



(86) **Date de dépôt PCT/PCT Filing Date:** 2011/02/01
(87) **Date publication PCT/PCT Publication Date:** 2011/08/04
(45) **Date de délivrance/Issue Date:** 2016/01/05
(85) **Entrée phase nationale/National Entry:** 2012/07/20
(86) **N° demande PCT/PCT Application No.:** JP 2011/052035
(87) **N° publication PCT/PCT Publication No.:** 2011/093509
(30) **Priorité/Priority:** 2010/02/01 (JP2010-019967)

(51) **Cl.Int./Int.Cl. C12N 15/52** (2006.01),
A23D 9/00 (2006.01), **A23L 1/30** (2006.01),
A61K 38/00 (2006.01), **A61K 8/36** (2006.01),
A61Q 19/10 (2006.01), **C11B 1/00** (2006.01),
C12N 1/19 (2006.01), **C12N 15/63** (2006.01),
C12N 9/00 (2006.01), **C12P 7/64** (2006.01)

(72) **Inventeur/Inventor:**
OCHIAI, MISA, JP

(73) **Propriétaire/Owner:**
SUNTORY HOLDINGS LIMITED, JP

(74) **Agent:** FETHERSTONHAUGH & CO.

(54) **Titre : POLYNUCLEOTIDE CODANT POUR UN HOMOLOGUE D'ACYL-COA SYNTHETASE ET SON PROCEDE D'UTILISATION**
(54) **Title: POLYNUCLEOTIDE ENCODING ACYL-COA SYNTHETASE HOMOLOG AND USE THEREOF**

(57) **Abrégé/Abstract:**

The present invention relates to an acyl-CoA synthetase homolog protein from microorganisms of the genus *Mortierella*, a polynucleotide encoding the protein, and so on. The invention provides polynucleotides comprising an acyl-CoA synthetase homolog protein gene from, e.g., *Mortierella alpina*, expression vectors comprising these polynucleotides and transformants thereof a method for producing lipids or fatty acids using the transformants, food products containing the lipids or fatty acids produced by the method, etc.



ABSTRACT

The present invention relates to an acyl-CoA synthetase homolog protein from microorganisms of the genus *Mortierella*, a polynucleotide encoding the protein, and so on. The invention provides polynucleotides comprising an acyl-CoA synthetase homolog protein gene from, e.g., *Mortierella alpina*, expression vectors comprising these polynucleotides and transformants thereof, a method for producing lipids or fatty acids using the transformants, food products containing the lipids or fatty acids produced by the method, etc.

DESCRIPTION

POLYNUCLEOTIDE ENCODING ACYL-COA SYNTHETASE HOMOLOG AND USE THEREOF

5

TECHNICAL FIELD

The present invention relates to a polynucleotide encoding an acyl-CoA synthetase homolog and use thereof.

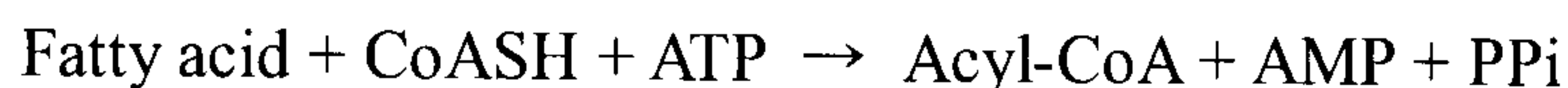
10 BACKGROUND ART

Fatty acids containing two or more unsaturated bonds are collectively referred to as polyunsaturated fatty acids (PUFAs) and known to specifically include arachidonic acid (ARA), dihomo- γ -linolenic acid (DGLA), eicosapentaenoic acid (EPA), docosahexaenoic acid (DHA), etc. Some of these polyunsaturated fatty acids cannot be synthesized in the animal body. It is therefore necessary to compensate these polyunsaturated fatty acids as essential amino acids from food.

Polyunsaturated fatty acids are widely distributed; for instance, arachidonic acid can be separated from lipids extracted from the adrenal glands and livers of animals. However, polyunsaturated fatty acids contained in animal organs are only in a small quantity and cannot be obtained sufficiently for large supplies when simply extracted or separated from animal organs. For this reason, microbial techniques have been developed for obtaining polyunsaturated fatty acids by cultivation of various microorganisms. Above all, microorganisms of the genus *Mortierella* are known to produce lipids containing polyunsaturated fatty acids such as arachidonic acid and the like.

Other attempts have also been made to produce polyunsaturated fatty acids in plants. Polyunsaturated fatty acids constitute storage lipids such as triacylglycerols and are known to be accumulated within microorganism mycelia or plant seeds.

Acyl-CoA synthetase (ACS) is an enzyme catalyzing the thioesterification of fatty acids and coenzyme A (CoA) and catalyzes the following reaction.



Acyl-CoA produced by ACS is involved in various life phenomena including the biosynthesis and remodeling of lipids, energy production by β -oxidation, acylation of proteins, expression regulation by fatty acids, etc. Furthermore, ACS is reportedly associated with extracellular uptake of fatty acids,

30179-221

intracellular transport of fatty acids, etc. (Non-Patent Documents 1 and 2). In view of the foregoing, it is considered to control the activity of ACS when polyunsaturated fatty acids or the like are produced by utilizing microorganisms or plants.

In the yeast *Saccharomyces cerevisiae* used as a model eukaryote, six (6)
 5 acyl-CoA synthetase genes (ScFAA1, ScFAA2, ScFAA3, ScFAA4, ScFAT1 and ScFAT2) are known (Non-Patent Document 1). The proteins encoded by these genes are different in substrate specificity, timing of expression, intracellular localization and function.

Patent Document 1 discloses nine (9) genes as the acyl-CoA synthetase gene
 10 (ScACS) derived from *Schizochytrium* sp. Patent Document 1 also discloses an increased production of DPA (n-6) (docosapentanoic acid (n-6)) or DHA when the gene encoding the *Schizochytrium* sp. PUFA synthase system is co-expressed with ScACS, as compared to the case where the co-expression with ScACS is not involved.

15 In addition, acyl-CoA synthetase genes derived from animals and plants are also reported (Non-Patent Document 2 and Patent Document 2).

[Patent Document 1] Japanese Unexamined Patent Application Publication
 (Translation of PCT Application) No. 2009-529890

20 [Patent Document 2] PCT International Publication Pamphlet WO 0209295

[Non-Patent Document 1] B. B. A. 1771, 286-298, 2007

[Non-Patent Document 2] Exp. Biol. Med., 233 (5), 507-521, 2008

DISCLOSURE OF THE INVENTION

25 Under the foregoing circumstances, it has been desired to isolate a novel gene that increases the amount of the fatty acids produced in a host cell or changes the composition of fatty acids produced, when the gene is expressed in the host cell.

As a result of extensive investigations, the present inventors have succeeded in cloning a gene encoding an ACS homolog of lipid-producing fungus *Mortierella alpina* (hereinafter "M. alpina") (MaACS), and accomplished the present invention.
 30 That is, the present invention provides the following polynucleotides, proteins, expression vectors, transformants, and a method for producing lipids or lipid compositions and foods, etc. using the transformants, as well as foods produced by the method, etc.

30179-221

The present invention as claimed relates to the following:

[1] A polynucleotide as defined in (a) or (b) below:

(a) a polynucleotide encoding a protein consisting of an amino acid sequence, wherein 1 to 10 amino acids are deleted, substituted, inserted and/or added in the amino acid
5 sequence set forth as SEQ ID NO: 47, wherein the protein has acyl-CoA synthetase activity or, when expressed in a host cell, functions to increase the amount of, or change the composition of, the fatty acids produced in the host cell; or

(b) a polynucleotide encoding a protein having an amino acid sequence having at least 90% identity to the amino acid sequence set forth as SEQ ID NO: 47, wherein the protein has
10 acyl-CoA synthetase activity or, when expressed in a host cell, functions to increase the amount of, or change the composition of, the fatty acids produced in the host cell;

[2] The polynucleotide according to [1] above, comprising the nucleotide sequence set forth as SEQ ID NO: 46;

[3] The polynucleotide according to [1] above, encoding a protein consisting of the
15 amino acid sequence set forth as SEQ ID NO: 47;

[4] The polynucleotide according to any one of [1] to [3] above, which is a DNA;

[5] A protein encoded by the polynucleotide according to any one of [1] to [4] above;

[6] A vector comprising the polynucleotide according to any one of [1] to [4] above;

[7] A non-human cell which is transformed with the polynucleotide according to any
20 one of [1] to [4] above, or into which the vector according to [6] above is introduced;

[8] A method for producing a lipid or fatty acid composition, which comprises collecting the lipid or fatty acid composition from the culture of the transformant according to [7] above, wherein the lipid is a triacylglycerol, and wherein the fatty acid is a polyunsaturated fatty acid having at least 18 carbon atoms.

30179-221

The polynucleotide of the present invention can be used for transformation of an appropriate host cell. The transformant thus produced can be used to produce fatty acid compositions, food products, cosmetics, pharmaceuticals, soaps, etc.

5 More specifically, the transformant of the present invention provides an extremely high production efficiency of lipids and fatty acids. Accordingly, the present invention can be effectively used to manufacture pharmaceuticals or health foods which require a large quantity of lipids or fatty acids.

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1 shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-1.

FIG. 2A shows the alignment between the genome sequence and CDS sequence of MaACS-1.

5 FIG. 2B is a continuation from FIG. 2A.

FIG. 3A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-2.

FIG. 3B is a continuation from FIG. 3A.

10 FIG. 4A shows the alignment between the genome sequence and CDS sequence of MaACS-2.

FIG. 4B is a continuation from FIG. 4A.

FIG. 4C is a continuation from FIG. 4B.

FIG. 5 shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-3.

15 FIG. 6A shows the alignment between the genome sequence and CDS sequence of MaACS-3.

FIG. 6B is a continuation from FIG. 6A.

FIG. 7A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-4.

20 FIG. 7B is a continuation from FIG. 7A.

FIG. 8A shows the alignment between the genome sequence and CDS sequence of MaACS-4.

FIG. 8B is a continuation from FIG. 8A

FIG. 8C is a continuation from FIG. 8B.

25 FIG. 9A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-5.

FIG. 9B is a continuation from FIG. 9A.

FIG. 10A shows the alignment between the genome sequence and CDS sequence of MaACS-5.

30 FIG. 10B is a continuation from FIG. 10A.

FIG. 11A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-6.

FIG. 11B is a continuation from FIG. 11A.

35 FIG. 12A shows the alignment between the genome sequence and CDS sequence of MaACS-6.

FIG. 12B is a continuation from FIG. 12A.

FIG. 13 shows the correspondence between the cDNA sequence and putative

amino acid sequence of MaACS-7.

FIG. 14A shows the alignment between the genome sequence and CDS sequence of MaACS-7.

FIG. 14B is a continuation from FIG. 14A.

5 FIG. 15A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-8.

FIG. 15B is a continuation from FIG. 15A.

FIG. 16A shows the alignment between the genome sequence and CDS sequence of MaACS-8.

10 FIG. 16B is a continuation from FIG. 16A.

FIG. 16C is a continuation from FIG. 16B.

FIG. 17 shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-9.

15 FIG. 18A shows the alignment between the genome sequence and CDS sequence of MaACS-9.

FIG. 18B is a continuation from FIG. 18A.

FIG. 19A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-10.

FIG. 19B is a continuation from FIG. 19A.

20 FIG. 20A shows the alignment between the genome sequence and CDS sequence of MaACS-10.

FIG. 20B is a continuation from FIG. 20A.

FIG. 20C is a continuation from FIG. 20B.

25 FIG. 21A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-11.

FIG. 21B is a continuation from FIG. 21A

FIG. 22A shows the alignment between the genome sequence and CDS sequence of MaACS-11.

FIG. 22B is a continuation from FIG. 22A.

30 FIG. 23A shows the correspondence between the cDNA sequence and putative amino acid sequence of MaACS-12.

FIG. 23B is a continuation from FIG. 23A

FIG. 24A shows the alignment between the genome sequence and CDS sequence of MaACS-12.

35 FIG. 24B is a continuation from FIG. 24A

FIG. 25A shows the alignment between MaACS having relatively high amino acid sequence homology to *S. cerevisiae*-derived FAA protein (FAA: fatty

30179-221

acid activation) and the FAA protein. The single underlined and double underlined sequences denote the ATP-AMP motif and the FACS/VLACS-FATP motif, respectively.

FIG. 25B is a continuation from FIG. 25A.

5 FIG. 25C is a continuation from FIG. 25B.

FIG. 26A shows the alignment between MaACS having relatively high amino acid sequence homology to *S. cerevisiae*-derived FAT protein (FAT: fatty acid transferase) and the FAT protein. The single underlined and double underlined sequences denote the ATP-AMP motif and the FACS/VLACS-FATP motif, respectively.

FIG. 26B is a continuation from FIG. 26A.

FIG. 27 shows changes with the passage of time in lipid production (FIG. 27A) and arachidonic acid production (FIG. 27B), per mycelia in MaACS-10-overexpressed *M. alpina*.

15 FIG. 28 shows changes with the passage of time in lipid production (FIG. 28A) and arachidonic acid production (FIG. 28B), per mycelia in MaACS-11-overexpressed *M. alpina*.

BEST MODE FOR CARRYING OUT THE INVENTION

20 Hereinafter, the present invention is described in detail. The embodiments described below are intended to be presented by way of example merely to describe the invention but not limited only to the following embodiments. The present invention may be implemented in various ways without departing from the gist of the invention.

25 This application claims priority to the Japanese Patent Application (No. 2010-19967) filed February 1, 2010.

30 As will be later described in detail in EXAMPLES below, the present inventors have succeeded for the first time in cloning the full-length cDNA of lipid-producing fungus *M. alpina*-derived ACS homolog genes (MaACS-1~12). The present inventors have also identified the nucleotide sequences of genomic DNAs of MaACS-1~12 from *M. alpina* and putative amino acid sequences thereof. The ORF sequences, putative amino acid sequences, CDS sequences, cDNA sequences and genome sequences of MaACS-1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 are SEQ ID NOs: 35 1, 6, 11, 16, 21, 26, 31, 36, 41, 46, 51 and 56 (hereinafter these sequences are

collectively referred to as "ORF sequences of MaACS-1~12"), SEQ ID NOs: 2, 7, 12, 17, 22, 27, 32, 37, 42, 47, 52 and 57 (hereinafter these sequences are collectively referred to as "amino acid sequences of MaACS-1~12"), SEQ ID NOs: 3, 8, 13, 18, 23, 28, 33, 38, 43, 48, 53 and 58 (hereinafter these sequences are collectively referred to as "CDS sequences of MaACS-1~12"), SEQ ID NOs: 4, 9, 14, 19, 24, 29, 34, 39, 44, 49, 54 and 59 (hereinafter these sequences are collectively referred to as "cDNA sequences of MaACS-1~12") and SEQ ID NOs: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 and 60 (hereinafter these sequences are collectively referred to as "genome sequences of MaACS-1~12"), respectively. These polynucleotides and proteins may be obtained by the methods described in EXAMPLES below, known genetic engineering techniques, known methods for synthesis, and so on.

1. Polynucleotide of the Invention

First, the present invention provides the polynucleotide described in any one selected from the group consisting of (a) to (g) below:

- (a) a polynucleotide comprising any one nucleotide sequence selected from the group consisting of the ORF sequences of MaACS-1~12;
- (b) a polynucleotide comprising any one nucleotide sequence selected from the group consisting of the cDNA sequences of MaACS-1~12;
- (c) a polynucleotide encoding a protein consisting of any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12;
- (d) a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 100 amino acids are deleted, substituted, inserted and/or added in any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12, and having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids produced in a host cell when expressed in the host cell;
- (e) a polynucleotide encoding a protein having an amino acid sequence having at least 60% identity to any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12, and having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids produced in a host cell when expressed in the host cell; and,
- (f) a polynucleotide which hybridizes to a polynucleotide consisting of a nucleotide sequence complementary to any one nucleotide sequence selected from the group consisting of the ORF sequences of MaACS-1~12 under stringent

conditions, and which encodes a protein having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids produced in a host cell when expressed in the host cell; and,

(g) a polynucleotide which hybridizes to a polynucleotide consisting of a nucleotide sequence complementary to any one nucleotide sequence selected from the group consisting of the cDNA sequences of MaACS-1~12 under stringent conditions, and which encodes a protein having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids produced in a host cell when expressed in the host cell.

As used herein, the term "polynucleotide" means a DNA or RNA.

As used herein, the term "polynucleotide which hybridizes under stringent conditions" refers to a polynucleotide obtained by the colony hybridization method, plaque hybridization method, Southern hybridization method or the like, using as a probe, for example, a polynucleotide consisting of a nucleotide sequence complementary to any one nucleotide sequence selected from the group consisting of the ORF sequences of MaACS-1~12 or any one nucleotide sequence selected from the group consisting of the cDNA sequences of MaACS-1~12, or the whole or part of a polynucleotide consisting of the nucleotide sequence encoding any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12. For the methods of hybridization, there are used the methods described in, e.g., "Sambrook & Russell, Molecular Cloning: A Laboratory Manual Vol. 3, Cold Spring Harbor, Laboratory Press 2001," "Ausubel, Current Protocols in Molecular Biology, John Wiley & Sons 1987-1997," etc.

As used herein, the term "stringent conditions" may be any of low stringent conditions, moderate stringent conditions and high stringent conditions. The term "low stringent conditions" are, for example, 5x SSC, 5x Denhardt's solution, 0.5% SDS, 50% formamide at 32°C. The term "moderate stringent conditions" are, for example, 5x SSC, 5x Denhardt's solution, 0.5% SDS, 50% formamide at 42°C, or 5x SSC, 1% SDS, 50 mM Tris-HCl (pH 7.5), 50% formamide at 42°C. The term "high stringent conditions" are, for example, 5x SSC, 5x Denhardt's solution, 0.5% SDS, 50% formamide at 50°C or 0.2 x SSC, 0.1% SDS at 65°C. Under these conditions, a DNA with higher identity is expected to be obtained efficiently at higher temperatures, though multiple factors are involved in hybridization stringency including temperature, probe concentration, probe length, ionic strength, time, salt concentration and others, and a person skilled in the art may appropriately select these factors to achieve similar stringency.

When commercially available kits are used for hybridization, for example,

an Alkphos Direct Labeling and Detection System (GE Healthcare) may be used. In this case, according to the attached protocol, after cultivation with a labeled probe overnight, the membrane is washed with a primary wash buffer containing 0.1% (w/v) SDS at 55°C to detect the hybridized DNA. Alternatively, in producing a probe
 5 based on the nucleotide sequence complementary to any one nucleotide sequence selected from the group consisting of the ORF sequences of MaACS-1~12 or any one nucleotide sequence selected from the group consisting of the cDNA sequences of MaACS-1~12, or based on the entire or part of the nucleotide sequence encoding any one amino acid sequence selected from the group consisting of the amino acid
 10 sequences of MaACS-1~12, hybridization can be detected with a DIG Nucleic Acid Detection Kit (Roche Diagnostics) when the probe is labeled with digoxigenin (DIG) using a commercially available reagent (e.g., a PCR Labeling Mix (Roche Diagnostics), etc.).

In addition to those described above, other polynucleotides that can be
 15 hybridized include DNAs having 50% or higher, 51% or higher, 52% or higher, 53% or higher, 54% or higher, 55% or higher, 56% or higher, 57% or higher, 58% or higher, 59% or higher, 60% or higher, 61% or higher, 62% or higher, 63% or higher, 64% or higher, 65% or higher, 66% or higher, 67% or higher, 68% or higher, 69% or higher, 70% or higher, 71% or higher, 72% or higher, 73% or higher, 74% or higher,
 20 75% or higher, 76% or higher, 77% or higher, 78% or higher, 79% or higher, 80% or higher, 81% or higher, 82% or higher, 83% or higher, 84% or higher, 85% or higher, 86% or higher, 87% or higher, 88% or higher, 89% or higher, 90% or higher, 91% or higher, 92% or higher, 93% or higher, 94% or higher, 95% or higher, 96% or higher, 97% or higher, 98% or higher, 99% or higher, 99.1% or higher, 99.2% or higher,
 25 99.3% or higher, 99.4% or higher, 99.5% or higher, 99.6% or higher, 99.7% or higher, 99.8% or higher or 99.9% or higher identity with the DNA for any one nucleotide sequence selected from the group consisting of the ORF sequences of MaACS-1~12 or for any one nucleotide sequence selected from the group consisting of the cDNA sequences of MaACS-1~12, or with the DNA encoding any one amino acid sequence
 30 selected from the group consisting of the amino acid sequences of MaACS-1~12, as calculated by a homology search software, such as FASTA, BLAST, etc. using default parameters.

Identity between amino acid sequences or nucleotide sequences may be determined using FASTA (Science 227 (4693): 1435-1441, (1985)), algorithm
 35 BLAST (Basic Local Alignment Search Tool) by Karlin and Altschul (Proc. Natl. Acad. Sci. USA, 87: 2264-2268, 1990; Proc. Natl. Acad. Sci. USA, 90: 5873, 1993). Programs called blastn, blastx, blastp, tblastn and tblastx based on the BLAST

algorithm have been developed (Altschul S. F. et al., J. Mol. Biol. 215: 403, 1990). When a nucleotide sequence is sequenced using blastn, the parameters are, for example, score=100 and wordlength=12. When an amino acid sequence is sequenced using blastp, the parameters are, for example, score=50 and wordlength=3. When
 5 BLAST and Gapped BLAST programs are used, default parameters for each of the programs are employed.

The polynucleotides of the present invention described above can be obtained by known genetic engineering techniques or known methods for synthesis.

10 2. Protein of the Invention

The present invention provides the proteins shown below.

- (i) A protein encoded by the polynucleotide of any one of (a) to (g) above.
- (ii) A protein comprising any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12.
- 15 (iii) A protein consisting of an amino acid sequence wherein one or more amino acids are deleted, substituted, inserted and/or added in any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12, and having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids produced
 20 in a host cell when expressed in the host cell.
- (iv) A protein having an amino acid sequence having at least 90% identity to any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12, and having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids
 25 produced in a host cell when expressed in the host cell.

The proteins described in (iii) or (iv) above are typically naturally occurring mutants of the protein consisting of any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12 and include those proteins which may be artificially obtained using site-directed mutagenesis described
 30 in, e.g., "Sambrook & Russell, Molecular Cloning: A Laboratory Manual Vol. 3, Cold Spring Harbor Laboratory Press 2001," "Ausubel, Current Protocols in Molecular Biology, John Wiley & Sons 1987-1997," "Nuc. Acids. Res., 10, 6487 (1982)," "Proc. Natl. Acad. Sci. USA, 79, 6409 (1982)," "Gene, 34, 315 (1985)," "Nuc. Acids. Res., 13, 4431 (1985)," "Proc. Natl. Acad. Sci. USA, 82, 488 (1985)," etc.

35 As used herein, the "protein consisting of an amino acid sequence wherein one or several acids are deleted, substituted, inserted and/or added in any one amino acid sequence selected from the group consisting of the amino acid sequences of

MaACS-1~12, and having an acyl-CoA synthetase activity or an activity of increasing the amount and/or changing the composition, of the fatty acids produced in a host cell when expressed in the host cell" includes proteins consisting of an amino acid sequence wherein, e.g., 1 to 100, 1 to 90, 1 to 80, 1 to 70, 1 to 60, 1 to 50, 1 to 40, 1 to 39, 1 to 38, 1 to 37, 1 to 36, 1 to 35, 1 to 34, 1 to 33, 1 to 32, 1 to 31, 1 to 30, 1 to 29, 1 to 28, 1 to 27, 1 to 26, 1 to 25, 1 to 24, 1 to 23, 1 to 22, 1 to 21, 1 to 20, 1 to 19, 1 to 18, 1 to 17, 1 to 16, 1 to 15, 1 to 14, 1 to 13, 1 to 12, 1 to 11, 1 to 10, 1 to 9 (1 to several), 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 1 to 2, or one amino acid is/are deleted, substituted, inserted and/or added in any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12, and having the acyl-CoA synthetase activity or the activity of increasing the amount and/or changing the composition, of the fatty acids produced in a host cell when expressed in the host cell. In general, the number of deletions, substitutions, insertions, and/or additions is preferably smaller.

Such proteins include a protein having an amino acid sequence having the identity of approximately 60% or higher, 61% or higher, 62% or higher, 63% or higher, 64% or higher, 65% or higher, 66% or higher, 67% or higher, 68% or higher, 69% or higher, 70% or higher, 71% or higher, 72% or higher, 73% or higher, 74% or higher, 75% or higher, 76% or higher, 77% or higher, 78% or higher, 79% or higher, 80% or higher, 81% or higher, 82% or higher, 83% or higher, 84% or higher, 85% or higher, 86% or higher, 87% or higher, 88% or higher, 89% or higher, 90% or higher, 91% or higher, 92% or higher, 93% or higher, 94% or higher, 95% or higher, 96% or higher, 97% or higher, 98% or higher, 99% or higher, 99.1% or higher, 99.2% or higher, 99.3% or higher, 99.4% or higher, 99.5% or higher, 99.6% or higher, 99.7% or higher, 99.8% or higher, or 99.9% or higher, to any one amino acid sequence selected from the group consisting of the amino acid sequences of MaACS-1~12, and having the diacylglycerol acyltransferase activity. As the identity percentage described above is higher, the protein is preferable in general.

The term deletion, substitution, insertion and/or addition of one or more amino acid residues in the amino acid sequence of the protein of the invention is intended to mean that one or more amino acid residues are deleted, substituted, inserted and/or added at optional and one or more positions in the same sequence. Two or more types of deletions, substitutions, insertions and additions may occur at the same time.

Examples of the amino acid residues which are mutually substitutable are given below. Amino acid residues in the same group are mutually substitutable. Group A: leucine, isoleucine, norleucine, valine, norvaline, alanine, 2-aminobutanoic

30179-221

acid, methionine, o-methylserine, t-butylglycine, t-butylalanine and cyclohexylalanine; Group B: aspartic acid, glutamic acid, isoaspartic acid, isoglutamic acid, 2-aminoadipic acid and 2-aminosuberic acid; Group C: asparagine and glutamine; Group D: lysine, arginine, ornithine, 2,4-diaminobutanoic acid and
 5 2,3-diaminopropionic acid; Group E: proline, 3-hydroxyproline and 4-hydroxyproline; Group F: serine, threonine and homoserine; and Group G: phenylalanine and tyrosine.

The protein of the present invention may also be produced by chemical synthesis methods such as the Fmoc method (fluorenylmethyloxycarbonyl method),
 10 the tBoc method (t-butyloxycarbonyl method), etc. In addition, peptide synthesizers available from Advanced Automation Peptide Protein Technologies, Perkin Elmer, Protein Technologies, PerSeptive, Applied Biosystems, SHIMADZU Corp., etc. may also be used for the chemical synthesis.

The protein encoded by the polynucleotide of the invention and the protein
 15 of the invention are both ACS homolog proteins and considered to have the acyl-CoA synthetase activity since the ATP-AMP motif and FACS/VLACS-FATP motif, which are important for the acyl-CoA synthetase activity, are conserved. As used herein, ATP, AMP, FACS, VLACS and FATP are intended to mean adenosine triphosphate, adenosine monophosphate, fatty acyl-CoA synthetase, very long chain acyl-CoA
 20 synthetase and fatty acid transport protein, respectively. Specific amino acid sequences of the ATP-AMP motif and FACS/VLACS-FATP motif contained in the protein of the present invention are shown in FIGS. 25 and 26 at the single underlined and double underlined sequences, respectively. With regard to representative amino acid sequences of the ATP-AMP motif and
 25 FACS/VLACS-FATP motif, reference may be made to databases including pfam, etc.

As used herein, the term "acyl-CoA synthetase activity (ACS activity)" is intended to mean the activity of promoting the acyl-CoA-forming reaction through formation of a thioester bond between a fatty acid and coenzyme A (chemical
 30 reaction equation below).



The acyl-CoA synthetase activity can be quantitatively confirmed, for
 35 example, by cultivating for a certain period of time host cells, into which the polypeptide of the present invention is introduced, preparing the lysate of the host cells, mixing the cell lysate with a labeled fatty acid (e.g., polyunsaturated fatty acid

labeled with a radioactive isotope, etc.) and coenzyme A, reacting them for a certain period of time, then extracting free fatty acids with n-heptane, and quantifying the fatty acyl-CoA which is formed during the above reaction and remained in the aqueous fraction, using a scintillation counter. For details of the method for
 5 confirming the acyl-CoA synthetase activity, reference may be made to Black P. N., et al. (J. B. C., 272 (8), 4896-4903, 1997). Alternatively, the acyl-CoA synthetase activity may also be assayed by the method described in "Evaluation of ACS Activity" of EXAMPLE 2, which involves no radioactive label.

The "activity of increasing the amount of the fatty acids produced in a host
 10 cell when expressed in the host cell" is intended to mean the activity that, when the polynucleotide of the present invention or the polynucleotide encoding the protein of the present invention is introduced (transformed) into a host cell and expressed in the host cell, increases the total fatty acid production, as compared to a reference cell (control) derived from the same strain as the host cell in which the polynucleotide
 15 described above is not introduced.

The "activity of changing the composition of the fatty acids produced in a host cell when expressed in the host cell" is intended to mean the activity that, when the polynucleotide of the present invention or the polynucleotide encoding the protein of the present invention is introduced (transformed) into a host cell and
 20 expressed in the host cell, changes the amount or ratio of various fatty acids produced, as compared to a reference cell (control) derived from the same strain as the host cell in which the polynucleotide described above is not introduced.

As used herein, the term "fatty acid" is intended to mean an aliphatic monocarboxylic acid (a carboxylic acid having one carboxylic residue and carbon
 25 atoms connected to each other in a chain) represented by general formula RCOOH (wherein R is an alkyl). The fatty acid includes a saturated fatty acid having no double bond and an unsaturated fatty acid containing a double bond(s) in the hydrocarbon chain. The fatty acid is preferably an unsaturated fatty acid, and more preferably, a polyunsaturated fatty acid containing a plurality of double bonds in the
 30 hydrocarbon chain. The polyunsaturated fatty acid includes preferably an unsaturated fatty acid having carbon atoms of 18 or more, e.g., an unsaturated fatty acid having carbon atoms of 18 or 20, and examples include, but not limited to, oleic acid, linoleic acid, linolenic acid (γ -linolenic acid, dihomo- γ -linolenic acid, etc.), arachidonic acid, and the like. The polyunsaturated fatty acids are particularly
 35 preferably linoleic acid, γ -linolenic acid, dihomo- γ -linolenic acid and arachidonic acid, more preferably, linoleic acid, dihomo- γ -linolenic acid and arachidonic acid,

and most preferably, dihomono- γ -linolenic acid and arachidonic acid.

In the present invention, the "host cell" is not particularly limited so long as the cell is capable of expressing the polynucleotide of the invention when the polynucleotide is introduced. The cells include cells derived from mammals
 5 (excluding human), insects, plants, fungi, bacteria, etc., preferably cells from plants and fungi, more preferably, cells from fungi, and most preferably, lipid-producing fungi or yeast.

The lipid-producing fungi which can be used are the lipid-producing fungi described in, e.g., MYCOTAXON, Vol. XLIV, No. 2, pp. 257-265 (1992). Specific
 10 examples include, but not limited to, microorganisms belonging to the genus *Mortierella* including microorganisms belonging to the subgenus *Mortierella*, e.g., *Mortierella elongata* IFO8570, *Mortierella exigua* IFO8571, *Mortierella hygrophila* IFO5941, *Mortierella alpina* IFO8568, ATCC16266, ATCC32221, ATCC42430, CBS 219.35, CBS224.37, CBS250.53, CBS343.66, CBS527.72, CBS528.72, CBS529.72,
 15 CBS608.70 and CBS754.68, etc., or microorganisms belonging to the subgenus *Micromucor*, e.g., *Mortierella isabellina* CBS194.28, IFO6336, IFO7824, IFO7873, IFO7874, IFO8286, IFO8308 and IFO7884, *Mortierella nana* IFO8190, *Mortierella ramanniana* IFO5426, IFO8186, CBS112.08, CBS212.72, IFO7825, IFO8184, IFO8185 and IFO8287, *Mortierella vinacea* CBS236.82, etc. Among others,
 20 *Mortierella alpina* is preferable.

Specific examples of the yeast include the genus *Saccharomyces*, the genus *Candida*, the genus *Zygosaccharomyces*, the genus *Pichia* and the genus *Hansenula*, and preferably, *Saccharomyces cerevisiae* in the genus *Saccharomyces*. In wild
 25 strains of yeast such as *Saccharomyces cerevisiae*, etc., saturated fatty acids or monovalent fatty acids having mainly 18 or less carbon atoms can be synthesized within the cells, but polyunsaturated fatty acids cannot be synthesized therein. For this reason, when yeast such as *Saccharomyces cerevisiae*, etc. is used as a host cell, it is preferred to impart the ability to synthesize polyunsaturated fatty acids to the yeast cells by genetic engineering, etc. The ability to synthesize polyunsaturated fatty
 30 acids can be imparted by introducing a gene encoding a protein derived from an organism that already possesses the ability to synthesize polyunsaturated fatty acids and takes part in fatty acid synthesis.

The "organism that already possesses the ability to synthesize polyunsaturated fatty acids" includes, for example, lipid-producing fungi. Specific
 35 examples of the lipid-producing fungi are the same as those given hereinabove.

Examples of the gene encoding a protein derived from an organism that already possesses the ability to synthesize polyunsaturated fatty acids and "gene

encoding the protein that takes part in fatty acid synthesis" include, but not limited to, $\Delta 12$ fatty acid desaturase gene, $\Delta 6$ fatty acid desaturase gene, GLELO fatty acid elongase gene and $\Delta 5$ fatty acid desaturase gene, etc. The nucleotide sequences of $\Delta 12$ fatty acid desaturase gene, $\Delta 6$ fatty acid desaturase gene, GLELO fatty acid elongase gene and $\Delta 5$ fatty acid desaturase gene are available by having access to databases including GenBank, etc. For example, in GenBank, Accession No. AB020033, No. AB020032, No. AB193123 and No. AB188307 are entered to access the respective sequences.

The genes for fatty acid synthesis-related proteins described above are inserted into appropriate vectors (e.g., pESC (Stratagene), pYES (Invitrogen), etc.), which are then introduced into yeast by the electroporation method, the spheroplast method (Proc. Natl. Acad. Sci. USA, 75 p1929 (1978)), the lithium acetate method (J. Bacteriology, 153, p163 (1983)), and the methods described in Proc. Natl. Acad. Sci. USA, 75 p1929 (1978), Methods in Yeast Genetics, 2000 Edition: A Cold Spring Harbor Laboratory Course Manual, etc.

Fatty acids can be extracted from the host cells transformed by the polynucleotide of the present invention or the polynucleotide encoding the protein of the present invention in the following manner. A host cell is cultured and then treated in a conventional manner, e.g., by centrifugation, filtration, etc. to obtain cultured cells. The cells are thoroughly washed with water and preferably dried. Drying may be accomplished by lyophilization, air-drying, etc. Depending upon necessity, the dried cells are disrupted using a Dymil or by ultrasonication, and then extracted with an organic solvent preferably in a nitrogen flow. Examples of the organic solvent include ether, hexane, methanol, ethanol, chloroform, dichloromethane, petroleum ether and so on. Alternatively, good results can also be obtained by alternating extraction with methanol and petroleum ether or by extraction with a single-phase solvent system of chloroform-methanol-water. Removal of the organic solvent from the extract by distillation under reduced pressure may give fatty acid-containing lipids. The fatty acids extracted may be converted into the methyl esters by the hydrochloric acid methanol method, etc.

The quantity or ratio of various fatty acids may be determined by analyzing the fatty acids extracted as described above using various chromatography techniques. Examples of the chromatography techniques include, but not limited to, high performance liquid chromatography and gas chromatography, and particularly preferably, gas chromatography.

3. Vector of the Invention and Vector-Introduced Transformants

In another embodiment, the present invention further provides the expression vector comprising the polynucleotide of the invention.

The vector of the invention is generally constructed to contain an expression cassette comprising:

- 5 (i) a promoter that can be transcribed in a host cell;
- (ii) any of the polynucleotides defined in (a) to (g) above that is linked to the promoter; and,
- (iii) an expression cassette comprising as a component a signal that functions in the host cell with respect to the transcription termination and
- 10 polyadenylation of RNA molecule.

The vector thus constructed is introduced into a host cell. Examples of host cells which may be appropriately used in the present invention are the same as described above.

In these host cells transformed by the vector of the present invention, the

15 ACS activity is more increased, fatty acids are more produced or the quantity or ratio of various fatty acids contained in the cells are changed, when compared to the host cells which are not transformed by the vector of the present invention.

Examples of the vectors available for introducing into lipid-producing fungi include, but not limited to, pDura5 (Appl. Microbiol. Biotechnol., 65, 419-425,

20 (2004)).

Any vector is available as the vector used to introduce into the yeast and not particularly limited so long as it is a vector capable of expressing the insert in the yeast cells. The vector includes, e.g., pYE22m (Biosci. Biotech. Biochem., 59, 1221-1228, 1995).

25 Promoters/terminators for regulating gene expression in host cells may be used in an optional combination as far as they function in the host cells. For example, a promoter of the histone H4.1 gene, a promoter of the glyceraldehyde-3-phosphate dehydrogenase, etc. may be used.

As selection markers used for the transformation, there may be utilized

30 auxotrophic markers (ura5, niaD), hygromycin-resistant gene, zeocin-resistant gene, genecitin-resistant gene (G418r), copper-resistant gene (CUP1) (Marin et al., Proc. Natl. Acad. Sci. USA, 81, 337 1984), cerulenin-resistant gene (fas2m, PDR4) (Junji Inokoshi, et al., Biochemistry, 64, 660, 1992; and Hussain et al., Gene, 101: 149, 1991, respectively), and the like.

35 For the transformation of host cells, generally known methods may be used. In lipid-producing fungi, the transformation may be performed, e.g., by the electroporation method (Mackenzie, D. A. et al., Appl. Environ. Microbiol., 66,

4655-4661, 2000) and the particle delivery method (the method described in JPA 2005-287403 "Method of Breeding Lipid-Producing Fungus"). On the other hand, the electroporation method, the spheroplast method (Proc. Natl. Acad. Sci. USA, 75 p1929 (1978)) and the lithium acetate method (J. Bacteriology, 153 p163 (1983)) as well as the methods described in Proc. Natl. Acad. Sci. USA, 75 p1929 (1978), Methods in yeast genetics, 2000 Edition: A Cold Spring Harbor Laboratory Course Manual, etc) may be used for the transformation of yeast. However, the method for transformation is not limited to those described above.

For general cloning techniques, reference may be made to "Sambrook & Russell, Molecular Cloning: A Laboratory Manual Vol. 3, Cold Spring Harbor Laboratory Press 2001", "Methods in Yeast Genetics, A laboratory manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)," etc.

4. Method for Producing the Lipid or Fatty Acid Composition of the Invention

In another embodiment, the present invention further provides a method for preparing a lipid or fatty acid composition which comprises using the transformant described above.

As used herein, the term "lipid" is intended to mean a simple lipid including a compound (e.g., a glyceride) which is composed of a fatty acid and an alcohol attached via an ester linkage, or its analog (e.g., a cholesterol ester), etc.; a complex lipid in which phosphoric acid, amino acid(s), saccharide(s) or the like are bound to a part of the simple lipid; or a derived lipid which is a hydrolysate of the lipid and is insoluble in water.

As used herein, the term "oil and fat" is intended to mean an ester of glycerol and a fatty acid (glyceride).

The term "fatty acid" is the same as defined above.

The method for extracting the lipid or fatty acid composition of the present invention is the same as the method for extracting fatty acids described above.

Fatty acids can be separated from the above fatty acid-containing lipids in a state of mixed fatty acids or mixed fatty acid esters by concentration and separation in a conventional manner (e.g., urea addition, separation under cooling, column chromatography, etc.).

The lipids produced by the method of the present invention include preferably unsaturated fatty acids, and more preferably, polyunsaturated fatty acids. Preferred examples of the polyunsaturated fatty acids are unsaturated fatty acids having 18 or more carbon atoms, e.g., unsaturated fatty acids having 18 to 20 carbon atoms, and include, but not limited to, oleic acid, linoleic acid, linolenic acid

(γ -linolenic acid and dihomogamma-linolenic acid, etc.), arachidonic acid, etc. Particularly preferred polyunsaturated fatty acids are linoleic acid, γ -linoleic acid, dihomogamma-linoleic acid and arachidonic acid, more preferably, linoleic acid, dihomogamma-linoleic acid and arachidonic acid, and most preferably, dihomogamma-linolenic acid and arachidonic acid.

The lipids produced by the method of the present invention and the composition of the fatty acids contained in the lipids may be confirmed by the lipid extraction method or fatty acid separation method described above, or a combination thereof.

The lipid or fatty acid composition obtained by the production method of the present invention can be provided for use in producing, e.g., food products, pharmaceuticals, industrial materials (raw materials for cosmetics, soaps, etc.), which contain oils and fats, in a conventional manner.

In a still other embodiment, the present invention provides a method for preparing food products, cosmetics, pharmaceuticals, soaps, etc. using the transformant of the present invention. The method involves the step of forming lipids or fatty acids using the transformant of the present invention.

Food products, cosmetics, pharmaceuticals, soaps, etc. containing the lipids or fatty acids produced are prepared in a conventional manner. As such, the food products, cosmetics, pharmaceuticals, soaps, etc. produced by the method of the present invention contain the lipids or fatty acids produced using the transformant of the present invention. The present invention further provides the food products, cosmetics, pharmaceuticals, soaps, etc. produced by such a method.

The form of the cosmetic (composition) or pharmaceutical (composition) of the present invention is not particularly limited and may be any form including the state of a solution, paste, gel, solid or powder. The cosmetic composition or pharmaceutical composition of the present invention may also be used as cosmetics or topical agents for the skin, including an oil, lotion, cream, emulsion, gel, shampoo, hair rinse, hair conditioner, enamel, foundation, lipstick, face powder, facial pack, ointment, perfume, powder, eau de cologne, tooth paste, soap, aerosol, cleansing foam, etc., an anti-aging skin care agent, anti-inflammatory agent for the skin, bath agent, medicated tonic, skin beauty essence, sun protectant, or protective and improving agent for skin troubles caused by injury, chapped or cracked skin, etc.

