYPVFQGGELTPSGGTAASPVVAAIVALLNDARL
 10 REGKPTLGFLNPLIYLHASKGFTDITSGQSEGCNGNNTQTGSPLPGAGFIAGAHWNAT
 KGWDPTTGFGVPNLKKLLALVRF

[18] >jgi|Trire2|81004|estExt_GeneWisePlus.C_220006 polypeptide (SEQ ID NO: 8)

MKFHAAALTLACCLASSASAGVAQPRADEVESAEQGKTFSLQIPNERYKGNIPAAYSIS
 ALAKYSPTIPDKIKHAIEINPDLHRKFSKLINAGNMTGTAVASPPPGADA EYVLPVKIGT
 15 PPQTLPLNLDTGSSDLWVISTDTYPPQVQGQTRYNVSASTTAQRLIGESWVIRYGDGSS
 ANGIVYKDRVQIGNTFNQQAVESAVNISNEISDDSFSSGLLGAASSAANTVRPDRQTT
 YLENIKSQ LARPVFTANLKKGKPGNYNFGYINGSEYIGPIQYAAINPSSPLWEVSVSGY
 RVGSNDTKYVPRVWNAIADTGTTLLVPNDIVSAYYAQVKGSTFSNDVGMMLVPCA
 ATLPDFAFGLGNYRGVIPGSYINYGRMNKTYCYGGIQSSEDAPFAVLGDIALKAQFVV
 20 FDMGNKVVG FANKNTNV

[19] >jgi|Trire2|110490|fgenesh5_pg.C_scaffold_20000149 polypeptide (SEQ ID NO: 9)

MNGLSFERAKALREPSADRATMHFSRLRGIKGRPGQQQGGQQHHHRS AVSALS RVRS
 SSSSSTSSGGFKQSFQNVSAVTDASTQYAIQCGWDGVPVWLLVDTGSSDTWATQTGF
 ECADLGGAHSEAACGFSRPLIDGFGGGQIDDLHFFLKYSGGERISGPMGYSDLSCGG
 25 VAVARQQVGLANSTYWHGNNVTVGILGLAYPAITSAYYGDVGSEAPWNAMSYPFL
 TSAISQGAIEPLFSVALVRNSTDGVIAWGGGLPPMDWQFHGFAKTDLIVANLIGSPETA
 WKHSFYTIVPDGMKWDQTTDTTKYPYIVDTGTTMLYLPPLAEAIANAFQPRAVYLY
 QWGTYFVECTAIPPHFAILIEGVEFWINPADLIYRDLVDPLTGYCAVGVASGGPGPYIL
 GDVFLQNVVAVFDVGAAEMRFYGRK

[20] >jgi|Trire2|121133|estExt_fgenesh5_pg.C_60052 polypeptide (SEQ ID NO: 10)

MEAILQAQAKFRLDRGLQKITAVRNKNYKRHGPKSYVYLLNRFGEPTKPGPYFQQH
 RIHQRLAHPDFKAAVGGRVTRQKVLAKKVKEDGTVDAGGSKTGEVDAEDQQNDSE
 YLCEVTIGTPGQKLMLDFDTGSSDLWVFSTELSKHLQENHAIFDPKKSSTFKPLKDQT
 30 WQISYGDGSSASGTCGSDTVTLGGLSIKNQTIELASKLAPQFAQGTGDGLLGLAWPQI

NTVQTDGRPTPANTPVANMIQQDDIPSDAQLFTAIFYSERDENAESFYTFGYIDQDLV
SASGQEIAWTDVDNSQGFWMFPSTKTTINGKDISQEGNTAIADTGTTLALVSDEVCEA
LYKAIPGAKYDDNQQGYVFPINTDASSLPKLVSVGNTQFVIQPEDLAFAPADDSNWY
GGVQSRGSPFDILGDVFLKSVYAIFDQGNQRFQAVPKIQAKQNLQPPQ

5 [21] >jgi|Trire2|106661|fgenesh5_pg.C_scaffold_7000462 polypeptide (SEQ ID NO: 11)
MDAIRARSAARRSNRFQAGSSKNVNGTADVESTNWAGAAITTSQVTEVSGTFTVPRPS
VPAGGSSREEYCGAAWVGIDGYSDADLIQTGVLWCVEDGEYLYEAWYEYLPAALVE
YSGISVTAGSVTVTATKTGTNSGVTTLTSGGKTVSHTFSRQNSPLPGTSAEWIVEDFT
SGSSLVPFADFGSVTFTGATAVVNGATVTAGGDSPIIDLEDSRGDILTSTTVSGSTVT
10 VEYE

[22] >jgi|Trire2|51365|estExt_Genewise1.C_220234 polypeptide (SEQ ID NO: 12)
MVRSAFLVSLLATFSGVIARVSGHGSKIIVPGAYIFEFEDSQDTADFYKKLNGEGSTRK
FDYKLFKGVSVQLKDLNHEAKAQQMAQLPAVKNVWPVTLIDAPNPKVEWVAGST
APTLESRAIKKPPIPNDSSDFPTHQMTQIDKLRAKGYTGKGVRAVIDTGIDYTHPALG
15 GCFGRGCLVSFGTDLVGDDYTGFNTVPDDDPVDCAGHGSHVAGIIAAQENPYGFTG
GAPDVTLGAYRVFGCDGQAGNDVLISAYNQAFEDGAQIITASIGGPSGWAEEPWAVA
VTRIVEAGVPCTVSAGNEGDSGLFFASTAANGKKVIAVASVDNENIPSVLSVASYKIDS
GAAQDFGYVSSSKAWDGVSKPLYAVSFDTTIPDDGCSPLPDSTPDLSYIVLVRRGTC
TFVQKAQNVAAGKAKYLLYNNIPGALAVDVSAPPEIEAVGMVDDKTGATWIAALK
20 DGKTVTLTLTDPIESEKQIQFSDNPTTGGALSGYTTWGPTWELDVKPKQISSPGGNILST
YPVALGGYATLSGTSMACPLTAAVALIGQARGTFDPALIDNLLATTANPQLFNDGEK
FYDFLAPVPQQGGGLIQAYDAAFATLLSPSSLSFNDDTHFIKKKQITLKNSTSKQRTY
KLNHVPTNTFYTLAPGNGYPAPFPNDAAHANLKFNLQQVTLPAGRSITVDVFPTTP
RDVDAKRLALWSGYITVNGTDGTSLSVPYQGLTGS LHKQKVLYPEDSWIADSTDESL
25 APVENGTVFTIPAPGNAGPDDKLPSLVVSPALGSRVVRVDLVLLSAPPHGTKLKTVKF
LDTTSIGQPAGSPLLWISRGANPIAWTGELSDNKFAPPGTYKAVFHALRIFGNEKKKED
WDVSESPAFTIKYA

[23] >jgi|Trire2|122076|estExt_fgenesh5_pg.C_100212 polypeptide (SEQ ID NO: 13)
MRASPLAVAGVALASAAQAQVVQFDIEKRHAPRLSRRDGTIDGTLSNQRVQGGYFIN
30 VQVGSPGQNITLQLDTGSSDVWVPSSTAACITQVSESNPGCQFGSFNPDDSDTFDEVG
QGLFDITYVDGSSSKGDYFQDNFQINGVTVKNLTMGLGLSSSIPNGLIGVGYMNDEAS
VSTTRSTYPNLPIVLQQKLINSVAFSLWLNDLDASTGSILFGGIDTEKYHGDLSIDIIS
PNGGKTFTEFAVNLYSVQATSPSGTDTLSTSEDTLIAVLDSGTTLTYLPQDMAEEAWN
EVGAEYSNELGLAVVPCSVGNTNGFFSFTFAGTDGPTINVTLSLVLDLFSGGPAPRFS

SGPNKGQSICEFGIQNGTGSPFLLGDTFLRSAFVVYDLVNNQIAIAPTNNSTRTNVVA
 ASSGAPIPSATAAPNQSRTGHSSSTHSGLSAASGFHDGDDENAGSLTSVFSGPGMAVV
 GMTICYTLLGSAIFGIGWL

[24] >|Trire2|123865|estExt_fgenesh5_pg.C_270147 polypeptide (SEQ ID NO: 14)

5 MRACLLFLGITALATAIPALKPPHGSPPDRAHTTQLAKVSIALQPECRELLEQALHHLS
 PSSPRYGRYLGREEAKALLRPRREATAAVKRWLARAGVPAHDVLTGQFIHVRTLAE
 KAQALLGFEYNSTLGSQTIAISTLPGKIRKHMVTVQYVPLWTEADWEECKTIITPCLK
 RLYHVDSYRAKYESSSLFGIVGFSGQAAQHDELDFLHDFAPYSTNANFSIESVNGGQ
 SPQGMNEPASEANGDVQYAVAMGYHVPVRYAVGGENHDIIPDLDLVDTTEEYLEPF
 10 LEFASHLLDLDDDELPRVVSISYGANEQLFPRSYAHQVCDMFGQLGARGVSIVVAAG
 DLGPGVSCQSNDSARPKFIPSPATCPYVTSVGSTRGIMPEVAASFSSGGFSDYFARP
 AWQDRAVGAYLGAHGEEWEGFYNPAGRGFPDVAQGVNFRFRAHGNESLSSGTSL
 SPVFAALIALNDHRKSGMPPMGFLNPWIYTVGSHAFTDIIEARSEGCPGQSVEYLAS
 PYIPNAGWSAVPGWDPVTGWGTPLFDRMLNLSLV

15 [25] >|Trire2|81517|estExt_GeneWisePlus.C_250157 polypeptide (SEQ ID NO: 15)

MKSALLWAAPLSLLAGLGACGKNFDELLEVPEGWTQLNDVNPSQHIRLSIAVKQPYI
 DSLEARMAEKGNRLSMQEVRELQTPAKKDIDNVLHWLSQNNLYGVVEKDFIRVWTT
 VAKAEPLLKMKLSRFSYEGKPAVLRTTKYTIPDSVADSISFINPINNFMSARHRERGLTF
 LLPPSKGAVLPGNNTAYCAGSVTPSCLSKLYNINYSANTSSPVIFGVAGFLEENANLQ
 20 DLRQFLNQSAPEVAKTGRTINVELVNGGVNSQELSESIEAALDVDYAVSLGFPTNVT
 FYSTGGRGVKLNDDGQPIEGEDDDNEPYLEFFQYLLAKPDGQVPHVLSLSYSDDELSV
 PRDYAKRVCSLFGLLTARGTSIIFSSGDGGARGGRASSCLTNDGTRQVTMATFPPTCP
 WVTSIGAVTNIAEPPNGARFSTGGFSQYFAQPRWQNEAVEGYVKALGGHLDGYYNES
 MRAIPDVSAVGTAFSIISGGYPRSVQGTASAPVFAAMIALINDARLRAGKKSGLFLNQ
 25 HLYSSEVRVAVLQDITAGQSASCIWNDADIPGGWPAAEGWDAITGLGVPKRFDKLMEV
 LVNHLPLRLATGFARFKMPTQHVEPESHALLQDVANSLLKARKVVVVTGAGISTNSGIP
 DFRSENGLYSLIQAQFDAANQPTRPAERSKADGTADSGEPRPTKRGMTSREASPDLD
 EVTRQLRDDIEARAESQRPAASSRAAGTQPAVAASDANASAEGVCLSTPRRKPALPST
 PLPTTSPLSSPPREDFLIPLPSWSSSSILRTEDRKRTDASHNVVSSPLSSPPPVLFDPFHPS
 30 SPSDEDRSRRSSTTPSEAGENPPNAMPASQTSSFGKANLPNMKGKDLFDASIWSDPTRT
 SVFYQFATSLRQKVRDAEPTSSHKFISHLRDRGKLVRCYTQNIQIEEKVGLSTSLEDG
 PGSRGRFSRKATANASQLNKMVQEVSSGEGGASSHVNASSQSSNGSSEQSSAESSQAN
 DRTEEEDSAGSSSTTAATTTTTTTTTPPDRPKPAPRKEPPQSGVECVFLHGSLLQLLRCFLC
 GQVCSWDDDDREVETLSGLQPECPHCVGATEARQERGKRALGVGKLRPDIVLYGEEH

PSAHLISPIVTHDLALYPDMLLILGTSLRVHGLKVLVREFAKTVHSRGGKVVFVNFTKP
 PESSWGDIIIDYWIQWDCDAWVSDLQVRIPKLWQDPEPPKPKKKRDSGGAADDNREEK
 KRPPAQNPAVALRDTKVNGAYCTLKILKELRRITYTRDSAAIGNPIITAPEPTTTEAISRA
 VAMDSSRTSTPRGKSKRSRRSATGAIDRPKRTPSTLNPNGHRSKKPTAEAKKKQQEEE
 5 EVPETPSQPPASTVEEFSSILHSVKSNNPRIRKRKMIDGEEFVFPVAVGKKRGAVDSLYKGP
 GDDTKELPPLRMP

[26] >jgi|Trire2|123244|estExt_fgenesh5_pg.C_190118 polypeptide (SEQ ID NO: 16)

MRSVVALSMAAVAQASTFQIGTIHEKSAPVLSNVEANAIPDAYIIKFKDHSVGEDDASK
 HHDWISQSIHTNVEQERLELRKRSNVFGADDVFDGLKHTFKIGDGFKG YAGHFHESVIE
 10 QVRNHPDVEYIERDSIVHTMLPLESKDSIIVEDSCNGETEKQAPWGLARISHRET LNFG
 SFNKYLYTADGGEGVDAYVIDTGTNIEHVDFEGRAKWGKTIPAGDEDEDGNGHGTH
 CSGTVAGKKYGVAKKAHVYAVKVLRSNGSGTMSDVVKGVEY AALSHIEQVKKAKK
 GKRKGFKGSVANMSLGGGKTQALDAAVNAAVRAGVHF A AAGNDNADACNYSPA
 AATEPLTVGASALDDSRAYFSNYGKCTDIFAPGLSIQSTWIGSKYAVNTISGTSMASPH
 15 ICGLLAYYLSLQPADSEFAVAPITPKKLKESVISVATKNALSDLPDSDTPNLLAWNGG
 GCSNFSQIVEAGSYTVKPKQNKQAKLPSTIEELEEAIEGDFEVVSGEIVKGAKSFGSKA
 EKFAKKIHDLVEEEIEEFISELSE

[27] >jgi|Trire2|81070|estExt_GeneWisePlus.C_220116 polypeptide (SEQ ID NO: 17)

MAVLSRLALTASFALCGVSAAGIQQLTAPESLPPSHEAVADYGSKPIIDSEALQSAISI
 20 DTLVKRAESFYKFAKASEEEYGHPTRVIGSAGHEQTLNYITNTLLDHGDYNNVSVQEF
 PVTLSNVFQFRLVLADEVSKSAIPMGLTPPTKDKEPVHGD LVLVQNSGCDASDYPQNV
 KGNI AFIRRGACSF GDKSIGAGKAGAKAAVIYNTDPEELHGT LGLPVEDHIATFGIDGV
 EGKKILAKLSNGESVDAIAYIDAEVKQIQTVNVLAQTEEGDPDNCV MLGGHSDGVAE
 GPGINDDGSGSISVLEVAVQLTKFKVNNCVRFAWWAEEEEGLLGSDFYAASLSDEEN
 25 QKIRLFMDYDMMASPNFAYQIYNATNAESPAGSEELRNLYVDWYKSQGLNYTFIPFD
 GRSDYDGFIRAGIPAGGIATGAEAVKTKEEAEMFGGRAGEWLDPCYHQLCDDLGNLN
 HTAWEVNTKLIHSAVATYALSFDGFPKRKLETEM SAYSQTTKHHGPKLIL

[28] >jgi|Trire2|53961|e_gw1.1.1743.1 polypeptide (SEQ ID NO: 18)

MQPSFGSFLVTVLSASMAAGSVIPSTNANPGSFEIKRSANKAFTGRNGPLALARTYAK
 30 YGVEVPKTLVDAIQLVKSIQLAKRDSATVTATPDHDDIEYLVPVKIGTPPQTLNLDFT
 GSSDLWVFSSDVPDPTSSQGHDIYTPSKSTSSKKLEGASWNITYGDRSSSSGDVYHDIVS
 VGNLTVKSQAVESARNVSAQFTQGNNDGLVGLAFSSINTVKPTPQKTWYDNIVGSLD
 SPV FVADLRHDTPGSYHFGSIPSEASKAFYAPIDNSKGFWQFSTSSNISGQFNAVADTG

TTLLLASDDLVKAYYAKVQGARVNVFLGGYVFNCTTQLPDFTFTVGEGNITVPGTLIN
YSEAGNGQCFCGGIQPSGGLPFAIFGDIALKAAYVIFDSGNKQVGWAQKK

[29] >|jgi|Trire2|69555|e_gw1.29.235.1 polypeptide (SEQ ID NO: 19)

MFIAGVALSALLCADTVLAGVAQDRGLAARLARRAGRRSAPFRNDTSHATVQSNWG
5 GAILEGSGFTAASATVNVPRGGGGSNAAGSAWVGIDGASCQTALQTGFDWYGDGTY
DAWYEWYPEFAADFSGIDIRQGDQIAMSVVATSLTGGSATLENLSTGQKVTQNFNRV
TAGSLCETSAEFIIEDFEECNSNGSNCPVPFASFSPAITFSSATATRSGRSVSLSGAEITE
VIVNNQDLTRCSVSGSSTLTCSYV

[30] >|jgi|Trire2|21668|estExt_fgenesh1_pm.C_30227 polypeptide (SEQ ID NO: 20)

10 MTITAQKLTPEVLLAAPRRSPGVPNATGELVLYTVSTYSFDSHSKTAQIRVLNLKEGTS
HLVSEDSAASEPIWIAEQEIAYVKSLDHGASALVAQHVFNPNESNTIQRFGGSINSLKA
KPLSTDKVAFCCAALTTPDGRMYSPAAEPKSYTSAKIYTSLFVRHWDSWNTENKNSL
WYGQLNKVDGKWTLGNSLTNLLAGTRLHSPVPPFGGTGDFDISTTGIVFVAKDPDL
NHARTTKTDLYFVPLNSYLDQPTFPQIVKTGALRGYSLSPVFSNDGKQVAFLRMKSQQ
15 YEADKTRLLLLIPDVTDLNSVQEFYATEDGKGGWDYKPDWLIWSHDDKELYVAAEKH
ARVVLWKLPPSPLEAKSLPTPIHEDGSVAEARVLGKGSSLLITRSRVESNYSILDPAS
KSTTISSSSSRQGKTFALSKSQCEIWFNGSKGYPIHALVTLPTFDSSKKYPLAFFIHGG
PQGAWGDDWSTRWNPVFAEQGYVVVSPNPTGSTGYGQDHTDAIQNNWGGDPYID
LVKCFEFLEEEVNYIDTTRAVALGASYGGYMINWIQGHDLGRKFALVCHDGVFSTL
20 NQWSTEELFFPEHDFGGALWENREGYEKWDPKXHVGNWATPQLVIHNELDYRLPISE
GLAMFNVLQARGVPSKFVMFPDEHHWVLKPENSLVWHREVLNWINKYSGISEKN

[31] >|jgi|Trire2|58698|e_gw1.5.263.1 polypeptide (SEQ ID NO: 21)