The cosmetic composition of the present invention may further be formulated appropriately with other oils and fats and/or dyes, fragrances, preservatives, surfactants, pigments, antioxidants, etc., if necessary. The formulation ratio of these materials may be appropriately determined by those skilled in the art,

depending upon purpose (for example, oils and fats may be contained in the composition in 1 to 99.99 wt %, preferably, 5 to 99.99 wt %, and more preferably, 10 to 99.95 wt%). If necessary, the pharmaceutical composition of the present invention may also contain other pharmaceutically active components (e.g., anti-inflammatory components) or aid components (e.g., lubricants or vehicle components). Examples of the other components commonly used in a cosmetic or a skin preparation for external use include an agent for acne, an agent for preventing dandruff or itching, an antiperspirant and deodorant agent, an agent for burn injury, an anti-mite and lice agent, an agent for softening keratin, an agent for xeroderma, an antiviral agent, a percutaneous absorption promoting agent, and the like.

The food product of the present invention includes a dietary supplement, health food, functional food, food product for young children, baby food, infant modified milk, premature infant modified milk, geriatric food, etc. As used herein, the food or food product is intended to mean a solid, fluid and liquid food as well as a mixture thereof, and collectively means an edible stuff.

The term dietary supplement refers to food products enriched with specific nutritional ingredients. The term health food refers to food products which are healthful or beneficial to health, and encompasses dietary supplements, natural foods, diet foods, etc. The term functional food refers to a food product for replenishing nutritional ingredients which assist body control functions and is synonymous with a food for specified health use. The term food for young children refers to a food product given to children up to about 6 years old. The term geriatric food refers to a food product treated to facilitate digestion and absorption when compared to untreated foods. The term infant modified milk refers to modified milk given to children up to about one year old. The term premature infant modified milk refers to modified milk given to premature infants until about 6 months after birth.

The form of these food products includes natural foods (treated with fats and oils) such as meat, fish and nuts; foods supplemented with fats and oils during cooking, e.g., Chinese foods, Chinese noodles, soups, etc.; foods prepared using fats and oils as heating media, e.g., tempura or deep-fried fish and vegetables, deep-fried foods, fried bean curd, Chinese fried rice, doughnuts, Japanese fried dough cookies or karinto; fat- and oil-based foods or processed foods supplemented with fats and oils during processing, e.g., butter, margarine, mayonnaise, dressing, chocolate, instant noodles, caramel, biscuits, cookies, cakes, ice cream; and foods sprayed or coated with fats and oils upon finishing, e.g., rice crackers, hard biscuits, sweet bean paste bread, etc. However, the food product is not limited to foods containing fats and oils, and other examples include agricultural foods such as bakery products,

30179-221

noodles, cooked rice, sweets (e.g., candies, chewing gums, gummies, sweet tablets, Japanese sweets), bean curd or tofu and processed products thereof; fermented foods such as Japanese rice wine or sake, medicinal liquor, sweet cooking sherry or mirin, vinegar, soy sauce and bean paste or miso, etc.; livestock food products such as
 5 yoghurt, ham, bacon, sausage, etc.; seafood products such as minced and steamed fish cake or kamaboko, deep-fried fish cake or ageten and puffy fish cake or hanpen, etc.; as well as fruit drinks, soft drinks, sports drinks, alcoholic beverages, tea, etc.

The food product of the present invention may also be in the form of pharmaceutical preparations such as capsules, etc., or in the form of a processed food
 10 such as natural liquid diets, defined formula diets and elemental diets formulated with the oil and fat of the present invention together with proteins, sugars, trace elements, vitamins, emulsifiers, aroma chemicals, etc., health drinks, enteral nutrients, and the like.

As described above, fatty acids can be efficiently produced by expressing
 15 the ACS homolog gene of the present invention in host cells.

Furthermore, the expression level of the gene can be used as an indicator to study conditions for cultivation, cultivation control, etc. for efficient fatty acid production.

20 EXAMPLES

Hereinafter, the present invention is described in more detail with reference to EXAMPLES but it should be understood that the invention is not deemed to limit the scope of the invention to these EXAMPLES.

25 [EXAMPLE 1]

Genome Analysis of *M. alpina*

The *M. alpina* 1S-4 strain was plated on 100 ml of GY2:1 medium (2% glucose and 1% yeast extract, pH 6.0) followed by shake culture at 28°C for 2 days. The mycelial cells were collected by filtration, and genomic DNA was prepared
 30 using DNeasy (QIAGEN). The nucleotide sequence of the genomic DNA described above was determined using a Roche 454 GS FLX Standard. On this occasion, nucleotide sequencing of a fragment library was performed in two runs and nucleotide sequencing of a mate paired library in three runs. The resulting nucleotide sequences were assembled to give 300 supercontigs.

35

Synthesis of cDNA and Construction of cDNA Library

The *M. alpina* strain 1S-4 was plated on 100 ml of medium (1.8% glucose,

30179-221

1% yeast extract, pH 6.0) and precultured for 3 days at 28°C. A 10 L culture vessel (Able Co., Tokyo) was charged with 5 L of medium (1.8% glucose, 1% soybean powder, 0.1% olive oil, 0.01% Adekanol, 0.3% KH₂PO₄, 0.1% Na₂SO₄, 0.05% CaCl₂·2H₂O and 0.05% MgCl₂·6H₂O, pH 6.0), and the whole amount of the

5 pre-cultured product was plated thereon, followed by aerobic spinner culture under conditions of 300 rpm, 1 vvm and 26°C for 8 days. On Days 1, 2 and 3 of the cultivation, glucose was added in an amount corresponding to 2%, 2% and 1.5%, respectively. The mycelial cells were collected at each stage on Days 1, 2, 3, 6 and 8 of the cultivation to prepare total RNA by the guanidine hydrochloride/CsCl method.

10 Using an Oligotex-dT30<Super>mRNA Purification Kit (Takara Bio Inc.), poly(A)+RNA was purified from the total RNA. A cDNA library was constructed for each stage using a ZAP-cDNA Gigapack III Gold Cloning Kit (STRATAGENE).

Search for ACS Homolog

15 Using as a query the amino acid sequences of ScFAA1 (YOR317W), ScFAA2 (YER015W), ScFAA3 (YIL009W), ScFAA4 (YMR246W), ScFAT1 (YBR041W) and ScFAT2 (YBR222C), which are ACS from yeast, a tblastn search was performed against the genome nucleotide sequence of the *M. alpina* strain 1S-4. As a result, hits were found in twelve (12) sequences. That is, hit was found on

20 supercontigs containing the sequence shown by SEQ ID NO: 5, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 20, SEQ ID NO: 25, SEQ ID NO: 30, SEQ ID NO: 35, SEQ ID NO: 40, SEQ ID NO: 45, SEQ ID NO: 50, SEQ ID NO: 55 or SEQ ID NO: 60. The genes bearing SEQ ID NO: 5, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 20, SEQ ID NO: 25, SEQ ID NO: 30, SEQ ID NO: 35, SEQ ID NO: 40, SEQ

25 ID NO: 45, SEQ ID NO: 50, SEQ ID NO: 55 and SEQ ID NO: 60 were designated respectively as MaACS-1, MaACS-2, MaACS-3, MaACS-4, MaACS-5, MaACS-6, MaACS-7, MaACS-8, MaACS-9, MaACS-10, MaACS-11 and MaACS-12.

Cloning of ACS Homolog

30 For cloning of the cDNAs corresponding to the MaACS-1~12 genes, screening of the cDNA library described above was performed. Probe labeling was performed by PCR using an ExTaqTM (Takara Bio Inc.). That is, digoxigenin (DIG)-labeled amplified DNA probes were prepared using a PCR Labeling Mix (Roche Diagnostics) instead of dNTP mix attached to ExTaq.

35 Conditions for hybridization were set as follows.

Buffer: 5x SSC, 1% SDS, 50 mM Tris-HCl (pH 7.5), 50% formaldehyde;

Temperature: 42°C (overnight);

Wash conditions: 0.2x SSC, in 0.1% SDS solution (65°C) for 20 mins. x 3

Detection was performed using a DIG Nucleic Acid Detection Kit (Roche
5 Diagnosits). Phage clones were obtained by screening and plasmids were excised
from the phage clones by in vivo excision to give the respective plasmid DNAs.

Primers for preparing the probes used for screening of the respective genes,
the number of nucleotides in CDS of the respective genes, the number of amino acids
in the amino acid sequences deduced from the nucleotide sequences of CDS, and the
10 number of exons and introns by comparison of genomic DNA sequences with CDS
sequences are given below.

(1) MaACS-1

Primer ACS-1-1F: 5'-GTCGGCTCCAAGCTTGCAATCC-3' (SEQ ID NO:
15 61)
Primer ACS-1-2R: 5'-GGACAGCTCCAGCACTGTGGTAAAG-3' (SEQ ID
NO: 62)
cDNA (SEQ ID NO: 4)
CDS (SEQ ID NO: 3): 1857 bp
20 ORF (SEQ ID NO: 1): 1854 bp
Amino acid sequence (SEQ ID NO: 2): 618 amino acids (see FIG. 1)
Number of exons: 5, number of introns: 4 (see FIG. 2)

(2) MaACS-2

25 Primer ACS-2-1F: 5'-GACCACGGGATTCCCCAAGGCTGC-3' (SEQ ID
NO: 63)
Primer ACS-2-2R: 5'-CTTGGTCGCGCTTGTTTCCTGGCCAC-3' (SEQ ID
NO: 64)
cDNA (SEQ ID NO: 9)
30 CDS (SEQ ID NO: 8): 1929 bp
ORF (SEQ ID NO: 6): 1926 bp
Amino acid sequence (SEQ ID NO: 7): 642 amino acids (see FIG. 3)
Number of exons: 8, number of introns: 7 (see FIG. 4)

35 (3) MaACS-3

Primer ACS-3-1F: 5'-TACAGCTTTGTTGCTGTCCCCATC-3' (SEQ ID
NO: 65)

Primer ACS-3-2R: 5'-GATGATGGGTGTGCTTGCAAAGATC-3' (SEQ ID NO: 66)

cDNA (SEQ ID NO: 14)

CDS (SEQ ID NO: 13): 1653 bp

5 ORF (SEQ ID NO: 11): 1650 bp

Amino acid sequence (SEQ ID NO: 12): 550 amino acids (see FIG. 5)

Number of exons: 9, number of introns: 8 (see FIG. 6)

(4) MaACS-4

10 Primer ACS-4-1F: 5'-AACCCAAAGCTGCGCCAGGCTGTCC-3' (SEQ ID NO: 67)

Primer ACS-4-2R: 5'-TTACAGCTTGGATTCCTTTTGATGG-3' (SEQ ID NO: 68)

cDNA (SEQ ID NO: 19)

15 CDS (SEQ ID NO: 18): 2067 bp

ORF (SEQ ID NO: 16): 2064 bp

Amino acid sequence (SEQ ID NO: 17): 688 amino acids (see FIG. 7)

Number of exons: 7, number of introns: 6 (see FIG. 8)

20 (5) MaACS-5

Primer ACS-5-1F: 5'-GTCGTGCCCCGATGCGGAGACGC-3' (SEQ ID NO: 69)

Primer ACS-5-2R: 5'-TCAGTGGATCCCGTTATACATCAG-3' (SEQ ID NO: 70)

25 cDNA (SEQ ID NO: 24)

CDS (SEQ ID NO: 23): 1980 bp

ORF (SEQ ID NO: 21): 1977 bp

Amino acid sequence (SEQ ID NO: 22): 659 amino acids (see FIG. 9)

Number of exons: 6, number of introns: 5 (see FIG. 10)

30

(6) MaACS-6

Primer ACS-6-1F: 5'-GCGTCCCCCTCTATGATACATTG-3' (SEQ ID NO: 71)

35 NO: 72) Primer ACS-6-2R: 5'-GTGGGATGCAGGACGGCAACATCG-3' (SEQ ID

cDNA (SEQ ID NO: 29)

CDS (SEQ ID NO: 28): 1980 bp

ORF (SEQ ID NO: 26): 1977 bp

Amino acid sequence (SEQ ID NO: 27): 659 amino acids (see FIG. 11)

Number of introns: at least 5 (see FIG. 12)

5 (7) MaACS-7

Primer ACS-7-1F: 5'-GGATGCCGAACAACAGCGCGTGG-3' (SEQ ID NO: 73)

Primer ACS-7-2R: 5'-GCACCCTCCTCAGAAACAGCCCTC-3' (SEQ ID NO: 74)

10 cDNA (SEQ ID NO: 34)

CDS (SEQ ID NO: 33): 1827 bp

ORF (SEQ ID NO: 31): 1824 bp

Amino acid sequence (SEQ ID NO: 32): 608 amino acids (see FIG. 13)

Number of exons: 5, number of introns: 4 (see FIG. 14)

15

(8) MaACS-8

Primer ACS-8-1F: 5'-CAGTCGAGTACATTGTCAACCACG-3' (SEQ ID NO: 75)

20 Primer ACS-8-2R: 5'-GCGGTTCAAGAGGCGAGGCACAGC-3' (SEQ ID NO: 76)

cDNA (SEQ ID NO: 39)

CDS (SEQ ID NO: 38): 2079 bp

ORF (SEQ ID NO: 36): 2076 bp

Amino acid sequence (SEQ ID NO: 37): 692 amino acids (see FIG. 15)

25 Number of exons: 8, number of introns: 7 (see FIG. 16)

(9) MaACS-9

Primer ACS-9-1F: 5'-GTTTCATCTTCTGCTGGCTGGGTCTC-3' (SEQ ID NO: 77)

30 Primer ACS-9-2R: 5'-GTTGCGTTGTTTACGCGGCAATCC-3' (SEQ ID NO: 78)

cDNA (SEQ ID NO: 44)

CDS (SEQ ID NO: 43): 1851 bp

ORF (SEQ ID NO: 41): 1848 bp

35 Amino acid sequence (SEQ ID NO: 42): 616 amino acids (see FIG. 17)

Number of exons: 5, number of introns: 4 (see FIG. 18)

(10) MaACS-10

Primer ACS-10-1F: 5'-ATGGAAACCTTGGTTAACGGAAAG-3' (SEQ ID NO: 79)

Primer ACS-10-2R: 5'-TCAGCAAAGATGGCCTTGGGCTGG-3' (SEQ ID NO: 80)

cDNA (SEQ ID NO: 49)

CDS (SEQ ID NO: 48): 2076 bp

ORF (SEQ ID NO: 46): 2073 bp

Amino acid sequence (SEQ ID NO: 47): 691 amino acids (see FIG. 19)

Number of exons: 8, number of introns: 7 (see FIG. 20)

(11) MaACS-11

Primer ACS-11-1F: 5'-GTCAAGGGCGAGACTCGCATCC-3' (SEQ ID NO: 81)

Primer ACS-11-2R: 5'-CGGTGACGATGGTCATGGACTGC-3' (SEQ ID NO: 82)

cDNA (SEQ ID NO: 54)

CDS (SEQ ID NO: 53): 2043 bp

ORF (SEQ ID NO: 51): 2040 bp

Amino acid sequence (SEQ ID NO: 52): 680 amino acids (see FIG. 21)

Number of exons: 3, number of introns: 2 (see FIG. 22)

(12) MaACS-12

Primer ACS-12-1F: 5'-GCGAGACCCGCATCCGCCGCTCC-3' (SEQ ID NO: 83)

Primer ACS-12-2R: 5'-GACCGTCCTCGCCCAGGGTGTCG-3' (SEQ ID NO: 84)

cDNA (SEQ ID NO: 59)

CDS (SEQ ID NO: 58): 2043 bp

ORF (SEQ ID NO: 56): 2040 bp

Amino acid sequence (SEQ ID NO: 57): 680 amino acids (see FIG. 23)

Number of exons: 3, number of introns: 2 (see FIG. 24)

Sequencing Analysis

The identity between the CDS nucleotide sequences of 12 ACS homologs from *M. alpina* is shown in TABLE 1 and the identity between the amino acid sequences is shown in TABLE 2. MaACS-11 and MaACS-12 showed high identity

of 80.2% in the nucleotide sequence and 84.3% in the amino acid sequence.

TABLE 1 Sequence identity among CDS nucleotide sequences of ACS homologs from *M. alpina*

	MaACS-1	MaACS-2	MaACS-3	MaACS-4	MaACS-5	MaACS-6	MaACS-7	MaACS-8	MaACS-9	MaACS-10	MaACS-11	MaACS-12
MaACS-1	-	51.3	42.9	45.4	44.7	46.6	46.0	45.6	69.5	44.7	46.2	45.8
MaACS-2		-	43.4	46.8	46.9	45.7	46.5	44.6	52.5	44.9	44.9	44.0
MaACS-3			-	38.0	38.2	38.9	43.7	37.5	42.8	41.5	39.0	39.1
MaACS-4				-	50.4	51.6	43.8	57.7	44.2	47.0	49.7	49.3
MaACS-5					-	70.8	44.7	53.0	44.9	46.6	48.9	47.2
MaACS-6						-	46.2	53.0	45.2	47.9	49.2	49.4
MaACS-7							-	44.2	45.9	42.3	45.0	44.6
MaACS-8								-	44.3	48.1	50.7	50.8
MaACS-9									-	42.7	46.2	47.8
MaACS-10										-	51.8	52.1
MaACS-11											-	80.2
MaACS-12												-

TABLE 2 Sequence identity among amino acid sequences of ACS homologs from *M. alpina*

	MaACS-1	MaACS-2	MaACS-3	MaACS-4	MaACS-5	MaACS-6	MaACS-7	MaACS-8	MaACS-9	MaACS-10	MaACS-11	MaACS-12
MaACS-1	-	36.6	11.8	13.9	15.1	15.8	18.0	14.8	71.9	13.5	14.6	15.0
MaACS-2		-	11.0	14.0	15.4	15.0	17.2	13.2	37.0	12.3	12.7	13.8
MaACS-3			-	21.7	21.5	20.8	13.1	21.1	10.5	17.7	18.5	17.9
MaACS-4				-	37.5	37.5	17.0	50.9	15.4	22.8	29.8	29.5
MaACS-5					-	77.9	17.0	41.2	16.4	25.2	29.1	29.8
MaACS-6						-	16.6	39.8	16.6	25.3	29.9	29.4
MaACS-7							-	15.5	17.0	15.3	16.2	16.7
MaACS-8								-	15.2	24.9	27.8	28.6
MaACS-9									-	14.1	14.5	14.7
MaACS-10										-	32.8	32.6
MaACS-11											-	84.3
MaACS-12												-

Using as query sequences the putative amino acid sequences for the CDS sequences of MaACS-1~12, BLASTp search was performed against the amino acid sequences registered in GenBank. The proteins having the amino acid sequence which matched the putative amino acid sequences of MaACS-1~12 with highest score and the identity between these proteins and the putative amino acid sequences of MaACS-1~12 are shown in TABLE 3. The identity of the putative amino acid sequences of MaACS-1~12 with the amino acid sequences of *S. cerevisiae*-derived acyl-CoA synthetases are also shown in TABLE 4.

TABLE 3 Sequence identity between the amino acid sequences of *M. alpina*-derived ACS homologs and known amino acid sequences

	identity(%)	gi	
MaACS-1	41.8	71014575	Putative protein from <i>Ustilago maydis</i>
MaACS-2	35.4	71014575	Putative protein from <i>Ustilago maydis</i>
MaACS-3	23.5	71895089	Chick ACS long-chain family member 5
MaACS-4	36.9	115487304	Putative protein from <i>Oryza sativa</i>
MaACS-5	42.5	168065128	Putative protein from <i>Physcomitrella patens</i>
MaACS-6	40.9	13516481	Long-chain acyl-CoA synthetase from <i>Arabidopsis thaliana</i>
MaACS-7	45.7	120612991	Putative protein from <i>Acidovorax avenae</i> subsp. <i>citrulli</i>
MaACS-8	40.0	13516481	Long-chain acyl-CoA synthetase from <i>Arabidopsis thaliana</i>
MaACS-9	37.8	67538044	Putative protein from <i>Aspergillus nidulans</i>
MaACS-10	33.2	171682488	Putative protein from <i>Podospira Anserina</i>
MaACS-11	48.8	169854433	Putative protein from <i>Coprinopsis atramentarius</i>
MaACS-12	45.1	156045509	Putative protein from <i>Sclerotinia sclerotiorum</i>

TABLE 4 Comparison of amino acid sequences of *M. alpina*-derived ACS homologs and amino acid sequences of *S. cerevisiae*-derived ACS

	ScFAA1	ScFAA2	ScFAA3	ScFAA4	ScFAT1	ScFAT2
MaACS-1	13.8	15.3	13.6	13.5	29.8	18.1
MaACS-2	12.5	13.6	13.4	13.5	26.3	17.5
MaACS-3	15.8	14.0	15.0	14.8	13.6	12.9
MaACS-4	26.3	28.3	23.9	24.2	14.0	16.0
MaACS-5	25.6	28.2	25.5	25.8	13.2	18.6
MaACS-6	25.3	28.4	25.8	25.5	13.0	18.1
MaACS-7	16.5	17.5	16.0	16.9	16.6	20.6
MaACS-8	23.0	28.0	21.3	22.8	12.2	14.8
MaACS-9	15.6	15.5	14.3	14.7	30.1	18.3
MaACS-10	30.8	20.6	30.6	30.6	14.0	14.2
MaACS-11	39.6	22.6	37.3	38.7	12.9	15.8
MaACS-12	41.3	22.3	39.8	39.0	14.4	16.2

5

FIG. 25 shows the alignment between MaACS from MaACS-1~12, which have relatively high amino acid sequence homology to the *S. cerevisiae*-derived FAA proteins, and the FAA proteins. FIG. 26 shows the alignment of the ACS homologs having relatively high amino acid sequence homology to *S. cerevisiae*-derived FAT proteins. The regions of the ATP-AMP motif and FACS/VLACS-FATP motif, which are important motifs for the ACS activity, are highly conserved in both groups shown in FIGS. 25 and 26.

15 Construction of Expression Vector

Vectors for expressing MaACS-1, MaACS-10, MaACS-11, MaACS-6, MaACS-8 and MaACS-9, respectively, in yeast were constructed as follows, using the expression vector pYE22m (Biosci. Biotech. Biochem., 59, 1221-1228, 1995).

The plasmid containing SEQ ID NO: 29, which was obtained by screening MaACS-6, was digested with restriction enzymes BamHI and XhoI. The resulting DNA fragment of approximately 2.1 kbp was ligated to the DNA fragment obtained

20

by digestion of vector pYE22m with restriction enzymes BamHI and Sall using a Ligation High (TOYOBO) to give plasmid pYE-ACS-6.

Using the plasmid containing cDNA of MaACS-8 as a template, PCR was performed with the primers below using ExTaq (Takara Bio Inc.). The thus amplified
5 DNA fragment was cloned by a TOPO-TA Cloning Kit (Invitrogen).

Primer EcoRI-ACS-8-F: 5'-GGATCCATGCCTTCCTTCAAAAAGTACAACC-3'
(SEQ ID NO: 85)

Primer SmaI-ACS-8-R: 5'-CCCGGGCAAAGAGTTTTCTATCTACAGCTT-3' (SEQ
10 ID NO: 86)

The nucleotide sequence of the insert was verified and the plasmid containing the correct nucleotide sequence was digested with restriction enzymes EcoRI and SmaI. Using a Ligation High (TOYOBO), the resulting DNA fragment of
15 approximately 2.1 kbp was ligated to the DNA fragment obtained by digesting vector pYE22m with restriction enzymes EcoRII and SmaI to give plasmid pYE-ACS-8.

Using the plasmid containing cDNA of MaACS-9 as a template, PCR was performed with the primers below using ExTaq (Takara Bio Inc.). The thus amplified
DNA fragment was cloned by a TOPO-TA Cloning Kit (Invitrogen).

20

Primer EcoRI-ACS-9-F: 5'-GAATTCATGGTTGCTCTCCCACTCG-3' (SEQ ID NO: 87)

Primer BamHI-ACS-9-R: 5'-GGATCCCTACTATAGCTTGGCCTTGCC-3' (SEQ ID NO: 88)

25

The nucleotide sequence of the insert was verified and the plasmid containing the correct nucleotide sequence was digested with restriction enzymes EcoRI and BamHI. Using a Ligation High (TOYOBO), the resulting DNA fragment of approximately 2.0 kbp was ligated to the DNA fragment obtained by digesting
30 vector pYE22m with restriction enzymes EcoRII and BamHI to give plasmid pYE-ACS-9.

Using the plasmid containing cDNA of MaACS-1 as a template, PCR was performed with the primers below using ExTaq (Takara Bio Inc.). The thus amplified
DNA fragment was cloned by a TOPO-TA Cloning Kit (Invitrogen).

35

Primer EcoRI-ACS-1-F: 5'-GGATCCATGTATGTCGGCTCCAAGCTTGC-3' (SEQ ID NO: 89)

Primer Sall-ACS-1-R: 5'-GTCGACTCAAAGCCTGGCTTTGCCGCTGACG-3'
(SEQ ID NO: 90)

The nucleotide sequence of the insert was verified and the plasmid
5 containing the correct nucleotide sequence was digested with restriction enzymes
EcoRI and Sall. Using a Ligation High (TOYOBO), the resulting DNA fragment of
approximately 1.9 kbp was ligated to the DNA fragment obtained by digesting vector
pYE22m with restriction enzymes EcoRI and Sall to give plasmid pYE-ACS-1.

10 Using the plasmid containing cDNA of MaACS-10 as a template, PCR was
performed with the primers below using ExTaq (Takara Bio Inc.). The thus amplified
DNA fragment was cloned by a TOPO-TA Cloning Kit (Invitrogen).

Primer ACS-10-1F: 5'- GGATCCATGGAAACCTTGGTTAACGGAAAG-3' (SEQ
ID NO: 91)

15 Primer KpnI-ACS-10-R: 5'- GGTACCTAGAACTTCTTCCACATCTCCTC-3' (SEQ
ID NO: 92)

The nucleotide sequence of the insert was verified and the plasmid
containing the correct nucleotide sequence was digested with restriction enzymes
20 EcoRI and KpnI. Using a Ligation High (TOYOBO), the resulting DNA fragment of
approximately 2.1 kbp was ligated to the DNA fragment obtained by digesting vector
pYE22m with restriction enzymes EcoRI and KpnI. Plasmid pYE-ACS-10 was
obtained by screening for the orientation that the GAPDH promoter of vector
pYE22m was located at its 5' end of CDS of MaACS-10.

25 Using the plasmid containing cDNA of MaACS-11 as a template, PCR was
performed with the primers below using ExTaq (Takara Bio Inc.). The thus amplified
DNA fragment was cloned by a TOPO-TA Cloning Kit (Invitrogen).

Primer SacI-ACS-11-F: 5'-GAGCTCATGCCAAAGTGCTTTACCGTCAACG-3'
30 (SEQ ID NO: 93)

Primer BamHI-ACS-11-R: 5'-GGATCCTTACTTGGAGCCATAGATCTGCTTG-3'
(SEQ ID NO: 94)

The nucleotide sequence of the insert was verified and the plasmid
35 containing the correct nucleotide sequence was digested with restriction enzymes
SacI and BamHI. Using a Ligation High (TOYOBO), the resulting DNA fragment of
approximately 2.0 kbp was ligated to the DNA fragment obtained by digesting vector

pYE22m with restriction enzymes SacI and BamHI to give plasmid pYE-ACS-11.

Expression in Yeast

Acquisition of Transformants

5 The yeast *S. cerevisiae* EH13-15 strain (trp1,MAT α) (Appl. Microbiol. Biotechnol., 30, 515-520, 1989) was transformed with plasmids pYE22m, pYE-MaACS-6, pYE-MaACS-8 and pYE-MaACS-9, respectively, by the lithium acetate method. The transformants were screened for the ability to grow on SC-Trp agar medium (2% agar) (per liter, 6.7 g Yeast Nitrogen Base w/o Amino Acids
10 (DIFCO), 20 g glucose, 1.3 g amino acid powders (a mixture of 1.25 g adenine sulfate, 0.6 g arginine, 3 g aspartic acid, 3 g glutamic acid, 0.6 g histidine, 1.8 g leucine, 0.9 g lysine, 0.6 g methionine, 1.5 g phenylalanine, 11.25 g serine, 0.9 g tyrosine, 4.5 g valine, 6 g threonine and 0.6 g uracil).

15 Cultivation of Yeast

One each from the transformants obtained using the respective plasmids was provided for the following cultivation experiment.

One platinum loop of the yeast was plated on 10 ml of SC-Trp and cultured with shaking for preincubation at 30°C for a day. After 1 ml of the preincubation was
20 added to the SC-Trp medium, main cultivation was performed by shake culturing at 30°C for a day.

Analysis of Fatty Acids in Mycelia

25 The yeast culture broth was centrifuged to recover the mycelial cells. After washing with 10 ml of sterile water, the mycelial cells were again centrifuged, recovered and lyophilized. The fatty acids in the mycelial cells were converted into the methyl esters by the hydrochloric acid-methanol method followed by extraction with hexane. After hexane was removed by distillation, the fatty acids were analyzed by gas chromatography.

30 The fatty acid production per medium is shown in TABLE 5. In the strains transformed by ppYE-MaACS-6, pYE-MaACS-8 or pYE-MaACS-9, the fatty acid production per medium was increased as compared to the control which was transformed by pYE22m.

35 TABLE 5 Fatty Acid Production by Transformant per Medium

	Control	MaACS-6	MaACS-8	MaACS-9
Fatty acid production (mg/L)	135	159	196	187

Expression in Arachidonic Acid-Producing Yeast

(1) Breeding of Arachidonic Acid-Producing Yeast Strains

To breed arachidonic acid-producing yeast strain (*S. cerevisiae*), the
5 following plasmids were constructed.

First, using the cDNA prepared from *M. alpina* strain 1S-4 as a template, PCR was performed with ExTaq using the primer pair of $\Delta 12$ -f and $\Delta 12$ -r, $\Delta 6$ -f and $\Delta 6$ -r, GLELO-f and GLELO-r, or $\Delta 5$ -f and $\Delta 5$ -r to amplify the $\Delta 12$ fatty acid desaturase gene (GenBank Accession No. AB020033) (hereinafter " $\Delta 12$ gene"), the
10 $\Delta 6$ fatty acid desaturase gene (GenBank Accession No. AB020032) (hereinafter " $\Delta 6$ gene"), the GLELO fatty acid elongase gene (GenBank Accession No. AB193123) (hereinafter "GLELO gene") and the $\Delta 5$ fatty acid desaturase gene (GenBank Accession No. AB188307) (hereinafter " $\Delta 5$ gene") in the *M. alpina* strain 1S-4.

15 $\Delta 12$ -f: 5'-TCTAGAATGGCACCTCCCAACACTATTG-3' (SEQ ID NO: 95)
 $\Delta 12$ -r: 5'-AAGCTTTTACTTCTTGAAAAAGACCACGTC-3' (SEQ ID NO: 96)
 $\Delta 6$ -f: 5'-TCTAGAATGGCTGCTGCTCCCAGTGTGAG-3' (SEQ ID NO: 97)
 $\Delta 6$ -r: 5'-AAGCTTTTACTGTGCCTTGCCCATCTTGG-3' (SEQ ID NO: 98)
GLELO-f: 5'-TCTAGAATGGAGTCGATTGCGCAATTCC-3' (SEQ ID NO: 99)
20 GLELO-r: 5'-GAGCTCTTACTGCAACTTCCTTGCCTTCTC-3' (SEQ ID NO: 100)
 $\Delta 5$ -f: 5'-TCTAGAATGGGTGCGGACACAGGAAAAACC-3' (SEQ ID NO: 101)
 $\Delta 5$ -r: 5'-AAGCTTTTACTCTTCCTTGGGACGAAGACC-3' (SEQ ID NO: 102)

These genes were cloned with the TOPO-TA-Cloning Kit. The clones were
25 confirmed by their nucleotide sequences. The clones containing the nucleotide sequences of the $\Delta 12$ gene, $\Delta 6$ gene, GLELO gene and $\Delta 5$ gene were designated as plasmids pCR-MA $\Delta 12$ DS (containing the nucleotide sequence of the $\Delta 12$ gene), pCR-MA $\Delta 6$ DS (containing the nucleotide sequence of the $\Delta 6$ gene), pCR-MAGLELO (containing the nucleotide sequence of the GLELO gene) and
30 pCR-MA $\Delta 5$ DS (containing the nucleotide sequence of the $\Delta 5$ gene), respectively.

On the other hand, the plasmid pURA34 (JPA 2001-120276) was digested with restriction enzyme HindIII. The resulting DNA fragment of approximately 1.2 kb was inserted into the HindIII site of the vector, which was obtained by digesting pUC18 vector (Takara Bio Inc.) with restriction enzymes EcoRI and SphI, then blunt
35 ending and self ligating said vector. The clone in which the EcoRI site of the vector was located at its 5' end of URA3 was designated as pUC-URA3. Also, the DNA fragment of approximately 2.2 kb, which was obtained by digesting YEp13 with

restriction enzymes Sall and XhoI, was inserted into the Sall site of vector pUC18. The clone in which the EcoRI site of the vector was located at its 5' end of LUE2 was designated as pUC-LEU2.

Next, the plasmid pCR-MA Δ 12DS was digested with restriction enzyme
 5 HindIII, followed by blunt ending and further digestion with restriction enzyme XbaI. The resulting DNA fragment of approximately 1.2 kbp was ligated to the DNA fragment of approximately 6.6 kbp, which was obtained by digesting vector pESC-URA (STRATAGENE) with restriction enzyme SacI, blunt ending and further digesting with restriction enzyme SpeI. Thus, the plasmid pESC-U- Δ 12 was obtained.
 10 The plasmid pCR-MA Δ 6DS was digested with restriction enzyme XbaI, followed by blunt ending and further digestion with restriction enzyme HindIII. The resulting DNA fragment of approximately 1.6 kbp was ligated to the DNA fragment of approximately 8 kbp, which was obtained by digesting the plasmid pESC-U- Δ 12 with restriction enzyme Sall, blunt ending and further digesting with restriction
 15 enzyme HindIII, thereby to give the plasmid pESC-U- Δ 12: Δ 6. This plasmid was partially digested with restriction enzyme PvuII. The resulting fragment of approximately 4.2 kb was inserted into the SmaI site of pUC-URA3 to give the plasmid pUC-URA- Δ 12: Δ 6.

Also, the plasmid pCR-MAGLELO was digested with restriction enzymes
 20 XbaI and SacI. The resulting DNA fragment of approximately 0.95 kbp was ligated to the DNA fragment of approximately 7.7 kbp, which was obtained by digesting vector pESC-LEU (STRATAGENE) with restriction enzymes XbaI and SacI. Thus, the plasmid pESC-L-GLELO was obtained. The plasmid pCR-MA Δ 5DS was digested with restriction enzyme XbaI, followed by blunt ending and further
 25 digestion with restriction enzyme HindIII. The resulting DNA fragment of approximately 1.3 kbp was ligated to the DNA fragment of approximately 8.7 kbp, which was obtained by digesting the plasmid pESC-L-GLELO with restriction enzyme ApaI, blunt ending and further digesting with restriction enzyme HindIII, thereby to give the plasmid pESC-L-GLELO: Δ 5. This plasmid was digested with
 30 restriction enzyme PvuII and the resulting fragment of approximately 3.2 kbp was inserted into the SmaI site of pUC-LEU2 to give plasmid pUC-LEU-GLELO: Δ 5. The *Saccharomyces cerevisiae* strain YPH499 (STRATAGENE) was co-transformed by the plasmid pUC-URA- Δ 12: Δ 6 and plasmid pUC-LEU-GLELO: Δ 5. The transformants were screened for the ability to grow on SC-Leu,Ura agar medium.
 35 Among the transformants thus obtained, random one strain was designated as the strain ARA3-1. By cultivating the strain in a galactose-supplemented medium, the strain became capable of expressing from the GAL1/10 promoter the Δ 12 fatty acid

desaturase gene, the $\Delta 6$ fatty acid desaturase gene, the GLELO gene and the $\Delta 5$ fatty acid desaturase gene.

(2) Transformation into Arachidonic Acid-Producing Yeast and Analysis

5 The ARA3-1 strain was transformed by plasmids pYE22m, pYE-ACS-1, pYE-ACS-10 and pYE-ACS-11, respectively. Transformants were screened for the ability to grow on SC-Trp,Leu,Ura agar medium (2 % agar) (per liter, 6.7 g Yeast Nitrogen Base w/o Amino Acids (DIFCO), 20 g glucose and 1.3 g amino acid powders (a mixture of 1.25 g adenine sulfate, 0.6 g arginine, 3 g aspartic acid, 3 g
10 glutamic acid, 0.6 g histidine, 0.9 g lysine, 0.6 g methionine, 1.5 g phenylalanine, 11.25 g serine, 0.9 g tyrosine, 4.5 g valine and 6 g of threonine). Random four strains from the respective plasmid-transfected strains were used for the subsequent cultivation.

These strains were cultivated at 30°C for a day in 10 ml of the
15 SC-Trp,Leu,Ura liquid medium described above. One milliliter of the culture was plated on 10 ml of SG-Trp,Leu,Ura liquid medium (per liter, 6.7 g Yeast Nitrogen Base w/o Amino Acids (DIFCO), 20 g galactose and 1.3 g amino acid powders (a mixture of 1.25 g adenine sulfate, 0.6 g arginine, 3 g aspartic acid, 3 g glutamic acid, 0.6 g histidine, 0.9 g lysine, 0.6 g methionine, 1.5 g phenylalanine, 11.25 g serine,
20 0.9 g tyrosine, 4.5 g valine and 6 g threonine) and then cultivated at 15°C for 6 days. The mycelial cells were collected, washed with water and then lyophilized. After the fatty acids in the dried mycelial cells were converted to the methyl esters by the hydrochloric acid-methanol method, the analysis of fatty acids was performed by gas chromatography. The ratio of each PUFA to the total fatty acids in the control strain
25 transformed by plasmid pYE22m, and in the strains transformed by each ACS homolog from *Mortierella* is shown in TABLE 6.

TABLE 6 % Ratio of PUFA in ACS homolog expression strains from *Mortierella*

	control	MaACS-1	MaACS-10	MaACS-11
18:2	7.23 $\pm$ 0.11	8.15 $\pm$ 0.29	14.87 $\pm$ 0.28	10.57 $\pm$ 0.30
18:3(n-6)	0.38 $\pm$ 0.01	0.44 $\pm$ 0.04	1.67 $\pm$ 0.10	0.92 $\pm$ 0.07
DGLA	0.41 $\pm$ 0.01	0.42 $\pm$ 0.02	0.30 $\pm$ 0.17	0.33 $\pm$ 0.03
ARA	0.42 $\pm$ 0.01	0.63 $\pm$ 0.04	0.47 $\pm$ 0.10	0.75 $\pm$ 0.10

Average $\pm$ Standard Deviation

30

As shown in TABLE 6, the ratio of fatty acids could be modified by expressing the ACS homolog from *Mortierella*. Particularly in the MaACS-11 expression strain, the ratios of arachidonic acid, linoleic acid and γ -linolenic acid were increased by about 1.8 times, about 1.5 times and about 2.4 times, respectively,

as compared to the control strain. In the MaACS-1 expression strain, the ratio of arachidonic acid was increased by about 1.5 times, as compared to the control strain. Further in the MaACS-10 expression strain, the ratios of linoleic acid and γ -linolenic acid were increased by about 2 times and about 4 times, respectively, as compared to the control strain.

[EXAMPLE 2]

Construction of Expression Vector

Expression Vector for Yeast

The vector pYE-ACS-12 for expressing MaACS-12 in yeast was constructed as follows. Using a plasmid containing the cDNA of MaACS-12 as a template, PCR was performed with the following primers using KOD-Plus-(TOYOBO).

Primer Eco-ACS-G-F: 5'-GAATTCATGACAAAGTGCCTCACCGTCG-3' (SEQ ID NO: 103)

Primer Sma-ACS-G-R: 5'-CCCGGGACTTAGGCCGTTCCATAAAGCTG-3' (SEQ ID NO: 104)

The amplified DNA fragment was cloned using a Zero Blunt TOPO PCR Cloning Kit (Invitrogen). The nucleotide sequence of the insert was verified and the plasmid containing the correct nucleotide sequence was digested with restriction enzymes EcoRI and SmaI. Using a Ligation High (TOYOBO), the resulting DNA fragment of approximately 2 kbp was ligated to the DNA fragment obtained by digesting vector pYE22m with restriction enzyme BamHI and then blunt ending with a Blunting Kit (TAKARA Bio) and further digesting with EcoRI, to give plasmid pYE-ACS-12.

Expression Vector for *M. alpina*

The vector for expressing MaACS-10 and MaACS-11 in *M. alpina* was constructed as follows.

First, pUC18 was digested with restriction enzymes EcoRI and HindIII and an adapter obtained by annealing oligo DNA MCS-for-pUC18-F2 with MCS-for-pUC18-R2 was inserted therein to construct plasmid pUC18-RF2.

MCS-for-pUC18-F2:

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

MCS-for-pUC18-R2:

5'-AGCTTGGCGCGCCGTCGACCTAGTCTAGAATAGTTTAGCGGCCGCATTC
TTATG-3' (SEQ ID NO: 106)

5 Using the genome DNA of *M. alpina* as a template, PCR was performed
with the primers Not1-GAPDHt-F and EcoR1-Asc1-GAPDHt-R using KOD-Plus-
(Toyobo). The amplified DNA fragment of about 0.5 kbp was cloned using a Zero
Blunt TOPO PCR Cloning Kit (Invitrogen). After the nucleotide sequence of the
insert was verified, the DNA fragment of about 0.9 kbp obtained by digesting with
10 restriction enzymes NotI and EcoRI was inserted into the NotI and EcoRI site of
plasmid pUC18-RF2 to construct plasmid pDG-1.

Not1-GAPDHt-F: 5'-AGCGGCCGCATAGGGGAGATCGAACC-3' (SEQ ID NO:
107)

15 EcoR1-Asc1-GAPDHt-R:

5'-AGAATTCGGCGCGCCATGCACGGGTCCTTCTCA-3' (SEQ ID NO: 108)

 Using the genome of *M. alpina* as a template, PCR was performed with the
primers URA5g-F1 and URA5g-R1 using KOD-Plus- (Toyobo). The amplified DNA
20 fragment was cloned using a Zero Blunt TOPO PCR Cloning Kit (Invitrogen). After
the nucleotide sequence of the insert was verified, the DNA fragment of about 2 kbp
obtained by digestion with Sall was inserted into the Sall site of plasmid pDG-1. The
plasmid that the 5' end of URA5 gene inserted was oriented toward the EcoRI side of
the vector was designated as the plasmid pDuraG.