MAWLKKLALVLLAIVPYATASPALSPRSREILSLEDLESEDKYVIGLKQGLSPTDLKKH
LLRVS AVQYRNKNSTFEGGTGVKRTYAIGDYRAYTAVLDRDTVREIWNNTLEKPPW
25 GLATLSNKKPHGFLYRYDKSAGEGTFAYVLDTGINSKHVDFEGRAYMGFSPPKTEPT
DINGHGTHVAGIIGGKTFGVAKKTQLIGVKVFLDDEATTSTLMEGLEWAVNDITTKGR
QGRSVINMSLGGPYSQLNDAIDHIADMILPVAAAGNKGIPATFISPASADKAMTVG
AINSDWQETNFSNFGPQVNILAPGEDVLSAYVSTNTATRVLSGTSMAPHVAGLALYL
MALEEFDSTQKLTDRILQLGMKNKVVNLMTDSPNLIIHNNVK

30 [32] >|jgi|Trire2|122703|estExt_fgenesh5_pg.C_140132 polypeptide (SEQ ID NO: 22)

MASRRALNLSQGLRARSGLSGLRRGFATPSTVGKTQTTTLKNGLTVATEYSPWAQT
STVGMWIDAGSRAETNETNGTAHFLEHLAFKGTAKRSQHLELEIENMGHNLNAYTS
RENTVYFAKAFNSDIPQTVDILADILQNSKLEQSAIERERDVILRESEEVEKQVEEVVFD
HLHATAFQHQLGRITLGRQNIRDITRTELVNYIKNNYTADRMVLAAGGVPHEQLV

ELAEKHFSGLASHGPETEAYVLSKQKADFIGSDVRVRDDTMPTANVAIAVEGVSWNS
 DDYYTALVAQAIVGNYDKAMGNAPHQGGKLSGYVHKHDLANSFMSFSTSYSDTGL
 WGIYLVTDNATRLDDLHVHFAIREWMRLCYNVSEAEVERAKAQLKASILLSLDGTTAV
 AEDIGRQLITTGRRASPGEIERKIDAITDKDVTDFANRYLWDKDIAISAVGKIEALFDYQ
 5 RLRNTMKPKF

[33] >jgi|Trire2|60581|e_gw1.7.573.1 polypeptide (SEQ ID NO: 23)

MSRRPVYFNPLAESWTAPSPDDPQLAYRFHSQLPAYSPQLIPLTDLAKELGVQSIHLK
 DETSRLGLPSFKILGASWGTFRAlAQRLDLPIDSSLGSVEQKLASSNITLYAATDGNHG
 RAVARMASILGVPAQIHVPTTMHQSTIDLIKSEGARVVISDGFYDDAVVDARVAAAK
 10 DDTALVIQDFASGDYVQIPQWIVDGYLTMMLEIDGQLGCTTPDLVVVPVGVGSFAQA
 VVTHFKKPGKQTKVLTVEPDTSASLWKSLSRGESSSTSEKSPSIMAGLDCGTPSSISWA
 VLRHGVDA SLTISDYEAHKACEYLKSQGVSAAGPCGAAPIAALRRLERADRERLGLTK
 NSVVVIFCTEGARDYDVPHSVASDDPVEITQTLVKINSANPFLGSVPGPGETAIARYITA
 WMEHRDIESHWIEPTSGRPSVVGIVRGLGGGKSLMLNGHIDTVTLMGYEGNPLSGDIQ
 15 DGKLYGRGAADMKGGVAAAMAALANAKKHSRLRGDVIFTGVADEEFESIGTQQVLEA
 GWTADGAIVSEPTNMEILYAHKGFVWFDVDIQGLAAHGSRYDLGIDAISKAGYFLVE
 LDKHASHLTAQSGDAVLGPGSIHASLIKGGEEVSSYPSRVQIQLERRTVNGETPETVRK
 ELEEILDGLTKTVPNFTYSLRTTFHRSPFKADLSHPFAKL VHKHVGNTL GREPAVLGAP
 YWTDCA LLDGAGIPALLWGPQGEGLHGKLEYADVESIKQVAEALTAIAVEFCS

20 [34] >jgi|Trire2|74129|estExt_GeneWisePlus.C_11150 polypeptide (SEQ ID NO: 24)

MPLVVPGIMTNSDDKTQEWANKLVGKTYSENESNETRDLPEVHRIIKPGSIVTKDFRP
 ERLNIHLNEDGTVSHVRHG

>jgi|Trire2|111915|fgenesh5_pg.C_scaffold_30000054 polypeptide (SEQ ID NO: 25)

MLFQTM LLA LITSLALAQSEVGRPCGFKMAPCPFDMKCVPDNAYCPHPSRCPGHCEF
 25 KNKYDQCGGFTPRPHVCRRGSRCQDDPRLPPNCGMACDAPGICIPENAPFCGGFMGL
 ACPKGLYCYDALDDCDPNNGGADCGGICL

[35] The nucleic acid sequences of the proteases and protease inhibitors are listed, below.

[36] >jgi|Trire2|73897|estExt_GeneWisePlus.C_10741 polynucleotide (SEQ ID NO: 26)

CTCTGCCACTTAGAAGAATCATCCCTCGCAGCTTTTCTCCATCGCAATGGCTCCC
 30 GCTTCCCAAGTCGTCTCAGCTCTCATGCTGCCCGCTCTCGCCTTGGGAGCCGCCAT
 CCAGCCCCGTGGCGCTGACATCGTGGGAGGAACCGCCGCTCGCTCGGCGAGTTC
 CCCTACATTGTCAGTCTGCAGAACCCCAACCAGGGCGGCCACTTCTGCGGTGGTGT
 CTTGGTCAACGCCAACACCGTCGTTACCGCCGCTCACTGCTCCGTTGTCTACCCTG
 CCTCGCAGATCCGCGTCCGCGCCGGTACTCTTACCTGGA ACTCTGGCGGTACCCTG

GTCGGCGTCTCCCAGATCATCGTGAACCCGTCTACAACGACCGCACCACCGACTT
TGACGTTGCCGTCTGGCACCTGTCCAGCCCTATCCGCGAGAGCTCCACCATTGGCT
ACGCCACTCTTCCCCGCCAGGGCTCCGACCCCGTGGCCGGCTCGACCGTCACCACC
GCTGGCTGGGGCACCACCAGCGAGAACTCCAACCTCCATCCCCTCCCGCCTGAACA
5 AGGTCTCCGTCCCCGTCTGTCGCCCCGCTCCACCTGCCAGGCCGACTACCGCAGCCAG
GGGCTCAGTGTACCAACAACATGTTCTGCGCCGGCCTCACCCAGGGCGGCAAGG
ACTCTTGCTCTGGCGACTCTGGCGGCCCCATCGTTGACGCCAACGGTGTCTCCAG
GGTGTCTGTTTCTTGGGGTATCGGCTGTGCTGAGGCCGGTTTCCCTGGTGTCTACAC
CAGAATCGGCAACTTTGTCAACTACATCAACCAGAACCTCGCATAAGCGATTCCCT
10 GTGTTGGCAACCAAAGCAGTAAAAAAAAAATGGAAAAGAAAGCATCTTCTAAGT
GGCAGGGTCTCTGTTGGCTCTTTTGTGATTGGTTATGTAGCTGAGGCTGTGGGCAA
AATCGGGCT

[37] >|gi|Trire2|77579|estExt_GeneWisePlus.C_80475 polynucleotide (SEQ ID NO: 27)
CATCGCCTTCCACCTGCCAACGAGCCCATCATCAACACCTCACCTCCGCGTCGCCCC
15 AACATGAAGAGCGCGTACTTGCCGCCGCGGCGCTTGTCGGCTCCGCCCAAGCCG
GCATTACAAAGATGAAGCTGCAGAAGGTCTCCCTGGAGCAGCAGCTGGAGGGTTC
GAGCATCGAGGCCCACGTCCAGCAGCTCGGCCAGAAGTACATGGGCGTCCGCCCT
ACTAGCCGTGCCGAGGTCATGTTCAACGACAAGCCGCCCAAGGTCCAGGGCGGGC
ACCCGGTTCCCGTCACCAACTTCATGAATGCCCAATACTTCTCTGAGATTACCATC
20 GGCACCCCCCTCAGTCGTTCAAGGTTGTCCTCGACACGGGAAGCTCTAACCTCTG
GGTTCCTCTCAGTCGTGCAACAGCATCGCCTGCTTCCTGCACTCCACGTACGATT
CGTCTTCATCGTCGACGTACAAGCCCAACGGCTCCGATTTTGAGATCCACTACGGA
TCAGGTAGCTTGACTGGCTTCATCTCCAACGATGTCTGACGATTGGCGACCTCAA
GATCAAGGGGCAGGACTTTGCCGAGGCAACCAGCGAGCCCGGCCTTGCCCTTGCT
25 TTCGGCCGCTTCGACGGCATTCTTGGCCTTGCTACGATACCATCTCGGTCAATGG
CATTGTCCCCCCTTTTACCAGATGGTCAACCAGAAGCTGATCGACGAGCCCGTCT
TTGCTTTCTACCTGGGAAGCAGCGACGAGGGTTCCGAGGCTGTCTTTGGCGGCGTC
GACGATGCTCACTACGAGGGCAAGATTGAGTACATTCCCCTGCGCCGCAAGGCCT
ACTGGGAGGTGGACCTTGACTCCATTGCCTTCGGTGACGAGGTGCGCGAGCTCGA
30 GAACACTGGCGCCATCCTTGACACCGGCACCTCTCTCAACGTCCTCCCCTCGGGCC
TCGCCGAGCTCCTGAACGCTGAGATTGGCGCCAAGAAGGGCTTTGGCGGTCAAGTA
CACTGTTGACTGCTCCAAGCGTGATTCCCTCCCCGACATCACCTTCAGCCTGGCCG
GCTCCAAGTACAGCCTTCCCGCCAGCGACTACATCATTGAGATGTCTGGCAACTGC
ATTCGTCTTCCAGGGCATGGACTTCCCCGAGCCCGTGGGCCCCCTGGTCATTCT

GGGTGATGCTTTCTTGCGCCGCTACTACTCCGTCTACGACCTTGGCAGGGACGCCG
TTGGTCTTGCCAAGGCCAAATAA

[38] >jgi|Trire2|22459|estExt_fgenesh1_pm.C_100158 polynucleotide (SEQ ID NO: 28)

ATGAAGACCTTTATCCCCGTTGCGCTTCTTGCGCTCTCCCAAGCCACCTCGGCGTG
5 TCTTCTTCCCCATGAGCGAGACGAGGCGAGTGGCCTGAAGCCAGTCGTCCGCCGC
CAGTCCAGCAACGGAACCTCCATCGGCACCGGCGACAGATTTAGCGGTGGCTCAA
CTGCGCCTCGAGGACTGGGAACTCAGTCCTCATCCACCTCCTTCAGCACAAATTCTC
AACGTCAAGGAGGTCACCAGCGGTCTGCAGGGCCTGGCCAACACGTACGGAGTTC
AAACTTTCAACACTCCTTACAAGACGGCCCAGGGCGCAACCGTCGTCTGGAGCCAA
10 GGTCGGCGGTACCGGAGGTAACGTGTACCGATGCCTACCGCGTCTTCTTCAACGGCA
ACATCCACGCCCAGAGCGCGGCTCCGCAGACAGTGTGCTCTACTTCATCTCGGAT
CTGCTGTATGCCAACGCCCACAACACCGGCCTCACCTACGGCTCCAAGACGTACA
GCAACGCCCAGGTGAAGACGGCGCTGGCCGCGGGAATCGTCTTCATCCCGCTCAG
CAACCCCGACGGCGTGGCCTATGACCAGTCCACCAACAGCTGCTGGCGAAAGAAC
15 CGCAACCCCAACTCGGGCCAGTCCCCCGGAGTCGACCTGAACCGGAACTTTGACT
TCCTCTGGGACTTCCGCAACCTCTTCGCCTCGTCGGCACAGTCCAGCGTTGGCTCG
ACCTCTCCTAGCTCCGAGACGTACCACGGTGCCAGCGCCTTCTCTGAGCCCGAGAC
CAAGAACATCAAGTGGGTGTTTCGACACGTACTCCAAAGTCCGCTGGTTCGTGGAC
CTTCACTCGTATGCCGGTGACGTCCTGTGGAAC TGGGGCAGCGATGAGAACCAGG
20 TCGACTACCCGACCATGAACTTCCTCAATGGCACCTACAACAAGGTCCGCGGAAT
CCTGACCGACACGCCGTCTCCCGGCCGAGGCTACGGCGAGTATGTCCCCCAGGCG
GACCTGGACGTTAAGGAGGCCGCGGCCAAGCGCGTGGCCTCAGCTCTGACTGCTG
GCGGCGGCCGCTCTTACACGGCCTTTCAGTCGGCGCAGCTGTATCCCACCTCGGGC
GCGAGTGATGACTACGTTACTCGAGGCAC TTTTCCGATCCTACCAAGAATCTCAT
25 CCACTCTTACACGATTGAGTTTGGCTTTGCCAACAACGCCGCGTCTTGCCCCCTTTTA
TCCCTCTGTTTCCCAGTACAACCTCGAACCTCAGGGCTACCAGTGCTGGTTTCATGG
AGCTGCTGTTGGCAGCCACCGACTACGGCCTTGGAGATGCAACGACGTGTTAG

[39] >jgi|Trire2|74156|estExt_GeneWisePlus.C_11223 polynucleotide (SEQ ID NO: 29)

AGCAACCTTCTCCGATATTCAAGATGCAGACCTTTGGAGCTTTTCTCGTTTCCTTCC
30 TCGCCGCCAGCGGCCTGGCCGCGGCCCTCCCCACCGAGGGTCAGAAGACGGCTTC
CGTCGAGGTCCAGTACAACAAGAACTACGTCCCCCACGGCCCTACTGCTCTCTTCA
AGGCCAAGAGAAAAGTATGGCGCTCCCATCAGCGACAACCTGAAGTCTCTCGTGGC
TGCCAGGCAGGCCAAGCAGGCTCTCGCCAAGCGCCAGACCGGCTCGGCGCCCAAC
CACCCCAGTGACAGCGCCGATTTCGGAGTACATCACCTCCGTCTCCATCGGCACTCC

GGCTCAGGTCCTCCCCCTGGACTTTGACACCGGCTCCTCCGACCTGTGGGTCTTTA
GCTCCGAGACGCCCCAAGTCTTCGGCCACCGGCCACGCCATCTACACGCCCTCCAA
GTCGTCCACCTCCAAGAAGGTGTCTGGCGCCAGCTGGTCCATCAGCTACGGCGAC
GGCAGCAGCTCCAGCGGCGATGTCTACACCGACAAGGTCACCATCGGAGGCTTCA
5 GCGTCAACACCCAGGGCGTCGAGTCTGCCACCCGCGTGTCCACCGAGTTCGTCCA
GGACACGGTCATCTCTGGCCTCGTCGGCCTTGCCTTTGACAGCGGCAACCAGGTCA
GGCCGCACCCGCAGAAGACGTGGTTCTCCAACGCCGCCAGCAGCCTGGCTGAGCC
CCTTTTCACTGCCGACCTGAGGCACGGACAGAACGGCAGCTACAACCTTTGGCTAC
ATCGACACCAGCGTCGCCAAGGGCCCCGTTGCCTACACCCCCGTTGACAACAGCC
10 AGGGCTTCTGGGAGTTCCTGCTCGGGCTACTCTGTCGGCGGCGGCAAGCTCAAC
CGCAACTCCATCGACGGCATTGCCGACACCGGCACCACCCTGCTCCTCCTCGACGA
CAACGTCGTCGATGCCTACTACGCCAACGTCCAGTCGGCCCAGTACGACAACCAG
CAGGAGGGTGTCTCTTCGACTGCGACGAGGACCTCCCTTCGTTTCAGCTTCGGTGT
TGGAAGCTCCACCATCACCATCCCTGGCGATCTGCTGAACCTGACTCCCCCTCGAGG
15 AGGGCAGCTCCACCTGCTTCGGTGGCCTCCAGAGCAGCTCCGGCATTGGCATCAA
CATCTTTGGTGACGTTGCCCTCAAGGCTGCCCTGGTTGTCTTTGACCTCGGCAACG
AGCGCCTGGGCTGGGCTCAGAAATAA

[40] >|gi|Trire2|119876|estExt_fgenes5_pg.C_10816 polynucleotide (SEQ ID NO: 30)

ACGGCGTGTTTCATAGCCTGCTCTGGGTGCTTGGGCATATTCTCAAGCCGCCATCAC
20 AACGTTGGGCTTGTCTATTCTAAATTTACCCCCATCTTCATCCACAATGCGTTTCG
TTCAGTACGTCTCGCTGGCCGGCCTCTTCGCCGCTGCCACAGTTTCTGCAGGCGTC
GTTACCGTCCCCCTTTGAGAAGCGCAACCTCAACCCCGACTTTGCCCCGTCCTGCT
GCGTCGCGATGGCAGCGTGAGCCTTGACGCCATCAATAACCTCACTGGAGGTGGC
TACTATGCCCAATTCAGCGTTGGCAGCGCCGCTCAGAAGCTGAGTTTCTCTCGA
25 TACCGGCAGCAGTGATACCTGGGTCAACTCCGTCACTGCGGACCTCTGCACAGAC
GAGTTCACTCAGCAGACTGTTGGAGAATACTGTTTCAGACAGTTCAACCCAAGGA
GGAGTAGTTCTATAAGGCGAGCACTGAAGTCTTCGACATCACCTACCTTGACGGC
CGCAGGATACGAGGCAACTACTTCACGGATACCGTCACCATCAACCAGGCGAACA
TCACGGGCCAGAAGATTGGCCTAGCCCTGCAGTCAGTCCGCGGCACAGGCATCCT
30 GGGCTTGGGATTCCGGGAAAACGAGGCAGCCGACACCAAGTATCCCACCGTCATC
GATAACCTGGTGTCTCAAAAGGTCATTCTGTTCCGGCATTCACTCTCTACCTCAA
CGACCTGCAAACTAGCCAGGGCATCCTCCTCTTTGGTGGTGTGATACCGACAAGT
TCCACGGCGGCCTTGCCACTCTGCCCCCTCAGTCGCTGCCGCCGTCCATTGCCGAG
ACCCAGGATATTGTCATGTACAGTGTTAACCTTGATGGCTTCTCGGCGTCTGACGT

TGATACGCCCCGACGTCAGCGCCAAAGCCGTTCTCGACTCTGGCTCAACAATCACCC
TCCTCCCAGATGCTGTCGTGCAGGAACCTTTTGGACGAGTACGACGTCCTCAACATT
CAGGGACTCCCCGTTCTTTTATCGACTGCGCCAAAGCAAACATCAAGGATGCCAC
CTTCAACTTCAAATTCGACGGCAAGACGATCAAGGTGCCCATTGACGAAATGGTC
5 TTGAACAACCTCGCCGCGGCTTCAGACGAGATTATGTCGGATCCCTCCTTGAGCAA
GTTTTTCAAGGGTTGGAGCGGAGTGTGCACCTTTGGCATGGGCTCGACCAAGACCT
TTGGCATCCAGTCTGACGAGTTTGTCTGCTGGGCGACACCTTCTTGCGGTCCGCC
TATGTCGTCTACGACCTGCAGAACAAGCAGATTGGCATTGCCAGGCGACGCTCA
ACTCGACCAGCAGCACCATTTGTCGAGTTCAAGGCGGGCTCCAAGACCATTCCCGG
10 CCCTGCTTCCACGGGCGACGACTCTGACGATTCGTCGGACGATTCAGATGAAGATT
CTGCCGGTGCTGCTCTTCATCCTACATTTTCTATTGCTCTGGCCGGCACCTTGTTCA
CGGCCGTTTCCATGATGATGAGTGTACTATAG

[41] >|gi|Trire2|111818|fgenesh5_pg.C_scaffold_29000088 polynucleotide (SEQ ID NO:
31)

15 ATGCATGCGGCCCTCGTGTCACTATCAATCGCGTCTCTTGCCGCAGCTGCTCTACC
CGAGGGCGTGGTCGAGGTTCCGCTCCAGCGCATCAAGAATCAGACTGCGTACGGC
GTCGAGTTCCAGGTCGGCAATCCGCCCCAAAAGGCTGTCATTTCTGCCGATACAGG
CAGTCCGACTTATGCCTTTGAGTCTCCGCGCAACACCGTCTGCCAGCAAGGCCTCT
GCTCAGCCTATGGAACCTACGATAACACAACCTCTACCACGTCTAAATGGCTCAGC
20 GACGGCTACTCCGACATGCTCATCGACCACGGCTTCGGCAGTTTTATCAACGACAC
TCTTCGAATTGGCGGCGTCACCCTCAACGACATGATGTTTGGCGTCGTTGAGCAAA
ACTTTGCCTCGTTCCCGATCAATACCCAGCAAACCGCTATTTTTGGCCTCGGGGCA
TTCTGTCAGACGCCTGCGTGCGACACATATTCGACGTTCTCAACCAGCTATACGA
GCACGGTGCGATTGCTCGGCGTGCTTCAGTGTGTATCTGGGCCCAAACGACCCGG
25 ATGCTACGGGGTCGCTGCTTTTCGGGGGGATTGATCTTGCGAAACGCCAGGGCCC
GGTCCATAAATTGAAGGTCCTCGATCCTACAGTAACAGCGGCGAATCTCCAGCCC
AACTGGGTCAGCCTGTCCAGCATCGAACTCCAGCTTTCCAATGGCACTACGATTAC
TTCGACGTATGACAACGGGACGTATGCTCTTTGGGACACGGGGTCCCCGGGCTGG
TATGTCCAGCAAGACATGTTTCGACGCCCTCACCAGCTACTGGGGCCTCACGAATCC
30 CAATCCCAACAATGATATCATCGTTGACTGCAAGTTCCGTGAGCCGTCCGACGACT
CGCTTGCCGTGAATATCGCCCCAGGAGTCACCATCAACGTGCCCTGAGTTCGTTG
CCTATTGATAACGGCGACGGCACTTGCACTGCGCGTGTTTCGCCTTGGGAAGAGT
GATGGGCGATACGTTTCTTCGAAATGTCTACTTTACGTTTGACTACGAGGATCTGA

CAGTTGAATTTGCTTTAGTCAAGTATACCGATGAGACGAACATTGTCAAAATTGAG
TAA

[42] >|Trire2|82623|estExt_GeneWisePlus.C_450028 polynucleotide (SEQ ID NO: 32)

ATGGCAAAGTTGAGCACTCTCCGGCTTGCGAGCCTTCTTTCCCTTGTCAGTGTGCA
5 GGTATCTGCCTCTGTCCATCTATTGGAGAGTCTGGAGAAGCTGCCTCATGGATGGA
AAGCAGCTGAAACCCCGAGCCCTTCGTCTCAAATCGTCTTGCAGGTTGCTCTGACG
CAGCAGAACATTGACCAGCTTGAATCGAGGCTCGCAGCTGTATCCACACCCACTTC
TAGCACCTACGGCAAATACTTGGATGTAGACGAGATCAACAGCATCTTCGCTCCA
AGTGATGCTAGCAGTTCTGCCGTCGAGTCTTGGCTTCAGTCCCACGGAGTGACGAG
10 TTACACCAAGCAAGGCAGCAGCATTGTTTCAAACAAACATCTCCACTGCAAAT
GCGATGCTCAGCACCAATTTCCACACGTACAGCGATCTCACCGGCGCGAAGAAGG
TGCGCACTCTCAAGTACTCGATCCCGGAGAGCCTCATCGGCCATGTCGATCTCATC
TCTCCCACGACCTATTTTGGCACGACAAAGGCCATGAGGAAGTTGAAATCCAGTG
GCGTGAGCCCAGCCGCTGATGCTCTAGCCGCTCGCCAAGAACCTTCCAGCTGCAA
15 AGGAACTCTAGTCTTTGAGGGAGAAACGTTCAATGTCTTTCAGCCAGACTGTCTCA
GGACCGAGTATAGTGTTGATGGATACACCCCGTCTGTCAAGTCTGGCAGCAGAAT
TGGGTTTGGTTCCCTTTCTCAATGAGAGCGCAAGCTTCGCAGATCAAGCACTCTTTG
AGAAGCACTTCAACATCCCCAGTCAAACTTCTCCGTTGTCCTGATCAACGGTGGA
ACGGATCTCCCTCAGCCGCTTCTGACGCCAACGATGGCGAAGCCAACCTGGACG
20 CTCAAACCATTTTGACCATCGCACATCCTCTCCCCATCACCGAATTCATCACCGCC
GGCAGTCCGCCATACTTCCCCGATCCAGTTGAACCTGCGGGAACACCCAACGAGA
ACGAGCCTTATTTACAGTATTACGAATTTCTGTTGTCCAAGTCCAACGCTGAAATT
CCGCAAGTCATTACCAACTCCTACGGCGACGAGGAGCAAACCTGTGCCGCGGTCAT
ATGCCGTTTCGAGTTTGCAATCTGATTGGTCTGCTAGGACTACGCGGTATCTCTGTC
25 CTTCAATTCCTCGGGCGACGAGGGTGTGGGCGCCTCTTGCGTTGCTACCAACAGCAC
CACGCCTCAGTTTAACCCCATCTTTCTGCTACATGTCCTTATGTTACAAGTGTG
CGGAACCGTGAGCTTCAATCCCGAGGTTGCCTGGGCTGGTTTCATCTGGAGGTTTCA
GCTACTACTTCTCTAGACCCTGGTACCAGCAGGAAGCTGTGGGTACTTACCTTGAG
AAATATGTCAGTGCTGAGACAAAGAAATACTATGGACCTTATGTCGATTTCTCCGG
30 ACGAGGTTTCCCCGATGTTGCAGCCCACAGCGTCAGCCCCGACTATCCTGTGTTTC
AGGGCGGTGAACTACCCCAAGCGGAGGCACTTCAGCAGCCTCTCCTGTCGTAGC
AGCCATCGTGCGCTGTTGAACGATGCCCGTCTCCGCGAAGGAAAACCCACGCTT
GGATTTCTCAATCCGCTGATTTACCTACACGCCTCCAAAGGGTTCACCGACATCAC
CTCGGGCCAATCTGAAGGGTGCAACGGCAATAACACCCAGACGGGCAGTCCTCTC

CCAGGAGCCGGCTTCATTGCAGGCGCACACTGGAACGCGACCAAGGGATGGGACC
CGACGACTGGATTTGGTGTTCACAAACCTCAAAAAGCTCCTCGCACTTGTCCGGTTC
TAAGAGGACTCGGGTGAATATAGGGCCGCGGCGGATGGCTCGCAATAGGACTGC
CATGGAATTGGTAGAAGTCACCATAGGATCATCATTCTCGTCC

5 [43] >|gi|Trire2|81004|estExt_GeneWisePlus.C_220006 polynucleotide (SEQ ID NO: 33)
ATGAAGACCTTTATCCCCGTTGCGCTTCTTGCGCTCTCCCAAGCCACCTCGGCGTG
TCTTCTTCCCCATGAGCGAGACGAGGCGAGTGGCCTGAAGCCAGTCGTCCGCCGC
CAGTCCAGCAACGGAACCTCCATCGGCACCGGCGACAGATTTAGCGGTGGCTCAA
CTGCGCCTCGAGGACTGGGAACTCAGTCCTCATCCACCTCCTTCAGCACAAATTCTC
10 AACGTCAAGGAGGTCACCAGCGGTCTGCAGGGCCTGGCCAACACGTACGGAGTTC
AAACTTTCAACACTCCTTACAAGACGGCCCAGGGCGCAACCGTCGTCTGGAGCCAA
GGTCGGCGGTACCGGAGGTAACGTACCGATGCCTACCGCGTCTTCTTCAACGGCA
ACATCCACGCCCCGAGAGCGCGGCTCCGCAGACAGTGTGCTCTACTTCATCTCGGAT
CTGCTGTATGCCAACGCCCACAACACCGGCCTCACCTACGGCTCCAAGACGTACA
15 GCAACGCCCAGGTGAAGACGGCGCTGGCCGCGGGAATCGTCTTCATCCCGCTCAG
CAACCCCGACGGCGTGGCCTATGACCAGTCCACCAACAGCTGCTGGCGAAAGAAC
CGCAACCCCAACTCGGGCCAGTCCCCCGGAGTCGACCTGAACCGGAACTTTGACT
TCCTCTGGGACTTCCGCAACCTCTTCGCCTCGTCGGCACAGTCCAGCGTTGGCTCG
ACCTCTCCTAGCTCCGAGACGTACCACGGTGCCAGCGCCTTCTCTGAGCCCAGAC
20 CAAGAACATCAAGTGGGTGTTTCGACACGTACTCCAAAGTCCGCTGGTTCGTGGAC
CTTCACTCGTATGCCGGTGACGTCCTGTGGAACCTGGGGCAGCGATGAGAACCAGG
TCGACTACCCGACCATGAACTTCCTCAATGGCACCTACAACAAGGTCCGCGGAAT
CCTGACCGACACGCCGTCTCCCGGCCGAGGCTACGGCGAGTATGTCCCCCAGGCG
GACCTGGACGTTAAGGAGGCCGCGGCCAAGCGCGTGGCCTCAGCTCTGACTGCTG
25 GCGGCGGCCGCTCTTACACGGCCTTTTCAGTCGGCGCAGCTGTATCCCACCTCGGGC
GCGAGTGATGACTACGCTTACTCGAGGCACTTTTCCGATCCTACCAAGAATCTCAT
CCACTCTTACACGATTGAGTTTGGCTTTGCCAACAACGCCGCGTCTTGCCCTTTTA
TCCCTCTGTTTCCCAGTACAACCTCGAACCTCAGGGCTACCAGTGCTGGTTTCATGG
AGCTGCTGTTGGCAGCCACCGACTACGGCCTTGAGATGCAACGACGTGTTAG

30 [44] >|gi|Trire2|110490|fgenesh5_pg.C_scaffold_20000149 polynucleotide (SEQ ID NO:
34)
ATGAACGGCCTCAGCTTTGAACGGGCCAAGGCCCTCCGGGAGCCCAGCGCGGACC
GCGCAACGATGCACTTTTCGCGCCTCAGGGGCATCAAGGGCCGCCCGGCCAGCA
GCAGGGCCAGCAGCAACACCACCACCGCTCTGCCGTCTCGGCGCTCAGCCGCGTC

CGCAGCAGCAGCAGCAGCAGCACGAGCAGCGGCGGCTTCAAGCAGTCGTTCCAAA
ACGTGTCGGCCGTGACGGACGCCTCGACGCAGTACGCCATCCAGTGCGGCTGGGA
CGGCGTGCCCGTCTGGCTGCTGGTCGACACGGGCAGCTCCGACACGTGGGCGACG
CAGACGGGCTTCGAGTGCGCGGACCTGGGCGGGCGACGCGCACAGCGAGGCGGCG
5 TGCGGCTTCTCGCGGCCGCTGATTGACGGCTTCGGCGGGCGGCCAGATCGACGACCT
GCACTTCTTCCTCAAGTACGGCTCGGGCGAGCGCATCTCGGGCCCGATGGGCTACA
GCGACCTGTCTGTCGGCGGGCGTCGCCGTGGCCCGGCAGCAGGTCGGCCTGGCCAA
CAGCACCTACTGGCACGGCAACAACGTCACCGTCGGCATCCTGGGGCTGGCGTAC
CCGGCCATCACGAGCGCCTACTACGGCGACGTCGGCTCCGAGGCGCCCTGGAACG
10 CAATGAGCTACACGCCCTTCCTGACGAGCGCCATCAGCCAGGGAGCCATCGAGCC
GCTCTTCAGCGTCGCCCTCGTGCGCAACTCGACCGACGGCGTCATTGCCTGGGGCG
GGCTGCCCCCATGGACTGGCAGTTTCACGGCTTTGCCAAGACGGATCTGATTGTG
GCAAACCTCATCGGGTCGCCCAGACGGCGTGGAAGCACTCCTTCTACACGATTG
TGCCCGACGGCATGAAATGGGACCAGACGACGGACACGACCAAGTATCCATACAT
15 TGTCGACACGGGCACGACGATGCTCTACCTGCCCCCTCCCCTCGCCGAGGCCATCG
CCAACGCCTTCCAGCCGCGCGCCGTCTACCTCTACCAATGGGGCACCTACTTCGTC
GAGTGCACCGCCATCCCGCCGCACTTTGCCATCCTCATCGAGGGCGTCGAGTTCTG
GATCAACCCGGCCGACCTCATCTACCGCGACCTCGTCGACCCGCTGACGGGCTACT
GCGCCGTTGGCGTCGCCAGCGGCGGGCCGGGGCCGTATATTCTCGGCGACGTCTTC
20 TTGCAAAACGTCGTCGCCGTGTTTGACGTCGGAGCCGCCGAGATGCGCTTCTATGG
GCGCAAGTGA

[45] >|gi|Trire2|121133|estExt_fgenes5_pg.C_60052 polynucleotide (SEQ ID NO: 35)

AGCCCCTCGACTCCAGAAACCTCGAGGCCTCTCCAATCATGAGATGCAGCTACAA
GCTGCTCTCTGTTGTTTTTCCTGCTTCTGCAGCTTTTTTCACAGTATCGGTTCCCTCTG
25 CATAGCTTTGCATCTGCCATTCCAGGGGTTGTGCCAGGATCTTGCGGACGACAGCT
GCTGCGCATTGCCTGTGTAGCGTCACACTGACGGAAGAGGCAAACCCGTGCATCA
TTCAACCTGCTGTGAGACACACGCGTCAGCTGTCAAGGGACACACGATCGATGTT
ACAGCGACACGCGCGTTGAGTATCATCATATGGCTGATAACTGGTTGATATTGTGA
CGATCTGATATACGCAGACATTCTTTGAAACGACGGTGGAATCAAGACAGGCCTA
30 CGTAAACCTGCCCTCACTGCTGTGTATAAATCGGCCTCAACTCCTCCCACTATTCC
CGGGAACAGAATCCCAAACCTCCATCAACCAATTCCCACACAGACTCAACAACCAG
CCTCAACCCCATCAAAACAGACAGGCAGACACACAGCTCCCGTCCAAAATGGAAG
CCATCCTCCAGGCCCAGGCAAAGTTCAGGCTCGACCGCGGCCTCCAGAAAATCAC
CGCCGTGCGCAACAAGAACTACAAGCGTCACGGCCCCAAGTCCTACGTCTACCTC

CTCAACCGCTTTGGCTTCGAGCCCACCAAGCCCGGCCCTTACTTCCAGCAGCACCG
CATCCACCAGCGCGGCCTCGCCACCCAGACTTCAAGGCCGCCGTCGGCGGGAGA
GTCACGAGGCAGAAGGTGCTCGCCAAGAAGGTCAAGGAGGATGGGACGGTCGAT
GCGGGCGGCAGTAAGACGGGAGAGGTGCTGCTGAGGACCAGCAGAATGACTCT
5 GAATACCTTTGCGAGGTGACCATCGGAACTCCTGGGCAGAAGCTTATGCTCGACTT
TGACACCGGCTCATCCGATCTCTGGGTCTTCTCTACAGAACTGAGCAAGCATCTTC
AAGAAAACACGCCATCTTCGACCCCAAAAAGTCGTCCACGTTCAAGCCCCTAAA
GGACCAGACATGGCAAATCTCCTACGGCGACGGCAGCTCCGCCTCGGGCACCTGC
GGCTCCGACACCGTCACCCTCGGCGGTCTCTCCATCAAGAACCAGACCATTGAACT
10 CGCCTCCAAGCTCGCGCCCCAGTTCGCCAGGGCACCGGCGATGGCCTCCTCGGTC
TCGCCTGGCCGCAGATCAACACCGTCCAGACGGACGGCCGCCGACGCCCGCCAA
CACGCCCGTCGCCAACATGATCCAGCAGGACGATATCCCCAGCGACGCCAGCTC
TTTACCGCCGCCTTCTACAGCGAGCGCGACGAAAACGCAGAGTCTTTCTACACCTT
TGGCTACATCGACCAGGACCTCGTGTCCGCCTCGGGCCAGGAGATTGCCTGGACG
15 GATGTCGACAACTCGCAGGGCTTCTGGATGTTCCCTTCCACTAAGACTACCATCAA
CGGCAAGGACATTTCTCAGGAGGGTAACACCGCCATTGCCGACACTGGAACCACC
CTCGCTCTGGTCTCCGACGAGGTCTGCGAGGCGCTGTACAAGGCCATCCCTGGCGC
CAAGTACGACGATAACCAGCAGGGCTACGTCTTCCCCATCAACACCGATGCTTCG
AGTCTGCCCCGAGCTCAAGGTCTCGGTTGGAAACACGCAGTTTGTTCATCCAGCCCGA
20 GGACCTGGCTTTTGCTCCTGCGGACGACAGCAACTGGTACGGCGGTGTTTCAGTCGA
GGGGAAGCAACCCCTTTGACATTCTGGGTGACGTATTCTCAAGTCGGTCTATGCG
ATCTTTGATCAGGGCAACCAGCGATTCGGCGCAGTCCCCAAGATCCAGGCAAAGC
AGAACCTGCAGCCCCCCCAGTAAACACCCCAACTCGTCTTATGCCTTCTCGTACCT
TTAGCTCAATAGCGAATGTGGAATGCCCCGTCAATTTAATGAACTGCTATATAATA
25 AACCC

[46] >|gi|Trire2|106661|fgenesh5_pg.C_scaffold_7000462 polynucleotide (SEQ ID NO: 36)

[47] ATGGATGCTATCCGAGCCAGGAGTGCTGCCCCGCCGCTCCAACAGGTTCCAA
GCCGGCTCATCTAAGAACGTCAATGGAACGGCCGATGTTGAGTCTACGAACTGGG
CAGGCGCCGCTATCACGACTTCGGGCGTTACCGAGGTCAGCGGCACCTTCACGGT
30 CCCCAGGCCCAGCGTCCCAGCTGGTGGCAGTTCTCGCGAGGAGTACTGTGGTGCT
GCTTGGGTGGAATTGACGGCTACTCTGATGCCGACCTCATCCAGACTGGCGTCCT
TTGGTGCGTCGAAGACGGCGAATACTTGTACGAGGCTTGGTATGAGTACCTTCCGG
CCGCGTTGGTTGAGTACTCTGGCATCTCTGTGACGGCTGGATCGGTTGTCACCGTT
ACTGCCACCAAGACTGGCACTAACAGCGGCGTTACTACCCTGACAAGCGGTGGGA