25

URA5g-F1: 5'-GTCGACCATGACAAGTTTGC-3' (SEQ ID NO: 109)

URA5g-R1: 5'-GTCGACTGGAAGACGAGCACG-3' (SEQ ID NO: 110)

 Subsequently, PCR was performed with KOD-Plus- (TOYOBO) using the
30 genome of *M. alpina* as a template and the primers hisHp+URA5-F and
hisHp+MGt-F. Using an In-Fusion (registered trade name) Advantage PCR Cloning
Kit (TAKARA Bio), the amplified DNA fragment of about 1.0 kbp was ligated to the
DNA fragment of about 5.3 kbp amplified by PCR with KOD-Plus- (TOYOBO)
using pDuraG as a template and the primers pDuraSC-GAPt-F and URA5gDNA-F,
35 to give plasmid pDura-RhG.

hisHp+URA5-F:

5'-GGCAAACCTTGTCATGAAGCGAAAGAGAGATTATGAAAACAAGC-3' (SEQ ID NO: 111)

hisHp+MGt-F:

5'-CACTCCCTTTTCTTAATTGTTGAGAGAGTGTTGGGTGAGAGT-3' (SEQ ID NO: 112)

pDuraSC-GAPt-F: 5'-TAAGAAAAGGGAGTGAATCGCATAGGG-3' (SEQ ID NO: 113)

URA5gDNA-F: 5'-CATGACAAGTTTGCCAAGATGCG-3' (SEQ ID NO: 114)

10 Using the plasmid pDUra-RhG as a template, the DNA fragment of about 6.3 kbp was amplified by PCR with KOD-Plus- (TOYOBO) using the primers pDuraSC-GAPt-F and pDurahG-hisp-R.

15 pDurahG-hisp-R: 5'-ATTGTTGAGAGAGTGTTGGGTGAGAGTG-3' (SEQ ID NO: 115)

Using the plasmid containing cDNA of MaACS-10, the DNA fragment of about 2.1 kbp was amplified by PCR with KOD-Plus- (TOYOBO), using the primers below.

20

Primer ACS-10+hisp-F:

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

Primer ACS-10+MGt-R:

25 5'-CACTCCCTTTTCTTACTAGAACTTCTTCCACATCTCCTCAATATC-3' (SEQ ID NO: 117)

30 The resulting DNA fragment was ligated to the 6.3 kbp DNA fragment described above using an In-Fusion (registered trade name) Advantage PCR Cloning Kit (TAKARA BIO) to give plasmid pDUraRhG-ACS-10.

Using the plasmid containing cDNA of MaACS-11 as a template, the 2.1 kbp DNA fragment was amplified by PCR with KOD-Plus- (TOYOBO) using the primers below.

35

Primer ACS-11+MGt-R:

5'-CACTCCCTTTTCTTATTACTTGGAGCCATAGATCTGCTTGA-3' (SEQ ID

30179-221

NO: 118)

Primer ACS-11+hisp-F:

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

5

The resulting DNA fragment was ligated to the 6.3 kbp DNA fragment described above using an In-Fusion (registered trade name) Advantage PCR Cloning Kit (TAKARA BIO) to give the plasmid pDuraRhG-ACS-11.

10 Evaluation of ACS Activity

The yeast EH13-15 was transformed by plasmids pYE22m, pYE-ACS-5, pYE-ACS-8, pYE-ACS-10, pYE-ACS-11 and pYE-ACS-12, respectively, and random two transformants obtained were cultivated as follows. One platinum loop of the mycelial cells were plated on 10 ml of SC-Trp medium and cultivated with shaking for preincubation at 30°C for a day. After 1 % of the preincubation was added to 100 ml of the SD-Trp medium, main cultivation was performed by shake culturing at 28°C for a day.

The crude enzyme solution was prepared as follows. The mycelial cells were collected by centrifugation, washed with water and temporarily stored at -80°C. The mycelial cells were suspended in 5 ml of Buffer B (50 mM sodium sulfate buffer (pH 6.0), 10% glycerol and 0.5 mM PMSF). The mycelial cells were then disrupted with a French press (16 kPa, 3 times). Centrifugation was carried out at 1,500 xg at 4°C for 10 minutes and centrifuged. The supernatant obtained was used as the crude enzyme solution.

The ACS activity was determined by the following procedures based on the description of a reference literature (J.B.C., 272 (8), 1896-4903, 1997). The reaction solution contained 200 mM Tris-HCl (pH7.5), 2.5 mM ATP, 8 mM MgCl₂, 2 mM EDTA, 20 mM NaF, 0.1% TritonTMX-100, 50 µg/ml fatty acids, 50 µM CoA and 100 µl of the crude enzyme solution (suitably diluted in Buffer B), and was made 500 µl in total. The reaction was carried out at 28°C for 30 minutes. After completion of the reaction, 2.5 ml of stop solution (isopropanol : n-heptane : 1 M sulfuric acid (40: 20: 1)) was added and the mixture was thoroughly agitated. Furthermore, 2 ml of n-heptane was added thereto. After thoroughly mixing them, the mixture was centrifuged to recover the upper layer. Further 2 ml of n-heptane was added to the lower layer and treated in the same manner to recover the upper layer. The upper layers recovered were combined and evaporated to dryness using a centrifugal

concentrator. Then, 50 μ l of 0.2 mg/ml tricosanoic acid (23: 0) was added thereto as an internal standard. The fatty acids were converted into the methyl esters by the hydrochloric acid-methanol method, followed by fatty acid analysis using gas chromatography. The amount of the fatty acids, which were changed to acyl-CoA and thus distributed into the lower layer by the procedures above, was calculated from the amount of fatty acids detected. The results are shown in the table below. The ACS activity is expressed as the amount of fatty acids distributed into the lower layer by the procedures above, per weight of the protein in the crude enzyme solution. The control is the strain transformed by pYE22m and the others are the transformants in which the expression vectors of the respective genes were introduced.

TABLE 7 ACS Activity on Palmitic Acid

	MaACS-5		MaACS-10		MaACS-11		MaACS-12		Control	
	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2
mg/mg·protein	0.26	0.20	0.41	0.34	0.49	0.43	0.31	0.40	0.11	0.12

When palmitic acid was used as substrate, MaACS-5, MaACS-10, MaACS-11 and MaACS-12 showed the ACS activity of approximately 2 to 4 times the control.

TABLE 8 ACS Activity on Oleic Acid

	MaACS-10		MaACS-11		MaACS-12		Control	
	#1	#2	#1	#2	#1	#2	#1	#2
mg/mg·protein	0.25	0.20	0.25	0.16	0.16	0.18	0.09	0.11

When oleic acid was used as substrate, MaACS-10, MaACS-11 and MaACS-12 showed the ACS activity of approximately twice the control.

TABLE 9 ACS Activity on Linoleic Acid

	MaACS-5		MaACS-8		MaACS-10		MaACS-11		MaACS-12		Control	
	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2
mg/mg·protein	0.47	0.42	0.42	0.38	5.24	5.67	3.46	3.20	0.95	0.78	0.14	0.14

When linoleic acid was used as substrate, MaACS-5, MaACS-8 and MaACS-12 showed the ACS activity of several times (approximately 3, 3 and 6 times, respectively) the control, whereas MaACS-10 and MaACS-11 showed the ACS activity of several tens times (approximately 40 and 20 times, respectively) the

control.

TABLE 10 ACS Activity on γ -Linoleic Acid

	MaACS-5		MaACS-8		MaACS-10		MaACS-11		MaACS-12		Control	
	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2
mg/mg*protein	0.26	0.28	0.16	0.31	0.63	0.59	0.90	0.75	0.52	0.63	0.07	0.09

5

When γ -linoleic acid was used as substrate, all of MaACS-5, MaACS-8, MaACS-10, MaACS-11 and MaACS-12 showed the ACS activity of approximately 2 to 10 times the control.

10 TABLE 11 ACS Activity on Dihomo- γ -Linoleic Acid

	MaACS-10		MaACS-11		MaACS-12		Control	
	#1	#2	#1	#2	#1	#2	#1	#2
mg/mg*protein	4.98	4.21	2.75	2.98	2.04	1.86	0.09	0.05

When dihom- γ -linoleic acid was used as substrate, all of MaACS-10, MaACS-11 and MaACS-12 showed the ACS activity of several tens times (approximately 60 times, 40 times and 30 times, respectively) the control.

15

TABLE 12 ACS Activity on Arachidonic Acid

	MaACS-10		MaACS-11		MaACS-12		Control	
	#1	#2	#1	#2	#1	#2	#1	#2
mg/mg*protein	8.12	7.19	2.73	2.87	1.08	0.87	0.13	0.03

20 When arachidonic acid was used as substrate, MaACS-10, MaACS-11 and MaACS-12 showed the ACS activity of several tens times (approximately 90 times, 30 times and 10 times, respectively) the control.

As above, MaACS-10, MaACS-11 and MaACS-12 in particular showed a higher activity on polyunsaturated fatty acids of 20 carbon atoms such as dihom- γ -linoleic acid or arachidonic acid.

25

Arachidonic Acid Uptake Activity of ACS-Expressed Yeast

The yeast EH13-15 was transformed by plasmids pYE22m, pYE -ACS-10, pYE -ACS-11 and pYE -ACS-12, respectively, and random two transformants obtained were cultivated as follows. One platinum loop of the cells were plated on 10 ml of SC-Trp medium and cultivated with shaking for preincubation at 30°C for a

30

30179-221

day. After 100 μ l of the preincubation was added to 10 ml of the SC-Trp medium in which 50 μ g/ml of arachidonic acid was supplemented, main cultivation was performed by shake culturing at 25°C for a day. The mycelial cells were collected, lyophilized and subjected to fatty acid analysis. The ratio of arachidonic acid taken up into the mycelial cells to the added arachidonic acid was determined. The results are shown in TABLE 14. The control is the strain transformed by pYE22m and the others are the transformants in which the expression vectors of the respective genes were introduced.

10 TABLE 13 Dry Mycelial Weight

	Control		MaACS-10		MaACS-11		MaACS-12	
	#1	#2	#1	#2	#1	#2	#1	#2
%	36.63	37.81	65.86	66.64	61.53	61.35	63.64	67.06

TABLE 14 Ratio of Arachidonic Acid Taken Up into Mycelia

	Control		MaACS-10		MaACS-11		MaACS-12	
	#1	#2	#1	#2	#1	#2	#1	#2
mg/10ml	15.30	15.80	19.60	18.10	16.70	17.40	16.80	16.20

15

Acquisition of M. Alpina Transformants

Using as a host the uracil-auxotrophic strain Δ ura-3 derived from M. alpina strain 1S-4 as described in PCT International Publication Pamphlet WO 2005/019437 entitled "Method of Breeding Lipid-Producing Fungus"), transformation was performed by the particle delivery method using the plasmids pDUraRhG-ACS-10 and pDUraRhG-ACS-11, respectively. For screening of the transformants, SC agar medium was used (0.5% Yeast Nitrogen Base w/o Amino Acids and Ammonium Sulfate (Difco), 0.17% ammonium sulfate, 2% glucose, 0.002% adenine, 0.003% tyrosine, 0.0001% methionine, 0.0002% arginine, 0.0002% histidine, 0.0004% lysine, 0.0004% tryptophan, 0.0005% threonine, 0.0006% isoleucine, 0.0006% leucine, 0.0006% phenylalanine, and 2% agar).

Evaluation of M. Alpina Transformants

The transformants obtained were plated on 4 ml of GY medium and cultured with shaking at 28°C for 2 days. The mycelial cells were collected by filtration, and RNA was extracted with an RNeasy Plant Kit (QIAGEN). A SuperScriptTM First Strand System for RT-PCR (Invitrogen) was used to synthesize cDNA. To confirm expression of the respective genes from the introduced constructs, RT-PCR was

performed with the following primer pairs.

ACS10-RT1: 5'-GTCCCGAATGGTTCCT-3' (SEQ ID NO: 120)

ACS10-RT2: 5'-AGCGGTTTTCTACTTGC-3' (SEQ ID NO: 121)

5 ACS11-RT1: 5'-AACTACAACCGCGTCG-3' (SEQ ID NO: 122)

ACS11-RT2: 5'-CGGCATAAACGCAGAT-3' (SEQ ID NO: 123)

10 In the transformants that overexpression was confirmed, one transformant each was plated on 10 ml of GY medium (2% glucose and 1% yeast extract) and cultured with shaking at 28°C at 300 rpm for 3 days. The whole volume of the culture was transferred to 500 ml of GY medium (2 L Sakaguchi flask) and shake cultured at 28°C and 120 rpm. Three, seven, ten and twelve days after this day, 5 ml each and 10 ml each were taken and filtered. After the mycelial cells were dried at 120°C, fatty acids were converted into the methyl esters by the hydrochloric acid-methanol method and analyzed by gas chromatography. The fatty acid production and the amount of arachidonic acid produced, per dried mycelial cells were monitored with the passage of time. The transformant host strain Δ ura-3 was used as control. The results are shown in FIG. 27 (MaACS-10) and FIG. 28 (MaACS-11).

20 As shown in FIGS. 27 and 28, when MaACS-10 and MaACS-11 were overexpressed in *M. alpina*, both the amount of fatty acids and the amount of arachidonic acid per mycelia were increased as compared to the control.

INDUSTRIAL APPLICABILITY

25 The polynucleotide of the present invention is expressed in an appropriate host cell to efficiently produce fatty acids, in particular, polyunsaturated fatty acids. The fatty acids produced in host cells according to the present invention can be used to produce fatty acid compositions, food products, cosmetics, pharmaceuticals, soaps, etc.

30

[Sequence Listing]

SEQUENCE LISTING IN ELECTRONIC FORM

In accordance with Section 111(1) of the Patent Rules, this description contains a sequence listing in electronic form in ASCII text format (file: 30179-221 Seq 18-07-12 v1.txt).

A copy of the sequence listing in electronic form is available from the Canadian Intellectual Property Office.

The sequences in the sequence listing in electronic form are reproduced in the following table.

SEQUENCE TABLE

<110> SUNTORY HOLDINGS LIMITED

<120> Acyl-CoA Synthetase Homologues and Use of the Same

<130> PCT11-0003

<150> JP 2010-19967

<151> 2010-02-01

<160> 123

<170> PatentIn version 3.4

<210> 1

<211> 1854

<212> DNA

<213> Mortierella alpina

<400> 1

atggatgctg	tccctgcagt	tgctgctgcg	gccatccccg	cagccatgta	tgctggctcc	60
aagcttgcaa	tcccccgtag	tgtaagttta	gctaaaggcc	tagtcagtgc	caagctaggt	120
tacaggtcct	acgagaagaa	cgactcgatc	aatatctctt	atcgttttga	agagacctgt	180
aagaagcacc	ctcatcgaga	agctttgggtg	tttgaaggca	aatcgtagac	cttccaggac	240
atccagcgag	aatcgaatag	ggtgggacac	tggtgtgtgt	ccaaaggcgt	caagcgagga	300
gagatcggtg	cgctcttcat	gcaaaataag	ccagagtttc	tcttcttctg	gcttggactc	360
aacaagatcg	gcgctacggg	agcattcatc	aacacgaacc	tctcgggcaa	acctctgacg	420
cactcattgc	gtaccgcgac	agcatccatt	ctgattatgg	atgcggaact	gccgacgcc	480
atttatagtg	tcctcgatga	agtccttgag	atgggatatc	agatatattc	ctacggagga	540
tcccagcaac	acgcctttgc	tacacaagtt	gaactttctc	aaatctcgga	tgcggccttg	600
cccaagagtc	tgcaaggaa	aaccactgca	aatgatattg	ccatgttgat	ttacacctcc	660
ggaacgacgg	gtttgcccaa	agctggacgg	ttctcccatg	ctcgagccaa	cgttgccgca	720
cttttctgga	cgtctttcta	ccacttcagc	gaaaaagacc	gcctgtacat	cgcttgcct	780
ctttaccaca	gtgctggagc	tgctcttgga	atatgtgtgg	cctgggtcac	cggtgctacg	840
gtggtcctgg	cgcgcaagtt	ttcaactact	tccttctggg	acgaatgcag	ggccaacaag	900
gtcacctgta	tccagtatat	tggagaaatc	tgccgatact	tactgaatgc	tcctccttct	960
cccttgga	agacacacac	gatccgaatg	gcgcatggca	acggcatgcg	tccggatgta	1020
tggaacagat	tcagagatcg	tttcggcatc	cctttgatcg	gagaatggta	tgcaagcact	1080
gagggcaccg	gaatcttgac	aaactataac	acaggacceca	atggcgctgg	tgcgatagga	1140
tacagaggct	ccttggccag	aactgtcgat	aagggtctga	agattgcgaa	gttcgacatc	1200

```

caaaccgagg aacttattcg tgacaaaaaat ggtcgatgca ttgagtgtgt cgcagatgag 1260
cccggcgagc tcttgacaat gattgattca agtgatccca ctgcgcgttt ccaaggggtac 1320
cataaaaaatg caggtgcaaa ctccaagaaa gtcgtccagg atgcattcag tgttggcgac 1380
caatactttc gtactggtga catccttcgt cgcgacgctg atggctatct ctattttggc 1440
gatcgtgttg gagatacttt ccgctggaaa tctgaaaacg tgtcaactgc ggaggtttct 1500
gaggtgctct cagcataccc ggactgcacg gaggtcaacg tttatggcgt tcaagtcctt 1560
ggacacgacg gccgcgcagg catggctgcc attgtctcca aggacaccat gaactgggat 1620
agtttcgcca agtttgact caaaaatctg ccgaagtact ctgtgcgat tttcatcgcg 1680
aaggtcccag agatggagat tacgggaacg ttcaagcaac gaaaggttga actggtgaac 1740
gagggcatgg acccgagcaa gatcaaagac gaaatgctgt ggttggatgg gcactcctac 1800
cggcccttca aagaggcgga gcatactaga gtcgtcagcg gcaaagccag gctt 1854

```

<210> 2

<211> 618

<212> PRT

<213> Mortierella alpina

<400> 2

```

Met Asp Ala Val Pro Ala Val Ala Ala Ala Ala Ile Pro Ala Ala Met
1          5          10          15
Tyr Val Gly Ser Lys Leu Ala Ile Pro Arg Asp Val Lys Leu Ala Lys
          20          25          30
Gly Leu Val Ser Ala Lys Leu Gly Tyr Arg Ser Tyr Glu Lys Asn Asp
          35          40          45
Ser Ile Asn Ile Ser Tyr Arg Phe Glu Glu Thr Cys Lys Lys His Pro
          50          55          60
His Arg Glu Ala Leu Val Phe Glu Gly Lys Ser Tyr Thr Phe Gln Asp
65          70          75          80
Ile Gln Arg Glu Ser Asn Arg Val Gly His Trp Leu Leu Ser Lys Gly
          85          90          95
Val Lys Arg Gly Glu Ile Val Ser Leu Phe Met Gln Asn Lys Pro Glu
          100         105         110
Phe Leu Phe Phe Trp Leu Gly Leu Asn Lys Ile Gly Ala Thr Gly Ala
          115         120         125
Phe Ile Asn Thr Asn Leu Ser Gly Lys Pro Leu Thr His Ser Leu Arg
          130         135         140
Thr Ala Thr Ala Ser Ile Leu Ile Met Asp Ala Glu Leu Pro Thr Pro
145         150         155         160
Ile Tyr Ser Val Leu Asp Glu Val Leu Glu Met Gly Tyr Gln Ile Tyr
          165         170         175
Ser Tyr Gly Gly Ser Gln Gln His Ala Phe Ala Thr Gln Val Glu Leu
          180         185         190
Ser Gln Ile Ser Asp Ala Ala Leu Pro Lys Ser Leu Arg Arg Lys Thr
          195         200         205
Thr Ala Asn Asp Ile Ala Met Leu Ile Tyr Thr Ser Gly Thr Thr Gly
          210         215         220
Leu Pro Lys Ala Gly Arg Phe Ser His Ala Arg Ala Asn Val Ala Ala
225         230         235         240
Leu Phe Trp Thr Ser Phe Tyr His Phe Ser Glu Lys Asp Arg Leu Tyr
          245         250         255
Ile Ala Leu Pro Leu Tyr His Ser Ala Gly Ala Val Leu Gly Ile Cys
          260         265         270
Val Ala Trp Val Thr Gly Ala Thr Val Val Leu Ala Arg Lys Phe Ser
          275         280         285
Thr Thr Ser Phe Trp Asp Glu Cys Arg Ala Asn Lys Val Thr Val Ile
290         295         300

```

Gln Tyr Ile Gly Glu Ile Cys Arg Tyr Leu Leu Asn Ala Pro Pro Ser
 305 310 315 320
 Pro Leu Asp Lys Thr His Thr Ile Arg Met Ala His Gly Asn Gly Met
 325 330 335
 Arg Pro Asp Val Trp Asn Arg Phe Arg Asp Arg Phe Gly Ile Pro Leu
 340 345 350
 Ile Gly Glu Trp Tyr Ala Ser Thr Glu Gly Thr Gly Ile Leu Thr Asn
 355 360 365
 Tyr Asn Thr Gly Pro Asn Gly Ala Gly Ala Ile Gly Tyr Arg Gly Ser
 370 375 380
 Leu Ala Arg Thr Val Asp Lys Gly Leu Lys Ile Ala Lys Phe Asp Ile
 385 390 395 400
 Gln Thr Glu Glu Leu Ile Arg Asp Lys Asn Gly Arg Cys Ile Glu Cys
 405 410 415
 Val Ala Asp Glu Pro Gly Glu Leu Leu Thr Met Ile Asp Ser Ser Asp
 420 425 430
 Pro Thr Arg Ala Phe Gln Gly Tyr His Lys Asn Ala Gly Ala Asn Ser
 435 440 445
 Lys Lys Val Val Gln Asp Ala Phe Ser Val Gly Asp Gln Tyr Phe Arg
 450 455 460
 Thr Gly Asp Ile Leu Arg Arg Asp Ala Asp Gly Tyr Phe Tyr Phe Gly
 465 470 475 480
 Asp Arg Val Gly Asp Thr Phe Arg Trp Lys Ser Glu Asn Val Ser Thr
 485 490 495
 Ala Glu Val Ser Glu Val Leu Ser Ala Tyr Pro Asp Cys Ile Glu Val
 500 505 510
 Asn Val Tyr Gly Val Gln Val Pro Gly His Asp Gly Arg Ala Gly Met
 515 520 525
 Ala Ala Ile Val Ser Lys Asp Thr Met Asn Trp Asp Ser Phe Ala Lys
 530 535 540
 Phe Ala Leu Lys Asn Leu Pro Lys Tyr Ser Val Pro Ile Phe Ile Arg
 545 550 555 560
 Lys Val Pro Glu Met Glu Ile Thr Gly Thr Phe Lys Gln Arg Lys Val
 565 570 575
 Glu Leu Val Asn Glu Gly Met Asp Pro Ser Lys Ile Lys Asp Glu Met
 580 585 590
 Leu Trp Leu Asp Gly His Ser Tyr Arg Pro Phe Lys Glu Ala Glu His
 595 600 605
 Thr Arg Val Val Ser Gly Lys Ala Arg Leu
 610 615

<210> 3
 <211> 1857
 <212> DNA
 <213> Mortierella alpina

<400> 3
 atggatgctg tccctgcagt tgctgctgcg gccatccccg cagccatgta tgtcggctcc 60
 aagcttgcaa tccccgtga tgtcaagtta gctaaaggcc tagtcagtgc caagctaggt 120
 tacaggctct acgagaagaa cgactcgatc aatatctctt atcgttttga agagacctgt 180
 aagaagcacc ctcatcgcca agctttggtg tttgaaggca aatcgtaacac cttccaggac 240
 atccagcgag aatcgaatag ggtgggacac tggctgttgt ccaaaggcgt caagcgagga 300
 gagatcgtgt cgctcttcat gcaaaataag ccagagtttc tcttcttctg gcttggactc 360
 aacaagatcg gcgctacggg agcattcatc aacacgaacc tctcgggcaa acctctgacg 420
 cactcattgc gtaccgcgac agcatccatt ctgattatgg atgcggaact gccgacgcc 480
 atttatagtg tcctcgatga agtccttgag atgggatatc agatatattc ctacggagga 540

tcccagcaac	acgcctttgc	tacacaagtt	gaactttctc	aaatctcgga	tgcggccttg	600
cccaagagtc	tgcgaaggaa	aaccactgca	aatgatattg	ccatgttgat	ttacacctcc	660
ggaacgacgg	gtttgcccaa	agctggacgg	ttctcccatg	ctcgagccaa	cgttgccgca	720
cttttctgga	cgtctttcta	ccacttcagc	gaaaaagacc	gcctgtacat	cgccttgcc	780
ctttaccaca	gtgctggagc	tgtccttgga	atatgtgtgg	cctgggtcac	cgggtgctacg	840
gtggctcctgg	cgcgcaagtt	ttcaactact	tccttctggg	acgaatgcag	ggccaacaag	900
gtcaccgtga	tccagtatat	tggagaaatc	tgccgatact	tactgaatgc	tcctccttct	960
cccttggaca	agacacacac	gatccgaatg	gcgcattggc	acggcatgcg	tccggatgta	1020
tggaaacagat	tcagagatcg	tttcggcatc	cctttgatcg	gagaatggta	tgcaagcact	1080
gagggcaccg	gaatcttgac	aaactataac	acaggaccca	atggcgctgg	tgcgatagga	1140
tacagaggct	ccttggccag	aactgtcgat	aagggctctga	agattgcgaa	gttcgacatc	1200
caaaccgagg	aacttattcg	tgacaaaaat	ggtcgatgca	ttgagtgtgt	cgcagatgag	1260
cccggcgagc	tcttgacaat	gattgattca	agtgatccca	ctcgcgcttt	ccaagggtag	1320
cataaaaatg	caggtgcaaa	ctccaagaaa	gtcgtccagg	atgcattcag	tggtggcgac	1380
caatactttc	gtactgggtga	catccttcgt	cgcgacgctg	atggctattt	ctattttggc	1440
gatcgtgttg	gagatacttt	ccgctggaaa	tctgaaaacg	tgtcaactgc	ggagggtttct	1500
gaggtgctct	cagcataccc	ggactgcac	gaggtcaacg	tttatggcgt	tcaagtccct	1560
ggacacgacg	gccgcgcagg	catggctgcc	attgtctcca	aggacaccat	gaactgggat	1620
agtttcgcca	agtttgcact	caaaaatctg	ccgaagtact	ctgtgccgat	tttcatccgc	1680
aaggtcccag	agatggagat	tacgggaacg	ttcaagcaac	gaaaggttga	actggtgaac	1740
gagggcatgg	acccgagcaa	gatcaaagac	gaaatgctgt	ggttggatgg	gcactcctac	1800
cggcccttca	aagaggcgga	gcatactaga	gtcgtcagcg	gcaaagccag	gctttga	1857

<210> 4

<211> 2066

<212> DNA

<213> *Mortierella alpina*

<400> 4

gctctttttt	gttctgttct	ttgccacccc	actctgctgt	gcctctccac	tccccctgcc	60
cgtcgaacgt	tcttctgtca	ctttgcacag	cagtatctcc	tctgatctcg	cttgrttata	120
ttccctctaa	tctcgtttgg	ttatatccc	tctgatctcg	ctcggttata	ttcttcagat	180
atggatgctg	tccctgcagt	tgtctgctgcg	gccatccccg	cagccatgta	tgtcggctcc	240
aagcttgcaa	tcccccgatg	tgtcaagtta	gctaaaggcc	tagtcagtgc	caagctaggt	300
tacaggctct	acgagaagaa	cgaactcgatc	aatatctctt	atcgttttga	agagacctgt	360
aagaagcacc	ctcatcgcg	agctttgggtg	tttgaaggca	aatcgtacac	cttccaggac	420
atccagcgag	aatcgaatag	ggtgggacac	tggctgttgt	ccaaaggcgt	caagcgagga	480
gagatcgtgt	cgtctttcat	gcaaaaataag	ccagagtttc	tcttcttctg	gcttggactc	540
aacaagatcg	gcgctacggg	agcattcatt	aacacgaacc	tctcgggcaa	acctctgacg	600
cactcattgc	gtaccgcgac	agcatccatt	ctgattatgg	atgcggaact	gccgacgcc	660
atttatagtg	tcctcgatga	agtccttgag	atgggatata	agatatattc	ctacggagga	720
tcccagcaac	acgcctttgc	tacacaagtt	gaactttctc	aaatctcgga	tgcggccttg	780
cccaagagtc	tgcgaaggaa	aaccactgca	aatgatattg	ccatgttgat	ttacacctcc	840
ggaacgacgg	gtttgcccaa	agctggacgg	ttctcccatg	ctcgagccaa	cgttgccgca	900
cttttctgga	cgtctttcta	ccacttcagc	gaaaaagacc	gcctgtacat	cgccttgcc	960
ctttaccaca	gtgctggagc	tgtccttgga	atatgtgtgg	cctgggtcac	cgggtgctacg	1020
gtggctcctgg	cgcgcaagtt	ttcaactact	tccttctggg	acgaatgcag	ggccaacaag	1080
gtcaccgtga	tccagtatat	tggagaaatc	tgccgatact	tactgaatgc	tcctccttct	1140
cccttggaca	agacacacac	gatccgaatg	gcgcattggc	acggcatgcg	tccggatgta	1200
tggaaacagat	tcagagatcg	tttcggcatc	cctttgatcg	gagaatggta	tgcaagcact	1260
gagggcaccg	gaatcttgac	aaactataac	acaggaccca	atggcgctgg	tgcgatagga	1320
tacagaggct	ccttggccag	aactgtcgat	aagggctctga	agattgcgaa	gttcgacatc	1380
caaaccgagg	aacttattcg	tgacaaaaat	ggtcgatgca	ttgagtgtgt	cgcagatgag	1440
cccggcgagc	tcttgacaat	gattgattca	agtgatccca	ctcgcgcttt	ccaagggtag	1500
cataaaaatg	caggtgcaaa	ctccaagaaa	gtcgtccagg	atgcattcag	tggtggcgac	1560
caatactttc	gtactgggtga	catccttcgt	cgcgacgctg	atggctattt	ctattttggc	1620

gatcgtgttg	gagatacttt	ccgctggaaa	tctgaaaacg	tgtcaactgc	ggagggtttct	1680
gaggtgctct	cagcataccc	ggactgcac	gaggtcaacg	tttatggcgt	tcaagtcct	1740
ggacacgacg	gccgcgcagg	catggctgcc	attgtctcca	aggacaccat	gaactgggat	1800
agtttcgcca	agtttgcact	caaaaatctg	ccgaagtact	ctgtgccgat	tttcatccgc	1860
aaggtcccag	agatggagat	tacgggaacg	ttcaagcaac	gaaagggttg	actgggtgaac	1920
gagggcatgg	acccgagcaa	gatcaaagac	gaaatgctgt	ggttggatgg	gcactcctac	1980
cggcccttca	aagaggcgga	gcatactaga	gtcgtcagcg	gcaaagccag	gctttgacga	2040
ataaaattat	ttcgttttgt	ccgttg				2066

<210> 5

<211> 2278

<212> DNA

<213> *Mortierella alpina*

<400> 5

atggatgctg	tccctgcagt	tgtgtgtg	gccatccccg	cagccatgta	tgtcggctcc	60
aagcttgcaa	tccccgtga	tgtcaagtta	gctaaaggcc	tagtcagtgc	caagctaggt	120
tacaggctct	acgagaagaa	cgactcgatc	aatatctctt	atcgttttga	agagacctgt	180
aagaagcacc	ctcatcgcca	agctttgggtg	tttgaaggca	aatcgtaac	cttccaggac	240
atccagcgag	gtaacaagaa	aaacaattgt	ccaaagtga	agtcgaacgc	atccgaaatt	300
tttactcaag	ataaattgga	tcatcgac	acacccctcc	gtcgatttca	ctaacccttg	360
acttggaatg	tagaatcgaa	taggggtgga	cactggctgt	tgtccaaagg	cgtcaagcga	420
ggagagatcg	tgtcgctctt	catgcaaaat	aagccagagt	ttctcttctt	ctggcttgga	480
ctcaacaaga	tgggcgctac	gggagcattc	atcaacacga	acctctcggg	caaacctctg	540
acgcactcat	tgcgtaccgc	gacagcatcc	attctgatta	tggatgcgga	actgccgacg	600
ccattttata	gtgtcctcga	tgaagtcctt	gagatgggat	atcagatata	ttcctacgga	660
ggatcccagc	aacacgcctt	tgtacacaa	gttgaacttt	ctcaaactct	ggatgcggcc	720
ttgcccaaga	gtctgcgaag	gaaaaccact	gcaaatgata	ttgccatgtt	gatttacacc	780
tccggaacga	cgggtttgcc	caaagctgga	cggttctccc	atgctcgagc	caacggtagg	840
attatacccc	ccctcctccc	ccccccctt	ttttttcatt	tgtgtgaag	ttattagctg	900
ttccactage	atattgactc	atattcacgt	tcctttttaca	cgtcgggatc	cagttgccgc	960
actttttctg	acgtctttct	accacttcag	cgaaaaagac	cgctgtaca	tgccttgcc	1020
tctttaccac	agtgtctggag	ctgtccttgg	aatatgtgtg	gcctgggtca	cgggtgctac	1080
ggtggtcctg	gcgcgcaagt	tttcaactac	ttccttctgg	gacgaatgca	gggccaacaa	1140
ggtcaccgtg	atccagtata	ttggagaaat	ctgccgatac	ttactgaatg	ctcctccttc	1200
tcccttgga	aagacacaca	cgatccgaat	ggcgcattgg	aacggcatgc	gtccggatgt	1260
atggaacaga	ttcagagatc	gtttcggcat	ccctttgatc	ggagaatgg	atgcaagcac	1320
tgagggcacc	ggaatcttga	caaactataa	cacaggaccc	aatggcgctg	gtgcgatagg	1380
atacagaggc	tccttgga	gaactgtcga	taagggtctg	aagattgcga	agttcgacat	1440
ccaaaccgag	gaacttattc	gtgacaaaaa	tggtcgatgc	attgaggtaa	agttgacagt	1500
attaagttga	acatatccca	cagtaccctt	tgtcttcggg	gtccaaaata	ctgactactt	1560
gcttgatgcc	ccttcaagtg	tgtcgcagat	gagcccggcg	agctcttgac	aatgattgat	1620
tcaagtgate	ccactcgcgc	tttccaaagg	taccataaaa	atgcagggtgc	aaactccaag	1680
aaagtcgtcc	aggatgcatt	cagtgttggt	aggtgtaatc	ttcaccttgc	gtgccttcaa	1740
cagcaaactt	tgtgatccgc	gctgactcga	tgaccctgtc	ctgcctacat	ttctaggcga	1800
ccaatacttt	cgtactgggtg	acatccttcg	tgcgcacgct	gatggctatt	tctattttgg	1860
cgatcgtgtt	ggagataact	tccgctggaa	atctgaaaac	gtgtcaactg	cggagggttc	1920
tgagggtgct	tcagcatacc	cggactgcat	cgagggtcaac	gtttatggcg	ttcaagtccc	1980
tggacacgac	ggccgcgcag	gcatggctgc	cattgtctcc	aaggacacca	tgaactggga	2040
tagtttcgcc	aagtttgcac	tcaaaaatct	gccgaagtac	tctgtgccga	ttttcatccg	2100
caaggtccca	gagatggaga	ttacgggaac	gttcaagcaa	cgaaagggtg	aactggtgaa	2160
cgagggcattg	gacccgagca	agatcaaaga	cgaaatgctg	tggttggatg	ggcactccta	2220
ccggcccttc	aaagaggcgg	agcatactag	agtcgtcagc	ggcaaagcca	ggctttga	2278

<210> 6
 <211> 1926
 <212> DNA
 <213> Mortierella alpina

<400> 6
 atggcaagta ccaaactcact aaggacttgg cgcttggttcg cgttgggtctc catgcacgca 60
 aaagacttga ggccttggtc gagaatcggg actgcagtct ttacaacagg ttcgaggaac 120
 aatgccagat cgggcctttc tctggtgccc ttgtttttga gaacacgtct tacacctgga 180
 gagacttggg gctggcgctc aacaggatgg ccatttggtt tggtgctcaa ggaatccaaa 240
 aaaaaggagc gtgtggcgat gatgatgcat aactcgcttc tgttcattat cacctggctg 300
 gcaatgctca agatcatggt tgtacctgct tttatcaata accagattgc aggacctgtt 360
 ctgggttcatt ctcttaaagt ggccgacgcc aagtttctct tgttcgatta cgagttggca 420
 cctgtcatcc aaaagtcgct caatgagatc aaggacatgg gttacaatct ctacactgtc 480
 acaccaagg atcaagttct aggtcaactt tacgccaatc tgcccgaggc tgctcgtcag 540
 gtgttggaat aggtccttc attctttggt tatgtcgaat ggcagaacct cagtaccgaa 600
 ggtttctcga acgagagtcg tcaggaggtg gtgatctccg accccgcagc cttgatttac 660
 accagcggga ccacgggatt cccaaggct gctatcatgg accatggacg ttgcaacttg 720
 gcttcgatct cttatggcac tctatgcggc atcaaaccag agaacaagg tttacatcaca 780
 ttgccgctct atcattctgc tggagccatc attggtctgg gccagagctt caccagcgga 840
 tgcaccattg tgctggcgcg aaagttctcc gtgacaaagt tttggcgtga ttgcgttgag 900
 tacgacgtaa ctcatcttca gtacattggc gaactctgcc gctacctct aaatgcccc 960
 gaaagtccac tggacaaaag gcataagggt cggtatggcg ttggcaacgg aatgcgccc 1020
 gatgtttggg caaagtttca ggaacgattc aatatcccca ttattggtga gtactacgc 1080
 atgagcgaag gaacatcgct gcttttgaat gtggccagga acaagcgcg ccaagggtgcg 1140
 gtgggattcc gtggccccgt cgtgagggcc ttgacgcctc ccgttcaact ggtcaagggtg 1200
 gactttgaca cggaggagct gatccgcgat aagaagacgg gactttgcgt cctatgccag 1260
 cctggtgaga ttggagaact ggtcacgcta gccgacaaca agacgactgg cgcacgctat 1320
 gctgggtatt tcaatcagcc agaggtttcg aaggcaaggc tgggtccagaa cgtggtagt 1380
 aaggacgaca tctacttccg gacgggtgac ctcttgact ccaaggacca gtactggtac 1440
 tttgctgac gcgcaggaga cacgtaccgg tggaaaggag agaacgtgtc gacagccag 1500
 attgcagaca ctatcggccg tgttgagggc gtggctagtt gtactgttta tggcgtatcg 1560
 gtcccgggca tggatggacg cgcgggcatg gctgctttgg tgctcaagaa ctcgattgtg 1620
 cagatggcag gtggaagcca ggcaaagttc catgtggatg aggtgcgct gaacgcgttt 1680
 ttgcgtgact tgagcaagga tgtggtcaaa aaactgccgg cgtatgcgat tcctcggttc 1740
 ttgcgcattg cagagcagga actggagacg acgggcacgt tcaagaacaa gaaggtggag 1800
 ctgaagaagg aagggttcga cctcggttaag gtcaaggagc ggctgtactg gtggacaccc 1860
 aagggtgaat atgccccctt tggcgtggcg gagaacgagc agatcctcgc aggacgcgct 1920
 cgtctt 1926

<210> 7
 <211> 642
 <212> PRT
 <213> Mortierella alpina

<400> 7
 Met Ala Ser Thr Lys Ser Leu Arg Thr Trp Arg Leu Phe Ala Leu Val
 1 5 10 15
 Ser Met His Ala Lys Asp Leu Arg Pro Trp Ser Arg Ile Gly Thr Ala
 20 25 30
 Val Phe Thr Thr Gly Ser Arg Asn Asn Ala Arg Ser Gly Leu Ser Leu
 35 40 45
 Leu Pro Leu Phe Leu Arg Thr Arg Leu Thr Pro Gly Glu Thr Trp Ser
 50 55 60
 Trp Arg Pro Thr Gly Trp Pro Ile Gly Leu Leu Leu Lys Glu Ser Lys
 65 70 75 80