AGACTGTATCGCACACCTTCTCCCGACAGAACTCACCCCTGCCCCGGCACGAGCGCT
GAGTGGATTGTCTGAGGACTTTACCTCTGGCAGCTCGTTGGTTCCCTTTGCAGACTT
TGGATCCGTTACCTTTACCGGAGCCACCGCTGTCGTGAACGGAGCAACTGTGACA
GCGGGCGGCGACTCTCCCGTCATTATCGACCTGGAGGACTCAAGGGGAGACATTC
5 TCACATCAACAACGGTCTCTGGCAGCACTGTGACTGTTGAGTATGAATAG711
[48] >jgi|Trire2|51365|estExt_Genewise1.C_220234 polynucleotide (SEQ ID NO: 37)
TCGTCTTGTCTGCGCAAAGCATTGTGTGTCACCTGTCCATACAGCCCTCCCACAG
TCAGACAAGCGGGATAGAGTGTGAGCCATGGTGCCTCCGCCCTATTCGTGTCGCT
GCTCGCGACCTTCTCCGGAGTCATTGCCCGTGTCTCCGGGCATGGGTCAAAGATCG
10 TTCCCGGCGCGTACATCTTCGAATTCGAGGATTCACAGGACACGGCCGATTTCTAC
AAGAAGCTCAACGGCGAGGGCTCAACGCGCCTGAAGTTCGACTACAAGCTGTTCA
AGGGCGTCTCCGTCCAGCTCAAGGACCTAGACAACCATGAGGCAAAGGCCAGCA
GATGGCCCAGCTGCCTGCTGTCAAGAACGTGTGGCCCGTCACCCTCATCGACGCCC
CCAACCCCAAGGTCGAGTGGGTGCGGCAGCACGGCGCCTACTCTGGAGAGCAG
15 GGCGATCAAGAAGCCACCGATCCCGAACGACTCGAGCGACTTCCCCACGCACCAG
ATGACCCAAATCGACAAGCTGCGAGCCAAGGGCTACACGGGCAAGGGCGTCAGG
GTTGCCGTCATTGATACAGGCATTGACTACACCCACCCTGCTCTCGGCGGCTGCTT
TGGTAGGGGCTGTCTGGTCTCCTTTGGCACCGATTTGGTCGGTGACGACTACACCG
GCTTTAACACGCCTGTCCCCGATGATGACCCCGTCGACTGCGCCGGCCACGGTTCT
20 CACGTTGCTGGTATCATTGCTGCGCAGGAGAATCCGTACGGCTTCACTGGCGGCGC
TCCCGATGTCACCCTCGGCGCTTATCGAGTCTTTGGCTGCGACGGCCAGGCCGGTA
ACGATGTCCTGATTTCCGCTTACAACCAGGCCTTTGAGGACGGTGCCCAGATCATC
ACTGCCTCCATTGGCGGTCCCTCTGGCTGGGCTGAGGAGCCGTGGGCGGTTGCCGT
CACCCGCATCGTTGAGGCAGGTGTTCCCTGCACGGTCTCTGCCGGCAACGAGGGC
25 GACTCTGGTCTCTTCTTTGCCAGCACGGCAGCCAATGGCAAGAAAGTCATTGCTGT
CGCCTCCGTCGACAACGAGAACATCCCTTCAGTGCTGTCCGTGGCCTCTTACAAAA
TTGACAGCGGCGCTGCCCAGGACTTTGGCTACGTCTCCTCCTCCAAGGCGTGGGAC
GGCGTGAGCAAGCCCCTGTATGCTGTGTCGTTTCGACACTACTATTCCCGACGATGG
CTGCTCGCCTCTCCCTGACAGCACTCCCGACCTCTCTGACTACATTGTCCTTGTCGG
30 CCGTGGCACCTGCACCTTTGTCCAGAAAGCCAAAATGTCGCTGCAAAGGGCGCC
AAGTACCTGCTCTATTATAACAACATTCCCGGTGCGCTGGCCGTCGATGTCAGCGC
CGTCCCCGAGATTGAGGCTGTCGGCATGGTTCGATGACAAGACGGGTGCTACCTGG
ATTGCCGCCCTCAAGGATGGAAAGACCGTCACCCTGACACTGACTGACCCGATCG
AGAGCGAGAAGCAAATTCAGTTCAGCGACAACCCGACAACCTGGCGGTGCTCTGAG

CGGCTACACAACCTGGGGCCCTACCTGGGAGCTGGACGTCAAGCCTCAGATCAGC
TCTCCCGGCGGCAACATTCTCTCCACGTACCCCGTGGCTCTCGGAGGATATGCCAC
CCTGTCCGGTACCTCCATGGCCTGCCCCCTGACGGCGGCTGCTGTTGCTCTGATTG
GACAAGCTCGTGGCACCTTTGACCCTGCCTTGATCGACAACCTTGTTGGCAACGACT
5 GCGAACCCCCAGCTGTTCAACGACGGCGAGAAGTTCTACGACTTCCTCGCCCCCGT
TCCCCAACAGGGCGGTGGCCTCATCCAGGCCTACGATGCCGCCTTTGCGACCACTC
TCCTGTCACCGTCCAGCCTGTCGTTCAACGACACTGACCACTTCATCAAGAAGAAG
CAGATCACCTCAAGAACACCAGCAAGCAGAGGGTCACCTACAAGCTCAACCACG
TCCCCACCAACACCTTTTACACTCTGGCACCCGGTAACGGCTATCCAGCTCCCTTT
10 CCTAACGACGCCGTTGCCGCTCACGCCAATCTCAAGTTTAATCTGCAGCAAGTGAC
CCTGCCCCGCCGGCAGGTCCATCACTGTGACGTCTTCCCTACTCCCCCAGGGACG
TCGACGCCAAGCGCCTGGCGCTTTGGTCGGGCTACATCACGGTCAACGGCACGGA
TGGCACCAAGTCTGTCTGTCCCGTACCAGGGCCTACCGGCTCCCTGCACAAGCAGA
AGGTGCTCTATCCGGAGGACTCCTGGATCGCCGATTCCACCGATGAAAGCCTGGC
15 CCCTGTTGAGAACGGCACCGTCTTCACCATTCCCGCGCCGGGCAACGCTGGCCCCG
ATGACAAGCTCCCATCGCTCGTCGTCAGCCCTGCCCTTGGCTCTCGTTATGTCCGC
GTTGATCTCGTCCTCCTGTCCGCGCCTCCTCATGGCACCAAGCTCAAGACGGTCAA
GTTCCCTCGACACCACCTCCATCGGCCAGCCTGCCGGATCACCGCTCCTCTGGATCA
GCCGTGGCGCCAACCCTATTGCTTGGACCGGCGAGCTGTCTGACAACAAGTTTGCT
20 CCCCCTGGAACGTACAAGGCCGTGTTCCATGCTCTGCGTATTTTCGGCAACGAGAA
GAAGAAGGAGGACTGGGATGTGAGCGAATCTCCTGCCTTCACCATCAAGTATGCG
TAGGGTAGATGTAAGAGGGTTTCTTGAGGGGTAATATTCAGATCA
[49] >|gi|Trire2|122076|estExt_fgenes5_pg.C_100212 polynucleotide (SEQ ID NO: 38)
TGTTGATAACGCCCTGTCTTGTGTACAAAGATTTCTGTTGACGTTCCCTCCTACCCC
25 GTTGTTCGATACACATTCACCTGCCTGGTAGCACCTCAATACCAAGGTTACTACTA
ACTCCACCAAGACGCGCGGCGATCCATCCCCCTGACCCGAGAAGCCCTACGAAAA
AGCTATCGACATTGTAATAATCACACGAGATTGTAGATTCTTCTACCTTACCTGCCT
ACACACCTATAGCCCAGTAGCTTTGCGCCATGAGAGCCTCGCCCCTGGCCGTTGCA
GGCGTCGCCCTGGCCTCTGCAGCCCAGGCGCAAGTCGTCCAGTTCGACATCGAGA
30 AGCGCCACGCGCCGCGGCTCAGCCGCCGAGACGGAACCATCGACGGCACTCTGTC
CAACCAAAGAGTCCAGGGCGGATACTTTATCAACGTCCAGGTTGGAAGTCCTGGC
CAGAACATTACGTTGCAGCTCGACACAGGCAGTAGCGATGTCTGGGTCCCGTCAT
CAACGGCTGCCATCTGCACCCAGGTCTCAGAACGGAATCCCGGATGCCAATTCGG
AAGCTTCAACCCTGATGACTCCGATACCTTTGACGAGGTCGGCCAAGGCCTGTTG

ACATCACCTACGTCGACGGGTCCTCGTCCAAAGGCGATTACTTCCAGGACAAC TTC
CAGATCAACGGCGTGACCGTCAAGAATCTCACCATGGGTCTCGGCCTGAGCTCAT
CGATCCCCAACGGCCTGATAGGAGTGGGCTACATGAACGACGAGGCCTCCGTCAG
CACGACCCGAAGCACCTACCCGAACCTCCCCATCGTCCTGCAGCAGCAGAAGCTC
5 ATCAACAGCGTGGCCTTTAGCCTGTGGCTCAACGATCTGGATGCCAGCACGGGGT
CCATCCTGTTTCGGCGGCATCGACACGGAAAAGTACCACGGCGACCTGACTAGCAT
CGACATCATCTCTCCCAACGGCGGCAAGACCTTTACCGAGTTTGCCGTCAACCTGT
ACTCGGTGCAGGCCACAAGCCCGAGCGGAACCGACACGCTCTCCACAAGCGAGGA
TACGCTCATCGCCGTGCTGGACTCTGGAACGACCCTGACCTATCTGCCGCAGGACA
10 TGGCCGAGGAGGCGTGGAACGAAGTGGGAGCCGAGTACAGCAACGAGCTCGGCC
TGGCGGTGGTCCCCTGCTCTGTGGGCAACACCAACGGCTTCTTCTCCTTTACCTTTG
CCGGAACCGACGGCCCTACGATCAACGTCACCCTGTCCGAGCTGGTCCTCGACCTC
TTCAGCGGCGGGCCCCGCGCTCGGTTCTCGTCCGGCCCCAACAAAGGGCCAGTCAAT
CTGCGAGTTTGGCATCCAGAACGGGACCGGCTCTCCGTTCTGCTGGGCGACACGT
15 TCCTGCGCTCCGCGTTCGTCTACGACCTCGTCAACAACCAGATTGCCATTGCC
CCGACCAACTTCAACTCCACCAGGACCAACGTCGTCGCCTTTGCGAGCAGCGGCG
CCCCGATCCCGTCTGCCACGGCGGGCCCCGAACCAGAGCAGGACGGGGCCATTCTC
GTCGACGCACTCGGGCCTGTCTGCTGCCAGTGGCTTCCACGATGGTGACGACGAG
AATGCTGGCTCCTTGACGAGCGTCTTTTCTGGTCCGGGAATGGCCGTGGTCGGCAT
20 GACCATCTGCTACACTCTGCTCGGAAGCGCCATATTCGGCATTGGCTGGCTGTAAA
TGTGTGAATCCTGGCTTTTCTCAGCCCTATGAGTGCTCTTTGACGAACTTTTGAATT
TTGCTTGTTATTGGGGCGCATTGGTTTGATTATGGGGTGTCTGGTTCAAATGTTCT
TTATTTCTTCTTCTTCCGTCTATTCATACTAGCAGGCATTTCATTGCTGGAGTTTGG
ATGGACTTACAGGACTTTTCG
25 [50] >|gi|Trire2|123865|estExt_fgenes5_pg.C_270147 polynucleotide (SEQ ID NO: 39)
CTCCAGCCTCTCCTCGAACGACATCTTGCTGCAAACGCTCAGTATGAGGGCCTGCC
TGCTCTTCCTGGGCATCACCGCCCTGGCCACAGCAATCCCAGCTTTGAAACCACCC
CATGGGAGTCCAGACAGAGCCACACAACCCAGCTTGCCAAAGTCTCCATCGCCC
TGCAGCCCGAGTGTGAGAGCTCCTCGAGCAAGCCCTCCACCACCTCTCCGACCCC
30 TCAAGCCCACGGTACGGCCGTTACCTTGGCCGCGAAGAAGCCAAAGCTCTGCTCC
GCCCTCGACGAGAGGCCACGGCCGCGTCAAGAGATGGCTGGCCAGGGCTGGTGT
TCCTGCTCATGATGTTTTGACTGATGGGCAGTTCATCCACGTGCGGACGCTTGCGG
AGAAGGCTCAGGCTCTGCTTGGATTTCGAGTACAACCTCGACGCTGGGTTCGCAGAC
AATCGCAATCTCTACCCTTCCAGGAAAGATACGGAAGCATGTCATGACCGTTCAGT

ATGTTCCGCTTTGGACGGAAGCAGACTGGGAGGAATGCAAGACGATAATCACGCC
TAGCTGCCTGAAAAGGCTGTATCACGTCGACAGTTACCGCGCAAAGTATGAGAGC
AGCAGTCTTTTTTGGCATCGTGGGCTTCAGCGGGCAAGCGGCTCAGCACGACGAAC
TGGACAAGTTTCTCCACGATTTTGGCGCCCTACTCTACCAACGCAAACCTTCTCCATC
5 GAGTCTGTCAATGGAGGGCAGAGTCCCCAGGGGATGAACGAGCCGGCAAGCGAG
GCCAACGGAGATGTCCAGTACGCCGTTGCCATGGGCTATCATGTTCTGTTCTGGTA
TTACGCAGTCGGTGGCGAGAACCATGACATTATCCCAGACCTGGACCTGGTTGAT
ACAACAGAGGAGTATCTTGAGCCGTTTCTCGAATTCGCCAGCCATCTCCTCGACCT
CGACGACGATGAGCTCCCAAGGGTAGTCTCCATCTCGTACGGCGCCAACGAGCAG
10 CTCTTCCCCCGATCCTACGCCCATCAAGTCTGCGACATGTTTGGCCAGCTCGGCGC
CCGCGGCGTCTCCATCGTCGTGGCCGCTGGCGACCTGGGCCCCGGCGTATCATGCC
AATCAAACGACGGCTCAGCAAGGCCAAAGTTCATCCCCTCGTTTCCCGCGACGTG
CCCCTACGTGACGAGCGTGGGGTCGACGCGCGGCATCATGCCCAGGTGGCGGGCG
AGCTTTTTCGTGCGGGCGGCTTCTCGGATTATTTTGCTCGTCCGGCGTGGCAGGATAG
15 AGCGGTGGGCGCGTACCTCGGGGCGCATGGCGAGGAGTGGGAGGGCTTTTATAAT
CCTGCTGGGAGAGGGTTTCCTGATGTTGCTGCTCAGGGGGTGAATTTTCGGTTTAG
AGCCCATGGGAATGAGAGCTTGAGTTCCGGGACGAGCCTCTCCTCGCCGGTGTTTG
CCGCTCTCATTGCTCTGCTCAACGACCATCGATCTAAGAGCGGCATGCCGCCCATG
GGATTCTGAATCCTTGATCTACACCGTCGGCAGTCACGCCTTCACAGACATCAT
20 CGAGGCCAGGTCCGAGGGATGCCCAGGACAGAGCGTCGAATATCTCGCATCGCCC
TATATCCCCAACGCGGGGCTGGAGCGCCGTGCCTGGATGGGATCCGGTGACGGGTT
GGGGCACGCCGCTGTTTGACAGGATGCTGAATCTGTCTTTGGTGTAGGAGGTGATG
ATGGGTGATAAGCTTGGGTTTTGGAGAAGAC

[51] >|Trire2|81517|estExt_GeneWisePlus.C_250157 polynucleotide (SEQ ID NO: 40)

25 ATGAAGTCGGCATTGCTCTGGGCGGCTCCCCTCTCGCTGCTGGCGGGCCTCGGTGC
CTGCGGGAAGAACTTTGACGAGCTTCTCGAGGTCCCGGAAGGATGGACGCAGCTC
AACGATGCCGTCAATCCCTCGCAGCACATAACGGCTGTCGATTGCCGTGAAGCAGC
CATATATCGACAGCCTGGAGGCCAGGATGGCAGAGAAAGGCAACCGTCTGTGAT
GCAAGAGGTCCGGGAGCTGCAGACTCCAGCCAAAAAGGACATTGACAATGTGCTG
30 CACTGGCTCTCGCAGAACAACCTCTACGGCGTCGTCGAAAAGGACTTTATTCGAGT
CTGGACGACCGTCGCCAAAGCCGAGCCGCTGCTCAAGATGAAGCTCAGCCGCTTC
TCGTACGAAGGCAAGCCGGCGGTGCTGCGAACGACCAAGTACACCATCCCCGACT
CCGTGGCCGACTCCATCAGCTTCATCAACCCCATCAACAACCTTCATGTCGGCGCGC
CATCGTGAACGGGGCCTGACTTTCCTACTTCCTCCGTCAAAGGGCGCCGTCCTGCC

AGGGAACACGACGGCCTATTGCGCCGGCAGCGTGACCCCGTCGTGTCTTTCCAAG
CTCTACAACATCAACTACTCGCCCGCCAACACCAGCTCGCCCGTCATCTTTGGCGT
GGCTGGCTTCCTCGAGGAAAACGCCAACCTCCAGGATCTCCGACAGTTCCTCAACC
AGTCGGCGCCCGAGGTGGCCAAGACGGGACGCACCATCAATGTGCGAGCTGGTCAA
5 CGGCGGGCGTCAACTCGCAGGAGCTGAGCGAGTCTGGCATCGAAGCGGCCCTGGAC
GTTGACTATGCCGTGTGCTGGGCTTCCCCACAAACGTCACCTTTTACTCTACCGG
AGGCCGCGGGCGTCAAGCTCAACGACGACGGACAGCCCATCGAGGGCGAGGACGA
CGACAACGAGCCGTACCTCGAGTTCTTCCAGTACCTGCTGGCCAAGCCCGACGGC
CAGGTGCCGCACGTCCTGTGCTGCTGCTACTCGGACGACGAGTTGTCCGTGCCCCG
10 CGACTATGCGAAGCGCGTCTGCAGCCTGTTTGGCCTGCTGACGGCCCGCGGAACG
TCCATCATCTTCTCGTCCGGCGATGGCGGGCGCTCGCGGGCGCCGAGCATCGAGCTG
CCTGACCAACGACGGCACGAAGCGTCAAGTCACCATGGCGACCTTCCCGCCCACC
TGTCCCTGGGTGACGTCGATTGGCGCCGTCACCAACATCGCAGAGCCTCCCAACG
GCGCCAGATTACGACCCGGCGGCTTCTCCAGTACTTTGCCCAACCTCGATGGCAG
15 AACGAGGCAGTCGAGGGCTACGTCAAGGCGCTCGGGGGGCACCTGGACGGCTACT
ACAACGAGTCCATGCGAGCCATCCCCGACGTTTCAGCAGTCGGCACCGCCTTCAG
CATCATCTCGGGCGGCTACCCGCGCTCTGTTTCAGGGCACCAAGCGCCAGCGCCCCCG
TCTTCGCCGCCATGATTGCCCTGATCAACGACGCCCCGCTGCGCGCCGGCAAGAA
GTCGCTGGGCTTCCTGAACCAGCACCTGTACTCGAGCGAGGTCAGAGCCGTGCTG
20 CAGGACATCACGGCGGGACAGAGCGCCTCTTGATCTGGAACGACGCAGACATAC
CGGGAGGCTGGCCGGCGGGCCGAGGGCTGGGATGCCATTACGGGGCTGGGCGTGCC
CAAGCGGTTTCGACAAGCTGATGGAGGTGCTGGTGAATCATCTTCCACGCCTGGCG
ACTGGCTTCGCCCCGCTTCAAGATGCCACGCAGCACGTCGAGCCCGAGTCTCACGC
TCTGCTCCAAGATGTCGCAAACCTCTGCTCAAGGCGCGAAAGGTGGTCGTGGTTA
25 CCGGCGCGGGCATCAGCACCAACTCGGGCATACCGGACTTCCGCTCCGAAAATGG
CCTCTACTCGCTCATCCAGGCTCAGTTCGACGCTGCGAACCAGCCTACGCGACCGG
CCGAACGAAGCAAGGCTGACGGCACCGCCGACAGCGGAGAGGAGCCTCGCCCGA
CGAAACGCGGAATGACGTCGCGGGAGGCTTCCCCAGACCTGGACGAGGTTACCCG
ACAGCTGAGAGACGACATTGAAGCCCCGCGCCGAGTCCCAGAGGCCCCGCTGCGTCC
30 AGTCGAGCCGCCGGCACGCAGCCAGCCGTGGCCGCCTCGGATGCAAACGCGTCGG
CAGAGGGCGTTTGTGTTGAGCACTCCGCGTCGCAAGCCCGCGCTTCCCTCCACCCCC
CTGCCGACGACGTCTCCCTCTCTTCTCCTCCCCGAGAAGACTTTCTGATTCTTTG
CCGTCATGGTCTTCGTCATCGATCCTGCGAACCGAAGATCGTAAGCGCATCACCGA
CGCCTCTACAATGTCGTTTCGTCTCCTCTCTCCTCGCCGCGCCGGTCTCTTCGA

TCCCTTTCACCCTTCCTCTCCTTCCGACGAAGACAGGAGTCGCCGTAGTAGCACCA
CGCCGTCGGAAGCGGGAGAGAACCCGCCCAACGCGATGCCCGCTTCCCAGACATC
CAGCTTCGGAAAGGCCAAACCTGCCCAACATGAAGGGCAAGGACCTCTTCGACGCT
TCGATATGGTCGGATCCTACGAGGACCTCCGTCTTCTACCAGTTTGCAACGTCTCT
5 ACGGCAAAAGGTCCGCGACGCCGAGCCTACAAGCTCCCACAAGTTCATCAGCCAC
CTGAGGGACCGCGGCAAGCTGGTACGATGCTATACGCAAAACATTGACCAAATTG
AGGAGAAGGTCGGCCTGTCGACGTCTTTGGAGGATGGGCCGGGCAGTCGAGGACG
CTTCTCGCGCAAAGCGACGGCAAACGCTTCGCAGCTCAACAAAATGGTGCAAGAA
GTATCCAGCGGTGAGGGCGGCGCATCCAGCCATGTCAATGCGTCCAGCCAATCAT
10 CGAATGGGTTCGAGTGAGCAATCATCAGCAGAGTCATCACAGGCCAATGATAGGAC
GGAAGAGGAAGACAGTGCTGGCAGCAGCAGCACCACCGCCGCCACCACCACCAC
CACCACCACACCGCCAGATCGACCCAAGCCGGCCCCGAGAAAGGAGCCGCCGCA
ATCGGGTGTGGAGTGCGTCTTCCTGCATGGCTCGCTGCAGCTTCTACGATGCTTCC
TCTGCGGGCAGGTGTGCTCTTGGGACGACGACGACCGCGAGGTGGAGACGCTGTC
15 CGGCCTGCAGCCTGAATGTCCTCACTGCGTTGGCGCCACGGAAGCGCGTCAGGAA
AGGGGCAAGCGAGCGCTGGGCGTGGGCAAGCTGCGTCCCGACATTGTGCTCTATG
GCGAGGAGCACCCCAGCGCCACCTCATCAGCCCCATCGTGACGCACGATCTTGC
CCTGTATCCGGACATGCTCCTGATCCTGGGGACCAGCCTGCGAGTCCACGGGCTCA
AGGTCCTCGTCAGGGAGTTTGCAAAGACGGTGCACAGCCGTGGCGGCAAGGTAGT
20 CTTTGTCAACTTCACCAAGCCGCCAGAGAGCTCCTGGGGCGACATCATTGACTACT
GGATCCAGTGGGACTGTGACGCGTGGGTGAGCGATCTCCAGGTGCGCATCCCCAA
GCTCTGGCAAGATCCGGAGCCGCCAAAGCCCAAGAAGAAGCGCGACTCGGGAGG
AGCCGCCGACGACAATCGCGAGGAGAAGAAGAGACCGCCTGCACAGAACCCGGT
TGCTTTCGCGGACACAAAGGTGAATGGCGCGTACTGCACGCTGAAAATTCTCAA
25 GAGCTGCGCAGAATCACATACACTCGTGACTCGGCCGCCATCGGGAACCCCATCA
TCACCGCGCCCGAGCCCCCGACTACCACCGAGGCGATTTCTCGAGCTGTGGCCATG
GATTCTTCGCGCACAAGCACGCCCAGGGGCAAGTCCAAACGATCGCGTAGGTCTG
CGACAGGAGCCATAGACAGGCCTAAGCGGACGCCATCGACACTGAATCCGAACCA
TGGCCGGTCCAAGAAACCCACGGCAGAGGCCAAGAAGAAACAGCAGGAGGAGGA
30 GGAAGTGCCTGAGACGCCTAGCCAGCCGCCAGCTAGCACGGTGGAGGAATTCAGC
TCGATCCTCCACTCTGTCAAGTCTAACCACGCATACGGAAGCGGAAGATGATTG
ACGGAGAAGAATTTGTATTTCCCGCGGTTGGGAAGAAGCGCGGCGCGGTGGATAG
CCTATACAAGGGGCCGGGGGACGACACAAAGGAGCTACCACCTCTGCGCCCTATG
CCG

[52] >|gi|Trire2|123244|estExt_fgenesh5_pg.C_190118 polynucleotide (SEQ ID NO: 41)

CTCTCCCTCCCATCCATCACCCATCATCTTCCCCCAGTGGTTCCCATCACGCAGGTC
GAAGCCGGAGTTTCTCTGTTTCGCTGCGTTTTTTTTCAGTCTCTTCTGGTGTGTTTGT
CGCAGGTTGACCAACTTCCGTCCTTCCCTCACCAGCCCATCACATCCGCCACCATG
5 CGGTCCGTTGTCGCCCTCTCCATGGCGGCCGTTGCCAGGCCAGCACATTCCAGAT
TGGCACCATCCACGAGAAGTCGGCCCCCGTGCTGAGCAACGTCGAGGCCAACGCC
ATCCCCGATGCCTACATCATCAAGTTCAAGGACCACGTGGGTGAGGATGATGCCT
CCAAGCACCACGACTGGATCCAGAGCATCCACACAAACGTTGAGCAGGAGCGCCT
TGAGCTCCGCAAGCGAAGCAACGTCTTTGGCGCCGACGACGTCTTTGACGGTCTG
10 AAGCACACTTTCAAGATTGGCGACGGCTTCAAGGGCTACGCCGGTCACTTCCACG
AGTCTGTCAATTGAGCAGGTCCGGAACCACCTGACGTTGAGTACATCGAGCGCGA
CAGCATTGTGCACACCATGCTTCCCCTCGAGTCCAAGGACAGCATCATCGTTGAGG
ACTCGTGCAACGGCGAGACGGAGAAGCAGGCTCCCTGGGGTCTTGCCCGTATCTC
TCACCGAGAGACGCTCAACTTTGGCTCCTTCAACAAGTACCTCTACACCGCTGATG
15 GTGGTGAGGGTGTTGATGCCTATGTCATTGACACCGGCACCAACATCGAGCACGT
CGACTTTGAGGGTTCGTGCCAAGTGGGGCAAGACCATCCCTGCCGGCGATGAGGAC
GAGGACGGCAACGGCCACGGCACTCACTGCTCTGGTACCGTTGCTGGTAAGAAGT
ACGGTGTTGCCAAGAAGGCCACGTCTACGCCGTCAAGGTGCTCCGATCCAACGG
ATCCGGCACCATGTCTGACGTCGTCAAGGGCGTCGAGTACGCTGCTCTCTCCCACA
20 TTGAGCAGGTGAAGAAGGCCAAGAAGGGCAAGCGGAAGGGCTTCAAGGGCTCCG
TCGCCAACATGTCCCTCGGTGGTGGCAAGACCCAGGCTCTTGACGCTGCCGTCAAC
GCCGCCGTCCGCGCCGGTGTCCACTTTGCCGTTGCTGCCGGCAACGACAACGCTGA
TGCTTGCAACTACTCCCCGCTGCCGCCACTGAGCCCCTCACCGTCGGTGCTTCTG
CTCTCGATGACAGCCGTGCTTACTTCTCCAACCTACGGCAAGTGCCTGACATCTTC
25 GCCCCTGGTCTGAGCATCCAGTCCACCTGGATTGGCTCCAAGTATGCCGTCAACAC
CATCTCTGGTACCTCCATGGCCTCTCCTCACATCTGCGGTCTCCTGGCCTACTACCT
GTCTCTCCAGCCCGCTGGTGACTCTGAGTTCGCTGTTGCCCCCATCACCCCCAAGA
AGCTCAAGGAGAGCGTCATCTCTGTCGCCACCAAGAACGCCCTCTCTGACCTGCCC
GACTCTGACACCCCCAACCTGCTCGCCTGGAACGGCGGTGGCTGCAGCAACTTCTC
30 CCAGATTGTCGAGGCCGGCAGCTACACTGTCAAGCCCAAGCAGAACAAGCAGGCC
AAGCTCCCCAGCACCATGAGGAGCTCGAGGAGGCCATCGAGGGTGACTTTGAGG
TCGTCTCTGGCGAGATCGTCAAGGGTGCCAAGAGCTTTGGCTCCAAGGCGGAGAA
GTTTGCCAAGAAGATCCACGATCTCGTCGAGGAGGAGATTGAGGAGTTCATCTCT
GAGCTCTCCGAGTAAGGTCTATCTCTTTTGTAGTAGTCCCCAATTTATTCCTTTGTG

TGATGGCCACGATGCGCTCTACGTCATGATTGAGTACGAGGGTCATAGTATGGGA
TCAGGAGAGACTTGTTTTGCTTTTTTGGGTATTTGGCACGGATATGTTATTTAACGG
GGACTTTACCCCTAGGTTGTACTCTACGCTGCATGGTGCTACGCTCTCAAGATGA
GCTTATACAGATCATGATAATAAGACAAACCAGCATTGTCAGATTGATGCTTT

5 2056

[53] >jgi|Trire2|81070|estExt_GeneWisePlus.C_220116 polynucleotide (SEQ ID NO: 42)

CTCCTCTCCTCTACACACCCTCCATATCCTCACAGAACCATCATCATGGCCGTCCTC
AGCAGATTGGCCCTCACGGCCTCGTTCGCCCTCTGCGGCGTCTCGGCAGCCGGCAT
CCAGCAGCCGCTGACTGCCCCAGAGTCGCTGCCTCCCTCACACGAGGCCGTCGCC
10 GACTACGGCTCCAAGCCGATCATTGATTCCGAGGCCCTGCAGAGCGCCATCAGCA
TCGACACGCTCGTGAAGCGTGCGGAGAGCTTCTACAAGTTCGCAAAGGCCTCGGA
GGAGGAATACGGCCATCCGACTCGAGTCATTGGCAGTGCCGGCCACGAGCAGACG
CTCAACTACATCACAAACACCCTGCTCGACCACGGCGACTACTACAATGTCTCTGT
CCAGGAATTCCCCGTCACGCTGTCCAACGTCTTCCAGTTCCGCTGGTCTTGGCCG
15 ATGAGGTTTCCAAGTCGGCCATCCCCATGGGGCTGACGCCGCCACCAAGGACAA
GGAGCCCGTCCACGGCGACCTGGTGCTCGTCCAAAACAGCGGCTGCGACGCGTCC
GACTACCCCCAGAACGTCAAGGGCAACATTGCCTTTATCCGCCGAGGAGCCTGCT
CCTTTGGCGACAAGTCGATTGGGGCCGGCAAGGCTGGTGCCAAGGCTGCCGTCAT
CTACAACACCGACCCGGAGGAGCTTCATGGCACGCTGGGACTGCCCGTTGAGGAC
20 CACATTGCCACCTTTGGCATTGACGGCGTCGAGGGCAAGAAGATTCTCGCCAAGC
TCAGCAATGGCGAGTCTGTTCGATGCCATTGCCTACATCGATGCCGAGGTCAAGCA
GATCCAGACGGTCAATGTCTGGCCCAGACGGAAGAGGGCGATCCCGACAACCTGC
GTCATGCTCGGCGGTCACAGCGACGGAGTTGCCGAGGGCCCCGGCATCAACGACG
ACGGCTCCGGCAGTATATCCGTTCTGGAGGTGCGCGTTTCAGCTGACCAAGTTCAAG
25 GTGAACAACCTGCGTCCGCTTCGCCTGGTGGGCTGCCGAGGAGGAGGGCCTGCTGG
GCTCCGACTTTTACGCCGCCAGCCTGTCTGACGAGGAGAACC AAAAGATCCGCCT
CTTCATGGACTATGACATGATGGCCAGTCCCAACTTTGCCTACCAGATCTACAACG
CAACCAACGCCGAGAGCCCTGCCGGCTCCGAGGAGCTGAGGAACCTCTACGTGGA
CTGGTACAAGTCCCAGGGCCTCAACTACACCTTCATTCCCTTTGACGGCCGCAGCG
30 ACTACGACGGCTTCATCCGCGCTGGCATCCCCGCCGGTGGTATTGCTACCGGCGCC
GAGGCTGTCAAGACCAAGGAGGAGGCCGAGATGTTTGGCGGACGCGCCGGCGAG
TGGCTTGACCCTTGCTACCACCAGCTGTGTGACGACCTGGGCAACCTGAACCACAC
CGCCTGGGAGGTCAACACCAAACTTATTGCCCACTCGGTTGCCACCTACGCCCTCT
CGTTTGACGGCTTCCCCAAGCGCAAGCTGGAGACGGAGATGAGCGCCTACAGCCA

GACGACCAAGCACCCACGGGCCCCAAGCTGATTCTGTAAAGTGGTGGCGAGGTGCGT
TTCTGTAGCATTTACGAGCATGCTGT

[54] >jgi|Trire2|53961|e_gw1.1.1743.1 polynucleotide (SEQ ID NO: 43)

ATGCAGCCCTCATTTGGCAGCTTCCTCGTCACCGTCCTGTCTGCCTCCATGGCAGC
5 AGGCAGTGTCAATCCCAGCACAAACGCCAACCCTGGCTCCTTCGAGATCAAGAGA
TCCGCCAACAAAGCCTTCACAGGCCGCAATGGCCCTCTAGCATTAGCCCGTACATA
CGCCAAGTACGGTGTTGAAGTCCCCAAAACCTCTGGTCGATGCTATTCAACTCGTCA
AGTCCATCCAGCTCGCAAAGCGGGACAGCGCCACCGTCACTGCCACGCCGGACCA
CGACGACATCGAGTATCTTGTCCCCGTCAAGATCGGAACTCCTCCCCAAACACTTA
10 ACCTGGATTTTGACACGGGCAGCTCCGATCTCTGGGTCTTCTCATCAGATGTGCGAC
CCGACCTCCTCCCAGGGCCATGACATCTACACCCCGTCCAAGAGCACATCTTCCAA
AAAGTTGGAAGGAGCCTCATGGAACATCACATATGGAGACCGCTCATCATCATCC
GGCGATGTCTACCACGATATTGTCTCCGTCGGAAACCTGACAGTAAAGTCCCAAG
CCGTCGAGTCCGCTCGAAACGTCTCGGCCAGTTACCCAGGGCAACAACGACGG
15 CCTCGTCGGCCTGGCGTTTAGCTCCATCAACACAGTCAAGCCCACGCCGCAAAAG
ACGTGGTACGACAACATCGTCGGCAGCCTTGACTCTCCCGTCTTTGTTGCTGATCT
GCGCCACGACACGCCCGGCAGCTACCACTTCGGCTCCATCCCCTCCGAAGCAAGC
AAAGCCTTCTACGCCCCCATCGACAACAGCAAGGGCTTCTGGCAATTCAGCACGA
GCAGCAACATTAGCGGCCAGTTCAACGCCGTTGCAGACACTGGCACTACTCTGCT
20 GCTCGCCAGCGACGACCTCGTCAAGGCCTACTACGCAAAGGTCCAGGGCGCCCGT
GTGAACGTCTTCCTGGGCGGCTACGTCTTCAACTGCACCACTCAGCTGCCCCGACTT
TACCTTTACTGTTGGAGAGGGCAACATCACTGTCCCCGGTACCTTGATAAACTATT
CCGAGGCTGGCAACGGCCAGTGTTTTGGCGGTATTCAGCCGTCGGGGGGTCTTCCT
TTTGCTATCTTTGGTGACATTGCTCTTAAGGCTGCGTATGTTATTTTTGACAGTGGC
25 AACAAGCAGGTTGGCTGGGCGCAGAAGAAATAG

[55] >jgi|Trire2|69555|e_gw1.29.235.1 polynucleotide (SEQ ID NO: 44)

ATGTTTCATCGCTGGCGTCGCCCTCTCCGCCCTTCTTTGCGCCGATACCGTCTTGGCT
GGTGTGTTGCCAGGACCGAGGCCTTGCCGCTCGTCTCGCTCGCCGTGCTGGCCGTCG
TTCGGCTCCTTTTCAGAAACGACACCAGCCATGCTACTGTTCAATCTAACTGGGGTG
30 GTGCCATACTTGAGGGATCTGGCTTTACCGCAGCTTCGGCTACTGTGAACGTGCCA
AGGGGTGGTGGCGGATCCAATGCTGCTGGCTCTGCTTGGGTGGCATTGATGGCGC
CAGCTGCCAGACTGCCATTCTCCAGACGGGATTTGACTGGTACGGTGATGGCACCT
ACGATGCCTGGTATGAGTGGTACCCTGAGTTTGCTGCCGACTTCTCGGGCATTGAT
ATCCGCCAAGGCGACCAAATTGCCATGTCTGTTGTTGCCACCTCTTTAACCGGCGG

CTCTGCTACGCTCGAGAACCTTTCCACTGGCCAGAAGGTTACTCAGAACTTCAACC
GTGTCACTGCTGGGTCTCTCTGCGAAACAAGCGCCGAGTTCATTATCGAGGACTTT
GAGGAGTGCAACTCGAACGGCAGCAACTGCCAGCCTGTGCCATTTGCTTCTTTCAG
CCCCGCCATTACTTTTCAGCTCTGCCACAGCCACTCGAAGCGGCCGGAGCGTTTCTC
5 TGAGCGGCGCTGAAATCACCGAGGTTATTGTCAACAACCAGGACCTTACCAGATG
CTCCGTCTCCGGTAGCAGCACCTTGACCTGCTCTTACGTTTAG

[56] >|gi|Trire2|21668|estExt_fgenesh1_pm.C_30227 polynucleotide (SEQ ID NO: 45)