Lys	Lys	Glu	Arg	Val	Ala	Met	Met	Met	His	Asn	Ser	Pro	Leu	Phe	Ile	85	90	95
Ile	Thr	Trp	Leu	Ala	Met	Leu	Lys	Ile	Met	Val	Val	Pro	Ala	Phe	Ile	100	105	110
Asn	Asn	Gln	Ile	Ala	Gly	Pro	Val	Leu	Val	His	Ser	Leu	Lys	Val	Ala	115	120	125
Asp	Ala	Lys	Phe	Leu	Leu	Phe	Asp	Tyr	Glu	Leu	Ala	Pro	Val	Ile	Gln	130	135	140
Lys	Ser	Leu	Asn	Glu	Ile	Lys	Asp	Met	Gly	Tyr	Asn	Leu	Tyr	Thr	Val	145	150	155
Thr	Pro	Lys	Asp	Gln	Val	Leu	Gly	Gln	Leu	Tyr	Ala	Asn	Leu	Pro	Glu	165	170	175
Ala	Ala	Arg	Gln	Val	Leu	Asp	Glu	Ala	Pro	Ser	Phe	Phe	Gly	Tyr	Val	180	185	190
Glu	Trp	Gln	Asn	Leu	Ser	Thr	Glu	Gly	Phe	Ser	Asn	Glu	Ser	Arg	Gln	195	200	205
Glu	Val	Val	Ile	Ser	Asp	Pro	Ala	Ala	Leu	Ile	Tyr	Thr	Ser	Gly	Thr	210	215	220
Thr	Gly	Phe	Pro	Lys	Ala	Ala	Ile	Met	Asp	His	Gly	Arg	Cys	Asn	Leu	225	230	235
Ala	Ser	Ile	Ser	Tyr	Gly	Thr	Leu	Cys	Gly	Ile	Lys	Pro	Glu	Asn	Lys	245	250	255
Val	Tyr	Ile	Thr	Leu	Pro	Leu	Tyr	His	Ser	Ala	Gly	Ala	Ile	Ile	Gly	260	265	270
Leu	Gly	Gln	Ser	Phe	Thr	Ser	Gly	Cys	Thr	Ile	Val	Leu	Ala	Arg	Lys	275	280	285
Phe	Ser	Val	Thr	Lys	Phe	Trp	Arg	Asp	Cys	Val	Glu	Tyr	Asp	Val	Thr	290	295	300
His	Phe	Gln	Tyr	Ile	Gly	Glu	Leu	Cys	Arg	Tyr	Leu	Leu	Asn	Ala	Pro	305	310	315
Glu	Ser	Pro	Leu	Asp	Lys	Arg	His	Lys	Val	Arg	Met	Ala	Phe	Gly	Asn	325	330	335
Gly	Met	Arg	Pro	Asp	Val	Trp	Ala	Lys	Phe	Gln	Glu	Arg	Phe	Asn	Ile	340	345	350
Pro	Ile	Ile	Val	Glu	Tyr	Tyr	Ala	Met	Ser	Glu	Gly	Thr	Ser	Ser	Leu	355	360	365
Leu	Asn	Val	Ala	Arg	Asn	Lys	Arg	Asp	Gln	Gly	Ala	Val	Gly	Phe	Arg	370	375	380
Gly	Pro	Val	Val	Arg	Ala	Leu	Thr	Pro	Pro	Val	Gln	Leu	Val	Lys	Val	385	390	395
Asp	Phe	Asp	Thr	Glu	Glu	Leu	Ile	Arg	Asp	Lys	Lys	Thr	Gly	Leu	Cys	405	410	415
Val	Leu	Cys	Gln	Pro	Gly	Glu	Ile	Gly	Glu	Leu	Val	Thr	Leu	Ala	Asp	420	425	430
Asn	Lys	Thr	Thr	Gly	Ala	Arg	Tyr	Ala	Gly	Tyr	Phe	Asn	Gln	Pro	Glu	435	440	445
Val	Ser	Lys	Ala	Arg	Leu	Val	Gln	Asn	Val	Val	Val	Lys	Asp	Asp	Ile	450	455	460
Tyr	Phe	Arg	Thr	Gly	Asp	Leu	Leu	Tyr	Ser	Lys	Asp	Gln	Tyr	Trp	Tyr	465	470	475
Phe	Ala	Asp	Arg	Ala	Gly	Asp	Thr	Tyr	Arg	Trp	Lys	Gly	Glu	Asn	Val	485	490	495
Ser	Thr	Ala	Glu	Ile	Ala	Asp	Thr	Ile	Gly	Arg	Val	Glu	Gly	Val	Ala	500	505	510
Ser	Cys	Thr	Val	Tyr	Gly	Val	Ser	Val	Pro	Gly	Met	Asp	Gly	Arg	Ala	515	520	525

Gly Met Ala Ala Leu Val Leu Lys Asn Ser Ile Val Gln Met Ala Gly
 530 535 540
 Gly Ser Gln Ala Lys Phe His Val Asp Glu Ala Ala Leu Asn Ala Phe
 545 550 555 560
 Leu Arg Asp Leu Ser Lys Asp Val Val Lys Lys Leu Pro Ala Tyr Ala
 565 570 575
 Ile Pro Arg Phe Leu Arg Ile Ala Glu Gln Glu Leu Glu Thr Thr Gly
 580 585 590
 Thr Phe Lys Asn Lys Lys Val Glu Leu Lys Lys Glu Gly Phe Asp Leu
 595 600 605
 Gly Lys Val Lys Glu Arg Leu Tyr Trp Trp Thr Pro Lys Gly Glu Tyr
 610 615 620
 Ala Pro Phe Gly Val Ala Glu Asn Glu Gln Ile Leu Ala Gly Arg Ala
 625 630 635 640
 Arg Leu

<210> 8
 <211> 1929
 <212> DNA
 <213> Mortierella alpina

<400> 8
 atggcaagta ccaaactact aaggacttgg cgcttggttcg cgttggtctc catgcacgca 60
 aaagacttga ggccttggtc gagaatcggg actgcagtct ttacaacagg ttcgaggaac 120
 aatgccagat ccggccttcc tctgttgccc ttgtttttga gaacacgtct tacacctgga 180
 gagacttgga gctggcgctc aacaggatgg ccatttggtt tggtgctcaa ggaatccaaa 240
 aaaaaggagc gtgtggcgat gatgatgcat aactcgctc tggtcattat cacctggctg 300
 gcaatgctca agatcatggt tgtacctgct tttatcaata accagattgc aggacctgtt 360
 ctggttcatt ctcttaaagt ggccgacgcc aagtttctct tggtcgatta cgagttggca 420
 cctgtcatcc aaaagtcgct caatgagatc aaggacatgg gttacaatct ctacactgtc 480
 acaccaagg atcaagttct aggtcaactt tacgccaatc tgcccagggc tgctcgtcag 540
 gtgttgatg aggtccttc attctttggt tatgtcgaat ggcagaacct cagtaccgaa 600
 gggttctcga acgagagtcg tcaggagggtg gtgatctccg accccgcagc cttgatttac 660
 accagcggga ccacgggatt cccaaggct gctatcatgg accatggacg ttgcaacttg 720
 gcttcgatct cttatggcac tctatgcggc atcaaaccag agaacaagggt ttacatcaca 780
 ttgccgctct atcattctgc tggagccatc attggtcttg gccagagctt caccagcgga 840
 tgcaccattg tgctggcgcg aaagttctcc gtgacaaagt tttggcgtga ttgcgttgag 900
 tacgacgtaa ctcatcttca gtacattggc gaactctgcc gctacctctt aaatgcccc 960
 gaaagtccac tggacaaaag gcataagggt cggtatggcg ttggcaacgg aatgcgcccc 1020
 gatgtttggg caaagtttca ggaacgattc aatatcccca ttattgttga gtactacgcc 1080
 atgagcgaag gaacatcgct gcttttgaat gtggccagga acaagcgcga ccaagggtgcg 1140
 gtgggattcc gtggccccgt cgtgagggcc ttgacgcctc ccgttcaact ggtcaagggtg 1200
 gactttgaca cggaggagct gatccgcgat aagaagacgg gactttgcgt cctatgccag 1260
 cctggtgaga ttggagaact ggtcacgcta gccgacaaca agacgactgg cgcacgctat 1320
 gctgggtatt tcaatcagcc agaggtttgc aaggcaaggc tggccagaa cgtggtagtg 1380
 aaggacgaca tctacttccg gacgggtgac ctcttgtagt ccaaggacca gtactggtac 1440
 tttgctgata gcgcaggaga cacgtaccgg tggaaaggag agaactgtgc gacagccgag 1500
 attgcagaca ctatcggccg tgttgagggc gtggctagtt gtactgttta tggcgtatcg 1560
 gtcccgggca tggatggacg cgcgggcatg gctgcttttg tgctcaagaa ctcgattgtg 1620
 cagatggcag gtggaagcca ggcaaagttc catgtggatg aggtgcgct gaacgcgttt 1680
 ttgcgtgact tgagcaagga tgtggtcaaa aaactgccgg cgtatgcgat tectcggttc 1740
 ttgcgcattg cagagcagga actggagacg acgggcacgt tcaagaacaa gaaggtggag 1800
 ctgaagaagg aagggttcga cctcggttaag gtcaaggagc ggctgtactg gtggacaccc 1860
 aagggtgaat atgccccttt tggcgtggcg gagaacgagc agatcctcgc aggacgcgct 1920
 cgtctttga 1929

<210> 9
 <211> 2135
 <212> DNA
 <213> Mortierella alpina

<400> 9
 atccgcccac cgcctctctt gccactgatc tcaaagecgtg atccaaagggt cattcttagg 60
 cagcactcac gcagctactt agaactctac ccacatatcc cttattgata caatggctcc 120
 cgctcgctgca ctgcgcgcgc ctctggcggc aggatactat ctcaatggca agtaccaaat 180
 cactaaggac ttggcgcttg ttgcgcttgg tctccatgca cgcaaaagac ttgaggcctt 240
 ggtcgagaat cgggactgca gtctttacaa cagggttcgag gaacaatgcc agatccggcc 300
 tttctctgtt gcccttggtt ttgagaacac gtcttacacc tggagagact tggagctggc 360
 gtccaacagg atggcccatt ggtttgttgc tcaaggaatc caaaaaaag gagcgtgtgg 420
 cgatgatgat gcataactcg cctctgttca ttatcacctg gctggcaatg ctcaagatca 480
 tggttgtacc tgcttttata aataaccaga ttgcaggacc tgttctgggt cattctctta 540
 aagtggccga cgccaagttt ctcttggtcg attacgagtt ggcacctgtc atccaaaagt 600
 cgctcaatga gatcaaggac atgggttaca atctctacac tgtcacaccc aaggatcaag 660
 ttctaggtca actttacgcc aatctgcccg aggtctctcg tcagggtgtg gatgaggctc 720
 cttcattctt tggttatgtc gaatggcaga acctcagtac cgaaggtttc tcgaacgaga 780
 gtcgtcagga ggtggtgatc tccgaccccg cagccttgat ttacaccagc gggaccacgg 840
 gattcccca ggtgctatc atggaccatg gacgttgcaa cttggcttcg atctcttatg 900
 gcactctatg cggcatcaaa ccagagaaca aggtttacat cacattgccg ctctatcatt 960
 ctgctggagc catcattggt ctgggccaga gcttcaccag cggatgcacc attgtgctgg 1020
 cgcgaaagtt ctccgtgaca aagttttggc gtgattgcgt tgagtacgac gtaactcatt 1080
 ttcagtacat tggcgaactc tgccgctacc ttctaaatgc ccccgaaagt ccactggaca 1140
 aaaggcataa ggttcggatg gcgtttggca acggaatgcg cccggatgtt tgggcaaagt 1200
 ttcaggaacg attcaatata cccattattg ttgagtacta cgccatgagc gaaggaacat 1260
 cgctcgctttt gaatgtggcc aggaacaagc gcgaccaagg tgcggtggga ttccgtggcc 1320
 ccgtcgtagg ggccttgacg cctcccgttc aactggtcaa ggtggacttt gacacggagg 1380
 agctgatccg cgataagaag acgggacttt gcgtcctatg ccagcctggg gagattggag 1440
 aactggtcac gctagccgac aacaagacga ctggcgcacg ctatgctggg tatttcaatc 1500
 agccagaggt ttcgaaggca aggtctgggtc agaacgtggg agtgaaggac gacatctact 1560
 tccggacggg tgacctcttg tactccaagg accagtactg gtactttgct gatcgcgag 1620
 gagacacgta ccggtggaaa ggagagaacg tgtcgacagc cgagattgca gacactatcg 1680
 gccgtgttga gggcgtggct agttgtactg tttatggcgt atcgggtccc ggcatggatg 1740
 gacgcgcggg catggctgct ttggtgctca agaactcgat tgtgcagatg gcagggtgaa 1800
 gccaggcaaa gttccatgtg gatgaggctg cgctgaacgc gtttttgctg gacttgagca 1860
 aggatgtggt caaaaaactg ccggcgtatg cgattcctcg gttcttgccg attgcagagc 1920
 aggaactgga gacgacgggc acgttcaaga acaagaagggt ggagctgaag aaggaagggt 1980
 tcgacctcg taaggatcaag gagcggtgtg actggtggac acccaagggt gaatatgcc 2040
 cttttggcgt ggcgagagaac gagcagatcc tcgcaggacg cgctcgtctt tgagcgatgt 2100
 ttgtcaatga agtcatcggc atcatcatca tcac 2135

<210> 10
 <211> 3041
 <212> DNA
 <213> Mortierella alpina

<220>
 <221> misc_feature
 <222> (118)..(459)
 <223> n is a, c, g, or t

<400> 10
 atggcaagta ccaaatcact aaggacttgg cgcttggttcg cgttggtctc catgcacgca 60
 aaaggatatga acgatgcact tgggggacgc atctgtctca atgtgatttg cttgttcnnn 120

nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	180
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	240
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	300
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	360
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	420
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnc	gtctactttt	tttttttttt	480
ttttttttag	agagaaaggt	cgaatgtggg	gaccgaacct	attttttttct	cgtgcctttt	540
aggtttgttg	cttcagcggt	cttttgagct	gatatctttt	cgccactctt	gcatacctta	600
gacttgaggc	cttggtcgag	aatcgggact	gcagtcttta	caacagggttc	gaggaacaat	660
gccagatccg	gcctttctct	gttgcccttg	tttttgagaa	cacgtcttac	acctggagag	720
acttgagact	gggtatgtgc	agtaggattt	tttggaacgcc	tgatgctgca	ttctgtgata	780
gaggacatac	attaaataat	tgataatcatt	tcgtctgtcg	acctgtctgt	ctatcgatat	840
acagcgtcca	acaggatggg	aagttggcaa	ctgcatacaa	gtacttcgtg	ctctttatca	900
tgcgttactc	acgcttcatt	gcctcaatgt	ttcctttgaa	ctaggcccat	tggtttggtg	960
ctcaaggaat	ccaaaaaaaa	ggagcgtgtg	gcgatgatga	tgcataactc	gcctctgttc	1020
attatcacct	ggctggcaat	gctcaagatc	atggttgtag	ctgcttttat	caataaccag	1080
attgcaggac	ctgttctggt	tcattctctt	aaagtggccg	acgccaaagt	tctcttggtc	1140
gattacgagt	tggcacctgt	catccaaaag	tcgctcaatg	agatcaagga	catgggttac	1200
aatctctaca	ctgtcacacc	caaggatcaa	gttctaggtc	aactttacgc	caatctgcc	1260
gaggtgctc	gtcaggtggt	ggatgaggct	ccttcattct	ttggttatgt	cgaatggcag	1320
aacctcagta	ccgaagggtt	ctcgaacgag	agtcgtcagg	aggtgggat	ctccgacccc	1380
gcagccttga	tttacaccag	cgggaccacg	ggattcccca	aggtgctat	catggaccat	1440
ggacgttgca	actgtaagca	atcgcatagg	atcgatagcg	ctgaatggct	ggcgagtgga	1500
tgcaaagtgt	cgagatgctt	accattatcg	tgggtgtgct	ttttatagtg	gcttcgatct	1560
cttatggcac	tctatgcggc	atcaaaccag	agaacaaggt	ttacatcaca	ttgccgctct	1620
atcattctgc	tggaggtacg	tgttccatc	tcaccctcaa	catctcttac	gacggtttga	1680
tcctgttctt	acactcatta	cttctgggca	tgggaacaaa	gccatcattg	gtctgggcca	1740
gagcttcacc	agcggatgca	ccattgtgct	ggcgcgaaag	ttctccgtga	caaagttttg	1800
gcgtgattgc	gttgagtacg	acgtaactca	ttttcaggta	caagtcctat	ccaatggtct	1860
acataaccgtc	cttgtgtatt	ttcaacgcgc	acgcgcca	acccgctttt	atatgtatac	1920
ccgcagtaca	ttggcgaaact	ctgccgctac	cttctaaatg	cccccgaaag	tccactggac	1980
aaaaggcata	aggttcggat	ggcgtttggc	aacggaatgc	gcccggatgt	ttgggcaaag	2040
tttcaggaac	gattcaatat	ccccattatt	gttgagtact	acgccatgag	cgaaggaaaca	2100
tcgtcgcttt	tgaatgtggc	caggaacaag	cgcgaccaag	gtgcggtggg	attccgtggc	2160
ccgctcgtga	ggtatgcggc	atctgggcgc	tttagttctt	cgtgttcaat	ggtttcatta	2220
taacatcttc	agctcaactt	ttgcccgctg	tttttccttt	caattttgtt	tccactaggg	2280
ccttgacgcc	tcccgttcaa	ctgggtcaagg	tggactttga	cacggaggag	ctgatccgcg	2340
ataagaagac	gggactttgc	gtcctatgcc	agcctggtga	gattggagaa	ctggtcacgc	2400
tagccgacaa	caagacgact	ggcgcacgct	atgctgggta	tttcaatcag	ccagagggtt	2460
cgaaggcaag	gctgggtccag	aacgtggtag	tgaaggacga	catctacttc	cggacgggtg	2520
acctcttgta	ctccaaggac	cagtactggt	actttgctga	tcgcgcagga	gacacgtacc	2580
ggtggaaagg	agagaacgtg	tcgacagccg	agattgcaga	cactatcggc	cgtgttgagg	2640
gcgtggctag	ttgtactggt	tatggcgat	cggccccggg	catggatgga	cgcgcgggca	2700
tggctgcttt	ggtgctcaag	aactcgattg	tgcagatggc	aggtggaagc	caggcaaagt	2760
tccatgtgga	tgaggctgcg	ctgaacgcgt	ttttgcgtga	cttgagcaag	gatgtggtca	2820
aaaaactgcc	ggcgtatgcg	attcctcggt	tcttgccgat	tgcagagcag	gaactggaga	2880
cgacgggcac	gttcaagaac	aagaagggtg	agctgaagaa	ggaagggttc	gacctcggtg	2940
aggccaagga	gcggctgtac	tgggtggacac	ccaagggtga	atatgccctt	tttggcgtgg	3000
cggagaacga	gcagatcctc	gcaggacgcg	ctcgtctttg	a		3041

```
<210> 11
<211> 1650
<212> DNA
<213> Mortierella alpina
```

<400> 11

```

atggaaacgg atgctcttac catcgctttg accatcgcca tcgccatcgt gctggctttg      60
gtcaaattca acgaaaaaga gctgacctg catccgctcc tgctcgggca gcaatcgtct      120
gtcacgccc aacggaacga gggcgagtcc gttatccata gatccaaaac ggtgccacac      180
gggacactgc tgacgaagcg cccgagcgag aaaatcaaga ctctgcacga tgtctggcag      240
actggagcag ctgtcaaccc agccggccga tcgttgatgt ttatgctgca gaaccagttt      300
gcgtttatcg aggccacgta tgagcaagtc aataggagga ttggcggcct cggaacaggt      360
ttcgtgaagg caacagggct aaagcccaag acggacacac cagtaggaat ctttatgccc      420
tactctcaag aatcgttcgt tgcccagcag gcattctatc gatacagctt tgttgctgtc      480
cccatccatg atctgaggaa caacgacctc ttgggtggagg tagtagacca gaccaagctc      540
aaggccatca tagtctcaca aaaggtgctc ccgttattgc tgcaatctct gaaggagtgt      600
ccaaccatca agacaatcat catggcagga atctacatct cacaggagca gctggaaatg      660
gcagcacagc atggagtaaa gctgctcaaa ttcgcggcag tggaatatga gggatcctcg      720
actctgatgg agcctgttca gctgatccg gaggatgttg ccatgatcaa ctataacaca      780
aagtcgtctt cgctctcgaa aggcgtcatg cttacccatg ccaacctgat cgcggcgatg      840
actgccttca cggagtcact tccggcaaaa aagcgtttct ccagcaaaga tcgtcttctc      900
tctcattttt ccaatggaga tgtcatctct gtcttcatgt cgagcgccat catcctgatg      960
ggaggttctt tggcttttcc atctggtttg atgaagaacg ttttgcatga ttcccaagct     1020
tctgcaccaa cgatctttgc aagcacaccc atcatcctgg aaaagattca cgaagcactt     1080
cagttgacgt atggccaagg ctccatgttc aggcgcggct ttgctgccaa attggccata     1140
cttcaagctg gacgaatcac tacaacaagc ctatgggact tgattggact gggcgaggtc     1200
cgcagcaaac ttgggtggaaa ggttcgaatg gttgtaacaa cacatcctac caaacctgag     1260
acgctggatt atatcagagc cgcgatgggc atccatgtca ttaccactta cggcaggaca     1320
gagacgtcgg gcattgtgac agcccgcac atgctggatt atgccaacgc acctcattta     1380
ggaccaccag tgggttgcaa cgaggttaag cttgtggacg atgttgacgc tggctttaca     1440
agtgcagacg agcccaaccc acgagcgag atccttatcc gagggcccaa tgtgatgaaa     1500
ggttattaca agaagccggg tgccacttca acggctatcg atgaggaagg gtggttccat     1560
tcaggagagc tgggcacatt ccaactccaac ggcactttag acgtgttggg caagaagaag     1620
aagacgaagt ctgcagttgg atcaccgtca                                     1650

```

<210> 12

<211> 550

<212> PRT

<213> Mortierella alpina

<400> 12

```

Met Glu Thr Asp Ala Leu Thr Ile Ala Leu Thr Ile Ala Ile Ala Ile
1          5          10          15
Val Leu Ala Leu Val Lys Phe Asn Glu Lys Glu Pro Asp Leu His Pro
20          25          30
Leu Leu Leu Gly Gln Gln Ser Ser Val Thr Pro Ile Arg Asn Glu Gly
35          40          45
Glu Ser Val Ile His Arg Ser Lys Thr Val Pro His Gly Thr Leu Leu
50          55          60
Thr Lys Arg Pro Ser Glu Lys Ile Lys Thr Leu His Asp Val Trp Gln
65          70          75          80
Thr Gly Ala Ala Val Asn Pro Ala Gly Arg Ser Leu Met Phe Met Leu
85          90          95
Gln Asn Gln Phe Ala Phe Ile Glu Ala Thr Tyr Glu Gln Val Asn Arg
100         105         110
Arg Ile Gly Gly Phe Gly Thr Gly Phe Val Lys Ala Thr Gly Leu Lys
115         120         125
Pro Lys Thr Asp Thr Pro Val Gly Ile Phe Met Pro Tyr Ser Gln Glu
130         135         140
Ser Phe Val Ala Gln Gln Ala Phe Tyr Arg Tyr Ser Phe Val Ala Val
145         150         155         160

```

Pro	Ile	His	Asp	Leu	Arg	Asn	Asn	Asp	Leu	Leu	Val	Glu	Val	Val	Asp	
				165					170						175	
Gln	Thr	Lys	Leu	Lys	Ala	Ile	Ile	Val	Ser	Gln	Lys	Val	Leu	Pro	Leu	
			180					185					190			
Leu	Leu	Gln	Ser	Leu	Lys	Glu	Cys	Pro	Thr	Ile	Lys	Thr	Ile	Ile	Met	
		195				200						205				
Ala	Gly	Ile	Tyr	Ile	Ser	Gln	Glu	Gln	Leu	Glu	Met	Ala	Ala	Gln	His	
	210					215					220					
Gly	Val	Lys	Leu	Leu	Lys	Phe	Ala	Ala	Val	Glu	Tyr	Glu	Gly	Ser	Ser	
225					230					235					240	
Thr	Leu	Met	Glu	Pro	Val	Gln	Pro	Asp	Pro	Glu	Asp	Val	Ala	Met	Ile	
			245					250						255		
Asn	Tyr	Asn	Thr	Lys	Ser	Ser	Ser	Leu	Ser	Lys	Gly	Val	Met	Leu	Thr	
			260					265					270			
His	Ala	Asn	Leu	Ile	Ala	Ala	Met	Thr	Ala	Phe	Thr	Glu	Ser	Leu	Pro	
	275						280					285				
Ala	Lys	Lys	Arg	Phe	Ser	Ser	Lys	Asp	Arg	Leu	Leu	Ser	His	Phe	Ser	
	290					295					300					
Asn	Gly	Asp	Val	Ile	Ser	Val	Phe	Met	Ser	Ser	Ala	Ile	Ile	Leu	Met	
305					310					315					320	
Gly	Gly	Ser	Leu	Val	Phe	Pro	Ser	Gly	Leu	Met	Lys	Asn	Val	Leu	His	
			325					330						335		
Asp	Ser	Gln	Ala	Ser	Ala	Pro	Thr	Ile	Phe	Ala	Ser	Thr	Pro	Ile	Ile	
			340					345					350			
Leu	Glu	Lys	Ile	His	Glu	Ala	Leu	Gln	Leu	Thr	Tyr	Gly	Gln	Gly	Ser	
		355					360					365				
Met	Phe	Arg	Arg	Gly	Phe	Ala	Ala	Lys	Leu	Ala	Ile	Leu	Gln	Ala	Gly	
	370					375					380					
Arg	Ile	Thr	Thr	Thr	Ser	Leu	Trp	Asp	Leu	Ile	Gly	Leu	Gly	Glu	Val	
385					390				395						400	
Arg	Ser	Lys	Leu	Gly	Gly	Lys	Val	Arg	Met	Val	Val	Thr	Thr	His	Pro	
			405					410						415		
Thr	Lys	Pro	Glu	Thr	Leu	Asp	Tyr	Ile	Arg	Ala	Ala	Met	Gly	Ile	His	
			420					425					430			
Val	Ile	Thr	Thr	Tyr	Gly	Arg	Thr	Glu	Thr	Ser	Gly	Ile	Val	Thr	Ala	
		435					440					445				
Arg	Asn	Met	Leu	Asp	Tyr	Ala	Asn	Ala	Pro	His	Leu	Gly	Pro	Pro	Val	
	450					455					460					
Gly	Cys	Asn	Glu	Val	Lys	Leu	Val	Asp	Asp	Val	Ala	Ala	Gly	Phe	Thr	
465					470					475					480	
Ser	Ala	Asp	Glu	Pro	Asn	Pro	Arg	Gly	Glu	Ile	Leu	Ile	Arg	Gly	Pro	
			485					490						495		
Asn	Val	Met	Lys	Gly	Tyr	Tyr	Lys	Lys	Pro	Gly	Ala	Thr	Ser	Thr	Ala	
		500						505					510			
Ile	Asp	Glu	Glu	Gly	Trp	Phe	His	Ser	Gly	Glu	Leu	Gly	Thr	Phe	His	
		515					520					525				
Ser	Asn	Gly	Thr	Leu	Asp	Val	Leu	Gly	Lys	Lys	Lys	Lys	Thr	Lys	Ser	
	530					535					540					
Ala	Val	Gly	Ser	Pro	Ser											
545					550											

<210> 13

<211> 1653

<212> DNA

<213> Mortierella alpina

<400> 13

atggaaacgg	atgctcttac	catcgctttg	accatcgcca	tgcgcatcgt	gctggctttg	60
gtcaaattca	acgaaaaaga	gcctgacctg	catccgctcc	tgctcgggca	gcaatcgtct	120
gtcacgcca	ttcggaacga	ggcgagtc	gttatccata	gatccaaaac	ggtgccacac	180
gggacactgc	tgacgaagcg	cccgagcgag	aaaatcaaga	ctctgcacga	tgtctggcag	240
actggagcag	ctgtcaaccc	agccggccga	tcgttgatgt	ttatgctgca	gaaccagttt	300
gcgtttatcg	aggccacgta	tgagcaagtc	aataggagga	ttggcggctt	cggaacagggt	360
ttcgtgaagg	caacagggct	aaagcccaag	acggacacac	cagtaggaat	ctttatgccc	420
tactctcaag	aatcgttcgt	tgcccagcag	gcattctatc	gatacagctt	tgttgctgtc	480
cccatccatg	atctgaggaa	caacgacctc	ttgggtggagg	tagtagacca	gaccaagctc	540
aaggccatca	tagtctcaca	aaagggtgctc	ccgttattgc	tgcaatctct	gaaggagtgt	600
ccaacctatca	agacaatcat	catggcagga	atctacatct	cacaggagca	gctggaaatg	660
gcagcacagc	atggagtaaa	gctgctcaaa	ttcgcggcag	tggaatatga	gggatcctcg	720
actctgatgg	agcctgttca	gcctgatccg	gaggatgttg	ccatgatcaa	ctataacaca	780
aagtcgtctt	cgctctcgaa	aggcgtcatg	cttacctatg	ccaacctgat	cgcggcgatg	840
actgccttca	cggagtcact	tccggcaaaa	aagcgtttct	ccagcaaaga	tcgtcttctc	900
tctcattttt	ccaatggaga	tgatcatctct	gtcttcatgt	cgagcgccat	catcctgatg	960
ggaggttctt	tggtctttcc	atctggtttg	atgaagaacg	ttttgcatga	ttcccaagct	1020
tctgcaccaa	cgatctttgc	aagcacaccc	atcatcctgg	aaaagattca	cgaagcactt	1080
cagttgacgt	atggccaagg	ctccatgttc	aggcgcggt	ttgctgccaa	attggccata	1140
cttcaagctg	gacgaatcac	tacaacaagc	ctatgggact	tgattggact	ggcgaggtc	1200
cgcagcaaac	ttgggtggaaa	ggttcgaatg	gttgtaacaa	cacatcctac	caaacctgag	1260
acgctggatt	atatcagagc	cgcgatgggc	atccatgtca	ttaccactta	cggcaggaca	1320
gagacgtcgg	gcattgtgac	agcccgcac	atgctggatt	atgccaacgc	acctcattta	1380
ggaccaccag	tgggttgcaa	cgaggttaag	cttgtggacg	atggtgcagc	tggctttaca	1440
agtgcagacg	agcccaaccc	acgaggcgag	atccttatcc	gaggcccaaa	tgtgatgaaa	1500
ggttattaca	agaagccggg	tgccacttca	acggctatcg	atgaggaagg	gtggttccat	1560
tcaggagagc	tgggcacatt	ccactccaac	ggcactttag	acgtgttggg	caagaagaag	1620
aagacgaagt	ctgcagttgg	atcaccgctca	tga			1653

<210> 14

<211> 1753

<212> DNA

<213> Mortierella alpina

<400> 14

gcctactttg	cgctcgccctc	atcgacccaa	aggcagcaat	ggaaacggat	gctcttacca	60
tcgctttgac	catcgccatc	gccatcgtgc	tggctttggg	caaattcaac	gaaaaagagc	120
ctgacctgca	tccgctcctg	ctcgggcagc	aatcgtctgt	cacgcccatt	cggaacgagg	180
gcgagtcggt	tatccataga	tccaaaacgg	tgccacacgg	gacactgctg	acgaagcgcc	240
cgagcgagaa	aatcaagact	ctgcacgatg	tctggcagac	tggagcagct	gtcaacccag	300
ccggccgatc	gttgatgttt	atgctgcaga	accagtttgc	gtttatcgag	gccacgtatg	360
agcaagtcaa	taggaggatt	ggcggcttcg	gaacagggtt	cgtgaaggca	acagggctaa	420
agcccaagac	ggacacacca	gtaggaatct	ttatgcccta	ctctcaagaa	tcgttcgttg	480
cccagcaggc	attctatcga	tacagctttg	ttgctgtccc	catccatgat	ctgaggaaca	540
acgacctctt	ggtggaggta	gtagaccaga	ccaagctcaa	ggccatcata	gtctcacaaa	600
aggtgctccc	gttattgctg	caatctctga	aggagtgtcc	aacctcaag	acaatcatca	660
tggcaggaat	ctacatctca	caggagcagc	tggaaatggc	agcacagcat	ggagtaaagc	720
tgctcaaatt	cgcggcagtg	gaatatgagg	gatactcgac	tctgatggag	cctgttcagc	780
ctgatccgga	ggatgttgcc	atgatcaact	ataacacaaa	gtcgtcttcg	ctctcgaaag	840
gcgtcatgct	tacctatgcc	aacctgatcg	cggcgatgac	tgccttcacg	gagtcacttc	900
cggcaaaaaa	gcgtttctcc	agcaaagatc	gtcttctctc	tcatttttcc	aatggagatg	960
tcctctctgt	cttcatgtcg	agcgccatca	tcctgatggg	aggttctttg	gtctttccat	1020
ctggtttgat	gaagaacggt	ttgcatgatt	cccaagcttc	tgcaccaacg	atctttgcaa	1080
gcacacccat	catcctggaa	aagattcacg	aagcacttca	gttgacgtat	ggccaaggct	1140
ccatgttcag	gcgcggcttt	gctgccaat	tggccatact	tcaagctgga	cgaatcacta	1200

caacaagcct	atgggacttg	attggactgg	gcgaggtccg	cagcaaactt	ggtggaaagg	1260
ttcgaatggg	tgtaacaaca	catcctacca	aacctgagac	gctggattat	atcagagccg	1320
cgatgggcat	ccatgtcatt	accacttacg	gcaggacaga	gacgtcgggc	attgtgacag	1380
cccgcaacat	gctggattat	gccaacgcac	ctcatttagg	accaccagtg	ggttgcaacg	1440
aggttaagct	tgtggacgat	gttgacagct	gctttacaag	tgacagacgag	cccaaccac	1500
gaggcgagat	ccttatccga	ggccccaatg	tgatgaaagg	ttattacaag	aagccgggtg	1560
ccacttcaac	ggctatcgat	gaggaagggt	ggttccattc	aggagagctg	ggcacattcc	1620
actccaacgg	cacttttagac	gtgttgggca	agaagaagaa	gacgaagtct	gcagttggat	1680
caccgtcatg	aaaggagatg	ctgcatgtgc	tacagaatat	aaaaaggagg	aagatacggt	1740
cggttaaccac	atc					1753

<210> 15

<211> 2391

<212> DNA

<213> Mortierella alpina

<400> 15

atggaaacgg	atgctcttac	catcgctttg	accatcgcca	tcgccatcgt	gctggctttg	60
gtcaaattca	acgaaaaaga	gcctgacctg	catccgctcc	tgctcgggca	gcaatcgtct	120
gtcacgcccc	ttcggaacga	gggcgagtc	gttatccata	gatccaaaac	ggtgccacac	180
gggacactgc	tgacgaagcg	cccgagcgag	aaaatcaaga	ctctgcacga	tgtctggcag	240
actggagcag	ctgtcaaccc	agccggccga	tcgttgatgt	ttatgctgca	gaaccagttt	300
gcgtttatcg	aggtacgatg	gacctcgtgt	agtaaccccg	ctgtctcttg	agcaatatcg	360
caggagtctc	accattagag	gattcattct	ccttcgcata	ggccacgtat	ggtaacgtgt	420
gattcggtag	cctttctgtc	tgttgaagat	gcggtgatgt	ggatctctaa	cagatcttgg	480
ttcaatggtg	acacagagca	agtcaatagg	aggattggcg	gcttcggaac	aggtttcgtg	540
aaggcaacag	ggctaaagcc	caagacggac	acaccagtag	gaatctttat	gccctactct	600
caaggtagcg	gaacaagcgc	gtgagtactg	atccagcaac	gcatagggac	tgacgcgagt	660
gagccacgtg	tttgaccttg	caccgcgcct	ctttactgta	gaatcgttcg	ttgcccagca	720
ggcattctat	cgatacagct	ttgtttgetgt	ccccatccat	gatctgagga	acaacgacct	780
cttgggtggag	gtagtagacc	agaccaagct	caaggccatc	atagtctcac	aaaagggtgt	840
cccgttattg	ctgcaatctc	tgaaggagtg	tccaaccatc	aagacaatca	tcatggcagg	900
aatctacatc	tcacaggagc	agctggaaat	ggcagcacag	catggagtaa	agctgctcaa	960
attcgcggca	gtggaatatg	agggatcctc	gactctgatg	gagcctgttc	agcctggtat	1020
gtgaagcaaa	agtcaaggaa	atgcggttgc	ttgatgttcg	ctgcgatgtt	ttgaccacca	1080
cgacctcttt	taatagatcc	ggaggatgtt	gccatgatca	actataacac	aaagtcgtct	1140
tcggtatgga	tgcttgtttt	tcagtggtag	ttttgtataa	gcgggcatgg	agatgattga	1200
gcttacttga	agtactttcg	cgctgcatca	ttgatccgat	agctctcgaa	aggcgtcatg	1260
cttacctatg	ccaacctgat	cgcggcgatg	actgccttca	cggagtcact	tccggcaaaa	1320
aagcgtttct	ccagcaaaga	tcgtcttctc	tctcattttt	ccaatggaga	tgtcatctct	1380
gtcttcatgt	cgagcgccat	catcctgatg	ggaggttctt	tggtctttcc	atctggtttg	1440
atgaagaacg	ttttgcatga	ttcccaagct	tctgcaccaa	cgatctttgc	aaggtacaag	1500
atgttccagt	gtgatctgtg	gattgtctcg	ttattcatgt	gcaagatact	tacggcgtcg	1560
ttccatttat	ttttgacccc	tcctagcaca	cccatcatcc	tggaagagat	tcacgaagca	1620
cttcagttga	cgtatggcca	aggctccatg	ttcaggcgcg	gctttgctgc	caaattggcc	1680
atacttcaag	ctggacgaat	cactacaaca	agcctatggg	acttgattgg	actgggcgag	1740
gtccgcagca	aacttggtgg	aaaggttcga	atggttgtaa	caacacgtaa	gtctcctttt	1800
ttaccatgcg	ggcttacgta	cttgacgcat	gatattcgaa	gatactaaca	gtcttttcat	1860
gacgtcctg	tagatcctac	caaacctgag	acgctggatt	atatcagagc	cgcgatgggc	1920
atccatgtca	ttaccactta	cggcaggaca	gagacgtcgg	gcattgtgac	agcccgcaac	1980
atgctggatt	atgccaacgc	acctcattta	ggaccaccag	tgggttgcaa	cgaggttaag	2040
cttgtggacg	atgttgacgc	tggctttaca	agtgcagacg	agcccaaccc	acgaggcgag	2100
gtaaggagtg	catatctctt	gctgcttcat	gggcgataaa	aaggctcaaa	tgccgttgat	2160
tctactctaa	cagcgttatt	ttattgttca	acaatatatt	gtgcctagat	ccttatccga	2220
ggccccaatg	tgatgaaagg	ttattacaag	aagccgggtg	ccacttcaac	ggctatcgat	2280

gaggaaggggt ggttccattc aggagagctg ggcacattcc actccaacgg cacttttagac 2340
 gtgttgggca agaagaagaa gacgaagtct gcagttggat caccgtcatg a 2391

<210> 16
 <211> 2064
 <212> DNA
 <213> Mortierella alpina

<400> 16
 atgtccctcg accagaacgc ccagtccggt gagctcccag gcacccggca accaggccag 60
 acaggcatct atcgccgcaa aggtttcgag aatgcccttc tcgccgtccc acccagcaga 120
 ccgcacatca agaccatcta cgatgccttc cagcacggac tgaagcttaa tcccaacgga 180
 gctgccctgg gcagccgagt gtacgaccgc gtgacggaca cctttggagg ctatgtctgg 240
 cagacgtatg cacaggtgaa cgaccgcctc actcgcttcg gcagtggatt ggtcaaaatt 300
 cataaggacg tccatgggtct tgccaccgtg ggccagaagt ggtctctcgg aatctgggccc 360
 atcaaccgac ccgagtggac catcgcgctc gaggtttgct cggcctacaa cctggtctcc 420
 gtgggtcttt acgatacttt gggacccgag gctgtgactt atggcattaa tcacgtctgag 480
 tgctctattg tcgtaacaag tgtggatcat atcgcgacgc tgctgaacga atcttccaag 540
 atgcctgggc tcaaaatcat catcagcatg gatgacctcg atactgggag agcaggccca 600
 ggactggctc ccaccggcac catcctcagg acttacgctc aggacaaagg ggtactactt 660
 tatgattggg ctgaggttga agccgtcggg attcagcatg gacgaaagca tacgccacca 720
 acctcctccg acgcataata gatctgctat accagcggga caacaggctt gccaaaagg 780
 gccatthtga cccatggaaa cttgatcgcc cttttggcct ccagtgatgt ggccacacca 840
 gtgctggctg acgattgcct catcagtttc ttgcccctgc ctacgtctt tggtcgggtc 900
 atggagctct tcgcgatggc cgcaggagga aagattggct acagcacggg agatcctttg 960
 cgtctcttgg aggacgtctc gcacctaaag cctccatct tccccgtgt gccagactg 1020
 ctgaaccgcg tgtatgcaa ggtgtatgcg gcaactgttg gagcgcttg actcacagg 1080
 gcactggcgc gacgaggatt ggccaccaag ctaccaatt tgagagaggg caaaggtttc 1140
 caccacccat tgtgggaccg aatcctcttc tcaaaggtea agcaagcgct cggcggcaat 1200
 gtgagactga tgttgactgc ctccgctccc atctcggccg agatcttggga attcgtccgt 1260
 gtcgctttct gctgcgaggt cgtggaggca tatggacaga ctgagggcgg tggagcggcc 1320
 acaaacaccg tgattggcga gaccgaggct ggacacgtcg gtcctcctca agcttggtgc 1380
 gagatcaaac tgggtggatgt acccgagctg aactactttg cgaccgataa accattccct 1440
 cgtgggtgaga tttgtgtccg tggacccggg gtcattcctg gttatctcaa ggatgaggcc 1500
 aagaccaagg agaccattga tgaggagggc tggctgcact cgggcgatat cgccatcatg 1560
 agtggcaaag gcaccgttac catcattgac aggaagaaga acgtgttcaa gctgagccaa 1620
 ggagaataca tcgcggcaga gaacattgaa gggcgtttcc tctccaagg tccattcctc 1680
 caacaaattc tgggtgcacgg cgactcgacc gagagctgtt tgggtggccat cttgatccca 1740
 gagcctgagg ccttcctccc ctttgtgaac aaagtgcctg agaacgtcaa tcttcaacct 1800
 ggagatcttg cagcctacag gaagatcggt aacaacccaa agctgcgcca ggctgtcctc 1860
 aaagagctga tcaaggctgg caaggatgct ggattgaaag gctttgagat tccaaaggcg 1920
 atcctcctcg aatctgaggc attcacggtc gaaaacgaca agatgacccc gactttcaag 1980
 atcaaaagac accctgtcgt ccaggcttac cgcgagcaac tgacagccct ctacaacgaa 2040
 atccatcaaa aggaatccaa gctg 2064

<210> 17
 <211> 688
 <212> PRT
 <213> Mortierella alpina

<400> 17
 Met Ser Leu Asp Gln Asn Ala Gln Ser Val Glu Leu Pro Gly Thr Arg
 1 5 10 15
 Gln Pro Gly Gln Thr Gly Ile Tyr Arg Arg Lys Gly Phe Glu Asn Ala
 20 25 30