TGATGACCATCACGGCCCAGAACTCACCCCAGAGGTTCTCCTCGCGGCGCCGCG
GCGCTCCCCGGGGGTGCCAATGCCACGGGCGAGCTGGTTCTGTATACGGTCTCG
10 ACATACTCCTTCGACTCCCCTCCAAGACGGCGCAGATCCGCGTCCTCAACCTCAA
AGAAGGCACGTGCGACCTCGTCTCCGAAGACTCCGCCGCCAGCGAACCAGATCTGG
ATTGCCGAGCAGGAGATTGCCTATGTCAAGAGCCTCGATCACGGTGCTTCGGCCCT
GGTGGCGCAGCACGTCTTCAACCCGAATGAGTCAAACACGATCCAGCGCTTCGGC
GGCAGCATTAAACAGCCTCAAGGCCAAGCCGCTGTCCACAGACAAGGTGGCTTTCT
15 GCTGTGCTGCCCTGACGACCCCCGATGGCCGCATGTACAGCCCGGCTGCGGAGCC
AAAGTCTTACACGTCGGCCAAGATCTACACATCCCTCTTTGTGCGGCACTGGGACT
CCTGGAACACCGAGAACAAGAACTCGCTCTGGTACGGCCAGCTTAACAAGGTCGA
CGGCAAGTGGACGCTTGGAAGTCCGAGCTCACCAACCTCCTGGCCGGCACGCGG
CTCCATTCTCCCGTGCCCCCTTTGGAGGTACCGGTGACTTTGACATATCCACGAC
20 CGGCATTGTGTTTGTGCGAAAGGATCCGGATCTGAACCATGCGAGAACAACCAAG
ACGGACCTGTACTTTGTTCCACTAAATTCATACTTGGACCAGCCAACATTTCCCCA
GATTGTCAAGACGGGAGCGCTGCGAGGCTATTCGCTCTCTCCCGTGTTTTCGAATG
ATGGGAAACAGGTAGCTTTCCTGCGCATGAAGTCCCAGCAGTACGAGGCCGACAA
AACGAGGCTGTTGCTAATTCCGGACGTTACCGACTTGAGCAACGTTACAGGAGTTTT
25 ATGCCACGGAAGATGGCAAGGGTGGTTGGGACTACAAGCCCGATTGGCTCATCTG
GAGCCATGACGATAAGGAACTGTACGTTGCGGCCGAGAAGCACGCTCGAGTCGTC
CTGTGGAAGTTGCCGTCATCGCCATTGGAGGCCAAGTCTCTGCCCACTCCTATTCA
CGAGGATGGATCTGTGGCCGAAGCTAGAGTCCTCGGGAAAGGCTCGTCTCTTTTG
ATCACGACGAGATCTCGCGTCGAAAGCTCCAATACTCGATATTAGACCCCGCCTC
30 CAAATCGACCACCATCATCTCATCCAGTTCAAGGCAAGGTAAGACCTTTGCTCTAA
GCAAGTCGCAGTGCCAGGAGATTTGGTTCAATGGGTCCAAGGGCTACCCCATCCA
CGCCCTAGTGACCCTGCCCTCAACATTGACTCATCCAAGAAGTATCCCTTGGCCT
TTTTCATTCATGGCGGGCCCCAGGGAGCATGGGGAGACGATTGGAGCACTCGATG
GAACCCGGCCGTCTTTGCTGAGCAAGGCTACGTTGTGGTGAGCCCCAATCCAACC

GGGAGCACAGGATATGGCCAAGACCATACCGATGCCATCCAGAACAACCTGGGGA
GGCGACCCCTATATCGACCTGGTCAAGTGCTTTGAGTTTCTGGAGGAGGAAGTGA
ACTACATTGATACAACCAGAGCGGTGGCTCTCGGGGCGTCCTATGGCGGCTACAT
GATTAACCTGGATCCAAGGCCATGACCTTGGGCGGAAATTCAAGGCATTGGTGTGC
5 CACGACGGCGTCTTTTCGACCCTGAACCAGTGGTCCACAGAGGAGCTCTTCTTCCC
GGAGCACGACTTTGGCGGCGCCCTCTGGGAGAATAGAGAAGGTTACGAGAAGTGG
GATCCCGCAAAGCACGTTGGAAACTGGGCTACGCCGCAACTGGTCATTACAAATG
AGCTTGATTACCGCCTGCCCATCTCCGAGGGTCTGGCCATGTTCAACGTCCTACAG
GCTCGTGGCGTGCCGAGCAAGTTTGTCTGTTCCCGGACGAGCATCACTGGGTGCT
10 CAAGCCTGAGAACTCTCTCGTCTGGCACAGGGAAGTGTTGAATTGGATCAACAAG
TACAGTGGAATATCTGAAAAAACTAG

[57] >jgi|Trire2|58698|e_gw1.5.263.1 polynucleotide (SEQ ID NO: 46)

ATGGCCTGGTTGAAAAAGCTGGCCCTCGTCTTGTTGGCAATCGTGCCATATGCCAC
GGCTTCTCCGGCCCTCAGCCCCAGGTCCCGTGAGATTCTGAGCCTCGAAGATCTCG
15 AATCCGAGGACAAGTACGTCATCGGCCTGAAGCAGGGCCTGTCGCCCACGGACTT
GAAGAAGCATCTACTGCGGTCTCCGCCGTCCAATATCGCAACAAGAACAGCACG
TTCGAAGGCGGCACTGGCGTCAAGAGGACGTATGCCATCGGCGACTACCGGGCAT
ACACTGCCGTCTCGATCGAGACACGGTTCGAGAGATTTGGAACGATACTCTTGA
GAAGCCGCCCTGGGGCCTTGCAACGCTCTCGAACAAGAAGCCACACGGATTCTTG
20 TACCGCTACGACAAGAGCGCGGGCGAGGGAACCTTTGCCTACGTGCTGGACACGG
GCATCAACTCCAAGCACGTGGACTTTGAGGGCCGCGCCTACATGGGGTTTCAGTCC
GCCCAAGACGGAGCCGACGGACATTAACGGACACGGGACTCATGTTGCGGGCATC
ATCGGCGGCAAGACCTTTGGCGTGGCCAAGAAGACGCAGCTCATCGGCGTCAAGG
TGTTTCTGGATGATGAGGCGACGACGTCGACGCTGATGGAGGGGCTCGAGTGGGC
25 AGTCAATGACATTACGACAAAGGGCCGCCAGGGCCGTTCTGTCAATTAACATGTCTC
TTGGAGGACCCTATTCGCAAGCGTTGAACGACGCCATCGACCACATCGCCGACAT
GGGGATCCTCCCGGTTGCTGCTGCGGGAAACAAGGGCATTCCGGCCACGTTTCATCT
CGCCCGCTTCCGCCGACAAGGCGATGACGGTGGGCGCCATCAACTCGGACTGGCA
AGAGACCAACTTTTCCAACCTTTGGCCCCCAGGTCAACATTTTGGCGCCCGGAGAAG
30 ACGTCTTGTCGGCGTATGTGAGCACAAACACGGCGACGCGCGTGCTATCGGGCAC
TTCGATGGCGGCGCCTCACGTTGCCGGACTGGCGCTGTACCTGATGGCGTTGGAGG
AGTTTGACTCGACGCAGAAGCTGACGGACCGCATATTGCAGCTGGGGATGAAAAA
CAAGGTGGTCAACCTGATGACGGACTCGCCCAACCTGATTATTCACAACAATGTC
AAATGA

[58] >jgi|Trire2|122703|estExt_fgenesh5_pg.C_140132 polynucleotide (SEQ ID NO: 47)

ATGGCGTCCCGCAGATTAGCTCTGAACCTGTCGCAAGGCCTGCGAGCCCGTTCCGG
CCTGTCAGGCTTGCGCCGGGGCTTTGCTACGCCCTCCACCGTCGGCAAGACGCAGA
CGACGACGCTCAAGAACGGCCTGACTGTGCTACTGAATACTCACCATGGGCTCA
5 GACCTCGACCGTCGGCATGTGGATCGACGCCGGCTCCCGTGCTGAGACCAACGAG
ACCAACGGCACTGCCCCTTCTCGAGCACCTGGCCTTCAAGGGCACCGCAAAGC
GATCACAGCACCAATTGGAACCTCGAGATTGAGAACATGGGTGGTCACCTCAACGC
CTACACCTCGCGCGAGAACACCGTCTACTTTGCCAAGGCCTTCAACTCCGATATTC
CCCAGACCGTCGACATCCTGGCCGATATTCTGCAGAACTCCAAGCTCGAGCAGTCC
10 GCCATCGAGCGCGAGCGCGACGTCATTCTCCGAGAGTCCGAGGAGGTCGAGAAGC
AGGTTGAGGAGGTTGTCTTCGACCACCTGCACGCCACTGCCTTCCAGCACCAAGCCC
CTTGGCCGCACCATCCTCGGCCCCCGCCAGAACATCCGTGACATTACCCGCACCGA
GCTCGTCAACTACATCAAGAACAACCTACACCGCCGACCGCATGGTTCTCGCCGCTG
CCGGTGGCGTTCCCCACGAGCAGCTGGTTCGAGCTCGCCGAGAAGCACTTCTCCGG
15 CCTTGCCAGCCACGGACCCGAGACCGAGGCCTATGTCCTGTCTAAGCAGAAGGCC
GACTTCATTGGCTCTGACGTGCGTGTCCGCGACGACACCATGCCCACTGCCAACGT
CGCCATTGCTGTTGAGGGTGTGAGCTGGAACCTCTGATGACTACTACACTGCTCTCG
TCGCTCAGGCCATTGTTGGCAACTACGACAAGGCCATGGGCAACGCCCCCTACCA
GGGCGGCAAGCTCAGCGGCTACGTTTACAAGCACGACCTGGCCAACAGCTTCATG
20 AGCTTCTCCACCAGCTACAGCGACACTGGTCTCTGGGGCATCTACCTCGTCACCGA
TAACGCCACCCGCTCGACGACCTCGTCCACTTCGCCATCCGTGAGTGGATGCGTC
TCTGCTACAACGTCAGCGAAGCCGAGGTTGAGCGCGCCAAGGCCAGCTGAAGGC
CTCCATCCTGCTGTCCCTGGACGGCACCAACCGCCGTTGCCGAGGACATTGGCCGCC
AGCTCATCACCACTGGCCGCCGCGCCAGCCCCGGCGAGATTGAGCGCAAGATCGA
25 CGCCATCACCGACAAGGATGTCACCGACTTCGCCAACCGATACTCTGGGACAAG
GACATTGCCATCAGCGCTGTTGGAAAGATTGAGGCTCTGTTTGACTACCAGCGACT
GCGGAACACCATGAAGCCCAAGTTTTTAAGGGTGGGAGAGCGGATGCGAAGCGAA
TGGAGCAGGGGCGGACTTGTCTGAA

[59] >jgi|Trire2|60581|e_gw1.7.573.1 polynucleotide (SEQ ID NO: 48)

30 ATGAGCCGTCGCCCCGTCTACTTCAATCCCTTGGCCGAATCATGGACGGCGCCCTC
ACCGGATGACCCTCAGTTGGCATAACCGCTTCCATTACAATTACCGGCTTACTCTC
CCACTCAATTGATCCCTCTCACGGACCTAGCCAAGGAGCTCGGCGTGCAATCCATC
CACCTCAAGGACGAGACAAGTCGACTCGGACTCCCCTCCTTCAAGATCCTCGGCG
CCTCATGGGGTACATTTGCGGGCCATTGCCAGCGACTTGATCTTCCTATAGACTCT

TCCTTGGGCTCCGTCGAGCAGAAGTTGGCATCTTCAAATATCACACTCTACGCGGC
CACTGACGGCAATCATGGCCGGGCGGTTGCCCCGCATGGCCTCCATCCTGGGTGTCC
CAGCCCAGATCCACGTCCCTACGACCATGCACCAAAGCACCATTGATCTGATAAA
GTCAGAGGGTGCGCGCGTCGTCATCTCGGACGGCTTCTACGACGATGCTGTTGTTG
5 ACGCTCGAGTGGCAGCTGCTAAAGACGATACCGCCCTTGTTATCCAGGACTTTGCT
TCGGGCGACTACGTTTCAGATTCCGCAGTGGATCGTCGATGGCTACCTCACCATGAT
GCTGGAAATCGACGGCCAGCTTGGCTGTACTACCCCTGACCTCGTCGTTGTGCCGG
TGGGAGTCGGGTCTTTGCGCAAGCCGTCGTAACGCACTTCAAGAAGCCCGGAAA
ACAAACAAAAGTCCTGACCGTCGAACCTGATACATCCGCCAGTCTCTGGAAAAGC
10 CTCAGAAGCGGCGAGTCCTCCTCCACGTCGGAAAAGTCCCCCAGCATAATGGCAG
GGCTGGATTGCGGGACGCCTTCATCGATATCTTGGGCTGTTCTCCGACACGGAGTC
GATGCCAGCCTCACAATATCTGACTATGAGGCGCACAAAGCTTGCGAATACCTAA
AGTCTCAAGGTGTGTCTGCAGGGCCCTGTGGAGCCGCACCTATCGCTGCGCTCAGG
AGATTAGAGCGAGCGGATAGAGAAAGGCTGGGTCTCACCAAGAAGTCCGTCGTAG
15 TAATCTTTTGCACAGAGGGTGCACGAGACTACGACGTTCTCACAGCGTGGCGAG
CGACGATCCCGTCGAAATAACGCAGACGTTGGTGAAGATCAACTCGGCCAATCCG
TTCCTTGGATCGGTTCCCGGTCCTGGAGAAACTGCTATTGCTCGCTATATCACTGCT
TGGATGGAGCACCGTGATATTGAGTCTCATTGGATTGAGCCAACGTCCGGACGAC
CTTCTGTGGTAGGCATCGTCCGCGGACTCGGCGGTGGCAAGAGCCTCATGCTCAAC
20 GGGCACATTGACACCGTGACCCTCATGGGATACGAAGGCAATCCCCTCAGCGGTG
ACATCCAAGACGGCAAGCTCTATGGCCGTGGAGCCGCTGACATGAAAGGCGGTGT
AGCAGCTGCCATGGCTGCCCTAGCCAACGCGAAAAAGCACAGCCTCAGGGGAGAT
GTCATCTTCACAGGCGTGGCAGACGAGGAGTTTGAGAGCATCGGCACGCAGCAGG
TACTGGAGGCAGGCTGGACAGCTGACGGCGCAATCGTCAGTGAGCCGACCAACAT
25 GGAGATTCTATACGCTCACAAGGGCTTTGTATGGTTTGATGTTGACATTCAGGGCC
TTGCAGCGCACGGCTCGAGATACGACCTCGGTATTGATGCCATTAGCAAGGCAGG
CTATTTCTCTCGTTGAGTTGGATAAACATGCAAGCCATTTGACAGCGCAAAGCGGCG
ATGCTGTCCTTGGCCCAGGTAGCATTACGCCTCCCTCATCAAGGGCGGCGAGGA
AGTCTCATCATACCTTCTCGTGTTTCAGATCCAGCTCGAGCGGCGAACCCTGAACG
30 GCGAGACACCGGAAACTGTACGGAAAGAGCTCGAAGAGATACTCGACGGCCTGA
CGAAAACAGTCCCCAACTTCACGTACAGTCTTCGCACCACTTTTCACCGGTCTCCC
TTCAAGGCCGACCTTTCCACCCCTTTTTCGAAACTCGTCCACAAGCACGTGGGAAA
CACTCTTGGCAGGGAGCCTGCTGTCTTGGGCGCCCCCTACTGGACTGACTGCGCAC
TACTAGATGGTGCTGGGATTCCGGCATTGCTCTGGGGGCCCCAGGGTGAAGGCTT

ACATGGAAAGTTGGAATACGCAGATGTTGAGTCGATAAAGCAGGTCGCAGAGGCC
TTGACTGCAATCGCTGTGGAGTTTTGTAGCTGA

[60] >jgi|Trire2|74129|estExt_GeneWisePlus.C_11150 polynucleotide (SEQ ID NO: 49)

TGTCAGACTCGGTTGCGTCGTAACAGAAACACTATCAATCATCGCGACAACCAAC
5 AGACTTGAGAATTTGCCATTACACCACACCTACCGTCTACAGCACCCCCCGTCAA
TCTTGCAACCCTTTCAAATAAGAAGCTCGAGATTACAGCCCACCATGCCTCTCGTC
GTGCCAGGCATCATGACCAACTCGGACGACAAGACGCAGGAATGGGCCAACAAG
CTCGTCGGCAAGACGTACTCTGAAAACGAGTCCAACGAGACGCGGGATCTGCCAG
AGGTGCACCGCATCATCAAGCCCGGCTCCATTGTGACAAAGGACTTTAGGCCGGA
10 GAGGTTGAATATTCACTTGAATGAGGATGGGACCGTGTCGCATGTGCGGCATGGC
TGA

[61] >jgi|Trire2|111915|fgenesh5_pg.C_scaffold_30000054 polynucleotide (SEQ ID NO: 50)

ATGCTCTTCCAAACCATGCTCCTCGCCCTCATCACCTCCCTGGCCCTCGCCCAATCC
15 GAAGTCGGCCGTCCCTGCGGCTTCAAGATGGCCCCCTGCCCCCTCGACATGAAGTG
CGTCCCCGACAACGCCTACTGCCCTCACCCCAGCCGCTGCCCCGGGCCACTGCGAGT
TCAAGAACAAGTACGACCAGTGCGGCGGCTTCACGCCCCGGCCGCACGTTTGCCG
CCGGGGGTCCCGGTGCCAGGACGATCCCCGGCTGCCGCCCAACTGCGGCATGGCG
TGCGATGCCCCGGGGATCTGCATTCCCGAGAATGCTCCCTTCTGTGGAGGGTTCAT
20 GGGGTTGGCTTGTCCGAAGGGATTGTACTGTTATGATGCGTTGGATGACTGTGATC
CGAACAATGGAGGTGCAGACTGTGGAGGAATCTGTCTGTAG

[62] In some embodiments, the recombinant filamentous fungal cells include a mutation in a gene that has at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or even at least 99% nucleic acid sequence identity to the

25 polynucleotide of SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49 and/or SEQ ID NO: 50.

30 [63] In some embodiments, the recombinant filamentous fungal cells include a mutation in a gene that encodes a polypeptide having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or even at least 99% amino acid sequence identity to the a polypeptide of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 1, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8,

SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24 and/or SEQ ID NO: 25.

5 [64] All patents and publications, including all sequences disclosed within such patents and publications, referred to herein are expressly incorporated by reference. Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs (*See e.g., Singleton et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2D ED., ©*
10 1994, John Wiley and Sons, New York; and Hale and Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, © 1991, Harper Perennial, NY, both of which provide one of skill with a general dictionary of many of the terms used herein). Any methods and materials similar or equivalent to the various embodiments described herein can be used in the practice or testing of the present invention.

15 [65] It is intended that every maximum (or minimum) numerical limitation disclosed in this specification includes every lower (or higher) numerical limitation, as if such lower (or higher) numerical limitations were expressly written herein. Moreover, every numerical range disclosed in this specification is intended to include every narrower numerical range that falls within such broader numerical range, as if such narrower numerical ranges were all expressly
20 written herein.

[66] As used herein, the singular “a”, “an” and “the” includes the plural reference unless the context clearly dictates otherwise. Thus, for example, reference to a “host cell” includes a plurality of such host cells.

[67] Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxyl orientation, respectively.
25 The headings provided herein are not limitations of the various aspects or embodiments of the invention that can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the Specification as a whole.

[68] As used herein, the term “impaired” or “impairment” refers to any method that
30 decreases, but does not abolish, the functional expression of one or more genes or the functional activity of the resulting gene product (*i.e.* protein), fragments or homologues thereof, wherein the gene or gene product exerts its known function to a lesser extent than in the corresponding parent strain. It is intended to encompass any means of gene impairment

include partial deletions, disruptions of the protein-coding sequence, non-coding sequences, or both, insertions, additions, mutations, gene silencing (*e.g.* RNAi genes antisense) and the like.

[69] As used herein, “deletion” of a gene refers to deletion of the entire coding sequence, deletion of part of the coding sequence, or deletion of the coding sequence including flanking regions.

[70] As used herein “disruption” refers to a change in a nucleotide or amino acid sequence by the insertion of one or more nucleotides or amino acid residues, respectively, as compared to the parent or naturally occurring sequence. Accordingly, a “disruption sequence” or “disruption mutant” as used herein refers to a nucleic acid or amino acid sequence, typically a coding region sequence, that comprises an insertion of nucleotides or amino acids.

[71] As used herein, “insertion” or “addition” in the context of a sequence refers to a change in a nucleic acid or amino acid sequence in which one or more nucleotides or amino acid residues have been added as compared to the endogenous chromosomal sequence or protein product.

[72] As used herein, “non-revertable” refers to a strain which will naturally revert back to its corresponding parent strain with a frequency of less than 10^{-7} .

[73] As used herein, the term “corresponding parent strain” refers to the host strain from which a mutant is derived (*e.g.*, the originating and/or wild-type strain).

[74] As used herein, “strain viability” refers to reproductive viability. In some embodiments, the impairment of a gene does not deleteriously affect division and survival of the mutant under laboratory conditions.

[75] As used herein “coding region” refers to the region of a gene that encodes the amino acid sequence of a protein.

[76] As used herein “amino acid” refers to peptide or protein sequences or portions thereof. The terms “protein,” “peptide,” and “polypeptide” are used interchangeably.

[77] As used herein, the term “heterologous protein” or “exogenous protein” refers to a protein or polypeptide that does not naturally occur in the host cell, and includes genetically engineered versions of naturally occurring endogenous proteins.

[78] As used herein, “endogenous protein” or “native protein” refers to a protein or polypeptide naturally occurring in a cell.

[79] As used herein, “host,” “host cell,” or “host strain” refer to a cell that can express a DNA sequence introduced into the cell. In some embodiments of the present invention, the host cells are *Aspergillus sp.*

[80] As used herein, “filamentous fungal cell” refers to a cell of any of the species of microscopic fungi that grow as multicellular filamentous strands including but not limited to: *Aspergillus sp.*, *Rhizopus sp.*, *Trichoderma sp.*, and *Mucor sp.*

[81] As used herein, “*Aspergillus*” or “*Aspergillus sp.*” includes all species within the genus “*Aspergillus*,” as known to those of skill in the art, including but not limited to *A. oryzae*, *A. niger*, *A. awamori*, *A. nidulans*, *A. sojae*, *A. japonicus*, *A. kawachi* and *A. aculeatus*.

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

[83] As used herein the term “gene” means a segment of DNA involved in producing a polypeptide and can include regions preceding and following the coding regions (*e.g.*, promoter, terminator, 5' untranslated (5' UTR) or leader sequences and 3' untranslated (3' UTR) or trailer sequences, as well as intervening sequence (introns) between individual coding segments (exons).

[84] As used herein, “homologous gene,” “gene homolog,” or “homolog” refers to a gene which has a homologous sequence and results in a protein having an identical or similar function. The term encompasses genes that are separated by speciation (*i.e.*, the development of new species) (*e.g.*, orthologous genes), as well as genes that have been separated by genetic duplication (*e.g.*, paralogous genes).

[85] As used herein, “homologous sequences” refers to a nucleic acid or polypeptide sequence having at least about 99%, at least about 98%, at least about 97%, at least about 96%, at least about 95%, at least about 94%, at least about 93%, at least about 92%, at least about 91%, at least about 90%, at least about 88%, at least about 85%, at least about 80%, at least about 75% or at least about 70% sequence identity to a subject nucleotide or amino acid sequence when optimally aligned for comparison. In some embodiments, homologous sequences have between about 80% and 100% sequence identity, in some embodiments between about 90% and 100% sequence identity, and in some embodiments, between about 95% and 100% sequence identity.

[86] Sequence homology can be determined using standard techniques known in the art (see *e.g.*, Smith and Waterman, *Adv. Appl. Math.*, 1981. 2:482; Needleman and Wunsch, *J. Mol. Biol.*, 1970. 48:443; Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 1988. 85:2444; programs such as GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software

Package (Genetics Computer Group, Madison, WI); and Devereux *et al.*, *Nucl. Acid Res.*, 1984. 12:387-395).

[87] Useful algorithms for determining sequence homology include: PILEUP and BLAST (Altschul *et al.*, *J. Mol. Biol.*, 1990. 215:403-410; and Karlin *et al.*, *Proc. Natl. Acad. Sci. USA* 5 1993. 90:5873-5787). PILEUP uses a simplification of the progressive alignment method of Feng and Doolittle (Feng and Doolittle, *J. Mol. Evol.*, 1987. 35:351-360). The method is similar to that described by Higgins and Sharp (Higgins and Sharp, *CABIOS* 1989. 5:151-153). Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps.

10 [88] A particularly useful BLAST program is the WU-BLAST-2 program (See, Altschul *et al.*, *Meth. Enzymol.*, 1996. 266:460-480). WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database 15 against which the sequence of interest is being searched. However, the values may be adjusted to increase sensitivity. A % amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the "longer" sequence in the aligned region. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored).

20 [89] As used herein, the term "vector" refers to any nucleic acid that can be replicated in cells and can carry new genes or DNA segments into cells. Thus, the term refers to a nucleic acid construct designed for transfer between different host cells. An "expression vector" refers to a vector that has the ability to incorporate and express heterologous DNA fragments (*i.e.*, non-native DNA) in a cell. Many prokaryotic and eukaryotic expression vectors are commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.

25 [90] As used herein, the term "DNA construct" refers to a nucleic acid molecule 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.*, vectors or vector elements, as described above). For example, DNA construct can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus, or nucleic acid fragment. In some embodiments, DNA constructs also include a series of specified nucleic acid elements that 30

permit transcription of a particular nucleic acid in a target cell. In some embodiments, a DNA construct of the invention comprises a selectable marker.

[91] Also as used herein, the term “DNA construct” (as well as “transforming DNA,” and “transforming sequence”) refers to DNA that is used to introduce sequences into a host cell or organism (*i.e.*, “transform a host cell”). The DNA construct may be generated *in vitro* by PCR or any other suitable techniques. In some embodiments, the transforming DNA can include an incoming sequence, and/or can include an incoming sequence flanked by homology boxes. In yet a further embodiment, the transforming DNA comprises other non-homologous sequences, added to the ends (*e.g.*, stuffer sequences or flanks). The ends can be closed such that the transforming DNA forms a closed circle (*i.e.*, a plasmid), such as, for example, insertion into a vector.

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

[93] As used herein, the terms “isolated” and “purified” are used to refer to a molecule (*e.g.*, a nucleic acid or polypeptide) or other component that is removed from at least one other component with which it is naturally associated.

[94] As used herein, the term “altered expression” is construed to include an increase or decrease in production of a protein of interest by an altered (*i.e.*, engineered) cell strain relative to the normal level of production from the corresponding unaltered parent strain (*i.e.*, when grown under essentially the same conditions).

[95] As used herein, the term “enhanced expression” is construed to include the increased production of a protein of interest by an altered (*i.e.*, engineered) cell strain above the normal level of production from the corresponding unaltered parent strain (*i.e.*, when grown under essentially the same conditions).

[96] As used herein, the term “expression” refers to a process by which a polypeptide is produced. The process includes both transcription and translation of the gene. In some embodiments, the process also includes secretion of the polypeptide.

[97] As used herein in the context of “introducing a nucleic acid sequence into a cell,” the term “introducing” (and in past tense, “introduced”) refers to any method suitable for transferring the nucleic acid sequence into the cell, including but not limited to transformation, electroporation, nuclear microinjection, transduction, transfection, (*e.g.*, lipofection mediated and DEAE-Dextrin mediated transfection), incubation with calcium phosphate DNA

precipitate, high velocity bombardment with DNA-coated microprojectiles, agrobacterium mediated transformation, and protoplast fusion.

[98] As used herein, the terms “stably transformed” refers to a cell that has a non-native (heterologous) polynucleotide sequence integrated into its genome or as an episomal plasmid that is maintained for at least two generations.

[99] As used herein “an incoming sequence” refers to a DNA sequence that is being introduced into a host cell. The incoming sequence can be part of a DNA construct, can encode one or more proteins of interest (*e.g.*, heterologous protein), can be a functional or non-functional gene and/or a mutated or modified gene, and/or can be a selectable marker gene(s).

10 [100] As used herein, “homology box” refers to a nucleic acid sequence, which is homologous to the sequence of gene in the chromosome of a filamentous fungal cell. 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
15 a gene to be impaired according to the invention. These sequences direct where in the chromosome a DNA construct or incoming sequence is integrated and directs what part of the chromosome is replaced by the DNA construct or incoming sequence. While not meant to limit the invention, a homology box may include between about 1 base pair (bp) to 200 kilobases (kb). Typically, a homology box includes about between 1 bp and 10.0 kb; between
20 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.

25 [101] In an alternative embodiment, the transforming DNA sequence comprises homology boxes without the presence of an incoming sequence. In this embodiment, it is desired to delete the endogenous DNA sequence between the two homology boxes. Furthermore, in some embodiments, the transforming sequences are wild-type, while in other embodiments, they are mutant or modified sequences. In addition, in some embodiments, the transforming
30 sequences are homologous, while in other embodiments, they are heterologous.

[102] As used herein, the term “target sequence” refers to a DNA sequence in the host cell that encodes the sequence where it is desired for the incoming sequence to be inserted into the host cell genome. In some embodiments, the target sequence encodes a functional wild-type

gene or operon, while in other embodiments the target sequence encodes a functional mutant gene or operon, or a non-functional gene or operon.

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

[104] As used herein, the term “chromosomally integrated” refers to a sequence, typically a mutant gene (*e.g.*, disrupted form of a native gene), that has become incorporated into the chromosomal DNA of a host cell. Typically, chromosomal integration occurs via the process of “homologous recombination,” wherein the homologous regions of the introduced (transforming) DNA align with homologous regions of the host chromosome. Subsequently, the sequence between the homologous regions is replaced by the incoming sequence in a double crossover. Thus, “chromosomally integrated” is used interchangeably herein with “homologously recombined” or “homologously integrated.”

[105] As used herein, the terms “selectable marker” and “selective marker” refer to a nucleic acid capable of expression in host cell, which allows for ease of selection of those hosts containing the marker. Thus, the term “selectable marker” refers to genes that provide an indication that a host cell has taken up (*e.g.*, has been successfully transformed with) an incoming nucleic acid 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. Selective markers useful with the

present invention include, but are not limited to, antimicrobial resistance markers (*e.g.*, *ampR*; *phleoR*; *specR*; *kanR*; *eryR*; *tetR*; *cmpR*; *hygroR* and *neoR*; see *e.g.*, Guerot-Fleury, *Gene*, 1995. 167:335–337; Palmeros *et al.*, *Gene* 2000. 247:255-264; and Trieu-Cuot *et al.*, *Gene*, 1983. 23:331-341), auxotrophic markers, such as *trpC*, *pyrG* and *amdS*, and detection markers, such as β -galactosidase.

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

[107] A nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA encoding a secretory leader (*i.e.*, a signal peptide), is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation.

Generally, “operably linked” means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

[108] As used herein, the term “hybridization” refers to the process by which a strand of nucleic acid joins with a complementary strand through base pairing, as known in the art.

[109] A nucleic acid sequence is considered to be “selectively hybridizable” to a reference nucleic acid sequence if the two sequences specifically hybridize to one another under moderate to high stringency hybridization and wash conditions. Hybridization conditions are based on the melting temperature (T_m) of the nucleic acid binding complex or probe. For example, “maximum stringency” typically occurs at about $T_m-5^\circ\text{C}$ (5° below the T_m of the probe); “high stringency” at about $5-10^\circ\text{C}$ below the T_m ; “intermediate stringency” at about $10-20^\circ\text{C}$ below the T_m of the probe; and “low stringency” at about $20-25^\circ\text{C}$ below the T_m . Functionally, maximum stringency conditions may be used to identify sequences having strict

identity or near-strict identity with the hybridization probe; while an intermediate or low stringency hybridization can be used to identify or detect polynucleotide sequence homologs.

[110] Moderate and high stringency hybridization conditions are well known in the art. An example of high stringency conditions includes hybridization at about 42°C in 50%

5 formamide, 5X SSC, 5X Denhardt's solution, 0.5% SDS and 100 µg/ml denatured carrier DNA followed by washing two times in 2X SSC and 0.5% SDS at room temperature and two additional times in 0.1X SSC and 0.5% SDS at 42°C. An example of moderate stringent conditions include an overnight incubation at 37°C in a solution comprising 20% formamide, 5 x SSC (150mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x
10 Denhardt's solution, 10% dextran sulfate and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1x SSC at about 37 - 50°C. Those of skill in the art know how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

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

[112] As used herein, the term "primer" refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a
25 primer extension product which is complementary to a nucleic acid strand is induced, (*i.e.*, in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH). Usually, the primer is single stranded for maximum efficiency in amplification. Most often, the primer is an oligodeoxyribonucleotide.

[113] As used herein, the term "polymerase chain reaction" ("PCR") refers to methods for
30 amplifying DNA strands using a pair of primers, DNA polymerase, and repeated cycles of DNA polymerization, melting, and annealing (see, *e.g.*, U.S. Patent Nos. 4,683,195 4,683,202, and 4,965,188, which are hereby incorporated by reference herein).

[114] As used herein, the terms “restriction endonucleases” and “restriction enzymes” refer to bacterial enzymes, each of which cut double-stranded DNA at or near a specific nucleotide sequence.

5 [115] A “restriction site” refers to a nucleotide sequence recognized and cleaved by a given restriction endonuclease and is frequently the site for insertion of DNA fragments. In certain embodiments of the invention restriction sites are engineered into the selective marker and into 5' and 3' ends of the DNA construct.

[116] In the present invention, the host cell is a filamentous fungal cell. Filamentous fungal cells useful with the present invention include, but are not limited to: *Aspergillus sp.*, (e.g., *A. oryzae*, *A. niger*, *A. awamori*, *A. nidulans*, *A. sojae*, *A. japonicus*, *A. kawachi* and *A. aculeatus*); *Rhizopus sp.*, *Trichoderma sp.* (e.g., *Trichoderma reesei* (previously classified as *T. longibrachiatum* and currently also known as *Hypocrea jecorina*), *Trichoderma viride*, *Trichoderma koningii*, and *Trichoderma harzianum*s)) and *Mucor sp.* (e.g., *M. miehei* and *M. pusillus*). In a preferred embodiment, the host cells are *Aspergillus niger* cells.

15 [117] In certain embodiments, the mutation leading to altered decreased protease activity or increased protease inhibitor activity in the cell is a deletion in a non-coding regulatory region flanking the coding sequence of the gene. In another embodiment, the mutation comprises an insertion mutation. In certain embodiments, the insertion mutation comprises insertion of a selectable marker. In some embodiments, wherein the genomic DNA is already known, the 5' flanking fragment and the 3' flanking fragment of the locus to be deleted is cloned by two PCR reactions, and in embodiments wherein the locus is disrupted or otherwise altered, the DNA fragment is cloned by one PCR reaction.

25 [118] In some embodiments, the coding region flanking sequences include a range of about 1 bp to 2500 bp; about 1 bp to 1500 bp, about 1 bp to 1000 bp, about 1 bp to 500 bp, and 1 bp to 250 bp. The number of nucleic acid sequences comprising the coding region flanking sequence may be different on each end of the gene coding sequence. For example, in some embodiments, the 5' end of the coding sequence includes less than 25 bp and the 3' end of the coding sequence includes more than 100 bp.

30 [119] In some embodiments, the incoming sequence comprises a disruption sequence that comprises a selective marker flanked on the 5' and 3' ends with a fragment of the gene sequence. In other embodiments, when the DNA construct comprising the selective marker and gene, gene fragment or homologous sequence thereto is transformed into a host cell, the location of the selective marker renders the gene non-functional for its intended purpose. In some embodiments, the incoming sequence comprises the selective marker located in the

promoter region of the gene. In other embodiments, the incoming sequence comprises the selective marker located after the promoter region of gene.

[120] In yet other embodiments, the incoming sequence is a disruption sequence comprising the selective marker located in the coding region of the gene. In further embodiments, the incoming sequence comprises a selective marker flanked by a homology box on both ends. In still further embodiments, the incoming sequence includes a sequence that interrupts the transcription and/or translation of the coding sequence. In yet additional embodiments, the DNA construct includes restriction sites engineered at the upstream and downstream ends of the construct.

[121] In one embodiment, the *A. nidulans amdS* gene provides a selectable marker system for the transformation of filamentous fungi useful with the present invention. The *amdS* gene codes for an acetamidase enzyme deficient in strains of *Aspergillus* and provides positive selective pressure for transformants grown on acetamide media. The *amdS* gene can be used as a selectable marker even in fungi known to contain an endogenous *amdS* gene or homolog, e.g., in *A. nidulans* (Tilburn *et al. Gene* 1983, 26: 205-221) and *A. oryzae* (Gomi *et al. Gene* 1991, 108:91-98). Background *amdS* activity of non-transformants can be suppressed by the inclusion of CsCl in the selection medium.

[122] Methods for using *amdS* marker system in the transformation of industrially important filamentous fungi are established in the art (e.g., in *Aspergillus niger* (see e.g., Kelly and Hynes *EMBO J.* 1985, 4:475-479; Wang *et al., Fungal Genet. Biol.* 2008, 45(1):17-27); in *Penicillium chrysogenum* (see e.g., Beri and Turner, *Curr. Genet.* 2987, 11:639-641); in *Trichoderma reesei* (see e.g., Pentilla *et al. Gene* 1987, 61:155-164); in *Aspergillus oryzae* (see e.g., Christensen *et al., Bio/technology* 1988, 6:1419-1422); in *Trichoderma harzianum* (see e.g., Pe'er *et al., Soil Biol. Biochem.* 1990, 23:1043-1046); and U.S. patent no. 6,548,285, each of which is hereby incorporated by reference herein).

[123] The DNA constructs comprising an incoming sequence may be incorporated into a vector (e.g., in a plasmid), or used directly to transform the filamentous fungal cell, thereby resulting in a mutant. Typically, the DNA construct is stably transformed resulting in chromosomal integration of the impaired gene which is non-revertable. Methods for *in vitro* construction and insertion of DNA constructs into a suitable vector are well known in the art.

[124] Deletion and/or insertion of sequences is generally accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide linkers can be prepared and used in accordance with conventional practice. (See, Sambrook (1989) *supra*, and Bennett and Lasure, MORE GENE MANIPULATIONS IN FUNGI, Academic Press, San Diego

(1991) pp 70 – 76.). Additionally, vectors can be constructed using known recombination techniques (*e.g.*, Invitrogen Life Technologies, Gateway Technology). Examples of suitable expression and/or integration vectors that may be used in the practice of the invention are provided in Sambrook *et al.*, (1989) *supra*, Ausubel (1987) *supra*, van den Hondel *et al.* (1991) in Bennett and Lasure (Eds.) MORE GENE MANIPULATIONS IN FUNGI, Academic Press pp. 396-428 and U.S. Patent No. 5,874,276. Exemplary vectors useful with the present invention include pBS-T, pFB6, pBR322, pUC18, pUC100 and pENTR/D.