Leu	Leu	Ala	Val	Pro	Pro	Ser	Arg	Pro	His	Ile	Lys	Thr	Ile	Tyr	Asp
		35					40					45			
Ala	Phe	Gln	His	Gly	Leu	Lys	Leu	Asn	Pro	Asn	Gly	Ala	Ala	Leu	Gly
	50					55					60				
Ser	Arg	Val	Tyr	Asp	Pro	Val	Thr	Asp	Thr	Phe	Gly	Gly	Tyr	Val	Trp
65					70					75					80
Gln	Thr	Tyr	Ala	Gln	Val	Asn	Asp	Arg	Ile	Thr	Arg	Phe	Gly	Ser	Gly
				85					90					95	
Leu	Val	Lys	Ile	His	Lys	Asp	Val	His	Gly	Leu	Ala	Thr	Val	Gly	Gln
			100					105					110		
Lys	Trp	Ser	Leu	Gly	Ile	Trp	Ala	Ile	Asn	Arg	Pro	Glu	Trp	Thr	Ile
	115					120						125			
Ala	Ser	Glu	Ala	Cys	Ser	Ala	Tyr	Asn	Leu	Val	Ser	Val	Gly	Leu	Tyr
	130					135					140				
Asp	Thr	Leu	Gly	Pro	Glu	Ala	Val	Thr	Tyr	Gly	Ile	Asn	His	Ala	Glu
145					150					155					160
Cys	Ser	Ile	Val	Val	Thr	Ser	Val	Asp	His	Ile	Ala	Thr	Leu	Leu	Asn
				165					170					175	
Glu	Ser	Ser	Lys	Met	Pro	Gly	Leu	Lys	Ile	Ile	Ile	Ser	Met	Asp	Asp
			180					185					190		
Leu	Asp	Thr	Gly	Arg	Ala	Gly	Pro	Gly	Leu	Ala	Pro	Thr	Gly	Thr	Ile
	195						200					205			
Leu	Arg	Thr	Tyr	Ala	Gln	Asp	Lys	Gly	Val	Leu	Leu	Tyr	Asp	Trp	Ser
	210					215					220				
Glu	Val	Glu	Ala	Val	Gly	Ile	Gln	His	Gly	Arg	Lys	His	Thr	Pro	Pro
225					230					235					240
Thr	Ser	Ser	Asp	Ala	Tyr	Thr	Ile	Cys	Tyr	Thr	Ser	Gly	Thr	Thr	Gly
				245					250					255	
Leu	Pro	Lys	Gly	Ala	Ile	Leu	Thr	His	Gly	Asn	Leu	Ile	Ala	Leu	Leu
			260					265					270		
Ala	Ser	Ser	Asp	Val	Ala	Thr	Pro	Val	Leu	Ala	Asp	Asp	Cys	Leu	Ile
	275						280				285				
Ser	Phe	Leu	Pro	Leu	Pro	His	Val	Phe	Gly	Arg	Val	Met	Glu	Leu	Phe
	290					295					300				
Ala	Met	Ala	Ala	Gly	Gly	Lys	Ile	Gly	Tyr	Ser	Thr	Gly	Asp	Pro	Leu
305					310					315					320
Arg	Leu	Leu	Glu	Asp	Val	Ser	His	Leu	Lys	Pro	Ser	Ile	Phe	Pro	Ala
				325					330					335	
Val	Pro	Arg	Leu	Leu	Asn	Arg	Val	Tyr	Ala	Lys	Val	Tyr	Ala	Ala	Thr
			340					345					350		
Val	Gly	Ala	Pro	Gly	Leu	Thr	Gly	Ala	Leu	Ala	Arg	Arg	Gly	Leu	Ala
	355						360					365			
Thr	Lys	Leu	Thr	Asn	Leu	Arg	Glu	Gly	Lys	Gly	Phe	His	His	Pro	Leu
	370					375					380				
Trp	Asp	Arg	Ile	Leu	Phe	Ser	Lys	Val	Lys	Gln	Ala	Leu	Gly	Gly	Asn
385					390					395					400
Val	Arg	Leu	Met	Leu	Thr	Ala	Ser	Ala	Pro	Ile	Ser	Ala	Glu	Ile	Leu
				405					410					415	
Glu	Phe	Val	Arg	Val	Ala	Phe	Cys	Cys	Glu	Val	Val	Glu	Ala	Tyr	Gly
			420					425					430		
Gln	Thr	Glu	Gly	Gly	Gly	Ala	Ala	Thr	Asn	Thr	Val	Ile	Gly	Glu	Thr
	435						440					445			
Glu	Ala	Gly	His	Val	Gly	Pro	Pro	Gln	Ala	Cys	Cys	Glu	Ile	Lys	Leu
	450					455					460				
Val	Asp	Val	Pro	Glu	Leu	Asn	Tyr	Phe	Ala	Thr	Asp	Lys	Pro	Phe	Pro
465					470					475					480

Arg	Gly	Glu	Ile	Cys	Val	Arg	Gly	Pro	Gly	Val	Ile	Pro	Gly	Tyr	Leu
				485					490					495	
Lys	Asp	Glu	Ala	Lys	Thr	Lys	Glu	Thr	Ile	Asp	Glu	Glu	Gly	Trp	Leu
			500					505					510		
His	Ser	Gly	Asp	Ile	Ala	Ile	Met	Ser	Gly	Lys	Gly	Thr	Val	Thr	Ile
		515					520					525			
Ile	Asp	Arg	Lys	Lys	Asn	Val	Phe	Lys	Leu	Ser	Gln	Gly	Glu	Tyr	Ile
	530					535					540				
Ala	Ala	Glu	Asn	Ile	Glu	Gly	Arg	Phe	Leu	Ser	Lys	Val	Pro	Phe	Ile
545					550					555					560
Gln	Gln	Ile	Leu	Val	His	Gly	Asp	Ser	Thr	Glu	Ser	Cys	Leu	Val	Ala
			565					570						575	
Ile	Leu	Ile	Pro	Glu	Pro	Glu	Ala	Phe	Ile	Pro	Phe	Val	Asn	Lys	Val
			580					585					590		
Leu	Glu	Asn	Val	Asn	Leu	Gln	Pro	Gly	Asp	Leu	Ala	Ala	Tyr	Arg	Lys
		595				600					605				
Ile	Val	Asn	Asn	Pro	Lys	Leu	Arg	Gln	Ala	Val	Leu	Lys	Glu	Leu	Ile
	610					615					620				
Lys	Ala	Gly	Lys	Asp	Ala	Gly	Leu	Lys	Gly	Phe	Glu	Ile	Pro	Lys	Ala
625					630					635					640
Ile	Leu	Leu	Glu	Ser	Glu	Ala	Phe	Thr	Val	Glu	Asn	Asp	Lys	Met	Thr
			645					650						655	
Pro	Thr	Phe	Lys	Ile	Lys	Arg	His	Pro	Val	Val	Gln	Ala	Tyr	Arg	Glu
			660					665					670		
Gln	Leu	Thr	Ala	Leu	Tyr	Asn	Glu	Ile	His	Gln	Lys	Glu	Ser	Lys	Leu
		675					680					685			

<210> 18

<211> 2067

<212> DNA

<213> Mortierella alpina

<400> 18

atgtccctcg	accagaacgc	ccagtcctgt	gagctcccag	gcacccggca	accaggccag	60
acaggcatct	atcgccgcaa	aggcttcgag	aatgcccttc	tcgccgtccc	accagcaga	120
ccgcacatca	agaccatcta	cgatgccttc	cagcacggac	tgaagcttaa	tcccaacgga	180
gctgccctgg	gcagccgagt	gtacgacccg	gtgacggaca	cctttggagg	ctatgtctgg	240
cagacgtatg	cacaggtgaa	cgaccgcctc	actcgcttcg	gcagtggatt	gggtcaaaatt	300
cataaggacg	tccatggtct	tgccaccgtg	ggccagaagt	ggtctctcgg	aatctgggccc	360
atcaaccgac	ccgagtggac	catcgcgctc	gaggettgct	cggcctacaa	cctgggtctcc	420
gtgggtcttt	acgatacttt	gggacccgag	gctgtgactt	atggcattaa	tcacgctgag	480
tgtctctattg	tcgtaacaag	tgtggatcat	atcgcgacgc	tgtctgaacga	atcttccaag	540
atgcctgggc	tcaaaatcat	catcagcatg	gatgacctcg	atactgggag	agcaggccca	600
ggactggctc	ccaccggcac	catcctcagg	acttacgctc	aggacaaagg	ggtactactt	660
tatgattggt	ctgaggttga	agccgtcggg	attcagcatg	gacgaaagca	tacgccacca	720
acctctctcg	acgcatatac	gatctgctat	accagcggga	caacaggctt	gccaaaaggt	780
gccattttga	cccatggaaa	cttgatcgcc	cttttgccct	ccagtgatgt	ggccacacca	840
gtgctggctg	acgattgcct	catcagtttc	ttgcccctgc	ctcacgtctt	tggtcgggtc	900
atggagctct	tcgcgatggc	cgcaggagga	aagattggct	acagcacggg	agatcctttg	960
cgtctcttgg	aggacgtctc	gcacctaaag	ccctccatct	tccccgctgt	gcccagactg	1020
ctgaaccgcg	tgtatgccaa	ggtgtatgcg	gcaactgttg	gagcgcttgg	actcacaggg	1080
gcactggcgc	gacgaggatt	ggccaccaag	ctcaccaatt	tgagagaggg	caaagggttc	1140
caccacccat	tgtgggaccg	aatcctcttc	tcaaagggtca	agcaagcgct	cggcggcaat	1200
gtgagactga	tgttgactgc	ctccgctccc	atctcgcccg	agatcttgga	attcgtccgt	1260
gtcgctttct	gctgcgaggt	cgtggaggca	tatggacaga	ctgagggcgg	tggagcggcc	1320
acaaacaccg	tgattggcga	gaccgaggct	ggacacgtcg	gtcctcctca	agcttggttc	1380

gagatcaaac	tggtggatgt	acccgagctg	aactactttg	cgaccgataa	accattccct	1440
cgtggtgaga	tttgtgtcog	tggacccggt	gtcattcctg	gttatctcaa	ggatgaggcc	1500
aagaccaagg	agaccattga	tgaggagggc	tggtgcact	cgggcgatat	cgccatcatg	1560
agtggcaaag	gcacogttac	catcattgac	aggaagaaga	acgtgttcaa	gctgagccaa	1620
ggagaataca	tcgcggcaga	gaacattgaa	gggcgtttcc	tctccaaggt	tccattcatc	1680
caacaaattc	tggtgcacgg	cgactcgacc	gagagctggt	tggtggccat	cttgatccca	1740
gagcctgagg	ccttcatccc	ctttgtgaac	aaagtgtctg	agaacgtcaa	tcttcaacct	1800
ggagatcttg	cagcctacag	gaagatcggt	aacaacccaa	agctgcgcca	ggctgtcctc	1860
aaagagctga	tcaaggctgg	caaggatgct	ggattgaaag	gctttgagat	tccaaaggcg	1920
atcctcctcg	aatctgaggc	attcacggtc	gaaaacgaca	agatgacccc	gactttcaag	1980
atcaaaagac	accctgtcgt	ccaggcttac	cgcgagcaac	tgacagccct	ctacaacgaa	2040
atccatcaaa	aggaatccaa	gctgtaa				2067

<210> 19

<211> 2309

<212> DNA

<213> Mortierella alpina

<400> 19

cacacgtca	cgttcgtct	cacccgaccc	actccccact	ctcgtctca	ttctctccct	60
tgtccttccc	ttgtcccttt	caaggtctaa	cagcatcaac	atcagcatca	gcataagct	120
tctcattcct	ccctcgtcta	aatctctgaa	agagttcgt	ttgcaattca	gcaatgtccc	180
tcgaccagaa	cgcccagtc	gttgagctcc	caggcaccgc	gcaaccaggc	cagacaggca	240
tctatcgccg	caaaggcttc	gagaatgcc	ttctcgccgt	cccaccagc	agaccgcaca	300
tcaagaccat	ctacgatgcc	ttccagcacg	gactgaagct	taatcccaac	ggagctgccc	360
tgggcagccg	agtgtacgac	ccggtgacgg	acacctttgg	aggctatgtc	tggcagacgt	420
atgcacaggt	gaacgaccgc	atcactcgct	tcggcagtg	attggtcaaa	attcataagg	480
acgtccatgg	tcttgccacc	gtgggcccaga	agtggctct	cggaatctgg	gccatcaacc	540
gacccgagtg	gaccatcgcg	tccgaggctt	gctcggccta	caacctggtc	tccgtgggtc	600
tttacgatac	tttgggaccc	gaggctgtga	cttatggcat	taatcacgct	gagtgtctta	660
ttgtcgtaac	aagtgtggat	catatcgoga	cgctgtctgaa	cgaatcttcc	aagatgcctg	720
ggctcaaaat	catcatcagc	atggatgacc	tcgatactgg	gagagcaggc	ccaggactgg	780
ctcccaccgg	caccatcctc	aggacttacg	ctcaggacaa	aggggtacta	ctttatgatt	840
ggtctgaggt	tgaagccgtc	ggtattcagc	atggacgaaa	gcatacgcca	ccaacctcct	900
ccgacgcata	tacgatctgc	tataccagcg	ggacaacagg	cttgccaaaa	ggtgccattt	960
tgacccatgg	aaacttgatc	gcccttttgg	cctccagtga	tgtggccaca	ccagtgtctg	1020
ctgacgattg	cctcatcagt	ttcttgcccc	tgcctcacgt	ctttggtcgg	gtcatggagc	1080
tcttcgcgat	ggccgcagga	ggaaagattg	gctacagcac	gggagatcct	ttgcgtctct	1140
tggaggacgt	ctcgcaacct	aagccctoca	tcttccccgc	tgtgccccaga	ctgctgaacc	1200
gcgtgtatgc	caaggtgtat	gcggcaactg	ttggagcgcc	tggactcaca	ggggcactgg	1260
cgcgacgagg	attggccacc	aagctcacca	atttgagaga	gggcaaagg	ttccaccacc	1320
cattgtggga	ccgaatcctc	ttctcaaagg	tcaagcaagc	gctcggcggc	aatgtgagac	1380
tgatgttgac	tgcctccgct	cccatctcgg	ccgagatctt	ggaattcgtc	cgtgtcgctt	1440
tctgctgcga	ggtcgtggag	gcataatggac	agactgagg	cggtggagcg	gccacaaaca	1500
ccgtgattgg	cgagaccgag	gctggacacg	tcggtcctcc	tcaagcttgt	tgcgagatca	1560
aactggtgga	tgtaccgag	ctgaactact	ttgcgacoga	taaaccattc	cctcgtggtg	1620
agatttgtgt	ccgtggaccc	ggtgtcattc	ctgggttatct	caaggatgag	gccaaagacca	1680
aggagaccat	tgatgaggag	ggctggctgc	actcgggcga	tatcgccatc	atgagtggca	1740
aaggcacctg	taccatcatt	gacaggaaga	agaacgtgtt	caagctgagc	caaggagaat	1800
acatcgcggc	agagaacatt	gaagggcgtt	tctctctcaa	ggttccattc	atccaacaaa	1860
ttctgggtgca	cggcgactcg	accgagagct	gtttggtggc	catcttgatc	ccagagcctg	1920
aggccttcat	cccctttgtg	aacaaagtgc	tcgagaacgt	caatcttcaa	cctggagatc	1980
ttgcagccta	caggaagatc	gttaacaacc	caaagctgcg	ccaggctgtc	ctcaaagagc	2040
tgatcaaggc	tggcaaggat	gctggattga	aaggctttga	gattccaaag	gcgatccctc	2100
tcgaatctga	ggcattcacg	gtcgaaaacg	acaagatgac	cccactttc	aagatcaaaa	2160
gacaccctgt	cgtccaggct	taccgcgagc	aactgacagc	cctctacaac	gaaatccatc	2220

aaaaggaatc caagctgtaa aaagaaaccc ttagaacctg cgggtgctcgc agcaattaaa 2280
 aaaaaaagag agatattact ctcacagct 2309

<210> 20
 <211> 2801
 <212> DNA
 <213> Mortierella alpina

<400> 20
 atgtccctcg accagaacgc ccagtcctgt gagctcccag gcacccggca accaggccag 60
 acaggtaaca caggccagca gtactgcacc cgttttccat taaggttcca accccagtct 120
 tggacatggg gcgagacctg cttgttggtta ctgtggattc acaccccgcg cgccccttcc 180
 cttctggtgc cttgacctgt gtgacacccc gcaaactgct cagctcttct cggetgacca 240
 accttcattc acatccacgt aggcattctat cgcgcgaaag gcttcgagaa tgcccttctc 300
 gccgtcccac ccagcagacc gcacatcaag accatctacg atgccttcca gcacggactg 360
 aagcttaatc ccaacggagc tgccctgggc agccgagtgt acgacccggt gacggacacc 420
 tttggaggct atgtctggca gacgtatgca cagggtgaac accgcattac tcgcttcggc 480
 agtggattgg tcaaaattca taaggacgtc catggtcttg ccaccgtggg ccagaagtgg 540
 tctctcggaa tctgggccat caaccgaccc gagtggacca tcgcgtccga ggcttgctcg 600
 gcctacaacc tgggtctcgt ggggtctttac gatactttgg gacccgaggc tgtgacttat 660
 ggcattaatc acgctgagtg ctctattgtc gtaacaagtg gtaaggacat gaagccataa 720
 cgacaacggc taaaaaaaa catggttctc atgacaagta tcgatagtaa cgttattctt 780
 ggcgcttgta tgtgttttcta gtggatcata tcgcgacgct gctgaacgaa tcttccaaga 840
 tgccctgggt caaaatcacc atcagcatgg atgacctcga tactgggaga gcaggcccag 900
 gactggctcc caccggcacc atcctcagga cttacgctca ggacaaaggg gtactacttt 960
 atgattggtc tgaggttgaa gccgtcggta ttcagcatgg acgaaagcat acgccaccaa 1020
 cctcctccga cgcataatcg atctgctata ccagcgggac aacaggcttg ccagtaatat 1080
 gttgctttta ttcccacgca tacagtgtgc cgatattttc aatgttcaat tgctctcatt 1140
 agatgcatga cacttatcat tacttaagcg actttccctt ggcgttcata gaaagggtgc 1200
 attttgaccc atggaaactt gatcgccctt ttggcctcca gtgatgtggc cacaccagtg 1260
 ctggctgacg attgcctcat cagtttcttg cccctgcctc acgtcttttg tcgggtcatg 1320
 gagctcttcg cgatggccgc aggaggaaag attggctaca gcacgggaga tcctttgcgt 1380
 ctcttgaggg acgtctcgca cctaaagccc tccatcttcc ccgctgtgcc cagactgctg 1440
 aaccgcgtgt atgccaaggt gtatgcggca actggtggag cgcctggact cacaggggca 1500
 ctggcgcgac gaggattggc caccaagctc accaatttga gagagggcaa aggtttccac 1560
 caccattgtt gggaccgaat cctcttctca aaggtaagc aagcgtcgg cggaatgtg 1620
 agactgatgt tgactggtaa gtgtgctttt ggaagatgaa atcacgttta tgtaaccccc 1680
 ccccccccc cttgtttata aacattaatc gttgtactgt cgtcgtctag cctccgctcc 1740
 catctcggcc gagatcttgg aattcgtccg tgctcgtttc tgctgcgagg tcgtggaggc 1800
 atatggacag actgagggcg gtggagcggc caciaacacc gtgattggcg agaccgagc 1860
 tggacacgtc ggtcctcctc aagcttggtg cgagatcaaa ctgggtggatg tacccgagct 1920
 gaactacttt gcgaccgata aaccattccc tcgtggtgag atttgtgtcc gtggaccgag 1980
 tgtcattcct ggttatctca aggatgaggg caagaccaag gagaccattg atgaggaggg 2040
 ctggctgcac tcgggcgata tcgccatcat gagtggcaaa ggcaccgtta ccatcattga 2100
 caggaagaag aacgtgttca aggtaacaaa aacagtgcgc tctccacaga tctgttagcg 2160
 cgctttttct gcgcacagta cactgaaacc accctgtttt gctttgttcc tgaactagct 2220
 gagccaagga gaatacatcg cggcagagaa cattgaaggg cgtttcctct ccaaggttcc 2280
 attcatccaa caaattctgg tgcacggcga ctcgaccgag agctgttttg tggccatctt 2340
 gatcccagag cctgaggcct tcatcccctt tgtgaacaaa gtgctcgaga acgtcaatct 2400
 tcaacctgga gatcttgag cctacaggaa gatcgttaac aacccaaagc tgcgccaggc 2460
 tgtcctcaaa gagctgatca aggttggtgaa ggatgctgga ttgaaagggt acgtacagaa 2520
 ctctcttttg taccgtagcg gaggccact agagttgagg tgatacagat cgacagaaaa 2580
 aaaaaaaaaa ctaaacaatc tctcttcaaa aacttggtgt tctcttgtac cacagctttg 2640
 agattccaaa ggcgatcctc ctcgaaatct aggcattcac ggctgaaaaa gacaagatga 2700
 ccccgacttt caagatcaaa agacaccctg tcgtccaggc ttaccgagag caactgacag 2760
 ccctctacaa cgaaatccat caaaagggaat ccaagctgta a 2801

<210> 21
 <211> 1977
 <212> DNA
 <213> Mortierella alpina

<400> 21
 atgaccaccc aattgtactc catcgaagtg gcaggcagcc cagagattcc gggcgagggc 60
 aaacctcggc gcagcgttct cagcccagac aaactcgtcc agagctatca gtctttcaag 120
 ggcgacggct ccatcaccac tctatatgag aacttttttg agggcatcca gcgctcagag 180
 ggaggagagt ttctcggaca ccgccccatc gtcgataatg tagctcagcc gtacgaatgg 240
 ctaagctaca cgcgcgttca ggaacgtgtc gccaaactttg gcgctgggtc catccagctg 300
 ggctgaaag tcgactcgaa ctttggcacc ttttccatca acaggcccga atggacaatg 360
 agtgagctgg caggctacat gtacaacttt acatctgtgc cgctttacga cactctgggc 420
 gtctcggcca togaatacat cgttaatcag accgagatgg agaccatcat cgcgtcggct 480
 gataaagcct cgatcctgtt aaacatgaaa tcaactctgc cgacactcaa gaacattgtc 540
 gttatgggct cgctcgaaga cgcgctcgtt gtcgagggta gggaaatgga tatccacatc 600
 gttgcgtgga gtgacgtcga acgcgatggc ttcaacaacc ccgcgccagc caacctcca 660
 acaccggacg acgtcgccac catctgctac acgtcaggaa caaccgggac accaaagggc 720
 gcaatcctga cccacaaaaa ctttgtggct ggcttgcct cgttccatat gatggcaaag 780
 caccaaaagt ttttcatccc ctgcagcgtt gacactcaca tatcttacct gccctggca 840
 catgtgttcg agcgtttgtc tcaggctgtt atgatttctg gcgcagctcg gattgggtat 900
 taccaaggag acactttgaa gctactcgat gatgtggcga tcttgcagcc caccatcttt 960
 gtgtccgttc cacgactctt taacaggatt tacgacaagg ttctagcagg tgtgaaagcc 1020
 aaggcggtc tcgcagcttt cttattcaac cgcgcttttg aaaccaagaa ggctaatttg 1080
 aaacgcggta tcctggagca cgccatctgg gatcgactgg tatttgggtc aattcgtgcg 1140
 cgactcggtg gcaaagttaa gcatattgtc tcaggatcag cccctatagc cccggacgtc 1200
 atggatttcc ttcgcatttg cttcagtgcc gacgtttatg aagggtatgg acagacggag 1260
 caggctgctg gtttgtgtat gagctacaga ggtgacttga cctcgggtca agtgggaccc 1320
 cctcagctgt gcgtcgaagt gaagctcaga gacgttccgg acatgcacta cacaagccag 1380
 gacaagcctc gccctcgcgg ggagatcatg cttcgaggcc attcagtttt caaaggctat 1440
 tacaaggctc caaagcaaac agaggagaca ctggacgcac agggatgggc aagcactgga 1500
 gacgttgggtg aatgggacga gcgtggccgc ttggtggtga tcgaccgtgt caaaaacatt 1560
 ttcaagttgg ctcaaggcga atacattgca cctgaaaaga tcgaagccgt cctggccaaa 1620
 cactaccttg tcgccaggt ctttgtctac ggagactcct tccaagcgac attggtggga 1680
 gttgtcgtgc ccgatgcgga gacgctaaag ccttggggccg atgaccatgg ccttggaggc 1740
 aagagctatg aagaactatg cgctcatccc gctgtcaaag aaactttgct gaaggagctc 1800
 aaagagtttg gtcgtgaaaa tgatctgaag ggctttgaga tattgaagaa cattcatgta 1860
 acggcggagc aattctcaat tgagaatgat cttttgacac ccacattcaa gctgaagaga 1920
 cacaccgcga aagagaagta catcgccgag attgagctga tgtataacgg gatccac 1977

<210> 22
 <211> 659
 <212> PRT
 <213> Mortierella alpina

<400> 22
 Met Thr Thr Gln Leu Tyr Ser Ile Glu Val Ala Gly Ser Pro Glu Ile
 1 5 10 15
 Pro Gly Glu Gly Lys Pro Arg Arg Ser Val Leu Ser Pro Asp Lys Leu
 20 25 30
 Val Gln Ser Tyr Gln Ser Phe Lys Gly Asp Gly Ser Ile Thr Thr Leu
 35 40 45
 Tyr Glu Asn Phe Leu Glu Gly Ile Gln Arg Ser Glu Gly Gly Glu Phe
 50 55 60
 Leu Gly His Arg Pro Ile Val Asp Asn Val Ala Gln Pro Tyr Glu Trp
 65 70 75 80

Leu	Ser	Tyr	Thr	Arg	Val	Gln	Glu	Arg	Val	Ala	Asn	Phe	Gly	Ala	Gly	
				85					90					95		
Leu	Ile	Gln	Leu	Gly	Leu	Lys	Val	Asp	Ser	Asn	Phe	Gly	Ile	Phe	Ser	
			100					105					110			
Ile	Asn	Arg	Pro	Glu	Trp	Thr	Met	Ser	Glu	Leu	Ala	Gly	Tyr	Met	Tyr	
		115					120					125				
Asn	Phe	Thr	Ser	Val	Pro	Leu	Tyr	Asp	Thr	Leu	Gly	Val	Ser	Ala	Ile	
	130					135					140					
Glu	Tyr	Ile	Val	Asn	Gln	Thr	Glu	Met	Glu	Thr	Ile	Ile	Ala	Ser	Ala	
145					150					155					160	
Asp	Lys	Ala	Ser	Ile	Leu	Leu	Asn	Met	Lys	Ser	Thr	Leu	Pro	Thr	Leu	
			165					170						175		
Lys	Asn	Ile	Val	Val	Met	Gly	Ser	Leu	Glu	Asp	Ala	Leu	Val	Val	Glu	
			180					185					190			
Gly	Arg	Glu	Met	Asp	Ile	His	Ile	Val	Ala	Trp	Ser	Asp	Val	Glu	Arg	
	195					200						205				
Asp	Gly	Phe	Asn	Asn	Pro	Ala	Pro	Ala	Asn	Pro	Pro	Thr	Pro	Asp	Asp	
	210					215					220					
Val	Ala	Thr	Ile	Cys	Tyr	Thr	Ser	Gly	Thr	Thr	Gly	Thr	Pro	Lys	Gly	
225					230					235					240	
Ala	Ile	Leu	Thr	His	Lys	Asn	Phe	Val	Ala	Gly	Leu	Ala	Ser	Phe	His	
			245					250						255		
Met	Met	Ala	Lys	His	Gln	Lys	Phe	Phe	Ile	Pro	Ser	Ser	Val	Asp	Thr	
			260					265					270			
His	Ile	Ser	Tyr	Leu	Pro	Leu	Ala	His	Val	Phe	Glu	Arg	Leu	Ser	Gln	
	275						280					285				
Ala	Val	Met	Ile	Ser	Gly	Ala	Ala	Arg	Ile	Gly	Tyr	Tyr	Gln	Gly	Asp	
	290					295					300					
Thr	Leu	Lys	Leu	Leu	Asp	Asp	Val	Ala	Ile	Leu	Gln	Pro	Thr	Ile	Phe	
305					310					315					320	
Val	Ser	Val	Pro	Arg	Leu	Phe	Asn	Arg	Ile	Tyr	Asp	Lys	Val	Leu	Ala	
			325					330						335		
Gly	Val	Lys	Ala	Lys	Gly	Gly	Leu	Ala	Ala	Phe	Leu	Phe	Asn	Arg	Ala	
			340					345					350			
Phe	Glu	Thr	Lys	Lys	Ala	Asn	Leu	Lys	Arg	Gly	Ile	Leu	Glu	His	Ala	
	355						360					365				
Ile	Trp	Asp	Arg	Leu	Val	Phe	Gly	Ala	Ile	Arg	Ala	Arg	Leu	Gly	Gly	
	370					375					380					
Lys	Val	Lys	His	Ile	Val	Ser	Gly	Ser	Ala	Pro	Ile	Ala	Pro	Asp	Val	
385					390					395				400		
Met	Asp	Phe	Leu	Arg	Ile	Cys	Phe	Ser	Ala	Asp	Val	Tyr	Glu	Gly	Tyr	
			405						410					415		
Gly	Gln	Thr	Glu	Gln	Ala	Ala	Gly	Leu	Cys	Met	Ser	Tyr	Arg	Gly	Asp	
		420					425						430			
Leu	Thr	Ser	Gly	Gln	Val	Gly	Pro	Pro	Gln	Leu	Cys	Val	Glu	Val	Lys	
	435						440					445				
Leu	Arg	Asp	Val	Pro	Asp	Met	His	Tyr	Thr	Ser	Gln	Asp	Lys	Pro	Arg	
	450					455					460					
Pro	Arg	Gly	Glu	Ile	Met	Leu	Arg	Gly	His	Ser	Val	Phe	Lys	Gly	Tyr	
465					470					475					480	
Tyr	Lys	Ala	Pro	Lys	Gln	Thr	Glu	Glu	Thr	Leu	Asp	Ala	Gln	Gly	Trp	
			485						490					495		
Ala	Ser	Thr	Gly	Asp	Val	Gly	Glu	Trp	Asp	Glu	Arg	Gly	Arg	Leu	Val	
		500						505				510				
Val	Ile	Asp	Arg	Val	Lys	Asn	Ile	Phe	Lys	Leu	Ala	Gln	Gly	Glu	Tyr	
	515						520					525				

Ile Ala Pro Glu Lys Ile Glu Ala Val Leu Ala Lys His Tyr Leu Val
 530 535 540
 Ala Gln Val Phe Val Tyr Gly Asp Ser Phe Gln Ala Thr Leu Val Gly
 545 550 555 560
 Val Val Val Pro Asp Ala Glu Thr Leu Lys Pro Trp Ala Asp Asp His
 565 570 575
 Gly Leu Gly Gly Lys Ser Tyr Glu Glu Leu Cys Ala His Pro Ala Val
 580 585 590
 Lys Glu Thr Leu Leu Lys Glu Leu Lys Glu Phe Gly Arg Glu Asn Asp
 595 600 605
 Leu Lys Gly Phe Glu Ile Leu Lys Asn Ile His Val Thr Ala Glu Gln
 610 615 620
 Phe Ser Ile Glu Asn Asp Leu Leu Thr Pro Thr Phe Lys Leu Lys Arg
 625 630 635 640
 His Thr Ala Lys Glu Lys Tyr Ile Ala Glu Ile Glu Leu Met Tyr Asn
 645 650 655
 Gly Ile His

<210> 23
 <211> 1980
 <212> DNA
 <213> Mortierella alpina

<400> 23
 atgaccaccc aattgtactc catcgaagtg gcaggcagcc cagagattcc gggcgagggc 60
 aaacctcggc gcagcggttct cagcccagac aaactcgtcc agagctatca gtctttcaag 120
 ggcgacggct ccatcaccac tctatatgag aacttttttg agggcatcca gcgctcagag 180
 ggaggagagt ttctcggaca ccgccccatc gtcgataatg tagctcagcc gtacgaatgg 240
 ctaagctaca cgcgcgttca ggaacgtgtc gccaaacttt gcgctgggtct catccagctg 300
 ggctgaaag tcgactcgaa ctttggcac ttttccatca acaggcccga atggacaatg 360
 agtgagctgg caggctacat gtacaacttt acatctgtgc cgctttacga cactctgggc 420
 gtctcggcca tcgaatacat cgttaatcag accgagatgg agaccatcat cgcgtcggct 480
 gataaagcct cgatcctggt aaacatgaaa tcaactctgc cgacactcaa gaacattgtc 540
 gttatgggct cgctcgaaga cgcgctcgtt gtcgagggta gggaaatgga tatccacatc 600
 gttgcgtgga gtgacgtcga acgcgatggc ttcaacaacc ccgcgcacgc caacctcca 660
 acaccggacg acgtcgccac catctgctac acgtcaggaa caaccgggac accaaagggc 720
 gcaatcctga cccacaaaaa ctttgtggct ggccttgcct cgttccatat gatggcaaag 780
 caccaaaagt ttttcatccc ctgcagcgtt gacactcaca tatcttacct gccctggca 840
 catgtgttcg agcgtttgtc tcaggctgtt atgatttctg gcgcagctcg gattgggtat 900
 taccaaggag acactttgaa gctactcgat gatgtggcga tcttgacgcc caccatcttt 960
 gtgtccgttc cactactctt taacaggatt tacgacaagg ttctagcagg tgtgaaagcc 1020
 aaggcggtc tcgcagcttt cttattcaac cgcgcttttg aaaccaagaa ggctaatttg 1080
 aaacgcggta tcctggagca cgccatctgg gatcgactgg tatttgggtgc aattcgtgcg 1140
 cgactcggtg gcaaagttaa gcatattgtc tcaggatcag cccctatagc cccggacgtc 1200
 atggatttcc ttgcatttg cttcagtgcc gacgtttatg aagggtatgg acagacggag 1260
 caggetgctg gtttgtgtat gagctacaga ggtgacttga cctcgggtca agtgggaccc 1320
 cctcagctgt gcgtcgaagt gaagctcaga gacgttcagg acatgcacta cacaagccag 1380
 gacaagcctc gccctcgcgg ggagatcatg cttcgaggcc attcagtttt caaaggctat 1440
 tacaaggctc caaagcaaac agaggagaca ctggacgcac agggatgggc aagcactgga 1500
 gacgttgggtg aatgggacga gcgtggccgc ttggtggtga tcgaccgtgt caaaaacatt 1560
 ttcaagttgg ctcaaggcga atacattgca cctgaaaaga tcgaagccgt cctggccaaa 1620
 cactaccttg tcgccaggt ctttgtctac ggagactcct tccaagcgac attggtggga 1680
 gttgtcgtgc ccgatgcgga gacgctaaag ccttgggccc atgaccatgg ccttggaggc 1740
 aagagctatg aagaactatg cgctcatccc gctgtcaaag aaactttgct gaaggagctc 1800
 aaagagtttg gtcgtgaaaa tgatctgaag ggctttgaga tattgaagaa cattcatgta 1860

acggcggagc aattctcaat tgagaatgat cttttgacac ccacattcaa gctgaagaga 1920
cacaccgcga aagagaagta catcgccgag attgagctga tgtataacgg gatccactga 1980

<210> 24
<211> 2196
<212> DNA
<213> *Mortierella alpina*

<400> 24
tttttttttt tttcttttct ctccaaccct ttcaccccca cgctcggct cgtactcaag 60
cctcacgtcc acactctcgt cctctagcct gctgcattca cgattcacat tcctcctcga 120
ctccagcatc gctactcctt cgtgctactt tcaccatgac cacccaattg tactccatcg 180
aagtggcagg cagcccagag attccggggc agggcaaacc tcggcgcagc gttctcagcc 240
cagacaaact cgtccagagc tatcagtctt tcaagggcga cggtccatc accactctat 300
atgagaactt tttggagggc atccagcgct cagagggagg agagtctctc ggacacccgc 360
ccatcgtcga taatgtagct cagccgtacg aatggctaag ctacacgcgc gttcaggaac 420
gtgtcgccaa ctttggcgct ggtctcatcc agctgggcct gaaagtcgac tcgaactttg 480
gcatcttttc catcaacagg cccgaatgga caatgagtga gctggcaggc tacatgtaca 540
actttacatc tgtgccgctt tacgacactc tgggcgtctc ggccatcgaa tacatcgtta 600
atcagaccga gatggagacc atcatcgctt cggctgataa agcctcgatc ctgttaaaca 660
tgaaatcaac tctgccgaca ctcaagaaca ttgtcgttat gggctcgctc gaagacgcgc 720
tcgttgctga gggtagggaa atggatatcc acatcgttgc gtggagtgaac gtcgaacgcg 780
atggcttcaa caaccccgcg ccagccaacc ctccaacacc ggacgacgtc gccaccatct 840
gctacacgtc aggaacaacc gggacaccaa agggcgcaat cctgaccac aaaaactttg 900
tggctggcct tgcctcgttc catatgatgg caaagcacca aaagttttct atcccctcga 960
gcgctgacac tcacatatct tacctgcccc tggcacatgt gttcgagcgt ttgtctcagg 1020
ctgttatgat ttctggcgca gctcggattg ggtattacca aggagacact ttgaagctac 1080
tcgatgatgt ggcgatcttg cagcccacca tctttgtgtc cgttccacga ctctttaaca 1140
ggatttacga caaggttcta gcaggtgtga aagccaaggg cggctctcga gctttcttat 1200
tcaaccgcgc ttttgaaacc aagaaggcta atttgaaacg cggtatcctg gagcacgcca 1260
tctgggatcg actggatttt ggtgcaattc gtgcgcgact cggtgggcaa gttaagcata 1320
ttgtctcagg atcagccctt atagccccgg acgtcatgga tttccttcgc atttgcttca 1380
gtgccgacgt ttatgaaggg tatggacaga cggagcaggc tgctggtttg tgtatgagct 1440
acagaggtga cttgacctcg ggtcaagtgg gacccctca gctgtgcgtc gaagtgaagc 1500
tcagagacgt tccggacatg cactacacaa gccaggacaa gcctcgccct cgcggggaga 1560
tcatgcttcg aggccattca gttttcaaag gctattacaa ggctccaaag caaacagagg 1620
agacactgga cgcacaggga tgggcaagca ctggagacgt tggatgaatgg gacgagcgtg 1680
gccgcttggt ggtgatcgac cgtgtcaaaa acattttcaa gttggctcaa ggcgaataca 1740
ttgcacctga aaagatcgaa gccgtccttg ccaaacacta ccttgtcgcc caggtctttg 1800
tctacggaga ctcttccaa gcgacattgg tgggagttgt cgtgcccgat gcggagacgc 1860
taaagccttg ggccgatgac catggccttg gaggcaagag ctatgaagaa ctatgcgctc 1920
atcccgtgt caaagaaact ttgctgaagg agctcaaaga gtttggtcgt gaaaatgac 1980
tgaagggctt tgagatattg aagaacattc atgtaacggc ggagcaattc tcaattgaga 2040
atgatctttt gacaccaca ttcaagctga agagacacac cgcgaaagag aagtacatcg 2100
ccgagattga gctgatgtat aacgggatcc actgaaagag tctagccaaa gcagatcttt 2160
ttattactgt cgttaaaaaa actactcgta accatc 2196

<210> 25
<211> 2586
<212> DNA
<213> *Mortierella alpina*

<400> 25
atgaccaccc aattgtactc catcgaagtg gcaggcagcc cagagattcc gggcgagggc 60
aaacctcggc gcagcgttct cagcccagac aaactcgtcc agagctatca gtctttcaag 120

ggcgacggct	ccatcaccac	tctatatgag	aacttttttg	agggcatcca	gcgctcaggt	180
actgacaccc	cacatggtca	gcgagacttc	tccacatgac	accagcgttt	ctctattaac	240
gttatccctt	ggcgcccttg	gctttgttgt	ggttaacgat	gtgcctaata	atagagggag	300
gagagtttct	cggacaccgc	cccatcgctc	ataatgtagc	tcagccgtac	gaatggctaa	360
gctacacgcg	cgttcaggaa	cgtgtcgcca	actttggcgc	tgggtctcat	cagctgggcc	420
tgaaagtcga	ctcgaacttt	ggcatctttt	ccatcaacag	gcccgaatgg	gtgagtgcag	480
tgggaggatc	ctttttcttt	tatgaaatgg	gcccgtgcga	gtctagattc	gatgatggct	540
gcggttggac	ccggcctcgg	cccccaacca	aacacgtgtg	ggaaacccca	accgactcta	600
atagaactcc	gttcctctga	tttggcgacg	aacatactat	tacctgtaga	caatgagtga	660
gctggcaggc	tacatgtaca	actttacatc	tgtgccgctt	tacgacactc	tgggcgtctc	720
ggccatcgaa	tacatcgtta	atcagaccga	gatggagacc	atcatcgctg	cggctgataa	780
agcctcgatc	ctgttaaaca	tgaaatcaac	tctgccgaca	ctcaagaaca	ttgtcgttat	840
gggctcgcct	gaagacgcgc	tcgttgtcga	gggtagggaa	atggatatcc	acatcgttgc	900
gtggagtgcg	gtcgaacgcg	atggcttcaa	caaccccgcg	ccagccaacc	ctccaacacc	960
ggacgacgtc	gccaccatct	gctacacgtc	aggaacaacc	gggacaccaa	agtaacgtta	1020
agatgtcgat	atgagaatat	ctctaaaatg	tggcatgcag	cttgaatgag	aattcactcc	1080
ttcatagaca	cttgttgtcc	tttctcaagt	cactcacctc	tcttcattta	tcctacgtgt	1140
atgtgtatgt	aggggcgcaa	tcctgaccca	caaaaacttt	gtggctggcc	ttgcctcggt	1200
ccatatgatg	gcaaagcacc	aaaagttttt	catcccctcg	agcgttgaca	ctcacatatc	1260
ttacctgccc	ctggcacatg	tgttcgagcg	tttgtctcag	gctgttatga	tttctggcgc	1320
agctcggatt	gggtattacc	aaggagacac	tttgaagcta	ctcgatgatg	tggcgatctt	1380
gcagcccacc	atctttgtgt	cgttccacg	actctttaac	aggatttacg	acaaggttct	1440
agcaggtgtg	aaagccaagg	gcggtctcgc	agctttctta	ttcaaccgcg	cttttgaaac	1500
caagaaggct	aatttgaaac	gcggtatcct	ggagcacgcc	atctgggatc	gactgggtatt	1560
tggtgcaatt	cgtgcgcgac	tcggtggcaa	agttaagcat	attgtctcag	gatcagcccc	1620
tatagccccg	gacgtcatgg	atttccttcg	catttgcttc	agtgccgacg	tttatgaagg	1680
gtatggacag	acggagcagg	ctgctggttt	gtgtatgagc	tacagagggt	acttgacctc	1740
gggtcaagtg	ggaccccctc	agctgtgcgt	cgaagtgaag	ctcagagacg	ttccggacat	1800
gcactacaca	agccaggaca	agcctcgccc	tcgcggggag	atcatgcttc	gaggccattc	1860
agttttcaaa	ggctattaca	aggctccaaa	gcaaacagag	gagacactgg	acgcacaggg	1920
atgggcaagg	tatggttttg	tgcaaccaac	tattttgtca	attacgctta	acactgggtg	1980
tcatcagccc	gcactctgact	aaggtcacat	tgtacgcgac	atagcactgg	agacgttggt	2040
gaatgggacg	agcgtggccg	cttgggtggt	atcgaccgtg	tcaaaaacat	tttcaagttg	2100
gctcaaggcg	aatacattgc	acctgaaaag	atcgaagccg	tcctggccaa	acactacctt	2160
gtcgcgccagg	tctttgtcta	cggagactcc	ttccaagcga	cattgggtggg	agttgtcgtg	2220
cccgatgcgg	agacgctaaa	gccttgggcc	gatgaccatg	gccttggagg	caagagctat	2280
gaagaactat	gcgctcatcc	cgtgtcaaaa	gaaaccttgc	tgaaggagct	caaagagttt	2340
ggtcgtgaaa	atgatctgaa	gggctttgag	atattgaaga	acattcatgt	aacggcggag	2400
caattctcaa	ttgagaatga	tcttttggtg	agtcgatcgg	acttgccgcac	cgaacagttt	2460
gagagtcaag	gatcgtatat	taactaagcc	ttgctttggt	agacaccac	attcaagctg	2520
aagagacaca	ccgcgaaaga	gaagtacatc	gccgagattg	agctgatgta	taacgggatc	2580
cactga						2586

<210> 26

<211> 1977

<212> DNA

<213> Mortierella alpina

<400> 26

atggctactc	aaatgtactc	ggtggctcgtc	cccaacagcc	ccgacattcc	cggcgaaggc	60
aagccccgcc	gtagtgtgct	ttgtccagac	aagctcctgg	agaactaccc	ctcagtga	120
gcaggctcaa	cgatcacgac	cctgtacgag	aacttccaag	aagggtgtct	ccgttcaggc	180
ggcgcccatt	ttttgggcca	tcgtcccatt	gtgaatggcc	agcctcaggc	ttacaagtgg	240
cagtcgtatg	tcgatgtcag	caagcgtgtt	acgcacttcg	gcgctggcct	ggctcatctc	300
ggcttgtctc	caaagcaaaa	ctttggaatt	ttctctatca	accggcctga	gtggaccatg	360
agtgagcttg	ctggctatat	gcacaactac	accagcgctc	ccctctatga	tacattggga	420

```

gtcgccgcga tcgagtatat cgttaaccag actgagatgc agatcatcat tgcttcgtcc 480
gacaaagctt ctatcctcct ccacatgaaa tcagcacttc caaccgttca gacgattgtc 540
gtcatggggg aatttactga cgcctctcgc gcagagggta aggagctcaa catcaacatt 600
gtatcctgga ccgatgtcga aaagagcggg cttgagcggc ctgtcgaagc cgtgcacccc 660
acagccgagg atatecgtac catctgttac acatctggaa ccactggaac gccaaaaggt 720
gctatcttga ccacaagaa ctttgttgcc actatcgctt cattccacat gatggcaaag 780
catggcaggt tcttcattcc ctgcctgcc gacacacatg tatectacct gcccttgcc 840
cacgtctttg agcgctttg ccaggctgtt atgatctcgg gcgctgcgcg tattggttac 900
taccaaggag atacgctgaa gctgctggac gatgttgccg tcctgcatcc caccattttt 960
gcctccgtcc ctgctctctt taaccgtatc tacgacaagg tgcttgctgg cgtcaaggcc 1020
aagggtggta tcgccgcctt cttgtttaac cgcgcatata attccaagaa ggccaacttg 1080
cgaaagggcg tacttgagca tccgctctgg gacaagctgg tctttggagc gattcgcgcg 1140
cgcttggttg gcaagggtta gcacatcgtg tcaggatctg ccccatctc tcctgatgtg 1200
atggatttcc tccgcatctg cttcagcgtc gatgtgtatg agggatatgg ccagacggaa 1260
caggcagccg gattaagtat gagctatcgc ggtgatttga ctccaggaca ggttggtcca 1320
cctcaactgt gcacagaggt caagttgaag gacatcccta gtatgaacta tagcagcgcg 1380
gacaagcctt tccccgtgg agaaatcatg cttcgcgga actctgtgtt caagggctat 1440
tacaaagcac caaagcagac tgaagaaaca ttggatgctg acggttggtc cagtaccgga 1500
gacgttggac agtgggatgc ccaaggccgt ctggtggtca ttgatcgcgt caagaacatc 1560
ttcaagttgg cgcaaggaga atatattgcg cctgaaaaga tcgaggctgt cctcgccaag 1620
cacttctcgt ttgccagat ttttgtctat gggcactcgc tccaggccac cattgtcgcg 1680
gtggttgctc ctgatgctga gacgctcaag ttgtgggcta aagaaaacaa gctgggtgac 1740
aagtcttacg aggagctgtg cgcctcctc cagcttcgca caaccctcca aaaggagttg 1800
gctacttttg gcaaagaatc ggatctgaag ggctttgaga ttcctaagaa cattcatgtt 1860
atctccgagc agttttcaat tgagaacgat cttttgacct ccacctcaa gctgaagaga 1920
catgctgcca aagagaagta taacgccgaa atcgaccgca tgtatgcaga aatcgct 1977

```

<210> 27

<211> 659

<212> PRT

<213> Mortierella alpina

<400> 27

```

Met Ala Thr Gln Met Tyr Ser Val Val Val Pro Asn Ser Pro Asp Ile
1          5          10          15
Pro Gly Glu Gly Lys Pro Arg Arg Ser Val Leu Cys Pro Asp Lys Leu
          20          25          30
Leu Glu Asn Tyr Pro Ser Val Lys Ala Gly Ser Thr Ile Thr Thr Leu
          35          40          45
Tyr Glu Asn Phe Gln Glu Gly Val Leu Arg Ser Gly Gly Ala His Phe
          50          55          60
Leu Gly His Arg Pro Ile Val Asn Gly Gln Pro Gln Ala Tyr Lys Trp
65          70          75          80
Gln Ser Tyr Val Asp Val Ser Lys Arg Val Thr His Phe Gly Ala Gly
          85          90          95
Leu Ala His Leu Gly Leu Ser Pro Lys Gln Asn Phe Gly Ile Phe Ser
          100         105         110
Ile Asn Arg Pro Glu Trp Thr Met Ser Glu Leu Ala Gly Tyr Met His
          115         120         125
Asn Tyr Thr Ser Val Pro Leu Tyr Asp Thr Leu Gly Val Ala Ala Ile
          130         135         140
Glu Tyr Ile Val Asn Gln Thr Glu Met Gln Ile Ile Ile Ala Ser Ser
145         150         155         160
Asp Lys Ala Ser Ile Ile Leu His Met Lys Ser Ala Leu Pro Thr Val
          165         170         175

```

Gln	Thr	Ile	Val	Val	Met	Gly	Glu	Phe	Thr	Asp	Ala	Leu	Val	Ala	Glu	
			180					185					190			
Gly	Lys	Glu	Leu	Asn	Ile	Asn	Ile	Val	Ser	Trp	Thr	Asp	Val	Glu	Lys	
		195					200					205				
Ser	Gly	Leu	Glu	Arg	Pro	Val	Glu	Ala	Val	His	Pro	Thr	Ala	Glu	Asp	
	210					215					220					
Ile	Ala	Thr	Ile	Cys	Tyr	Thr	Ser	Gly	Thr	Thr	Gly	Thr	Pro	Lys	Gly	
225					230					235					240	
Ala	Ile	Leu	Thr	His	Lys	Asn	Phe	Val	Ala	Thr	Ile	Ala	Ser	Phe	His	
				245					250					255		
Met	Met	Ala	Lys	His	Gly	Arg	Phe	Phe	Ile	Pro	Ser	Pro	Ala	Asp	Thr	
		260					265						270			
His	Val	Ser	Tyr	Leu	Pro	Leu	Ala	His	Val	Phe	Glu	Arg	Leu	Cys	Gln	
	275						280					285				
Ala	Val	Met	Ile	Ser	Gly	Ala	Ala	Arg	Ile	Gly	Tyr	Tyr	Gln	Gly	Asp	
	290					295					300					
Thr	Leu	Lys	Leu	Leu	Asp	Asp	Val	Ala	Val	Leu	His	Pro	Thr	Ile	Phe	
305					310					315					320	
Ala	Ser	Val	Pro	Arg	Leu	Phe	Asn	Arg	Ile	Tyr	Asp	Lys	Val	Leu	Ala	
				325					330					335		
Gly	Val	Lys	Ala	Lys	Gly	Gly	Ile	Ala	Ala	Phe	Leu	Phe	Asn	Arg	Ala	
			340				345						350			
Tyr	Asn	Ser	Lys	Lys	Ala	Asn	Leu	Arg	Lys	Gly	Val	Leu	Glu	His	Pro	
	355					360						365				
Leu	Trp	Asp	Lys	Leu	Val	Phe	Gly	Ala	Ile	Arg	Ala	Arg	Leu	Gly	Gly	
	370					375					380					
Lys	Val	Lys	His	Ile	Val	Ser	Gly	Ser	Ala	Pro	Ile	Ser	Pro	Asp	Val	
385					390					395				400		
Met	Asp	Phe	Leu	Arg	Ile	Cys	Phe	Ser	Ala	Asp	Val	Tyr	Glu	Gly	Tyr	
			405						410					415		
Gly	Gln	Thr	Glu	Gln	Ala	Ala	Gly	Leu	Ser	Met	Ser	Tyr	Arg	Gly	Asp	
		420					425						430			
Leu	Thr	Pro	Gly	Gln	Val	Gly	Pro	Pro	Gln	Leu	Cys	Thr	Glu	Val	Lys	
	435						440					445				
Leu	Lys	Asp	Ile	Pro	Ser	Met	Asn	Tyr	Ser	Ser	Ala	Asp	Lys	Pro	Phe	
	450					455					460					
Pro	Arg	Gly	Glu	Ile	Met	Leu	Arg	Gly	Asn	Ser	Val	Phe	Lys	Gly	Tyr	
465					470					475					480	
Tyr	Lys	Ala	Pro	Lys	Gln	Thr	Glu	Glu	Thr	Leu	Asp	Ala	Asp	Gly	Trp	
			485						490					495		
Ser	Ser	Thr	Gly	Asp	Val	Gly	Gln	Trp	Asp	Ala	Gln	Gly	Arg	Leu	Val	
		500					505						510			
Val	Ile	Asp	Arg	Val	Lys	Asn	Ile	Phe	Lys	Leu	Ala	Gln	Gly	Glu	Tyr	
	515					520						525				
Ile	Ala	Pro	Glu	Lys	Ile	Glu	Ala	Val	Leu	Ala	Lys	His	Phe	Leu	Val	
	530					535					540					
Ala	Gln	Ile	Phe	Val	Tyr	Gly	His	Ser	Leu	Gln	Ala	Thr	Ile	Val	Ala	
545					550					555					560	
Val	Val	Val	Pro	Asp	Ala	Glu	Thr	Leu	Lys	Leu	Trp	Ala	Lys	Glu	Asn	
			565						570					575		
Lys	Leu	Gly	Asp	Lys	Ser	Tyr	Glu	Glu	Leu	Cys	Ala	Leu	Pro	Gln	Leu	
		580					585						590			
Arg	Thr	Thr	Leu	Gln	Lys	Glu	Leu	Ala	Thr	Phe	Gly	Lys	Glu	Ser	Asp	
	595					600						605				
Leu	Lys	Gly	Phe	Glu	Ile	Pro	Lys	Asn	Ile	His	Val	Ile	Ser	Glu	Gln	
	610					615					620					

Phe Ser Ile Glu Asn Asp Leu Leu Thr Pro Thr Phe Lys Leu Lys Arg
 625 630 635 640
 His Ala Ala Lys Glu Lys Tyr Asn Ala Glu Ile Asp Arg Met Tyr Ala
 645 650 655
 Glu Ile Ala

<210> 28
 <211> 1980
 <212> DNA
 <213> Mortierella alpina

<400> 28
 atggctactc aaatgtactc ggtgggtcgtc cccaacagcc cggacattcc cggcgaaggc 60
 aagccccgcc gtagtgtgct ttgtccagac aagctcctgg agaactaccc ctcagtgaag 120
 gcaggctcaa cgatcacgac cctgtacgag aacttccaag aagggtgttct ccgttcaggc 180
 ggcgcccatt ttttgggcca tcgtcccatt gtgaatggcc agcctcaggc ttacaagtgg 240
 cagtcgtatg tcgatgtcag caagcgtggt acgcacttcg gcgctggcct ggctcatctc 300
 ggcttgtctc caaagcaaaa ctttggaatt ttctctatca accggcctga gtggaccatg 360
 agtgagcctt ctggctatat gcacaactac accagcgtcc cctctatga tacattggga 420
 gtcgcgcgca tcgagtatat cgttaaccag actgagatgc agatcatcat tgcttcgtcc 480
 gacaaagctt ctatcactct ccacatgaaa tcagcacttc caaccgttca gacgattgtc 540
 gtcattgggg aatttactga cgctctcgtc gcagagggta aggagctcaa catcaacatt 600
 gtatcctgga ccgatgtcga aaagagcggg cttgagcggc ctgtcgaagc cgtgcacccc 660
 acagccgagg atatcgctac catctgttac acatctggaa ccactggaac gccaaaaggt 720
 gctatcttga cccacaagaa ctttgttgcc actatcgctt cattccacat gatggcaaag 780
 catggcagggt tcttcattcc ctgcctgcc gacacacatg tatectacct gccccttgcc 840
 cacgtctttg agcgcctttg ccaggctggt atgatctcgg gcgctgcgcg tattgggttac 900
 taccaaggag atacgctgaa gctgctggac gatgttgccg tcctgcatcc caccattttt 960
 gcctccgtcc ctgctctctt taaccgtatc tacgacaagg tgcttgctgg cgtcaaggcc 1020
 aagggtggta tcgcgcctt cttgtttaac cgcgcatata attccaagaa ggccaacttg 1080
 cgaaaggggc tacttgagca tccgctctgg gacaagctgg tctttggagc gattcgcgcg 1140
 cgcttggttg gcaagggtta gcacatcgtg tcaggatctg ccccatctc tcctgatgtg 1200
 atggatttcc tccgcatctg cttcagcgtc gatgtgtatg agggatatgg ccagacggaa 1260
 caggcagccg gattaagtat gagctatcgc ggtgatttga ctccaggaca ggttggccca 1320
 cctcaactgt gcacagagggt caagttgaag gacatcccta gtatgaacta tagcagcgcg 1380
 gacaagcctt tccccgtgg agaaatcatg cttcgcggaa actctgtgtt caagggtat 1440
 taaaagcac caaagcagac tgaagaaaca ttggatgctg acggttggtc cagtaccgga 1500
 gacgttggac agtgggatgc ccaaggccgt ctggtggtca ttgatcgcgt caagaacatc 1560
 ttcaagttgg cgcaaggaga atatattgcg cctgaaaaga tcgaggctgt cctcgccaag 1620
 cacttcctcg ttgccagat ttttgtctat gggcactcgc tcaggccac cattgtcgcg 1680
 gtggttgtcc ctgatgtga gacgtcaag ttgtgggcta aagaaaacaa gctgggtgac 1740
 aagtcttacg aggagctgtg cgctctccct cagcttcgca caaccctcca aaaggagtgt 1800
 gctacttttg gcaaagaatc ggatctgaag ggctttgaga ttctaagaa cattcatgtt 1860
 atctccgagc agttttcaat tgagaacgat cttttgaccc ccacctcaa gctgaagaga 1920
 catgctgcca aagagaagta taacgccgaa atcgaccgca tgtatgcaga aatcgcttaa 1980

<210> 29
 <211> 2113
 <212> DNA
 <213> Mortierella alpina

<400> 29
 tttcctcacc ttccctccgc tgcctctcgc tgcacaactc tctggcttat accatccacc 60
 cctctagccc cgccacttcg ccgccaacct catccgactc acaccgcaat ggctactcaa 120
 atgtactcgg tggctcgtccc caacagcccc gacattcccc gcgaaggcaa gcccgcgcgt 180

```

agtgtgcttt gtccagacaa gctcctggag aactaccct cagtgaagc aggetcaacg 240
atcacgaccc tgtacgagaa cttccaagaa ggtgttctcc gttcaggcgg cgcctatctt 300
ttgggccatc gtcccatgtt gaatggccag cctcaggctt acaagtggca gtcgtatgtc 360
gatgtcagca agcgtgttac gcacttcggc gctggcctgg ctcatctcgg cttgtctcca 420
aagcaaaact ttggaatttt ctctatcaac cggcctgagt ggaccatgag tgagcttgct 480
ggctatatgc acaactacac cagcgtcccc ctctatgata cattgggagt cgcgcgac 540
gagtatatcg ttaaccagac tgagatgcag atcatcattg cttcgtccga caaagcttct 600
atcatcctcc acatgaaatc agcacttcca accgttcaga cgattgtcgt catgggggaa 660
tttactgacg ctctcgtcgc agagggtaag gagctcaaca tcaacattgt atcctggacc 720
gatgtcgaag agagcggctc tgagcggcct gtcgaagccg tgcacccac agccgaggat 780
atcgtacca tctgttacac atctggaacc actggaacgc caaagggtgc tatcttgacc 840
cacaagaact ttgttgccac tatcgttca ttccacatga tggcaaagca tggcaggttc 900
ttcattccct cgcctgccga cacacatgta tcctacctgc cccttgccca cgtctttgag 960
cgcttttgcc aggtgttat gatctcgggc gctgcgcgta ttggttacta ccaaggagat 1020
acgtgaagc tgctggacga tgttgccgct ctgcatccca ccatttttgc ctccgtccct 1080
cgtctcttta accgtatcta cgacaagggtg cttgctggcg tcaaggocaa ggggtggtatc 1140
gccgccttct tgtttaaccg cgcataataat tccaagaagg ccaacttgcg aaagggcgta 1200
cttgagcatc cgtctggga caagctggct tttggagcga ttcgcgcgcg cttgggtggc 1260
aaggttaagc acatcgtgtc aggatctgcc cccatctctc ctgatgtgat ggatttctc 1320
cgcctctgct tcagcgtga tgtgtatgag ggatatggcc agacggaaca ggcagccgga 1380
ttaagtatga gctatcgcgg tgatttgact ccaggacagg ttggcccacc tcaactgtgc 1440
acagaggtca agttgaagga catccctagt atgaactata gcagcgcgga caagcctttc 1500
ccccgtggag aaatcatgct tcgcggaaac tctgtgttca agggctatta caaagcacca 1560
aagcagactg aagaaacatt ggatgctgac ggttggtcca gtaccggaga cgttggaacg 1620
tgggatgccc aaggccgtct ggtggtcatt gatcgcgtca agaacatctt caagttggcg 1680
caaggagaat atattgcgc tgaaaagatc gaggtgtcc tcgccaagca ctctctcgtt 1740
gccagattt ttgtctatgg gcatcgtc caggccacca ttgtcgcggt ggttgctcc 1800
gatgctgaga cgtcaagtt gtgggctaaa gaaaacaagc tgggtgacaa gtcttacgag 1860
gagctgtgcg ctctccctca gcttcgcaca accctccaaa aggagttggc tacttttggc 1920
aaagaatcgg atctgaaggg ctttgagatt cctaagaaca ttcattgttat ctccgagcag 1980
ttttcaattg agaacgatct tttgaccccc accttcaagc tgaagagaca tgctgcaaaa 2040
gagaagtata acgcccgaat cgaccgcatg tatgcagaaa tcgcttaata taaataatgg 2100
ttgtactcaa tat 2113

```

```

<210> 30
<211> 2581
<212> DNA
<213> Mortierella alpina

```

```

<220>
<221> misc_feature
<222> (1538)..(1937)
<223> n is a, c, g, or t

```

```

<400> 30
atggctactc aaatgtactc ggtggctcgc cccaacagcc ccgacattcc cggcgaaggc 60
aagccccgcc gtagtgtgct ttgtccagac aagctcctgg agaactaccc ctcaagtga 120
gcaggctcaa cgatcacgac cctgtacgag aacttccaag aagggtgttct ccgttcagg 180
aacaacgctt accacgtcga cgttcgcctc gcaatggcac ctttcccttc ggtcaaactt 240
cttaaatggt tcctttacaa tcgtaggcgg cgcctatctt ttgggccatc gtcccatgt 300
gaatggccag cctcaggctt acaagtggca gtcgtatgtc gatgtcagca agcgtgttac 360
gcacttcggc gctggcctgg ctcatctcgg cttgtctcca aagcaaaact ttggaatttt 420
ctctatcaac cggcctgagt gggtaggtga tgccttgctt tcttttgccc actcgtcgtc 480
aaggtaacgc agggcctgcg ccgattttta acagtacatt ctatgcactg tcgttacata 540
gaccatgagt gagcttgctg gctatatgca caactacacc agcgtcccc tctatgatac 600
attgggagtc gccgcgacg agtatatcgt taaccagact gagatgcaga tcatcattgc 660

```

ttcgtccgac	aaagcttcta	tcattcctcca	catgaaatca	gcacttccaa	ccgttcagac	720
gattgtcgtc	atgggggaat	ttactgacgc	tctcgtcgca	gagggtaagg	agctcaacat	780
caacattgta	tcctggaccg	atgtcgaaaa	gagcgggtctt	gagcggcctg	tcgaagccgt	840
gcaccccaca	gccgaggata	tcgctaccat	ctgttacaca	tctggaacca	ctggaacgcc	900
aaagtaagtc	aagatcatta	catggtgagc	ctccattgct	tggactgaac	agtctactca	960
cgcaggttct	tcgtttactt	tgacatgcgc	agaggtgcta	tcttgacca	caagaacttt	1020
gttgccacta	tcgcttcatt	ccacatgatg	gcaaagcatg	gcaggttctt	cattccctcg	1080
cctgccgaca	cacatgtatc	ctacctgccc	cttgcccacg	tctttgagcg	cctttgccag	1140
gctgttatga	tctcggggcgc	tgcgcgtatt	ggttactacc	aaggagatac	gctgaagctg	1200
ctggacgatg	ttgccgtcct	gcattcccacc	atttttgctt	ccgtccctcg	tctctttaac	1260
cgtatctacg	acaaggtgct	tgctggcgtc	aaggccaagg	gtggtatcgc	cgccttcttg	1320
tttaaccgcg	catataattc	caagaaggcc	aacttgcgaa	agggcgctact	tgagcatccg	1380
ctctgggaca	agctgggtctt	tggagcgatt	cgcgcgcgct	tgggtggcaa	ggttaagcac	1440
atcgtgtcag	gatctgcccc	catctctcct	gatgtgatgg	atttccctccg	catctgcttc	1500
agcgtgatg	tgtatgaggg	atatggccag	acggaacnnn	nnnnnnnnnn	nnnnnnnnnn	1560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1920
nnnnnnnnnn	nnnnnnntta	ccccacgact	ttcttttgtc	tcggcctggt	ccgcaatcat	1980
agtaccggag	acgttggaca	gtgggatgcc	caaggccgct	tgggtggcat	tgatcgcgct	2040
aagaacatct	tcaagttggc	gcaaggagaa	tatattgcgc	ctgaaaagat	cgaggctgtc	2100
ctcgccaagc	acttcctcgt	tgcccagatt	tttgtctatg	ggcactcgct	ccaggccacc	2160
attgtcgcgg	tggttgtccc	tgatgctgag	acgctcaagt	tgtgggctaa	agaaaacaag	2220
ctgggtgaca	agtcttacga	ggagctgtgc	gctctccctc	agcttcgcac	aacctccaa	2280
aaggagttgg	ctacttttgg	caaagaatcg	gatctgaagg	gctttgagat	tcctaagaac	2340
attcatgtta	tctccgagca	gttttcaatt	gagaacgata	ttttgggtgag	tgtgcttccg	2400
agtatgaacc	actgtgtcgt	atgtacgttc	gcattctgaa	agctaactct	ccgtaccctt	2460
ctacttcaat	tgtgaattct	cctcttgctg	catacagacc	cccaccttca	agctgaagag	2520
acatgctgcc	aaagagaagt	ataacgccga	aatcgaccgc	atgtatgcag	aaatcgctta	2580
a						2581

<210> 31
 <211> 1824
 <212> DNA
 <213> Mortierella alpina

<400>	31					
atgcacattc	tgaatgccac	aagaccattc	tccaggetgt	ctccaaccgt	aaggagacct	60
tggctaggac	tcggccagac	gcgcccttat	gctatcgcg	agaccgaggc	cagtcctagg	120
ctgtcatatg	tccgaggcac	caccgtcggc	acccagctat	gcgaggatcc	catcggtgcg	180
tactgggaca	gggtcgtcaa	tcgtcacggg	gaccgcctcg	gacttgctcg	caagcacgag	240
caggacctgc	actggacctt	ccgtcagttt	ggcggggcagg	ttgatagcct	ctgccgtggg	300
ctctatgagt	ctggcttgcg	aaaggggtgat	cgactcgctg	tctggatgcc	gaacaacagc	360
gcgtgggcca	cgctccagta	tgtactgcc	aagtctggca	tcattctggg	aactctcaac	420
cctgcgtacc	ggaggcaaga	gctactgcag	acattgtctt	tggctcgagt	caagtcattg	480
gtctatgtgc	caagtctaaa	gacttcgaat	tatagcgaga	tgttgctcga	cctcctacca	540
gaactccagt	accagtcgcc	aatcagctc	ttgaccgaga	agctaccctc	acttcgtcaa	600
gtcatcgtgt	ttgacaatgg	ctcgcaagtc	ccagagacag	caaaattgaa	gggattgaca	660
aagtatcagg	atttggtgat	caagaatccc	tcgaccgctg	tcgacggagc	tcttgaaaag	720
gaacggctcg	ctatcgacaa	cagggatata	atcaatctcc	agtttactag	cggaactaca	780
ggccttccca	agggcgctct	gctgtcgcat	cgaaacatct	tgaataacgg	cattcatatt	840
ggagataaca	tgcgactgac	ggaaaaggat	ttgctttgct	gcccgggtccc	gctctttcac	900
tgctttggac	tgggtgctggc	aagcttgggt	gcaatgaccc	atggcgcagg	aattatttac	960

```

ccttcgcagt cctttgatgc tgaggccaca ctgagggctg tttctgagga ggggtgctaca 1020
gcgctgcatg gcgtgccgac tatgctgttg gaagagatga accaccccaa ctttgcaaag 1080
tacaaccttt cgacacttcg gacaggaatt gcagctggat cccctgtgcc cattgaggtc 1140
atgaagaacg tgcagacaaa gatgaacctg aaggagctga ctatctgtta cggcatgacc 1200
gagacctcgc ccgtgtcctt catgacactc acaacggatg aattacggga tcgatgtgag 1260
actgttggac gaattatgcc acatctcgag gccaaagtcg tcaaccctga gacgggagag 1320
actttgccag tgaattcatc aggagagttg tgcacgcgcg ggtatgctgt gatggagggg 1380
ggttactggc gatcccagga gcagacagat gcagtgggtg acaaggatgg ctggatgcac 1440
actggcgaca ctgccgtgct cgatgaccgt ggcttttgca ggatcgacgg acgcatcaag 1500
gacatgggtg tccgaggagg cgaaaaaatc catcctgtag aggtcgagaa ctgtctcttt 1560
gagatggacg gcgtcaagaa cgtgtctgtg attggcgctt ccgacaagcg gtatggcgag 1620
caggtgtgtg cgtggatctc gaccaaggac ggggaagacg tcagtctgga ggcagtgcaa 1680
aagttctgtg agggcaagat tgcgcactac aaggtgccgc ggtatgtggt tgtgggtggag 1740
tccaatgagt tcccgactac cccctcgggc aagatccaaa agaattgtgat gcgcgagctg 1800
accaaggcga agctgcagct gcct 1824

```

```

<210> 32
<211> 608
<212> PRT
<213> Mortierella alpina

```

```

<400> 32
Met His Ile Leu Asn Ala Thr Arg Pro Phe Ser Arg Leu Ser Pro Thr
1          5          10          15
Val Arg Arg Pro Trp Leu Gly Leu Gly Gln Thr Arg Pro Tyr Ala Ile
          20          25          30
Ala Gln Thr Glu Ala Ser Pro Arg Leu Ser Tyr Val Arg Gly Thr Thr
          35          40          45
Val Gly Thr Gln Leu Cys Glu Asp Pro Ile Gly Ala Tyr Trp Asp Arg
          50          55          60
Val Val Asn Arg His Gly Asp Arg Leu Gly Leu Val Val Lys His Glu
65          70          75          80
Gln Asp Leu His Trp Thr Phe Arg Gln Phe Gly Gly Gln Val Asp Ser
          85          90          95
Leu Cys Arg Gly Leu Tyr Glu Ser Gly Leu Arg Lys Gly Asp Arg Leu
          100          105          110
Ala Val Trp Met Pro Asn Asn Ser Ala Trp Ala Thr Leu Gln Tyr Ala
          115          120          125
Thr Ala Lys Ser Gly Ile Ile Leu Val Thr Leu Asn Pro Ala Tyr Arg
          130          135          140
Arg Gln Glu Leu Leu Gln Thr Leu Ser Leu Val Glu Cys Lys Ser Leu
145          150          155          160
Val Tyr Val Pro Ser Leu Lys Thr Ser Asn Tyr Ser Glu Met Leu Leu
          165          170          175
Asp Leu Leu Pro Glu Leu Gln Tyr Gln Ser Pro Asn Gln Leu Leu Thr
          180          185          190
Glu Lys Leu Pro Ser Leu Arg Gln Val Ile Val Phe Asp Asn Gly Ser
          195          200          205
Gln Val Pro Glu Thr Ala Lys Leu Lys Gly Leu Thr Lys Tyr Gln Asp
          210          215          220
Leu Leu Ile Lys Asn Pro Ser Thr Ala Val Asp Gly Ala Leu Glu Lys
225          230          235          240
Glu Arg Leu Ala Ile Asp Asn Arg Asp Ile Ile Asn Leu Gln Phe Thr
          245          250          255
Ser Gly Thr Thr Gly Leu Pro Lys Gly Val Ser Leu Ser His Arg Asn
          260          265          270

```

Ile Leu Asn Asn Gly Ile His Ile Gly Asp Asn Met Arg Leu Thr Glu
 275 280 285
 Lys Asp Leu Leu Cys Cys Pro Val Pro Leu Phe His Cys Phe Gly Leu
 290 295 300
 Val Leu Ala Ser Leu Ala Ala Met Thr His Gly Ala Gly Ile Ile Tyr
 305 310 315 320
 Pro Ser Gln Ser Phe Asp Ala Glu Ala Thr Leu Arg Ala Val Ser Glu
 325 330 335
 Glu Gly Ala Thr Ala Leu His Gly Val Pro Thr Met Leu Leu Glu Glu
 340 345 350
 Met Asn His Pro Asn Phe Ala Lys Tyr Asn Leu Ser Thr Leu Arg Thr
 355 360 365
 Gly Ile Ala Ala Gly Ser Pro Val Pro Ile Glu Val Met Lys Asn Val
 370 375 380
 Gln Thr Lys Met Asn Leu Lys Glu Leu Thr Ile Cys Tyr Gly Met Thr
 385 390 395 400
 Glu Thr Ser Pro Val Ser Phe Met Thr Leu Thr Thr Asp Glu Leu Arg
 405 410 415
 Asp Arg Cys Glu Thr Val Gly Arg Ile Met Pro His Leu Glu Ala Lys
 420 425 430
 Val Val Asn Pro Glu Thr Gly Glu Thr Leu Pro Val Asn Ser Ser Gly
 435 440 445
 Glu Leu Cys Thr Arg Gly Tyr Ala Val Met Glu Gly Gly Tyr Trp Arg
 450 455 460
 Ser Gln Glu Gln Thr Asp Ala Val Val Asp Lys Asp Gly Trp Met His
 465 470 475 480
 Thr Gly Asp Thr Ala Val Leu Asp Asp Arg Gly Phe Cys Arg Ile Asp
 485 490 495
 Gly Arg Ile Lys Asp Met Val Ile Arg Gly Gly Glu Lys Ile His Pro
 500 505 510
 Val Glu Val Glu Asn Cys Leu Phe Glu Met Asp Gly Val Lys Asn Val
 515 520 525
 Ser Val Ile Gly Val Pro Asp Lys Arg Tyr Gly Glu Gln Val Cys Ala
 530 535 540
 Trp Ile Ser Thr Lys Asp Gly Lys Thr Val Ser Leu Glu Ala Val Gln
 545 550 555 560
 Lys Phe Cys Glu Gly Lys Ile Ala His Tyr Lys Val Pro Arg Tyr Val
 565 570 575
 Val Val Val Glu Ser Asn Glu Phe Pro Thr Thr Pro Ser Gly Lys Ile
 580 585 590
 Gln Lys Asn Val Met Arg Glu Leu Thr Lys Ala Lys Leu Gln Leu Pro
 595 600 605

<210> 33
 <211> 1827
 <212> DNA
 <213> Mortierella alpina

<400> 33
 atgcacattc tgaatgccac aagaccattc tccaggctgt ctccaaccgt aaggagacct 60
 tggctaggac tcggccagac gcgcccttat gctatcgcg agaccgaggc cagtcctagg 120
 ctgtcatatg tccgaggcac caccgtcggc acccagctat gcgaggatcc catcggtgcg 180
 tactgggaca gggtcgtcaa tcgtcacggg gaccgcctcg gacttgctcg caagcacgag 240
 caggacctgc actggacctt ccgtcagttt ggccggcagg ttgatagcct ctgccgtggg 300
 ctctatgagt ctggcttgcg aaaggggtgat cgactcgctg tctggatgcc gaacaacagc 360
 gcgtgggcca cgctccagta tgctactgcc aagtctggca tcattctggt aactctcaac 420

cctgcgtacc	ggaggcaaga	gctactgcag	acattgtctt	tggtcgagtg	caagtcattg	480
gtctatgtgc	caagtctaaa	gacttcgaat	tatagcgaga	tgttgctcga	cctcctacca	540
gaactccagt	accagtcgcc	aaatcagctc	ttgaccgaga	agctaccctc	acttcgtcaa	600
gtcatcgtgt	ttgacaatgg	ctcgcaagtc	ccagagacag	caaaattgaa	gggattgaca	660
aagtatcagg	atttggtgat	caagaatccc	tcgaccgctg	tcgacggagc	tcttgaaaag	720
gaacggctcg	ctatcgacaa	cagggatatc	atcaatctcc	agtttactag	cggaaactaca	780
ggccttccca	agggcgtctc	gctgtcgcat	cgaaacatct	tgaataacgg	cattcatatt	840
ggagataaca	tgcgactgac	ggaaaaggat	ttgctttgct	gcccggctcc	gctctttcac	900
tgctttggac	tgggtgctggc	aagcttggct	gcaatgaccc	atggcgagc	aattatttac	960
ccttcgcagt	cctttgatgc	tgaggccaca	ctgagggctg	tttctgagga	gggtgctaca	1020
gcgctgcatg	gcgtgccgac	tatgctgttg	gaagagatga	accaccccaa	ctttgcaaag	1080
tacaaccttt	cgacacttcg	gacaggaatt	gcagctggat	cccctgtgcc	cattgaggtc	1140
atgaagaacg	tgcagacaaa	gatgaacctg	aaggagctga	ctatctgtta	cggcatgacc	1200
gagacctcgc	ccgtgtcctt	catgacactc	acaacggatg	aattacggga	tcgatgtgag	1260
actgttggac	gaattatgcc	acatctcgag	gccaaagtcg	tcaaccctga	gacgggagag	1320
actttgccag	tgaattcatc	aggagagttg	tgcacgcgcg	ggtatgctgt	gatggagggt	1380
ggttactggc	gatcccagga	gcagacagat	gcagtgggtg	acaaggatgg	ctggatgcac	1440
actggcgaca	ctgcggtgct	cgatgaccgt	ggcttttgca	ggatcgacgg	acgcatcaag	1500
gacatggtga	tccgaggagg	cgaaaaaatc	catcctgtag	aggtcgagaa	ctgtctcttt	1560
gagatggacg	gcgtcaagaa	cgtgtctgtg	attggcgctc	ccgacaagcg	gtatggcgag	1620
caggtgtgtg	cgtggatctc	gaccaaggac	gggaagacgg	tcagtctgga	ggcagtgcaa	1680
aagttctgtg	agggcaagat	tgcgcactac	aaggtgccgc	ggtatgtggt	tgtggtggag	1740
tccaatgagt	tcccgactac	cccctcgggc	aagatccaaa	agaatgtgat	gcgcgagctg	1800
accaaggcga	agctgcagct	gccttga				1827

<210> 34
 <211> 2087
 <212> DNA
 <213> Mortierella alpina

<400> 34						
ggcacgaggc	tctactctcc	attgcccact	cactcattgc	ccctctgtcc	atcaccggca	60
ttgctcgttc	gcgccttcgc	ccactccact	ctttctttca	tctctctttt	acaacggcca	120
tctccccctc	gctctgcgct	tctcccatec	acgctaacaa	tgcacattct	gaatgccaca	180
agaccattct	ccaggetgtc	tccaaccgta	aggagacctt	ggctaggact	cggccagacg	240
cgcccttatg	ctatcgcgca	gaccgagggc	agtcctaggc	tgtcatatgt	ccgaggcacc	300
accgtcggca	cccagctatg	cgaggatccc	atcgggtgcg	actgggacag	ggtcgtcaat	360
cgtcacgggtg	accgcctcgg	acttgtcgtc	aagcacgagc	aggacctgca	ctggaccttc	420
cgtcagtttg	gcgggcagggt	tgatagcctc	tgccgtgggc	tctatgagtc	tggcttgcca	480
aaggggtgatc	gactcgctgt	ctggatgccg	aacaacagcg	cgtgggccac	gctccagtat	540
gctactgcca	agtcctggcat	cattctggta	actctcaacc	ctgcgtaccg	gaggcaagag	600
ctactgcaga	cattgtcttt	ggtcgagtg	aagtcattgg	tctatgtgcc	aagtctaaag	660
acttcgaatt	atagcgagat	gttgctcgac	ctcctaccag	aactccagta	ccagtcgcca	720
aatcagctct	tgaccgagaa	gctaccctca	cttcgtcaag	tcacgtgtgt	tgacaatggc	780
tcgcaagtc	cagagacagc	aaaattgaag	ggattgacaa	agtatcagga	tttggtgatc	840
aagaatccct	cgaccgctgt	cgacggagct	cttgaaaagg	aacggctcgc	tatcgacaac	900
agggatatca	tcaatctcca	gtttactagc	ggaactacag	gccttcccaa	gggcgtctcg	960
ctgtcgcata	gaaacatctt	gaataacggc	attcatattg	gagataacat	gcgactgacg	1020
gaaaaggatt	tgctttgctg	cccggctccc	ctctttcact	gctttggact	ggtgctggca	1080
agcttggtctg	caatgaccca	tggcgcagga	attatttacc	cttcgcagtc	ctttgatgct	1140
gaggeccacac	tgagggctgt	ttctgaggag	ggtgctacag	cgctgcatgg	cgtgccgact	1200
atgctgttgg	aagagatgaa	ccaccccaac	tttgcaaagt	acaacctttc	gacacttcgg	1260
acaggaattg	cagctggatc	ccctgtgccc	attgaggtca	tgaagaacgt	gcagacaaag	1320
atgaacctga	aggagctgac	tatctgttac	ggcatgaccg	agacctcgcc	cgtgtccttc	1380
atgacactca	caacggatga	attacgggat	cgatgtgaga	ctggttgagc	aattatgcca	1440
catctcgagg	ccaaagtcgt	caaccctgag	acgggagaga	ctttgccagt	gaattcatca	1500

ggagagttgt	gcacgcgcgg	gtatgctgtg	atggagggtg	gttactggcg	atcccaggag	1560
cagacagatg	cagtgggtgga	caaggatggc	tggatgcaca	ctggcgacac	tgccgtgctc	1620
gatgaccgtg	gcttttgcag	gatcgacgga	cgcatacaagg	acatgggtgat	ccgaggaggc	1680
gaaaaaatcc	atcctgtaga	ggtcgagAAC	tgtctctttg	agatggacgg	cgtcaagaac	1740
gtgtctgtga	ttggcggtcc	cgacaagcgg	tatggcgagc	aggtgtgtgc	gtggatctcg	1800
accaaggacg	ggaagacggg	cagtctggag	gcagtgcAAA	agttctgtga	gggcaagatt	1860
gcgcactaca	aggtgcccg	gtatgtggtt	gtgggtggagt	ccaatgagtt	cccgactacc	1920
ccctcgggca	agatccAAAA	gaatgtgatg	cgcgagctga	ccaaggcgaa	gctgcagctg	1980
ccttgatggg	actaggatat	ggagccgacg	aaagtaataa	aggcgtatgc	tggcatggcg	2040
caagatctga	gccctgcggg	gaggtgcatt	cagtgcgcgc	attagaa		2087