[125] In some embodiments, at least one copy of a DNA construct is integrated into the host chromosome. In some embodiments, one or more DNA constructs of the invention are used to transform host cells.

[126] Impairment occurs via any suitable means, including deletions, substitutions (*e.g.*, mutations), disruptions, insertions in the nucleic acid gene sequence, and/or gene silencing mechanisms, such as RNA interference (RNAi). In one embodiment, the expression product of an impaired gene is a truncated protein with a corresponding change in the biological activity of the protein. In some embodiments, the impairment results in an attenuation of biological activity of the gene. In some embodiments, remaining residual activity will be less than 25%, 20%, 15%, 10%, 5%, or 2% compared to the biological activity of the same or homologous gene in a corresponding parent strain.

[127] In some embodiments, impairment is achieved by deletion and in other embodiments impairment is achieved by disruption of the protein-coding region of the gene. In some embodiments, the gene is altered by homologous recombination.

[128] In the cases a deletion is used to impair a gene, typically the deletion is partial. In some embodiments, a deletion mutant comprises deletion of one or more genes that results in a stable and non-reverting deletion. Flanking regions of the coding sequence may include from about 1bp to about 500 bp at the 5' and 3' ends. The flanking region may be larger than 500 bp but typically does not include other genes in the region which may be impaired or deleted according to the invention.

[129] In some embodiments, the disruption sequence comprises an insertion of a selectable marker gene into or near the protein-coding region. Typically, this insertion is performed *in vitro* by reversely inserting a gene sequence into or near the coding region sequence of the gene to be impaired. Flanking regions of the coding sequence may include about 1 bp to about 500 bp at the 5' and 3' ends. The flanking region may be larger than 500 bp, but will typically not include other genes in the region. The DNA construct aligns with the homologous

sequence of the host chromosome and in a double crossover event the translation or transcription of the gene is disrupted.

[130] While not meant to limit the methods used for impairment or increased expression, in some embodiments decreased protease expression and/or increased protease inhibitor
5 expression is accomplished by homologous recombination or DNA editing, for example, using the CRISPR method, optionally in by insertion of a selectable marker in the coding or non-coding region of the target gene (s).

[131] In some embodiments, impairment of the gene is by insertion in a single crossover event with a plasmid as the vector. For example, the vector is integrated into the host cell
10 chromosome and the gene is altered by the insertion of the vector in the protein-coding sequence of the gene or in the regulatory region of the gene.

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

[133] In some embodiments a mutant encompassed by the invention will exhibit altered expression and translation (*i.e.*, protein production) of one or more endogenous and/or heterologous proteins of interest in comparison to the expression and translation of the same
20 protein(s) by the corresponding parent strain of filamentous fungus.

[134] In some embodiments, the mutants of filamentous fungal cells encompassed by the invention will produce the endogenous and/or heterologous proteins of interest in an amount at least about 5% to about 200% (or more) greater than the production of the same protein(s) in the corresponding parent strain. Accordingly, in some embodiments, the production of the
25 protein(s) of interest by the mutant is at least about 0% to 100% greater, and in some embodiments is at least about 10% to 60% greater, including embodiments wherein production at least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, and 55% greater, than the production of the endogenous and/or heterologous protein(s) in the corresponding parent strain.

[135] In some embodiments of the present invention, the protein of interest produced by the
30 mutant of a filamentous fungal cell is an intracellularly produced protein (*i.e.*, an intracellular, non-secreted polypeptide). In other embodiments, the protein of interest is a secreted polypeptide. In addition, the protein of interest may be a fusion or hybrid protein. In some embodiments, the mutant exhibits altered production of a plurality of proteins, some of which are intracellular and some of which are secreted.

[136] Proteins of interest useful with the present invention include enzymes known in the art, including, but not limited to those chosen from amylolytic enzymes, proteolytic enzymes, cellulytic enzymes, oxidoreductase enzymes and plant cell-wall degrading enzymes. More particularly, these enzyme include, but are not limited to amylases, glucoamylases, proteases, xylanases, lipases, laccases, phenol oxidases, oxidases, cutinases, cellulases, hemicellulases, esterases, peroxidases, catalases, glucose oxidases, phytases, pectinases, glucosidases, isomerases, transferases, galactosidases and chitinases. In some embodiments, enzymes include but are not limited to amylases, glucoamylases, proteases, phenol oxidases, cellulases, hemicellulases, glucose oxidases and phytases. In some embodiments, the polypeptide of interest is a protease, cellulase, glucoamylase or amylase.

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

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

[139] In some embodiments of the present invention, the polypeptide of interest is chosen from hormones, antibodies, growth factors, receptors, cytokines, etc. Hormones encompassed by the present invention include but are not limited to, follicle-stimulating hormone, luteinizing hormone, corticotropin-releasing factor, somatostatin, gonadotropin hormone, vasopressin, oxytocin, erythropoietin, insulin and the like. Growth factors include, but are not limited to platelet-derived growth factor, insulin-like growth factors, epidermal growth factor, nerve growth factor, fibroblast growth factor, transforming growth factors, cytokines, such as interleukins (*e.g.*, IL-1 through IL-13), interferons, colony stimulating factors, and the like. Antibodies include but are not limited to immunoglobulins obtained directly from any species from which it is desirable to produce antibodies. In addition, the present invention

encompasses modified antibodies. Polyclonal and monoclonal antibodies are also encompassed by the present invention. In some embodiments, the antibodies or fragments thereof are chimeric or humanized antibodies, including but not limited to: anti-p185^{Her2}, HulD10-, trastuzumab, bevacizumab, palivizumab, infliximab, daclizumab, and rituximab. In some embodiments, the antibodies include one or more antibodies, as in the case of an antibody cocktail for treating a disease such as caused by Ebola virus.

[140] In a further embodiment, the nucleic acid encoding the protein of interest will be operably linked to a suitable promoter, which shows transcriptional activity in a fungal host cell. The promoter may be derived from genes encoding proteins either endogenous or heterologous to the host cell. The promoter may be a truncated or hybrid promoter. Further the promoter may be an inducible promoter. Typically, the promoter is useful in a *Trichoderma* host or an *Aspergillus* host. Suitable nonlimiting examples of promoters include *cbh1*, *cbh2*, *egl1*, *egl2*, and *xyn1*. In one embodiment, the promoter is one that is native to the host cell. Other examples of useful promoters include promoters from the genes of *A. awamori* and *A. niger* glucoamylase genes (*glaA*) (Nunberg *et al.*, *Mol. Cell Biol.* 1984. 4:2306-2315 and Boel *et al.*, *EMBO J.* 1984. 3:1581-1585); *Aspergillus oryzae* TAKA amylase; *Rhizomucor miehei* aspartic proteinase; *Aspergillus niger* neutral alpha-amylase; *Aspergillus niger* acid stable alpha-amylase; *Trichoderma reesei* *stp1* and the *cellobiohydrolase 1* gene promoter (*see e.g.*, EP 0 137 280 A1, which is hereby incorporated by reference herein) and mutant, truncated and hybrid promoters thereof.

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

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

techniques are known in the art (see, *e.g.*, Ausubel *et al.*, (1987), *supra*, chapter 9; and Sambrook (1989) *supra*, Campbell *et al.*, *Curr. Genet.* 1989. 16:53-56 and THE BIOTECHNOLOGY OF FILAMENTOUS FUNGI, © 1992, Chap. 6. Eds. Finkelstein and Ball, Butterworth and Heinenmann, each of which is hereby incorporated by reference herein).

5 [143] Production of heterologous proteins in filamentous fungal cell expression systems are also known in the art. For example, the expression of heterologous proteins in *Trichoderma* is described in Harkki *et al.*, *Enzyme Microb. Technol.* 1991. 13:227-233; Harkki *et al.*, *Bio Technol.* 1989. 7:596-603; EP 244,234; EP 215,594; and Nevalainen *et al.*, "The Molecular Biology of *Trichoderma* and its Application to the Expression of Both Homologous and
10 *Heterologous Genes*", in MOLECULAR INDUSTRIAL MYCOLOGY, © 1992, Eds. Leong and Berka, Marcel Dekker Inc., NY, pp. 129 – 148; and U.S. Patent Nos. 6,022,725 and 6,268,328, each of which is hereby incorporated by reference herein.

[144] The expression of heterologous proteins in *Aspergillus sp.* is described in Cao *et al.*, *Sci.* 2000. 9:991-1001; and U.S. Patent No. 6,509,171, each of which is hereby incorporated by
15 reference herein.

[145] Transformants of the present invention can be purified using known techniques.

[146] The filamentous fungal cells may be grown in conventional culture medium. The culture media for transformed cells may be modified as appropriate for activating promoters and selecting transformants. The specific culture conditions, such as temperature, pH and the
20 like will be apparent to those skilled in the art. Typical culture conditions for filamentous fungi useful with the present invention are well known and may be found in the scientific literature such as Sambrook, (1982) *supra*, and from the American Type Culture Collection. Additionally, fermentation procedures for production of heterologous proteins are known *per se* in the art. For example, proteins can be produced either by solid or submerged culture,
25 including batch, fed-batch and continuous-flow processes. Fermentation temperature can vary somewhat, but for filamentous fungi such as *Aspergillus niger* the temperature generally will be within the range of about 20°C to 40°C, typically in the range of about 28°C to 37°C, depending on the strain of microorganism chosen. The pH range in the aqueous microbial ferment (fermentation admixture) should be in the exemplary range of about 2.0 to 8.0. With
30 filamentous fungi, the pH normally is within the range of about 2.5 to 8.0; with *Aspergillus niger* the pH normally is within the range of about 4.0 to 6.0, and typically in the range of about 4.5 to 5.5. While the average retention time of the fermentation admixture in the fermentor can vary considerably, depending in part on the fermentation temperature and culture employed, generally it will be within the range of about 24 to 500 hours, typically

about 24 to 400 hours. Any type of fermentor useful for culturing filamentous fungi may be employed in the present invention. One useful embodiment with the present invention is operation under 15L Biolafitte (Saint-Germain-en-Laye, France).

[147] Various assays are known to those of ordinary skill in the art for detecting and measuring activity of intracellularly and extracellularly expressed polypeptides. Means for determining the levels of secretion of a protein of interest in a host cell and detecting expressed proteins include the use of immunoassays with either polyclonal or monoclonal antibodies specific for the protein. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), fluorescence immunoassay (FIA), and fluorescent activated cell sorting (FACS). However, other methods are known to those in the art and find use in assessing the protein of interest (*See e.g.*, Hampton *et al.*, SEROLOGICAL METHODS, A LABORATORY MANUAL, © 1990, APS Press, St. Paul, MN; and Maddox *et al.*, *J. Exp. Med.*, 1983, 158:1211, each of which is hereby incorporated by reference herein).

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

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

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

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

CLAIMS

What is claimed is:

1. A filamentous fungal cell comprising at least one mutation that decreases the amount of active protease in the cell, wherein the mutation is in a gene encoding a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24 and SEQ ID NO: 25.
2. The filamentous fungal cell of claim 1, wherein the mutation is located in a non-coding region of the gene.
3. The filamentous fungal cell of claim 1, wherein the mutation is located in a coding region of the gene.
4. The filamentous fungal cell of any of claims 1-3, wherein the mutation results in reduced expression of a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22 and SEQ ID NO: 23.
5. The filamentous fungal cell of claim 4, wherein the mutation is a polynucleotide homologous to or identical to a polynucleotide selected from the group consisting of the polynucleotide of SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47 and SEQ ID NO: 48.
6. The filamentous fungal cell of any of claims 1-3, wherein the mutation results in overexpression of the polypeptide homologous to a polypeptide of SEQ ID NO: 24 and/or SEQ ID NO: 25.
7. The filamentous fungal cell of claim 6, wherein the mutation is a polynucleotide homologous to or identical to a polynucleotide of SEQ ID NO: 49 or SEQ ID NO: 50.

8. The filamentous fungal cell of any of claims 1-7, wherein the mutation comprises an insertion mutation.

9. The filamentous fungal cell of claim 8, wherein the insertion mutation comprises insertion of a selectable marker.

5 10. The filamentous fungal cell of claim 8 or 9, wherein the insertion mutation comprises insertion of an expression cassette for overexpressing the polypeptide of SEQ ID NO: 24 and/or SEQ ID NO: 25.

11. The filamentous fungal cell of any of claims 1-10, wherein the filamentous fungal cell is an *Aspergillus* species, a *Rhizopus* species, a *Trichoderma* species or a *Mucor* species.

10 12. The filamentous fungal cell of claim 11, wherein the *Trichoderma* species is selected from the group consisting of *Trichoderma reesei*, *Trichoderma viride*, *Trichoderma koningii*, and *Trichoderma harzianum*.

13. The filamentous fungal cell of any of claims 1-12, wherein the mutation results in increased production of a protein of interest compared to otherwise identical parental
15 filamentous fungal cells that lack the deletion.

14. The filamentous fungal cell of any of claims 1-13, wherein the protein of interest is an antibody or fragment, thereof.

15. A method for increasing expression of a protein of interest in a filamentous fungal host, the method comprising:

20 (i) introducing a mutation is in a gene encoding a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18,
25 SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24 and SEQ ID NO: 25 into filamentous fungal host cells capable of expressing a protein of interest;

(ii) cultivating the filamentous fungal cell having the deletion under conditions conducive for production of the protein of interest; and

30 (iii) recovering the protein of interest;

wherein the presence of the mutation results in increased production of the protein of interest compared to the production in otherwise identical parental filamentous fungal cells that lack the deletion.

16. The method of claim 15, wherein the mutation is located in a non-coding region of the gene, a coding region of the gene, or both.

17. The method of claim 15 or 16, wherein the mutation results in reduced expression of a polypeptide homologous to or identical to a polypeptide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, 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, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22 and SEQ ID NO: 23, and/or overexpression of a polypeptide homologous to or identical to a polypeptide of SEQ ID NO: 24 and/or SEQ ID NO: 25.

18. The method of any of claims 15-17, wherein the mutation is a polynucleotide homologous to or identical to a polynucleotide selected from the group consisting of the polynucleotide of SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49 and SEQ ID NO: 50.

19. The method of any of claims 15-18, wherein the mutation comprises an insertion mutation, optionally having a selectable marker.

20. The method of claim 19, wherein the filamentous fungal cell is an *Aspergillus* species, a *Rhizopus* species, a *Trichoderma* species or a *Mucor* species.

21. The method of any of claims 15-20, wherein the mutation results in increased production of a protein of interest compared to otherwise identical parental filamentous fungal cells that lack the deletion.

22. The method of any of claims 15-21, wherein the protein of interest is an antibody or fragment, thereof.

23. A method for increasing expression of a protein of interest in a filamentous fungal host, the method comprising:

- (i) cultivating the filamentous fungal cell of any of claims 1-14 under conditions conducive for production of a protein of interest; and
- (ii) recovering the protein of interest;

wherein the presence of the mutation results in increased production of the protein of interest compared to the production in otherwise identical parental filamentous fungal cells that lack the deletion.

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2017/061475

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N15/80 C12P21/02 C12N9/58 C12N15/52
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 C12P

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

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

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/075677 A2 (NOVOZYMES INC [US]; MAIYURAN SUCHINDRA [US]; JANG ABIGALL [US]; BROWN) 23 June 2011 (2011-06-23) abstract; sequences 5,6 page 17 page 10 page 12	1-5,8,9, 11-23
X	WO 2013/028912 A2 (NOVOZYMES INC [US]; YAYER DEBBIE [US]; JIN QIMING [US]) 28 February 2013 (2013-02-28) abstract; sequences 25,26 page 24 page 103 page 24 - page 25 ----- -/-	1-5,8,9, 11-23



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

31 January 2018

Date of mailing of the international search report

03/04/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Weinberg, Suzanna

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2017/061475

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2006/110677 A2 (GENENCOR INT [US]; WANG HUAMING [US]) 19 October 2006 (2006-10-19) abstract page 14 -----	1-5,8,9, 11-23
A	DIEGO MARTINEZ ET AL: "Genome sequencing and analysis of the biomass-degrading fungus <i>Trichoderma reesei</i> (syn. <i>Hypocrea jecorina</i>)", NATURE BIOTECHNOLOGY (ADVANCE ONLINE PUBLICATION), GALE GROUP INC, vol. 26, no. 5, 1 May 2008 (2008-05-01), pages 553-560, XP002689190, ISSN: 1087-0156, DOI: 10.1038/NBT1403 [retrieved on 2008-05-04] the whole document -----	1-5,8,9, 11-23
A	ANDRĂ CR SCHUSTER ET AL: "Biology and biotechnology of <i>Trichoderma</i> ", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, BERLIN, DE, vol. 87, no. 3, 12 May 2010 (2010-05-12), pages 787-799, XP019841630, ISSN: 1432-0614 the whole document -----	1-5,8,9, 11-23

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/061475

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-5, 8, 9, 11-23(all partially)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-5, 8, 9, 11-23(all partially)

A filamentous fungal cell comprising at least one mutation that decreases the amount of active protease in the cell, wherein the mutation is in a gene encoding a polypeptide homologous to or identical to a polypeptide of SEQ ID NO.1; a method for increasing expression of a protein of interest in a filamentous fungal host, the method comprising cultivating said filamentous fungal cell; a method for increasing expression of a protein of interest in a filamentous fungal host, the method comprising introducing a mutation into a gene encoding said polypeptide.

2-23. claims: 1-5, 8, 9, 11-23(all partially)

As set out for Invention 1, but wherein the polypeptide corresponds to one from the group consisting of SEQ ID NOs. 2-23 respectively.

24. claims: 1-23(partially)

As set out for Invention 1, but wherein the polypeptide corresponds to SEQ ID NO. 24.

25. claims: 1-23(partially)

As set out for Invention 1, but wherein the polypeptide corresponds to SEQ ID NO. 25.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2017/061475

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011075677	A2	23-06-2011	
		AU 2010330835 A1	19-04-2012
		AU 2016201422 A1	24-03-2016
		AU 2017254852 A1	16-11-2017
		BR 112012013205 A2	24-11-2015
		CA 2784599 A1	23-06-2011
		CN 102762714 A	31-10-2012
		CN 105368717 A	02-03-2016
		DK 2513290 T3	25-01-2016
		EP 2513290 A2	24-10-2012
		EP 3000871 A1	30-03-2016
		MX 350399 B	06-09-2017
		US 2012288892 A1	15-11-2012
		US 2015225761 A1	13-08-2015
		WO 2011075677 A2	23-06-2011
WO 2013028912	A2	28-02-2013	
		AU 2012298799 A1	13-02-2014
		BR 112014004190 A2	01-03-2017
		CN 103890180 A	25-06-2014
		EP 2748322 A2	02-07-2014
		US 2014212977 A1	31-07-2014
		US 2017015986 A1	19-01-2017
		WO 2013028912 A2	28-02-2013
WO 2006110677	A2	19-10-2006	
		CN 101146907 A	19-03-2008
		DK 2123746 T3	27-02-2017
		DK 2314667 T3	06-03-2017
		DK 2314668 T3	13-02-2017
		DK 2333045 T3	20-02-2017
		EP 1869160 A2	26-12-2007
		EP 2123746 A2	25-11-2009
		EP 2314667 A2	27-04-2011
		EP 2314668 A2	27-04-2011
		EP 2333045 A2	15-06-2011
		US 2006246545 A1	02-11-2006
		US 2011070612 A1	24-03-2011
		US 2014287462 A1	25-09-2014
		WO 2006110677 A2	19-10-2006