<210> 35
 <211> 2242
 <212> DNA
 <213> Mortierella alpina

<400> 35		
atgcacattc	tgaatgccac aagaccattc tccaggctgt ctccaaccgt aaggagacct 60	
tggctaggac	tgggccagac gcgcccttat gctatcgcg agaccgaggc cagtcctagg 120	
ctgtcatatg	tccgaggcac caccgtcggc acccagctat gcgaggatcc catcggtgcg 180	
tactgggaca	gggtcgtcaa tcgtcacggg gaccgcctcg gacttgctgt caagcacgag 240	
caggacctgc	actggacctt ccgtcagttt ggcgggcagg ttgatagcct ctgccgtggg 300	
ctctatgagt	ctggcttgcg aaagggtgat cgactcgcg atgtacactc tcgtcgtgcc 360	
tgaagcgctt	gaagggaaaa actgtattga acccccggcc gctgcattag cgtcgtttca 420	
acatttggct	aattttttct ttcctcgtc ctttgccgat gcaaacagtg tctggatgcc 480	
gaacaacagc	gcgtgggcca cgtccagta tgctactgcc aagtctggca tcattctggt 540	
aactctcaac	cctgcgtatg tactgcatgt cataatcata ctgcgctcgc ctcaaaaata 600	
tgcttttttg	acgtgtttac ttatcctgcg cgttctcata acaggtaccg gaggcaagag 660	
ctactgcaga	cattgtcttt ggtcgagtgc aagtcattgg tctatgtgcc aagtctaaag 720	
acttcgaatt	atagcgagat gttgctcgac ctctaccag aactccagta ccagtcgcca 780	
aatcagctct	tgaccgagaa gctaccctca ctctcgtaag tcctcgtgtt tgacaatggc 840	
tcgcaagtc	cagagacagc aaaattgaag ggattgacaa agtatcagga tttgttgatc 900	
aagaatccct	cgaccgctgt cgacggagct cttgaaaagg aacggctcgc tategacaac 960	
agggatatca	tcaatctcca gtttactagc ggaactacag gccttcccaa gtatgtgtca 1020	
gaaattgaac	ttgacacgcg atgcataatc gacgttcaac cacccttgtc ctgacaccgt 1080	
cttgcattca	caatactatg gtccaaattt taggggcgtc tcgtgtcgc atcgaaacat 1140	
cttgaataac	ggcattcata ttggagataa catgcgactg acggaaaagg atttgctttg 1200	
ctgcccggtc	ccgtctcttc actgcttttg actggtgctg gcaagcttg ctgcaatgac 1260	
ccatggcgca	ggaattatct acccttcgca gtcctttgat gctgaggcca cactgagggc 1320	
tgtttctgag	gaggggtgcta cagcgtgca tggcggtgcc actatgctgt tggaagagat 1380	
gaaccacccc	aactttgcaa agtacaacct ttcgacactt cggacaggaa ttgcagctgg 1440	
atcccctgtg	cccattgagg tcatgaagaa cgtgcagaca aagatgaacc tgaaggagct 1500	
gactatctgt	tacggcatga ccgagacctc gcccggtgtcc ttcattgacac tcacaacgga 1560	
tgaattacgg	gatcgatgtg agactgttgg acgaattatg ccacatctcg aggccaaagg 1620	
taaaacggct	gtcagcctgg tgttcataat tttgtcttga atctcgggtg tctgttgcta 1680	
attttgggta	cattgcttgc acaatacggg agtcgtcaac cctgagacgg gagagacttt 1740	
gccagtgaat	tcattcaggag agttgtgcac gcgcgggtat gctgtgatgg aggggtggtta 1800	
ctggcgatcc	caggagcaga cagatgcagt ggtggacaag gatggctgga tgcacactgg 1860	
cgacactgcc	gtgctcgatg accgtggctt ttgcaggatc gacggacgca tcaaggacat 1920	
ggtgatccga	ggaggcgaaa aaatccatcc tgtagaggtc gagaactgtc tctttgagat 1980	
ggacggcgtc	aagaacgtgt ctgtgattgg cgttcccgc aagcggtatg gcgagcaggt 2040	
gtgtgcgtgg	atctcgacca aggacgggaa gacggtcagt ctggaggcag tgcaaaagtt 2100	
ctgtgagggc	aagattgcgc actacaaggt gccgcggtat gtggttgtgg tggagtccaa 2160	
tgagttccc	actacccct cgggcaagat ccaaaagaat gtgatgcgcg agctgaccaa 2220	
ggcgaagctg	cagctgcctt ga	2242

<210> 36
 <211> 2076
 <212> DNA
 <213> Mortierella alpina

<400> 36
 atgccttcct tcaaaaagta caacctcgac aagcagagtg ttgaggtccc tggcactcgg 60
 aagcctggcg ettcaggcca ctacagacat gcgcctacg gcgatgctct tgtcaccaac 120
 atccgtgagg cccctcatat cgaaactctt tacgacatgt ggcagaactc tgtgacaaag 180
 tatggcggca atgacttttt gggtcaccgt cccttcaaca ctggtgcca gacctatggt 240
 ggctacagtt gggagacgta ccgccagatt aaccagcgcg ttaatgcgtt cggcagcggc 300
 atcatgcacc tgaacgaggt gatcctcggc aaccgccagc ttaaccgctg ggcgttgggc 360
 atctgggtccc acggtcgccc tgagtgggtc attacggaga tgagctgcaa ctgctacaac 420
 ctcatctctg ttgcattgta cgacaccctt ggacctgatg cagtcgagta cattgtcaac 480
 cacgccgaga ttgagattgt tgtctcaagt gccaaaccata tcgcctcttt gctcgagaac 540
 gccgagaagc tccccaaagt caaggccatt gtcagcatgg atgctcttca cgataccgtc 600
 cccgtccccg gcgccacctc tgccgcacag gttcttcgtg cctgggggtgc aaaaaagggc 660
 atcaaggctct atgactttta cgagattgag tccctcgggtg ccgagttccc tcgcaagcac 720
 ctgcctccca ccgtgatga ggtcgcctcc atctgctaca cttccggcac caccggtcag 780
 cctaaaggag ccatgctcac ccacagaaac tttgttgcta ctggttggtac caaccgcgag 840
 ggcatgcttc tcaccgagga cgacgttttg atcagtttct tgcccttggc tcacattatg 900
 ggacgcgtca ttgacacttg ctcgatgtac agcgggtggca agattgggta cttccgtgga 960
 gatattcttt tgcttctcga ggacgttgct gagctccgtc ccacattctt cccagctgtg 1020
 cctcgcctct tgaaccgcat ttatgccaaag ctcgttgcct ctaccattga ggcccccggt 1080
 ttgggtcgggtg ccttgggccc tgccgggtgc gccgccaaga tggccaacct tgctgccgga 1140
 aaggggtgtca accacgctct ctgggacaga ctgctgttca acaagggtcaa gatggccctg 1200
 ggtgggtcgcg ttcaggtcat cctgactgga tctgcgcccc ttgccaagga ggttctcagc 1260
 ttcttgagaa ttgctttcgg atgcgtgggt ttggagggat acggctccac tgagggcatg 1320
 gctaccgcca ccatcacaat ggctgatgag tacattcctg gtcacattgg ctgccctcgt 1380
 gctggatgcg agctcaagct ggtggatgtg cccgcgatga actacctctc taccgaccag 1440
 ccctaccccc gtggagagat ctggatccgt ggtgacactg ttttcaaagg atacttcaag 1500
 gacgagaaga acactagtga gactatcgac tctgaaggct ggctcgctac cgggtgatatt 1560
 ggattttgtg ataagcgtgg atgctttacg atcattgacc gcaagaagaa catcttcaag 1620
 ttggcacaag gtgaatacat tgctcctgaa aagattgaga acgtcttggg cgcacgctgc 1680
 aatcttgtcc agcagatcta tgttcattgg gattcgcttg agtccacctt ggtcgcagtt 1740
 cttattcccc agcccagagac cttcctgccc ttcgcgaatg ccattgctgg tgcctccgtc 1800
 actgctggag atgttgaggg tttgaacaag ctgtgccaaag atcccaaggt caagatcgcg 1860
 gttctgaagg agttggagaa ggccggaaag gccggtgcga tgcgcggatt cgagttcgtg 1920
 aagcgtgtcc acttgaccac ggatgcattc tcggtcgaca acggcatgat gacacctacc 1980
 ttcaaggtcc gtcgcccaca agtagccgag catttcaggg agcaaatcac ggccatgtat 2040
 aaggagatca atgcctcgac cctgtttgcc aagctg 2076

<210> 37
 <211> 692
 <212> PRT
 <213> Mortierella alpina

<400> 37
 Met Pro Ser Phe Lys Lys Tyr Asn Leu Asp Lys Gln Ser Val Glu Val
 1 5 10 15
 Pro Gly Thr Arg Lys Pro Gly Ala Ser Gly His Tyr Arg His Ala Ala
 20 25 30
 Tyr Gly Asp Ala Leu Val Thr Asn Ile Arg Glu Ala Pro His Ile Glu
 35 40 45
 Thr Leu Tyr Asp Met Trp Gln Asn Ser Val Thr Lys Tyr Gly Gly Asn
 50 55 60

Asp	Phe	Leu	Gly	His	Arg	Pro	Phe	Asn	Thr	Val	Ala	Gln	Thr	Tyr	Gly
65					70					75					80
Gly	Tyr	Ser	Trp	Glu	Thr	Tyr	Arg	Gln	Ile	Asn	Gln	Arg	Val	Asn	Ala
				85					90					95	
Phe	Gly	Ser	Gly	Ile	Met	His	Leu	Asn	Glu	Val	Ile	Leu	Gly	Asn	Arg
			100					105					110		
Gln	Leu	Asn	Arg	Trp	Ala	Leu	Gly	Ile	Trp	Ser	His	Gly	Arg	Pro	Glu
		115					120					125			
Trp	Phe	Ile	Thr	Glu	Met	Ser	Cys	Asn	Cys	Tyr	Asn	Leu	Ile	Ser	Val
	130					135					140				
Ala	Leu	Tyr	Asp	Thr	Leu	Gly	Pro	Asp	Ala	Val	Glu	Tyr	Ile	Val	Asn
145					150					155					160
His	Ala	Glu	Ile	Glu	Ile	Val	Val	Ser	Ser	Ala	Asn	His	Ile	Ala	Ser
				165					170					175	
Leu	Leu	Glu	Asn	Ala	Glu	Lys	Leu	Pro	Lys	Leu	Lys	Ala	Ile	Val	Ser
			180					185					190		
Met	Asp	Ala	Leu	His	Asp	Thr	Val	Pro	Val	Pro	Gly	Ala	Thr	Ser	Ala
	195						200					205			
Ala	Gln	Val	Leu	Arg	Ala	Trp	Gly	Ala	Gln	Lys	Gly	Ile	Lys	Val	Tyr
	210					215					220				
Asp	Phe	Asn	Glu	Ile	Glu	Ser	Leu	Gly	Ala	Glu	Phe	Pro	Arg	Lys	His
225					230					235					240
Leu	Pro	Pro	Thr	Ala	Asp	Glu	Val	Ala	Ser	Ile	Cys	Tyr	Thr	Ser	Gly
				245				250						255	
Thr	Thr	Gly	Gln	Pro	Lys	Gly	Ala	Met	Leu	Thr	His	Arg	Asn	Phe	Val
			260				265						270		
Ala	Thr	Val	Gly	Thr	Asn	Arg	Glu	Gly	Met	Leu	Leu	Thr	Glu	Asp	Asp
		275					280					285			
Val	Leu	Ile	Ser	Phe	Leu	Pro	Leu	Ala	His	Ile	Met	Gly	Arg	Val	Ile
	290					295					300				
Asp	Thr	Cys	Ser	Met	Tyr	Ser	Gly	Gly	Lys	Ile	Gly	Tyr	Phe	Arg	Gly
305					310					315					320
Asp	Ile	Leu	Leu	Leu	Leu	Glu	Asp	Val	Ala	Glu	Leu	Arg	Pro	Thr	Phe
				325					330					335	
Phe	Pro	Ala	Val	Pro	Arg	Leu	Leu	Asn	Arg	Ile	Tyr	Ala	Lys	Leu	Val
			340					345					350		
Ala	Ser	Thr	Ile	Glu	Ala	Pro	Gly	Leu	Val	Gly	Ala	Leu	Ala	Arg	Arg
		355					360					365			
Gly	Val	Ala	Ala	Lys	Met	Ala	Asn	Leu	Ala	Ala	Gly	Lys	Gly	Val	Asn
	370					375					380				
His	Ala	Leu	Trp	Asp	Arg	Leu	Leu	Phe	Asn	Lys	Val	Lys	Met	Ala	Leu
385					390					395					400
Gly	Gly	Arg	Val	Gln	Val	Ile	Leu	Thr	Gly	Ser	Ala	Pro	Ile	Ala	Lys
				405					410					415	
Glu	Val	Leu	Ser	Phe	Leu	Arg	Ile	Ala	Phe	Gly	Cys	Val	Val	Leu	Glu
			420					425					430		
Gly	Tyr	Gly	Ser	Thr	Glu	Gly	Met	Ala	Thr	Ala	Thr	Ile	Thr	Met	Ala
		435					440					445			
Asp	Glu	Tyr	Ile	Pro	Gly	His	Ile	Gly	Cys	Pro	Arg	Ala	Gly	Cys	Glu
	450					455					460				
Leu	Lys	Leu	Val	Asp	Val	Pro	Ala	Met	Asn	Tyr	Leu	Ser	Thr	Asp	Gln
465					470					475					480
Pro	Tyr	Pro	Arg	Gly	Glu	Ile	Trp	Ile	Arg	Gly	Asp	Thr	Val	Phe	Lys
				485					490					495	
Gly	Tyr	Phe	Lys	Asp	Glu	Lys	Asn	Thr	Ser	Glu	Thr	Ile	Asp	Ser	Glu
			500					505					510		

Gly Trp Leu Ala Thr Gly Asp Ile Gly Phe Val Asp Lys Arg Gly Cys
 515 520 525
 Phe Thr Ile Ile Asp Arg Lys Lys Asn Ile Phe Lys Leu Ala Gln Gly
 530 535 540
 Glu Tyr Ile Ala Pro Glu Lys Ile Glu Asn Val Leu Gly Ala Arg Cys
 545 550 555 560
 Asn Leu Val Gln Gln Ile Tyr Val His Gly Asp Ser Leu Glu Ser Thr
 565 570 575
 Leu Val Ala Val Leu Ile Pro Glu Pro Glu Thr Phe Leu Pro Phe Ala
 580 585 590
 Asn Ala Ile Ala Gly Ala Ser Val Thr Ala Gly Asp Val Glu Gly Leu
 595 600 605
 Asn Lys Leu Cys Gln Asp Pro Lys Val Lys Ile Ala Val Leu Lys Glu
 610 615 620
 Leu Glu Lys Ala Gly Lys Ala Gly Ala Met Arg Gly Phe Glu Phe Val
 625 630 635 640
 Lys Arg Val His Leu Thr Thr Asp Ala Phe Ser Val Asp Asn Gly Met
 645 650 655
 Met Thr Pro Thr Phe Lys Val Arg Arg Pro Gln Val Ala Glu His Phe
 660 665 670
 Arg Glu Gln Ile Thr Ala Met Tyr Lys Glu Ile Asn Ala Ser Thr Pro
 675 680 685
 Val Ala Lys Leu
 690

<210> 38
 <211> 2079
 <212> DNA
 <213> Mortierella alpina

<400> 38
 atgccttctt tcaaaaagta caacctcgac aagcagagtgt ttgaggtccc tggcactcgg 60
 aagcctggcg cttcaggcca ctacagacat gccgcctacg gcgatgctct tgtcaccaac 120
 atcogtgagg cccctcatat cgaaactctt tacgacatgt ggcagaactc tgtgacaaag 180
 tatggcggca atgacttttt gggtcaccgt cccttcaaca ctggttgcca gacctatggt 240
 ggctacagtt gggagacgta ccgccagatt aaccagcgcg ttaatgcgtt cggcagcggg 300
 atcatgcacc tgaacgaggt gatcctcggc aaccgccagc ttaaccgctg ggcgttgggc 360
 atctgggtccc acggtcgccc tgagtgggtc attacggaga tgagctgcaa ctgctacaac 420
 ctcatctctg ttgcattgta cgacaccctt ggacctgatg cagtcgagta cattgtcaac 480
 cacgccgaga ttgagattgt tgtctcaagt gcccaaccata tcgcctcttt gctcgagaac 540
 gccgagaagc tccccaaagt caaggccatt gtcagcatgg atgctcttca cgataccgtc 600
 cccgtccccg gcgccacctc tgccgcacag gttcttcgtg cctgggggtgc acaaaagggc 660
 atcaagggtct atgactttta cgagattgag tccctcgggtg ccgagttccc tcgcaagcac 720
 ctgcctccca ccgctgatga ggtcgcctcc atctgctaca cttccggcac caccggtcag 780
 cctaaaggag ccatgctcac ccacagaaac tttgttgcta ctggttggtac caaccgcgag 840
 ggcatgcttc tcaccgagga cgacgttttg atcagtttct tgcccttggc tcacattatg 900
 ggacgcgtca ttgacacttg ctcgatgtac agcgggtggca agattgggta cttccgtgga 960
 gatattcttt tgcttctcga ggacgttget gagctccgtc ccacattctt cccagctgtg 1020
 cctcgcctct tgaaccgcat ttatgccaaag ctcgttgcct ctaccattga ggcccccggt 1080
 ttgggtcgggt ccttggcccg tcgcgggtgc gccgccaaga tggccaacct tgetgcggga 1140
 aagggtgtca accacgctct ctgggacaga ctgctgttca acaagggtcaa gatggccctg 1200
 ggtgggtcgcg ttcagggtcat cctgactgga tctgcgcca ttgccaagga ggttctcagc 1260
 ttcttgagaa ttgctttcgg atgcgtggtt ttggagggat acggctccac tgagggcatg 1320
 gctaccgcca ccatcacaat ggctgatgag tacattcctg gtcacattgg ctgccctcgt 1380
 gctggatgcg agctcaagct ggtggatgtg cccgcgatga actacctctc taccgaccag 1440
 ccctaccccc gtggagagat ctggatccgt ggtgacactg ttttcaaagg atacttcaag 1500

gacgagaaga	acactagtga	gactatcgac	tctgaaggct	ggctcgctac	cggtgatatt	1560
ggatttgtgg	ataagcgtgg	atgctttacg	atcattgacc	gcaagaagaa	catcttcaag	1620
ttggcacaag	gtgaatacat	tgctcctgaa	aagattgaga	acgtcttggg	cgcacgctgc	1680
aatcttgtcc	agcagateta	tgttcatggg	gattcgcttg	agtccacctt	ggtcgcagtt	1740
cttattcccg	agcccagagac	cttcctgccc	ttcgcgaaatg	ccattgctgg	tgcctccgtc	1800
actgctggag	atgttgaggg	tttgaacaag	ctgtgccaaag	atcccaaggt	caagatcgcg	1860
gttctgaagg	agttggagaa	ggccggaaag	gccggtgcca	tgcgcggatt	cgagttcgtg	1920
aagcgtgtcc	acttgaccac	ggatgcattc	tccgtcgaca	acggcatgat	gacacctacc	1980
ttcaagggtcc	gtcgcccaca	agtagccgag	catttcaggg	agcaaatcac	ggccatgtat	2040
aaggagatca	atgcctcgac	ccctgttgcc	aagctgtag			2079

<210> 39

<211> 2288

<212> DNA

<213> Mortierella alpina

<400> 39

tgcttttctc	ttcttcgtca	ccctccttct	tcccattcct	ccggtcctcc	tccgttctca	60
atcagtttct	cagacctgt	ccattcctct	ggcctccaca	cacacccac	tctcccttga	120
acaaatacct	tatccagatc	aaagacatgc	cttccttcaa	aaagtacaac	ctcgacaagc	180
agagtgttga	ggtccttggc	actcggaagc	ctggcgcttc	aggccactac	agacatgccg	240
cctacggcga	tgctcttgtc	accaacatcc	gtgaggcccc	tcatatcgaa	actctttacg	300
acatgtggca	gaactctgtg	acaaagtatg	gcggaatga	ctttttgggt	caccgtccct	360
tcaacactgt	tgcccagacc	tatggtggct	acagttggga	gacgtaccgc	cagattaacc	420
agcgcgttaa	tgcgttcggc	agcggtatca	tgcacctgaa	cgaggtgatc	ctcggcaacc	480
gccagcttaa	ccgctgggcg	ttgggcattc	ggtcccacgg	tgcacctgag	tggttcatta	540
cggagatgag	ctgcaactgc	tacaacctca	tttctgttgc	attgtacgac	acccttggac	600
ctgatgcagt	cgagtacatt	gtcaaccacg	ccgagattga	gattgttgtc	tcaagtgcc	660
accatategc	ctctttgtc	gagaacgcgc	agaagctccc	caagctcaag	gccattgtca	720
gcatggatgc	tcttcacgat	accgtccccc	tccccggcgc	cacctctgcc	gcacagggtc	780
ttcgtgcctg	gggtgcacaa	aagggcattc	aggtctatga	ctttaacgag	attgagtcct	840
tccgtgccga	gttccctcgc	aagcacctgc	ctcccaccgc	tgatgaggtc	gcctccatct	900
gctacacttc	cggcaccacc	ggtcagccta	aaggagccat	gctcaccac	agaaactttg	960
ttgctactgt	tggtagcaac	cgcgagggca	tgttctcac	cgaggacgac	gttttgatca	1020
gtttcttgcc	cttggctcac	attatgggac	gcgtcattga	cacttgctcg	atgtacagcg	1080
gtggcaagat	tggttacttc	cgtggagata	ttcttttgc	tctcgaggac	gttgetgagc	1140
tccgtccac	attcttccca	gctgtgcctc	gcctcttgaa	ccgcatttat	gccaagctcg	1200
ttgcctctac	cattgagggc	cccggtttgg	tccgtgcctt	ggcccgtcgc	ggtgtcgcgc	1260
ccaagatggc	caaccttget	gccggaaagg	gtgtcaacca	cgtctcttgg	gacagactgc	1320
tgttcaacaa	ggtcaagatg	gccctgggtg	gtcgcgttca	ggtcatcctg	actggatctg	1380
cgcaccttgc	caaggagggt	ctcagcttct	tgagaattgc	tttcggatgc	gtggttttgg	1440
agggatacgg	ctccactgag	ggcatggcta	ccgccaccat	cacaatgggt	gatgagtaca	1500
ttcctggtea	cattggctgc	cctcgtgctg	gatgcgagct	caagctgggt	gatgtgcccg	1560
cgatgaacta	cctctctacc	gaccagccct	acccccgtgg	agagatctgg	atccgtgggtg	1620
acactgtttt	caaaggatac	ttcaaggacg	agaagaacac	tagtgagact	atcgactctg	1680
aaggctggct	cgctaccggg	gatattggat	ttgtggataa	gcgtggatgc	tttacgatca	1740
ttgaccgcaa	gaagaacatc	ttcaagttgg	cacaaggtga	atacattgct	cctgaaaaga	1800
ttgagaacgt	cttgggcgca	cgtgcgaatc	ttgtccagca	gatctatggt	catgggtgatt	1860
cgtttgagtc	caccttggtc	gcagttctta	ttcccagacc	cgagaccttc	ctgcccttcg	1920
cgaatgccat	tgttgggtgc	tccgtcactg	ctggagatgt	tgagggtttg	aacaagctgt	1980
gccaagatcc	caaggatcaag	atcgcggttc	tgaaggaggt	ggagaaggcc	ggaaaggccg	2040
gtgcgatgcg	cggattcgag	ttcgtgaagc	gtgtccactt	gaccacggat	gcattctcgg	2100
tcgacaacgg	catgatgaca	cctaccttca	aggtccgtcg	cccacaagta	gccgagcatt	2160
tcaggggagca	aatcacggcc	atgtataagg	agatcaatgc	ctcgacctct	gttgccaagc	2220
tgtagataga	aaactctttg	ccccttatta	ccctttgaat	agaagggtgac	acgttggttg	2280
attcacac						2288

<210> 40
 <211> 2816
 <212> DNA
 <213> Mortierella alpina

<400> 40
 atgccttcct tcaaaaagta caacctcgac aagcagagtg ttgagggtccc tggcactcgg 60
 aagcctggcg cttcaggcaa gttggatagt ggctcatgag atccaactgt tgggccacgc 120
 taccatgaaa ctagtcgcta atgcagattc tctattgcct ttcaccttct cacaggccac 180
 tacagacatg ccgcctacgg cgatgctctt gtcaccaaca tccgtgaggc ccctcatatc 240
 gaaactcttt acgacatgtg gcagaactgt aagttgattc ccgaggacct gggaaatctt 300
 tactcggatc tgtctcacia gaagactcac actctgctct ccacaattac cttgtaacct 360
 tcagctgtga caaagtatgg cggcaatgac tttttgggtc accgtccctt caacactggt 420
 gccagacct atggtggcta cagttgggag acgtaccgcc agattaacca gcgcgttaat 480
 gcgttcggca gcggtatcat gcacctgaac gaggtgatcc tcggcaaccg ccagcttaac 540
 cgctgggctg tgggcatctg gtcccacggg cgccctgagt ggttcattac ggagatgagc 600
 tgcaactgct acaacctcat ttctgttgca ttgtacgaca cccttggacc tgatgcagtc 660
 gagtacattg tcaaccacgc cgagattgag attgttgtct caagtggtaa gtcgcttttt 720
 tttttttgct attggcgctt gcttgcttgc ctgcacgtga aaactataat ggtttctgat 780
 ctcccttttg tgtacctcct cctttcgate aacttctttg taattttatt cccgggcatt 840
 gactctatat gcagccaacc atatcgctc tttgctcgag aacgccgaga agctcccaa 900
 gctcaaggcc attgtcagca tggatgctct tcacgatacc gtcccgcgtc ccggcgccac 960
 ctctgccgca caggttcttc gtgcctgggg tgcacaaaag ggcatacaag tctatgactt 1020
 taacgagatt gagtccctcg gtgccgagtt ccctcgcaag cacctgcctc ccaccgctga 1080
 tgaggtcgcc tccatctgct acaattccgg caccacggg cagcctgtaa gtgtgttttc 1140
 ccttcactga cgatgtcggc tgaggatgca ttttgcctgat tgaaacttcc catctaacat 1200
 cgcttgatct atagaaagga gccatgctca cccacagaaa ctttgttgct actgttggt 1260
 ccaaccgcca gggcatgctt ctccaccgag acgacgtttt gatcaggat aactcatctt 1320
 gaagctgatc aatgacacgc gatgtacaac gcgacgatgg agcgagatct acaatgcgga 1380
 atggctcacc tccggtttac aattaccaca cttctagttt cttgcccttg gctcacatta 1440
 tgggacgcgt cattgacact tgctcgatgt acagcgggtg caagattggt tacttccgtg 1500
 gagatattct tttgcttctc gaggacgttg ctgagctccg tcccacattc tcccagctg 1560
 tgctcgcct cttgaaccgc atttatgcca agctcgttgc ctctaccatt gagggccccg 1620
 gtttggtcgg tgcttggcc cgtcgcgggtg tcgccgcaa gatggccaac cttgctgccg 1680
 gaaaggggtg caaccacgct ctctgggaca gactgctggt caacaaggte aagatggccc 1740
 tgggtggctg cgttcaggte atcctgactg gatctgcgcc cattgccaa gaggttctca 1800
 gcttcttgag aattgctttc ggatgcgtgg ttttgagggg atacggctcc actgagggca 1860
 tggctaccgc caccatcaca atggctgagt acgtgaacct agttatttta ttgaaatgtc 1920
 gtggagcctg tcgactgtag catttcaatc taaccattag taatcttttc aaaaatagtg 1980
 agtacattcc tggtcacatt ggctgcctc gtgctggatg cgagctcaag ctggtggatg 2040
 tgccgcgat gaactacctc tctaccgacc agccctacc ccgtggagag atctggatcc 2100
 gtggtgacac tgttttcaaa ggatacttca aggacgagaa gaacactagt gagactatcg 2160
 actctgaagg ctggctcgt accggtgata ttggatttgt ggataagcgt ggatgcttta 2220
 cgatcattga ccgcaagaag aacatcttca aggtatgagc aaaaagtggg attgatctga 2280
 tcgtttctct tctcgttttc ccgttaagga attccgctca tactaacgtt ctctcaatg 2340
 gtttgcatgt attatagttg gcacaagggt aatacattgc tcctgaaaag attgagaacg 2400
 tcttgggcgc acgctgcaat cttgtccagc agatctatgt tcatggtgat tcgcttgagt 2460
 ccaccttggt cgcagttctt attcccagac ccgagacctt cctgcccttc gcgaatgcca 2520
 ttgctggtgc ctccgtcact gctggagatg ttgagggttt gaacaagctg tgccaagatc 2580
 ccaagggtca gatcgcggtt ctgaaggagt tggagaaggc cggaaaggcc ggtgcgatgc 2640
 gcggattcga gttcgtgaag cgtgtccact tgaccacgga tgcatctctg gtcgacaacg 2700
 gcatgatgac acctaccttc aaggtccgtc gccacaagt agccgagcat ttcagggagc 2760
 aaatcacggc catgtataag gagatcaatg cctcgacccc tgttgccaag ctgtag 2816

<210> 41
 <211> 1848

<212> DNA

<213> Mortierella alpina

<400> 41

```

atggttgctc tcccactcgt cgcagcagct gtcccagctg ccatgtatgt gagctcaaag      60
ctggcacttc ctcgggatat gaagttgatt aagagcttga tcggagccaa gatggcctac      120
agtgccatgg aaaagaacga cgccctcaac ctgacactcc gcttcgacga gtgctaccgc      180
aagtatcctg accgtgaagc cctgggtcttt gagggcaaat cctattcatt cctgatatt      240
cagcttgect ccaacaggtg cggcaactgg ttgctggcca aagggatcaa gcgaggagat      300
atcgtctcgc ttttcatgtt gaacaggcca gagttcatct tctgctggct gggctctaac      360
aagattggag ccaactggtgc cttcatcaat accaacctta cgggcaaacc cctgacacat      420
tccctccgga cagccacgtc gtcaatgttg atcatggaca cggagttgac agacgcgac      480
gccaactccc tggatgagat tcaggagatg ggctattcaa tttactctta cggacccgaa      540
gccgtggact ttgctacccc gatggatata tcgcagggtcc cagacaccga tacacccgaa      600
cacctgcgcc ggaacacgac cgcggatgac attgcgatgc tcctctacac ctctggaact      660
actggtcttc ccaaggccgg tegtgtctct catgcgcgtg cctctatggg acctcagttt      720
tggaaccgat tctatcactt cagtgaagac gacagggtct atctgtcctt gcccttgtac      780
cacagtgtcg gcgccatctt gggagtgttg gcttggttga cctcgggagc aaccttgatc      840
ctggcccgcg agttctccgc gacacatttc tgggaggatt gccgcgtgaa caacgcaact      900
gtgattcaat acattggaga aatttgcaga tatctgtctc acacgccaga atcacccctg      960
gacaaggcac actcgatacg actggcacat ggtaatggaa tgcgacccga tgtctggact     1020
cgcttcagag atcgggttcgg catcccgttg attggcgagt ggtatgcac gactgaggga     1080
actggagcct tgtcgaatta taacacaggc ccaggcggcg ctggagcgat tggataccgc     1140
ggtacccttg ccagagcatt ggataaagga ctcaggattg cgagatttga tgtccagaca     1200
gaggagttag ttccgggaaa aaacgggttat tgcattgagt gcaaacctgg cgagcccgga     1260
gaattgctga cgcttggtga tgctaaagag ccgaacaaag acttcaaagg ataccatcaa     1320
aaccaggcag cgaccaaaaa aaagattgtc aaagatgttt tcaaagccgg cgacatgtac     1380
ttccgtaccg gagatatcct tcggcgcgat agcgatgggt acttttactt tggcgaccgt     1440
gtgggcgata cattccggtg gaagtccgag aatgtgtcta cggccgaggt gtctgaagtc     1500
ctctcgcatg atccggactg tatcgaagtc aatgtgtatg gagttcagat cccagggcag     1560
gacggacgcg ccggtatggc agcgattgtg tccaagagca cgatggattg ggagaaattt     1620
gcggcgtatg cactcaagaa cctgccgcgg tattctgttc cgatctttat ccgcaagatg     1680
cccagatagg agatcacagg gacgttcaag cagcgcaaag tcgagttggg gaatgaggga     1740
atcgacccca agacgattgc caacgagatg ctgtggttgg acggacacca ctataagccg     1800
ttcaaggcgg ccgagcacca gcgcgtcatc agcggcaagg ccaagcta      1848

```

<210> 42

<211> 616

<212> PRT

<213> Mortierella alpina

<400> 42

```

Met Val Ala Leu Pro Leu Val Ala Ala Val Pro Ala Ala Met Tyr
1          5          10          15
Val Ser Ser Lys Leu Ala Leu Pro Arg Asp Met Lys Leu Ile Lys Ser
          20          25          30
Leu Ile Gly Ala Lys Met Ala Tyr Ser Ala Met Glu Lys Asn Asp Ala
          35          40          45
Leu Asn Leu Thr Leu Arg Phe Asp Glu Cys Tyr Arg Lys Tyr Pro Asp
          50          55          60
Arg Glu Ala Leu Val Phe Glu Gly Lys Ser Tyr Ser Phe Arg Asp Ile
65          70          75          80
Gln Leu Ala Ser Asn Arg Cys Gly Asn Trp Leu Leu Ala Lys Gly Ile
          85          90          95
Lys Arg Gly Asp Ile Val Ser Leu Phe Met Leu Asn Arg Pro Glu Phe
          100          105          110

```

Ile	Phe	Cys	Trp	Leu	Gly	Leu	Asn	Lys	Ile	Gly	Ala	Thr	Gly	Ala	Phe
		115					120					125			
Ile	Asn	Thr	Asn	Leu	Thr	Gly	Lys	Pro	Leu	Thr	His	Ser	Leu	Arg	Thr
		130				135					140				
Ala	Thr	Ser	Ser	Met	Leu	Ile	Met	Asp	Thr	Glu	Leu	Thr	Asp	Ala	Ile
145					150					155					160
Ala	Asn	Ser	Leu	Asp	Glu	Ile	Gln	Glu	Met	Gly	Tyr	Ser	Ile	Tyr	Ser
				165					170					175	
Tyr	Gly	Pro	Glu	Ala	Val	Asp	Phe	Ala	Thr	Pro	Met	Asp	Ile	Ser	Gln
			180					185					190		
Val	Pro	Asp	Thr	Asp	Thr	Pro	Glu	His	Leu	Arg	Arg	Asn	Thr	Thr	Ala
		195					200					205			
Asp	Asp	Ile	Ala	Met	Leu	Ile	Tyr	Thr	Ser	Gly	Thr	Thr	Gly	Leu	Pro
	210					215					220				
Lys	Ala	Gly	Arg	Val	Ser	His	Ala	Arg	Ala	Ser	Met	Gly	Pro	Gln	Phe
225					230					235					240
Trp	Asn	Arg	Phe	Tyr	His	Phe	Ser	Glu	Ser	Asp	Arg	Val	Tyr	Leu	Ser
				245					250					255	
Leu	Pro	Leu	Tyr	His	Ser	Ala	Gly	Ala	Ile	Leu	Gly	Val	Ile	Ala	Cys
			260					265					270		
Trp	Thr	Ser	Gly	Ala	Thr	Leu	Ile	Leu	Ala	Arg	Lys	Phe	Ser	Ala	Thr
		275					280					285			
His	Phe	Trp	Glu	Asp	Cys	Arg	Val	Asn	Asn	Ala	Thr	Val	Ile	Gln	Tyr
	290					295					300				
Ile	Gly	Glu	Ile	Cys	Arg	Tyr	Leu	Leu	Asn	Thr	Pro	Glu	Ser	Pro	Leu
305					310					315					320
Asp	Lys	Ala	His	Ser	Ile	Arg	Leu	Ala	His	Gly	Asn	Gly	Met	Arg	Pro
				325					330					335	
Asp	Val	Trp	Thr	Arg	Phe	Arg	Asp	Arg	Phe	Gly	Ile	Pro	Leu	Ile	Gly
			340					345					350		
Glu	Trp	Tyr	Ala	Ser	Thr	Glu	Gly	Thr	Gly	Ala	Leu	Ser	Asn	Tyr	Asn
		355					360					365			
Thr	Gly	Pro	Gly	Gly	Ala	Gly	Ala	Ile	Gly	Tyr	Arg	Gly	Thr	Leu	Ala
	370					375					380				
Arg	Ala	Leu	Asp	Lys	Gly	Leu	Arg	Ile	Ala	Arg	Phe	Asp	Val	Gln	Thr
385					390					395					400
Glu	Glu	Leu	Val	Arg	Asp	Lys	Asn	Gly	Tyr	Cys	Ile	Glu	Cys	Lys	Pro
			405						410					415	
Gly	Glu	Pro	Gly	Glu	Leu	Leu	Thr	Leu	Val	Asp	Ala	Lys	Glu	Pro	Asn
			420					425					430		
Lys	Asp	Phe	Lys	Gly	Tyr	His	Gln	Asn	Gln	Ala	Ala	Thr	Asn	Lys	Lys
		435					440						445		
Ile	Val	Lys	Asp	Val	Phe	Lys	Ala	Gly	Asp	Met	Tyr	Phe	Arg	Thr	Gly
	450					455					460				
Asp	Ile	Leu	Arg	Arg	Asp	Ser	Asp	Gly	Tyr	Phe	Tyr	Phe	Gly	Asp	Arg
465					470					475					480
Val	Gly	Asp	Thr	Phe	Arg	Trp	Lys	Ser	Glu	Asn	Val	Ser	Thr	Ala	Glu
			485						490					495	
Val	Ser	Glu	Val	Leu	Ser	Gln	Tyr	Pro	Asp	Cys	Ile	Glu	Val	Asn	Val
			500					505					510		
Tyr	Gly	Val	Gln	Ile	Pro	Gly	Gln	Asp	Gly	Arg	Ala	Gly	Met	Ala	Ala
		515				520						525			
Ile	Val	Ser	Lys	Ser	Thr	Met	Asp	Trp	Glu	Lys	Phe	Ala	Ala	Tyr	Ala
	530					535					540				
Leu	Lys	Asn	Leu	Pro	Arg	Tyr	Ser	Val	Pro	Ile	Phe	Ile	Arg	Lys	Met
545					550					555					560

Pro Glu Met Glu Ile Thr Gly Thr Phe Lys Gln Arg Lys Val Glu Leu
 565 570 575
 Val Asn Glu Gly Ile Asp Pro Lys Thr Ile Ala Asn Glu Met Leu Trp
 580 585 590
 Leu Asp Gly His His Tyr Lys Pro Phe Lys Ala Ala Glu His Gln Arg
 595 600 605
 Val Ile Ser Gly Lys Ala Lys Leu
 610 615

<210> 43
 <211> 1851
 <212> DNA
 <213> Mortierella alpina

<400> 43
 atggttgctc tcccactcgt cgcagcagct gtcccagctg ccatgtatgt gagctcaaag 60
 ctggcacttc ctcgggatat gaagttgatt aagagcttga tcggagccaa gatggcctac 120
 agtgccatgg aaaagaacga cgccctcaac ctgacactcc gcttcgacga gtgctaccgc 180
 aagtatcctg accgtgaagc cctggtcttt gagggcaaat cctattcatt ccgtgatatt 240
 cagcttgccct ccaacaggtg cggcaactgg ttgctggcca aagggatcaa gcgaggagat 300
 atcgtctcgc ttttcatgtt gaacaggcca gagttcatct tctgctggct gggctctaac 360
 aagattggag ccactgggtg cttcatcaat accaacctta cgggcaaacc cctgacacat 420
 tccctccgga cagccacgtc gtcaatgttg atcatggaca cggagttgac agacgcgac 480
 gccaaactccc tggatgagat tcaggagatg ggctattcaa tttactctta cggacccgaa 540
 gccgtggact ttgctacccc gatggatata tcgcagggtc cagacaccga tacacccgaa 600
 cacctgcgcc ggaacacgac cgcggatgac attgcgatgc tcatctacac ctctggaact 660
 actggtcttc ccaaggccgg tcgtgtctct catgcgcgtg cctctatggg acctcagttt 720
 tggaaccgat tctatcactt cagtgaagac gacaggggtc atctgtcctt gcccttgtac 780
 cacagtgtc ggcctatctt gggagtgtt gcttggttga cctcgggagc aaccttgatc 840
 ctggcccgcga agttctccgc gacacatttc tgggaggatt gccgcgtgaa caacgcaact 900
 gtgattcaat acattggaga aatttgcaga tatctgtctc acacgccaga atcacccctg 960
 gacaaggcac actcgatacg actggcacat ggtaatggaa tgcgaccoga tgtctggact 1020
 cgcttcagag atcgggttcgg catcccgttg attggcgagt ggtatgcac gactgaggga 1080
 actggagcct tgtcgaatta taacacaggg ccaggcggcg ctggagcgat tggataccgc 1140
 ggtacccttg ccagagcatt ggataaagga ctcaggattg cgagatttga tgtccagaca 1200
 gaggagttgg ttcgggacaa aaacggttat tgcattgagt gcaaacctgg cgagcccgga 1260
 gaattgctga cgcttggttga tgctaaagag ccgaacaaag acttcaaagg ataccatcaa 1320
 aaccaggcag cgaccaacaa aaagattgtc aaagatgttt tcaaagccgg cgacatgtac 1380
 ttccgtaccg gagatatact tcggcgcgat agcgatgggt acttttactt tggcgaccgt 1440
 gtgggcgata cattccggtg gaagtcogag aatgtgtcta cggccgaggt gtctgaagtc 1500
 ctctcgcagt atccggactg tatcgaagtc aatgtgtatg gagttcagat cccagggcag 1560
 gacggacgcg ccggtatggc agcgattgtg tccaagagca cgatggattg ggagaaattt 1620
 ggcggcgtatg cactcaagaa cctgccgcgg tattctgttc cgatctttat ccgcaagatg 1680
 cccgagatgg agatcacagg gacgttcaag cagcgcaaag tcgagttggg gaatgaggga 1740
 atcgacccca agacgattgc caacgagatg ctgtgggttg acggacacca ctataagccg 1800
 ttcaaggcgg ccgagcacca gcgcgtcatc agcggcaagg ccaagctata g 1851

<210> 44
 <211> 2017
 <212> DNA
 <213> Mortierella alpina

<400> 44
 tcgctateta tcacccctca ctcccactc cgcactctgc tcttcttttt tcctttctct 60
 ctctcaccgt cgccactgtc tctactttct ttaccacca cgcatacgtc acagcatgg 120

tgcctctcca	ctcgtcgcag	cagctgtccc	agctgccatg	tatgtgagct	caaagctggc	180
acttcctcgg	gatatgaagt	tgattaagag	cttgatcgga	gccaagatgg	cctacagtgc	240
catggaaaag	aacgacgccc	tcaacctgac	actccgcttc	gacgagtgtc	accgcaagta	300
tcctgaccgt	gaagccctgg	tctttgaggg	caaatcctat	tcattccgtg	atattcagct	360
tgcctccaac	aggtgcggca	actggttgct	ggccaaaggg	atcaagcgag	gagatatact	420
ctcgtctttc	atgttgaaca	ggccagagtt	catcttctgc	tggctgggtc	tcaacaagat	480
tggagccact	ggtgccttca	tcaataccaa	ccttacgggc	aaaccctga	cacattccct	540
ccgacagcc	acgtcgtcaa	tgttgatcat	ggacacggag	ttgacagacg	cgatcgccaa	600
ctccctggat	gagattcagg	agatgggcta	ttcaatttac	tcttacggac	ccgaagccgt	660
ggactttgct	accccgatgg	atatctcgca	ggtcccagac	accgatacac	ccgaacacct	720
gcgcgggaac	acgaccgcgg	atgacattgc	gatgctcatc	tacacctctg	gaactactgg	780
tcttcccaag	gccggtcgtg	tctctcatgc	gcgtgcctct	atgggacctc	agttttggaa	840
ccgattctat	cacttcagtg	agagcgacag	ggtctatctg	tccttgccct	tgtaccacag	900
tgctggcgcc	atcttgggag	tgattgcttg	ttggacctcg	ggagcaacct	tgatcctggc	960
ccgcaagttc	tccgcgacac	atttctggga	ggattgccgc	gtgaacaacg	caactgtgat	1020
tcaatacatt	ggagaaatth	gcagatatct	gctcaacacg	ccagaatcac	ccctggacaa	1080
ggcacactcg	atacgactgg	cacatggtaa	tggaatgcga	cccgatgtct	ggactcgctt	1140
cagagatcgg	ttcggcatcc	cgttgattgg	cgagtgggat	gcacgactg	agggaactgg	1200
agccttgctg	aattataaca	caggcccagg	cggcgctgga	gcgattggat	accgcggtac	1260
ccttgccaga	gcattggata	aaggactcag	gattgcgaga	tttgatgtcc	agacagagga	1320
gttggttcgg	gacaaaaacg	gttattgcat	tgagtgcaaa	cctggcgagc	ccggagaatt	1380
gctgacgctt	gttgatgcta	aagagccgaa	caaagacttc	aaaggatacc	atcaaaacca	1440
ggcagcgacc	aacaaaaaga	ttgtcaaaga	tgttttcaaa	gccggcgaca	tgtacttcgg	1500
taccggagat	atccttcggc	gcgatagcga	tgggtacttt	tactttggcg	accgtgtggg	1560
cgatacattc	cgggtggaagt	ccgagaatgt	gtctacggcc	gaggtgtctg	aagtcctctc	1620
gcagtatccg	gactgtatcg	aagtcaatgt	gtatggagtt	cagatcccag	ggcaggacgg	1680
acgcgcgggt	atggcagcga	ttgtgtccaa	gagcacgatg	gattgggaga	aatttgcggc	1740
gtatgcactc	aagaacctgc	cgcggtatte	tgttccgata	tttatccgca	agatgcccga	1800
gatggagatc	acagggacgt	tcaagcagcg	caaagtcgag	ttggtgaatg	agggaatcga	1860
ccccaagacg	attgccaacg	agatgctgtg	gttggaagg	caccactata	agccgttcaa	1920
ggcggccgag	caccagcgcg	tcacagcgcg	caaggccaag	ctatagtagg	gcgcgtgcgc	1980
caatgcagta	gcaatactat	tccccgcttt	gtccatt			2017

<210> 45
 <211> 2345
 <212> DNA
 <213> Mortierella alpina

atggttggtg	agtaaacagc	gatcgccctg	ggtccaatcg	actcatctgc	taatctgata	60
tcacacgtcc	cttatcaatt	cacaaaagaa	aaaaaagaga	caagagagag	aattactaac	120
attgctctct	ccctcgtgtc	gatgcagctc	tcccactcgt	cgcagcagct	gtcccagctg	180
ccatgtatgt	gagctcaaag	ctggcacttc	ctcgggatat	gaagttgatt	aagagcttga	240
tcggagccaa	gatggcctac	agtgccatgg	aaaagaacga	cgccctcaac	ctgacactcc	300
gcttcgacga	gtgctaccgc	aagtatcctg	accgtgaagc	cctgggtcttt	gagggcaaatt	360
cctattcatt	ccgtgatatt	cagcttggtg	agccattggt	acaaggcact	cacccacgcg	420
ttcctgcttt	cagagctcac	gactgctcga	accttcaatt	ttttgtatcg	atcgatcacc	480
gatcgatcat	gcgcccacat	tgcctagagc	tatggcacac	gcgctaattg	atgccttctt	540
gttatgaagc	ctccaacagg	tgcggcaact	ggttgctggc	caaagggatc	aagcgaggag	600
atategtctc	gcttttcatg	ttgaacaggc	cagagttcat	cttctgctgg	ctgggtctca	660
acaagattgg	agccactggg	gccttcatca	ataccaacct	tacgggcaaa	cccctgacac	720
attccctccg	gacagccacg	tcgtcaatgt	tgatcatgga	cacggagtgt	acagacgcga	780
tcgccaaactc	cctggatgag	attcaggaga	tgggctattc	aatttactct	tacggacccg	840
aagccgtgga	ctttgctacc	ccgatggata	tctcgcaggt	cccagacacc	gatacaccgg	900
aacacctgcg	ccggaacacg	accgcggatg	acattgcgat	gctcatctac	acctctggaa	960
ctactggtct	tcccaaggcc	ggtcgtgtct	ctcatgcgcg	tgcctctagt	aagttgagag	1020

tcttcagcct	ttgacatacg	tatTTTTTTga	gcgtgctact	aacagttctc	gttgccgtta	1080
tctgcatatt	tttagtgga	cctcagtttt	ggaacgatt	ctatcacttc	agtgagagcg	1140
acagggctta	tctgtccttg	cccttgatcc	acagtgcctg	cgccatcttg	ggagtgattg	1200
cttggttgac	ctcgggagca	accttgatcc	tggcccgcga	gttctccgcg	acacatttct	1260
gggaggattg	ccgcgtgaac	aacgcaactg	tgattcaata	cattggagaa	atttgcagat	1320
atctgctcaa	cacgccagaa	tcacccctgg	acaaggcaca	ctcgatacga	ctggcacatg	1380
gtaatggaat	gcgaaccgat	gtctggactc	gcttcagaga	tcggttcggc	atcccgttga	1440
ttggcgagtg	gtatgcatcg	actgaggga	ctggagcctt	gtcgaattat	aacacaggcc	1500
caggcggcgc	tggagcgatt	ggataccgcg	gtacccttgc	cagagcattg	gataaaggac	1560
tcaggattgc	gagatttgat	gtccagacag	aggagttggt	tcgggacaaa	aacggttatt	1620
gcattgaggt	aaaacataag	gcgcattgtt	gaagtctaaa	tcacttcaga	tgctttgttc	1680
ggcatgctta	ccaaacgcac	cgaccatctt	ttcactggtg	cggcatataa	tagtgcaaac	1740
ctggcgagcc	cggagaattg	ctgacgcttg	ttgatgctaa	agagccgaac	aaagacttca	1800
aaggatacca	tcaaaaccag	gcagcgacca	acaaaaagat	tgtcaaagat	gttttcaaag	1860
ccggcgacat	gtacttccgt	accggagata	tccttcggcg	cgatagcgat	gggtactttt	1920
actttggcga	ccgtgtgggc	gatacattcc	ggtggaagtc	cgagaatgtg	tctacggccg	1980
aggtgtctga	agtcctctcg	cagtatccgg	actgtatcga	agtcaatgtg	tatggagttc	2040
agatcccagg	gcaggacgga	cgcgcgggta	tggcagcgat	tgtgtccaag	agcacgatgg	2100
attgggagaa	atttgcggcg	tatgactcea	agaacctgcc	gcggtattct	gttccgatct	2160
ttatccgcaa	gatgcccag	atggagatca	cagggacgtt	caagcagcgc	aaagtcgagt	2220
tggatgaatga	gggaatcgac	cccaagacga	ttgccaacga	gatgctgtgg	ttggacggac	2280
accactataa	gccgttcaag	gcggccgagc	accagcgcgt	catcagcggc	aaggccaagc	2340
tatag						2345

<210> 46
 <211> 2073
 <212> DNA
 <213> Mortierella alpina

<400> 46	
atggaaacct	60
cgcaacgtca	120
aaggagcgca	180
gccatgggct	240
ggcgagaagc	300
cgccaagcca	360
aagggagact	420
ggatgcttct	480
atccagttta	540
cctgtggtgt	600
ggccaagagt	660
tttgagctgg	720
tctgccggaa	780
gaggtcatat	840
ggtcagccca	900
gcggccatgg	960
ttggcccatg	1020
ggattcggac	1080
aggtccaagg	1140
acgatctggg	1200
atccgtacac	1260
tctgagaact	1320
gttgagggca	1380
cggttcttga	1440
ggtgttgccg	1500
gcgcccagtc	1560

```

ggatcgggcg aaatctgggt ccgtggacct gcagtgatga agggatactt caaactcgag 1620
gaagagacca agaaggtgat gaccgggggat gggtgggttca agacaggcga cattggcacg 1680
atgaacccag acggcacact gtcaatcaag gacaggggtca agaatctggt caagctgtct 1740
catggagaat atgtcgccct ggagaaatgt gaagccgttt atcgcgattc caaggagatc 1800
aagagcattt gcatcgttgc ggacaatggg tgcctgtgtg tgctggccgt tgtggaaccg 1860
agccacgcag gggcgtctga caaggagatt ttggatatcc tgaagagcca agccaaggcg 1920
gcgggcctct ccaagtccga gactgtgcaa ggcgttatca ttgatgattc ggactggatg 1980
acgaatgggt tcatgacctc gagcagcaag gtcaagagac gcgagggtccg caaggcacac 2040
aacaaggata ttgaggagat gtggaagaag ttc 2073

```

<210> 47
 <211> 691
 <212> PRT
 <213> Mortierella alpina

<400> 47

Met	Glu	Thr	Leu	Val	Asn	Gly	Lys	Tyr	Ala	Val	Glu	Tyr	Asp	Glu	Val	1	5	10	15
Asp	His	Ile	Tyr	Arg	Asn	Val	Met	Ala	Thr	Gly	Gly	Leu	Leu	Asp	Arg	20	25	30	
Pro	Met	Pro	Pro	Tyr	Tyr	Asp	Ile	Lys	Glu	Arg	Thr	Met	Ala	His	Leu	35	40	45	
Phe	Glu	Tyr	Met	Ala	Asn	Thr	Tyr	Glu	Asp	Lys	Asp	Ala	Met	Gly	Trp	50	55	60	
Arg	Asp	Ile	Ile	Lys	Val	His	Lys	Val	Glu	Lys	Gln	Ala	Ala	Asn	Pro	65	70	75	80
Gly	Glu	Lys	Pro	Lys	Thr	Trp	Ile	Thr	Tyr	Glu	Leu	Ser	Asp	Tyr	Asn	85	90	95	
Trp	Met	Ser	Tyr	Arg	Gln	Ala	Lys	Asn	Tyr	Ala	Asp	Arg	Val	Gly	Leu	100	105	110	
Gly	Ile	Thr	Arg	Leu	Gly	Val	Glu	Lys	Gly	Asp	Phe	Val	Met	Ile	Phe	115	120	125	
Ala	Ser	Thr	Cys	Pro	Glu	Trp	Phe	Leu	Thr	Ala	His	Gly	Cys	Phe	Ser	130	135	140	
Gln	Ser	Val	Thr	Ile	Val	Thr	Ala	Tyr	Asp	Ser	Met	Asp	Glu	Lys	Ser	145	150	155	160
Ile	Gln	Phe	Ile	Val	Asp	Gln	Ser	Gln	Pro	Lys	Ala	Ile	Phe	Ala	Asp	165	170	175	
Ala	His	Thr	Leu	Pro	Val	Val	Ser	Lys	Leu	Met	Gln	Lys	Gly	Asn	Ser	180	185	190	
Gly	Val	Lys	Ala	Val	Ile	Tyr	Thr	Gly	Gln	Glu	Trp	Glu	Val	Thr	Asp	195	200	205	
Ala	Ile	Lys	Lys	Met	Glu	Gln	Val	Glu	Asn	Arg	Ser	Phe	Glu	Leu	Val	210	215	220	
His	Ile	Asp	Glu	Leu	Lys	Lys	Thr	Lys	Ser	Ala	Ser	Asn	Gly	Glu	Gln	225	230	235	240
Ser	Ala	Gly	Lys	Gly	Lys	Gln	Arg	Ser	Ser	Glu	Asp	Ala	Glu	Gly	Ala	245	250	255	
Gln	Asp	Glu	Ile	Glu	Val	Ile	Tyr	Pro	Lys	Ala	Asp	Asp	Leu	Ala	Cys	260	265	270	
Ile	Met	Tyr	Thr	Ser	Gly	Ser	Thr	Gly	Gln	Pro	Lys	Gly	Ala	Gln	Leu	275	280	285	
Thr	His	Gly	Asn	Leu	Met	Ala	Ala	Ile	Gly	Ser	Ala	Ala	Ala	Met	Glu	290	295	300	
Gly	Asp	Gln	Leu	Asp	Lys	Glu	Thr	Asp	Ile	Val	Ile	Ser	Tyr	Leu	Pro	305	310	315	320

Leu Ala His Val Leu Glu Phe Val Ile Ser His Phe Val Val Ser Met
 325 330 335
 Gly Cys Arg Leu Gly Phe Gly Arg Ala Arg Thr Leu Met Asp Asp Ala
 340 345 350
 Val Ala Pro Thr Ala Gly Ser Gly Arg Ser Lys Gly Leu Gly Asp Leu
 355 360 365
 Lys Ala Leu Gln Pro Thr Leu Met Ala Gly Val Pro Thr Ile Trp Glu
 370 375 380
 Arg Ile Arg Lys Gly Ile Leu Ala Glu Val Asn Lys Gln Ser Phe Pro
 385 390 395 400
 Ile Arg Thr Leu Phe Phe Ala Ala Leu Asn Thr Lys Trp Ala Ile Val
 405 410 415
 Gln Ala Thr Gly Ser Glu Asn Phe Val Thr Lys Thr Ile Asp Ser Leu
 420 425 430
 Val Phe Ser Lys Ala Lys Glu Leu Val Gly Gly Lys Leu Arg Leu Thr
 435 440 445
 Leu Thr Gly Gly Ala Gly Ile Ser Asp Glu Thr His Arg Phe Leu Ser
 450 455 460
 Met Val Met Cys Tyr Val Ile Ser Gly Tyr Gly Leu Thr Glu Val Cys
 465 470 475 480
 Gly Val Ala Ala Val Thr Leu Pro Arg Met Gly His Arg Leu Arg Thr
 485 490 495
 Val Gly Pro Pro Ala Pro Ser Leu Glu Leu Lys Leu Val Asn Val Pro
 500 505 510
 Asp Thr Glu Tyr Thr Gly Asp Asn Gly Ser Gly Glu Ile Trp Phe Arg
 515 520 525
 Gly Pro Ala Val Met Lys Gly Tyr Phe Lys Leu Glu Glu Glu Thr Lys
 530 535 540
 Lys Val Met Thr Gly Asp Gly Trp Phe Lys Thr Gly Asp Ile Gly Thr
 545 550 555 560
 Met Asn Pro Asp Gly Thr Leu Ser Ile Lys Asp Arg Val Lys Asn Leu
 565 570 575
 Val Lys Leu Ser His Gly Glu Tyr Val Ala Leu Glu Lys Cys Glu Ala
 580 585 590
 Val Tyr Arg Asp Ser Lys Glu Ile Lys Ser Ile Cys Ile Val Ala Asp
 595 600 605
 Asn Gly Cys Pro Val Leu Leu Ala Val Val Glu Pro Ser His Ala Gly
 610 615 620
 Ala Ser Asp Lys Glu Ile Leu Asp Ile Leu Lys Ser Gln Ala Lys Ala
 625 630 635 640
 Ala Gly Leu Ser Lys Ser Glu Thr Val Gln Gly Val Ile Ile Asp Asp
 645 650 655
 Ser Asp Trp Met Thr Asn Gly Phe Met Thr Ser Ser Ser Lys Val Lys
 660 665 670
 Arg Arg Glu Val Arg Lys Ala His Asn Lys Asp Ile Glu Glu Met Trp
 675 680 685
 Lys Lys Phe
 690

<210> 48
 <211> 2076
 <212> DNA
 <213> Mortierella alpina

<400> 48
 atggaaacct tgggtaacgg aaagtatgcg gtcgagtacg acgaggtcga tcacatctat

cgcaacgtca	tggctacagg	cgggctcctc	gacaggccta	tgcctccata	ctacgacatc	120
aaggagcgca	ccatggccca	cctctttgag	tatatggcca	acacctacga	agacaaagac	180
gccatgggct	ggcgagacat	tatcaaggtc	cacaaggctc	agaagcaggc	tgccaatcct	240
ggcgagaagc	caaagacctg	gatcaattat	gagctctcgg	actacaactg	gatgtcgtac	300
cgccaagcca	agaactatgc	agatcgagtt	ggcttgggca	tcacacgcct	tggagttgag	360
aagggagact	ttgtcatgat	ctttgctagc	acatgtcccc	aatggttcct	gacagcgcat	420
ggatgcttct	cgcagtcagt	gactatcgtg	acagcctacg	actcgatgga	cgagaagtcg	480
atccagttta	ttgttgacca	gtcccagccc	aaggccatct	ttgctgatgc	gcacacgctc	540
cctgtgggtg	ccaaactcat	gcagaagggc	aacagtgggt	tcaaggcagt	catttacaca	600
ggccaagagt	gggaagtgac	cgatgcaatc	aagaagatgg	agcaagtaga	aaaccgctca	660
tttgagctgg	ttcatatcga	cgaactcaag	aagaccaagt	cagcatctaa	cggcgaacag	720
tctgccggaa	aggggaagca	gagatcatct	gaggatgccg	aaggcgctca	ggacgagatc	780
gaggtcatat	accctaaggc	ggatgatctg	gcctgtatta	tgtatacctc	tgggtcgacg	840
ggtcagccca	agggcgcgca	attgacacat	ggcaacttga	tggcggccat	tgggaagtgt	900
gcggccatgg	agggcgacca	gctggacaag	gaaacagaca	ttgttatttc	atatctgccca	960
ttggeccatg	tcctcgagtt	tgtcatttcc	cactttgtgg	tatccatggg	ctgccgtctt	1020
ggattcggac	gagcacgcac	tctgatggat	gatgcagtcg	ctcccacgcg	aggaagtggc	1080
aggtccaagg	gccttgggtga	tctgaaggcg	ctccagccaa	cattgatggc	tgggtgtgcca	1140
acgatctggg	agcgtatccg	caagggcatc	ctggccgagg	tcaacaagca	atccttccct	1200
atccgtacac	tcttctttgc	tgcactcaac	accaagtggg	ctatcgtcca	ggctaccgga	1260
tctgagaact	ttgtcaccaa	gactattgac	tcgttgggtc	ttagtaaggc	taaggagctc	1320
gttggaggca	agctgcgcct	taccttgact	ggagggggcg	gaatcagtga	tgagacgcac	1380
cggttcttga	gcatggtaat	gtgctacgtt	atctcgggat	atggtctcac	tgaagtctgt	1440
ggtgttgccg	ctgtcaccct	gccacgtatg	ggtcaccgtc	tcaggaccgt	tggaccaccc	1500
gcgcccagtc	ttgagctgaa	gttgggtgaat	gtgcccgaca	ccgagtacac	aggagacaat	1560
ggatcgggcg	aaatctgggt	ccgtggacct	gcagtgatga	agggatactt	caaactcgag	1620
gaagagacca	agaaggtgat	gaccggggat	ggttggttca	agacaggcga	cattggcacg	1680
atgaaccag	acggcacact	gtcaatcaag	gacagggtca	agaatctggg	caagctgtct	1740
catggagaat	atgtcgcctt	ggagaaatgt	gaagccgttt	atcgcgattc	caaggagatc	1800
aagagcattt	gcatcgttgc	ggacaatggg	tgcctgtgt	tgttgccgt	tgtggaaccg	1860
agccacgcag	gggcgtctga	caaggagatt	ttggatatcc	tgaagagcca	agccaaggcg	1920
gcgggcctct	ccaagtccga	gactgtgcaa	ggcgttatca	ttgatgattc	ggactggatg	1980
acgaatgggt	tcatgacctc	gagcagcaag	gtcaagagac	gcgaggtcgg	caaggcacac	2040
aacaaggata	ttgaggagat	gtggaagaag	ttctag			2076

<210> 49

<211> 2170

<212> DNA

<213> Mortierella alpina

<400> 49

ttcacgcact	accttctttt	ttcaccatc	cagcagtaaa	gacaacatgg	aaaccttggg	60
taacggaaag	tatgcggtcg	agtacgacga	ggtcgatcac	atctatcgca	acgtcatggc	120
tacaggcggg	ctcctcgaca	ggcctatgcc	tccatactac	gacatcaagg	agcgcaccat	180
ggcccacctc	tttgagtata	tggccaacac	ctacgaagac	aaagacgcca	tgggctggcg	240
agacattatc	aagggtccaca	aggtcgagaa	gcaggctgcc	aatcctggcg	agaagccaaa	300
gacctggatc	acttatgagc	tctcggacta	caactggatg	tcgtaccgcc	aagccaagaa	360
ctatgcagat	cgagttggct	tgggcatcac	acgccttggg	gttgagaagg	gagactttgt	420
catgatcttt	gctagcacat	gtcccgaatg	gttcttgaca	gcgcatggat	gcttctcgca	480
gtcagtgact	atcgtgacag	cctacgactc	gatggacgag	aagtcgatec	agtttattgt	540
tgaccagtcc	cagcccaagg	ccatctttgc	tgatgcgcac	acgctccctg	tgggtgtccaa	600
actcatgcag	aagggcaaca	gtgggtgtcaa	ggcagtcatt	tacacaggcc	aagagtggga	660
agtgaccgat	gcaatcaaga	agatggagca	agtagaaaac	cgctcatttg	agctgggttca	720
tatcgacgaa	ctcaagaaga	ccaagtcagc	atctaacggc	gaacagtctg	ccggaaaggg	780
gaagcagaga	tcatctgagg	atgccgaagg	cgctcaggac	gagatcgagg	tcatataccc	840
taaggcggat	gatctggcct	gtattatgta	tacctctggg	tcgacgggtc	agcccaaggg	900

cgcgcaattg	acacatggca	acttgatggc	ggccattgga	agtgctgcgg	ccatggaggg	960
cgaccagctg	gacaaggaaa	cagacattgt	tatttcatat	ctgccattgg	cccatgtcct	1020
cgagtttgtc	atttcccact	ttgtgggtatc	catgggctgc	cgtcttggat	tcggacgagc	1080
acgcactctg	atggatgatg	cagtcgctcc	caccgcagga	agtggcaggt	ccaagggcct	1140
tggtgatctg	aaggcgctcc	agccaacatt	gatggctggg	gtgccaacga	tctgggagcg	1200
tatccgcaag	ggcatcctgg	ccgaggtcaa	caagcaatcc	ttccctatcc	gtacactctt	1260
ctttgctgca	ctcaacacca	agtgggctat	cgtccagggt	accggatctg	agaactttgt	1320
caccaagact	attgactcgt	tggtcttttag	taaggctaag	gagctcgttg	gaggcaagct	1380
gcgccettacc	ttgactggag	gggcccgaat	cagtgatgag	acgcaccggg	tcttgagcat	1440
ggtaatgtgc	tacgttatct	cgggatatgg	tctcactgaa	gtctgtgggtg	ttgccgctgt	1500
caccctgcca	cgtatgggtc	accgtctcag	gaccgttgga	ccacccgcgc	ccagtcttga	1560
gctgaagttg	gtgaatgtgc	ccgacaccga	gtacacagga	gacaatggat	cgggcgaaat	1620
ctggttccgt	ggacctgcag	tgatgaaggg	atacttcaaa	ctcgaggaag	agaccaagaa	1680
ggtgatgacc	ggggatgggt	ggttcaagac	aggcgacatt	ggcacgatga	acccagacgg	1740
cacactgtca	atcaaggaca	gggtcaagaa	tctggtcaag	ctgtctcatg	gagaatatgt	1800
cgccttgag	aaatgtgaag	ccgtttatcg	cgattccaag	gagatcaaga	gcatttgcac	1860
cgttgccggac	aatgggtgcc	ctgtgttgct	ggccgttggtg	gaaccgagcc	acgcaggggc	1920
gtctgacaag	gagattttgg	atatcctgaa	gagccaagcc	aaggcggcgg	gcctctccaa	1980
gtccgagact	gtgcaaggcg	ttatcattga	tgattcggac	tggatgacga	atgggttcat	2040
gacctcgagc	agcaagggtca	agagacgcga	ggtccgcaag	gcacacaaca	aggatattga	2100
ggagatgtgg	aagaagttct	agagaagcgt	gggaagggca	tgaataaac	atacgcaatg	2160
gatttattgg						2170

<210> 50

<211> 3236

<212> DNA

<213> Mortierella alpina

<400> 50

atggaaacct	tggttaacgg	aaagtatgcg	gtcgagtacg	acgaggtcga	tcacatctat	60
cgcaacgtca	tggtctacagg	cgggctcctc	gacaggccta	tgcctccata	ctacgacatc	120
aaggagcgca	ccatggccca	cctcttttgag	tatatggcca	acacctacga	agacaaagac	180
gccatgggct	ggcgagacat	tatcaaggta	ttgactgccc	cggcctatca	cttttttacc	240
cacacgatcc	tccctttttt	ttctctccca	ttcttctatc	ctgaccgtat	cgctatcgaa	300
cgagtcaacg	agatcagttc	ccacgcttac	tttactccct	cgtctgattc	tgattatttt	360
ctctctctgc	ctcttcgtga	tttgttcgca	ggtccacaag	gtcgagaagc	aggctgccaa	420
tcctggcgag	aagccaaaga	cctggatcac	ttatgagctc	tcggactaca	actggatgtc	480
gtaccgccaa	gccaagaact	atgcagatcg	agttggcttg	ggcatcacac	gccttgaggt	540
tgagaaggga	gactttgtca	tgatctttgc	tagcacatgg	tataacctca	tgagacaaac	600
tttcgcatcc	atgacagcat	cgcaagaaaa	aaaaaagaga	aggaaattat	gataattcgg	660
gccaagcgac	taaaacaccg	ctatcgcacg	cttttttttt	ttttttcgct	atcttgcatt	720
tcttatcgtt	caaatagtcc	cgaatggttc	ctgacagcgc	atggtaagct	ttttcttttt	780
gcggatgatg	atttcttctt	atgacagcat	gatgagattc	aaccgggatt	aggtgacggc	840
tgttgattgt	gcaaaagggg	ggccggacaa	tttgaagaga	cttggggagg	tgtttgatg	900
gacgacaaat	ccaagattca	acagaagact	ggtgccaggg	aagagtagag	agggtgccct	960
gttctttggc	atatgaaagt	ggattacgat	atggctgagt	taggagtcta	attcatcgcg	1020
gaaaaagagt	cgcgactgaa	gtcccaagtc	ggggactgta	gaaagtattc	cactcgtgtc	1080
tggtgaaatg	aggaggactg	ggttgggggt	tggagtgccg	acgagaaatc	atggaatacg	1140
cttctcggtc	ttcagtccac	tgatcactca	tggcgcaatt	gactctacaa	taataggatg	1200
cttctcgcag	tcagtgaact	tcgtgacagc	ctacgactcg	atggacgaga	agtcgatcca	1260
gtttattgtt	gaccagtccc	agcccaaggc	catctttgct	gatgcgcaca	cgctccctgt	1320
ggtgtccaaa	ctcatgcaga	agggaacacg	tggtgtcaag	gcagtcattt	acacaggcca	1380
agagtgggaa	gtgaccgatg	caatcaagaa	gatggagcaa	gtagaaaacc	gctcatttga	1440
gctggttcat	atcgacgaac	tcaagaagac	caagtcagca	tctaaccggc	aacagtctgc	1500
cggaaagggg	aagcagagat	catctgagga	tgccgaaggc	gctcaggacg	agatcgaggt	1560
catataccct	aaggcggatg	atctggcctg	tattatgtat	acctctgggt	cgacgggtca	1620

gcccgaagggc	gcgcaattga	cacatggcaa	cttgatggcg	gccattggaa	gtgctgcggc	1680
catggagggc	gaccagctgg	acaaggaaac	agacattgtt	atttcatatc	tgccattggc	1740
ccatgtcctc	gagtttgtca	tttcccactt	tgtggatatcc	atggtaagtc	gaacatccct	1800
ttacatctgc	atccccaaat	gcggctaaag	tcaagttgtt	gacctgaacg	tttatattca	1860
tcttagggct	gccgtcttgg	attcggacga	gcacgcactc	tgatggatga	tgcatctcgt	1920
cccaccgcag	gaagtggcag	gtccaagggc	cttgggtgatc	tgaaggcgct	ccagccaaca	1980
ttgatgggta	tgattggcat	gaagccagag	actaatatga	cgtagagtgg	gacaaaacct	2040
tattttgacg	tattgcatat	tgtgtcgatt	ttcaaaaacg	atagctgggtg	tgccaacgat	2100
ctgggagcgt	atccgcaagg	gcatcctggc	cgagggtcaac	aagcaatcct	tccctatccg	2160
tacactcttc	tttgctgcac	tcaacaccaa	gtgggctatc	gtccaggcta	ccggatctga	2220
gaactttgtc	accaagacta	ttgactcgtt	ggtctttagt	aaggctaagg	agctcgtttg	2280
aggcaagctg	cgccttacct	tgactggagg	ggccggaatc	agtgatgaga	cgcaccggtt	2340
cttgagcatg	gtaatgtgct	acgttatctc	gggatatggg	ctcactgaag	tctgtggtgt	2400
tgccgctgtc	accctgccac	gtatgggtca	ccgtctcagg	accgttggac	cacccgtaag	2460
tccgctcgca	tatcttctcg	cgatctgaat	atgcgagctg	tttttttttg	tctttcaaaa	2520
gctaacactt	tgctttttcg	cgacaacagg	cgcccagctc	tgagctgaag	ttggtgaatg	2580
tgcccgcacac	cgagtacaca	ggagacaatg	gatcgggcga	aatctgggtc	cgtggacctg	2640
cagtgatgaa	gggatacttc	aaactcgagg	aagagaccaa	gaaggatgatg	accgggggatg	2700
gttggttcaa	gacaggcgac	attggcacga	tgaaccacga	cggcacactg	tcaatcaagg	2760
acaggggtcaa	gaatctggtc	aagctgtctc	atggagaata	tgctgcacctg	gagaaatgtg	2820
aagccgttta	tcgcgattcc	aaggagatca	agagcatttg	catcgttgcg	gacaatgggt	2880
gccctgtgtt	gctggccgtt	gtggaaccga	gccacgcagg	ggtgagtga	acgcttttgg	2940
cctgacaagg	tctcgttctt	atgggaatgt	ggatcattac	tctatcata	acgagagagg	3000
ctgtattatt	ctacttgctg	aacgtggcat	aggcgtctga	caaggagatt	ttggatatcc	3060
tgaagagcca	agccaaggcg	gcgggcctct	ccaagtccga	gactgtgcaa	ggcgttatca	3120
ttgatgattc	ggactggatg	acgaatgggt	tcatgacctc	gagcagcaag	gtcaagagac	3180
gcgagggtccg	caaggcacac	aacaaggata	ttgaggagat	gtggaagaag	ttctag	3236

<210> 51
 <211> 2040
 <212> DNA
 <213> Mortierella alpina

<400> 51						
atgccaaagt	gctttaccgt	caacgtcggc	cccaggagacg	tcaagggcga	gactcgcac	60
cgctcgtcca	tccaggccgt	cgacaaactc	atggactcac	cctcaagcga	catcaagacc	120
ttgtacgatg	tcatccagta	ctctgccaa	gtccgcccc	acctcaacgc	catcggctac	180
cgcaagattg	tcaagatgat	cgaagaggaa	aaggagatca	ccaagatggg	cagcggcgag	240
cctgtcaagg	agaaaaagac	gtggaaatac	ttcaagctct	ccggtacca	ctatctgacc	300
tacaaggaca	ccaaggccgt	catcgacagc	attggaagtg	gcctgcgcaa	gtgggggtgtt	360
gagcccaagg	agaggatcac	cgtctttggg	tccacaagtg	ccaactggct	gctggctcgt	420
catgggtgct	tcacgcagtc	catgaccatc	gtcaccgtgt	atgacacttt	gggcgaggaa	480
ggattgctgc	actcgatgaa	cgaggccgag	gtgggaacgg	cctacacgaa	cgtgatattg	540
atcaagacaa	tgaccaacgt	ttcaggacgc	tgccccaccc	tcaagaggat	cgtctatgac	600
ggcgaagcca	acgcagcaga	cgtgatcgcc	cttcagacgg	cccatcctca	ccttcagctt	660
atcactctgg	aggagctgaa	gcagctcggg	gtggatcacc	ctgtggagcc	cactcctccc	720
accgccgagg	attgctcctg	catcatgtac	acttctggat	cgaccggaaa	ccctaaggga	780
gtcatcctca	ctcacggaaa	cctcattgcc	gccattggcg	gagttaacaa	gatgctggaa	840
aagtacattc	gcgaaggcga	tgtcttgctt	gcctaccttc	ccttggctca	cgttctggaa	900
ttcatggttg	agaacctctg	tctcttctgg	gggtgaaccc	ttggatatgg	tactgtccgc	960
acgtgacgg	atgcctctgt	gcgtgagtgc	caggggtgata	tcaaggaatt	gcggcctacg	1020
cttatgaccg	gcgttccagc	agtgtgggag	accatccgca	aagggtgttct	cgcccaagta	1080
aaccagggtt	cacctctggt	tcaatccgtc	ttcaacgcgg	ctctgaacgc	caaggcctgg	1140
tgcatggacc	gcaaactagg	cgctttgact	ggaattttcg	acactgtggg	gttcaacaag	1200
gtccgtcagc	aaactggagg	tcgtcttcgc	tacgcgctct	cgggcggcgc	gcctatctcc	1260
caggagaccc	agcgcttctt	gaccacagca	ctgtgcctta	tccttcaagc	ctatggcatg	1320

<400>	52															
Met	Pro	Lys	Cys	Phe	Thr	Val	Asn	Val	Gly	Pro	Glu	Asp	Val	Lys	Gly	
1				5					10					15		
Glu	Thr	Arg	Ile	Arg	Arg	Ser	Ile	Gln	Ala	Val	Asp	Lys	Leu	Met	Asp	
			20					25					30			
Ser	Pro	Ser	Ser	Asp	Ile	Lys	Thr	Leu	Tyr	Asp	Val	Ile	Gln	Tyr	Ser	
		35				40						45				
Ala	Lys	Val	Arg	Pro	Asn	Leu	Asn	Ala	Ile	Gly	Tyr	Arg	Lys	Ile	Val	
	50					55					60					
Lys	Met	Ile	Glu	Glu	Glu	Lys	Glu	Ile	Thr	Lys	Met	Val	Ser	Gly	Glu	
65					70					75					80	
Pro	Val	Lys	Glu	Lys	Lys	Thr	Trp	Lys	Tyr	Phe	Lys	Leu	Ser	Gly	Tyr	
				85					90					95		
His	Tyr	Leu	Thr	Tyr	Lys	Asp	Thr	Lys	Ala	Val	Ile	Asp	Ser	Ile	Gly	
			100					105					110			
Ser	Gly	Leu	Arg	Lys	Trp	Gly	Val	Glu	Pro	Lys	Glu	Arg	Ile	Thr	Val	
		115					120					125				
Phe	Gly	Ser	Thr	Ser	Ala	Asn	Trp	Leu	Leu	Val	Ala	His	Gly	Ala	Phe	
	130					135					140					
Thr	Gln	Ser	Met	Thr	Ile	Val	Thr	Val	Tyr	Asp	Thr	Leu	Gly	Glu	Glu	
145					150					155					160	
Gly	Leu	Leu	His	Ser	Met	Asn	Glu	Ala	Glu	Val	Gly	Thr	Ala	Tyr	Thr	
				165					170					175		
Asn	Ala	Asp	Leu	Ile	Lys	Thr	Met	Thr	Asn	Val	Ser	Gly	Arg	Cys	Pro	
			180					185					190			
Thr	Leu	Lys	Arg	Ile	Val	Tyr	Asp	Gly	Glu	Ala	Asn	Ala	Ala	Asp	Val	
		195					200					205				
Ile	Ala	Leu	Gln	Thr	Ala	His	Pro	His	Leu	Gln	Leu	Ile	Thr	Leu	Glu	
	210					215					220					
Glu	Leu	Lys	Gln	Leu	Gly	Val	Asp	His	Pro	Val	Glu	Pro	Thr	Pro	Pro	
225					230					235					240	
Thr	Ala	Glu	Asp	Cys	Ser	Cys	Ile	Met	Tyr	Thr	Ser	Gly	Ser	Thr	Gly	
				245					250					255		
Asn	Pro	Lys	Gly	Val	Ile	Leu	Thr	His	Gly	Asn	Leu	Ile	Ala	Ala	Ile	
			260					265					270			
Gly	Gly	Val	Asn	Lys	Met	Leu	Glu	Lys	Tyr	Ile	Arg	Glu	Gly	Asp	Val	
		275					280					285				

Leu	Leu	Ala	Tyr	Leu	Pro	Leu	Ala	His	Val	Leu	Glu	Phe	Met	Val	Glu
290						295					300				
Asn	Leu	Cys	Leu	Phe	Trp	Gly	Val	Thr	Leu	Gly	Tyr	Gly	Thr	Val	Arg
305					310					315					320
Thr	Leu	Thr	Asp	Ala	Ser	Val	Arg	Glu	Cys	Gln	Gly	Asp	Ile	Lys	Glu
				325					330					335	
Leu	Arg	Pro	Thr	Leu	Met	Thr	Gly	Val	Pro	Ala	Val	Trp	Glu	Thr	Ile
			340					345					350		
Arg	Lys	Gly	Val	Leu	Ala	Gln	Val	Asn	Gln	Gly	Ser	Pro	Leu	Val	Gln
		355					360					365			
Ser	Val	Phe	Asn	Ala	Ala	Leu	Asn	Ala	Lys	Ala	Trp	Cys	Met	Asp	Arg
		370				375					380				
Lys	Leu	Gly	Ala	Leu	Thr	Gly	Ile	Phe	Asp	Thr	Val	Val	Phe	Asn	Lys
385					390					395					400
Val	Arg	Gln	Gln	Thr	Gly	Gly	Arg	Leu	Arg	Tyr	Ala	Leu	Ser	Gly	Gly
				405					410					415	
Ala	Pro	Ile	Ser	Gln	Glu	Thr	Gln	Arg	Phe	Leu	Thr	Thr	Ala	Leu	Cys
			420					425					430		
Pro	Ile	Leu	Gln	Ala	Tyr	Gly	Met	Thr	Glu	Ser	Cys	Gly	Met	Cys	Ser
		435					440					445			
Ile	Met	Thr	Pro	Glu	Ala	Phe	Asn	Tyr	Asn	Arg	Val	Gly	Ser	Pro	Val
	450					455					460				
Pro	Cys	Thr	Glu	Val	Lys	Leu	Val	Asp	Val	Pro	Asp	Ala	Gly	Tyr	Phe
465					470					475					480
Ser	Thr	Asp	Ser	Pro	Arg	Pro	Arg	Gly	Glu	Ile	Trp	Ile	Arg	Gly	Pro
				485					490					495	
Ser	Ile	Thr	Ser	Gly	Tyr	Phe	Lys	Asn	Ala	Glu	Glu	Thr	Ser	Ala	Ala
			500					505					510		
Ile	Thr	Glu	Asp	Arg	Trp	Leu	Lys	Thr	Gly	Asp	Ile	Gly	Glu	Trp	His
		515					520					525			
Ala	Asp	Gly	Thr	Leu	Ser	Val	Ile	Asp	Arg	Lys	Lys	Asn	Leu	Val	Lys
	530					535					540				
Leu	Ser	His	Gly	Glu	Tyr	Ile	Ala	Leu	Glu	Lys	Leu	Glu	Ser	Val	Tyr
545					550					555					560
Lys	Ser	Thr	Ala	Tyr	Cys	Asn	Asn	Ile	Cys	Val	Tyr	Ala	Asp	Ser	Met
				565					570					575	
Gln	Asn	Lys	Pro	Val	Ala	Leu	Ile	Val	Ala	Ser	Glu	Pro	Arg	Ile	Leu
			580					585					590		
Glu	Leu	Ala	Lys	Ala	Lys	Gly	Leu	Glu	Ser	Arg	Asp	Phe	Ala	Val	Leu
		595					600				605				
Cys	His	Asp	Lys	Val	Ile	Ile	Lys	Ala	Val	Leu	Asp	Ala	Cys	Leu	Ala
	610					615					620				
Thr	Ala	Lys	Lys	Ala	Gly	Leu	Lys	Pro	Ala	Glu	Leu	Leu	Gln	Gly	Val
625					630					635					640
Tyr	Leu	Glu	Ser	Glu	Glu	Trp	Thr	Ala	Gln	Gly	Gly	Leu	Leu	Thr	Ala
				645					650					655	
Ala	Gln	Lys	Leu	Lys	Arg	Lys	Glu	Ile	Asn	Gln	Ala	Tyr	Ala	Asp	Gln
			660					665					670		
Ile	Lys	Gln	Ile	Tyr	Gly	Ser	Lys								
	675						680								

<210> 53
 <211> 2043
 <212> DNA
 <213> Mortierella alpina

<400> 53

atgccaaagt	gctttaccgt	caacgtcggc	cccaggagacg	tcaagggcga	gactcgcac	60
cgtcgctcca	tccaggccgt	cgacaaactc	atggactcac	cctcaagcga	catcaagacc	120
ttgtacgatg	tcattccagta	ctctgccaaag	gtccgccccca	acctcaacgc	catcggctac	180
cgcaagattg	tcaagatgat	cgaagaggaa	aaggagatca	ccaagatggg	cagcggcgag	240
cctgtcaagg	agaaaaagac	gtggaaatac	ttcaagctct	ccggctacca	ctatctgacc	300
tacaaggaca	ccaaggccgt	catcgacagc	attggaagtg	gcctgcgcaa	gtgggggtgt	360
gagcccaagg	agaggatcac	cgtctttggg	tccacaagtg	ccaactggct	gctggctcgt	420
catggtgcct	tcacgcagtc	catgaccatc	gtcaccgtgt	atgacacttt	gggcgaggaa	480
ggattgctgc	actcgatgaa	cgaggccgag	gtgggaacgg	cctacaogaa	cgctgatttg	540
atcaagacaa	tgaccaacgt	ttcaggacgc	tgccccaccc	tcaagaggat	cgtctatgac	600
ggcgaagcca	acgcagcaga	cgtgatcgcc	cttcagacgg	cccattcctca	ccttcagctt	660
atcactctgg	aggagctgaa	gcagctcggg	gtggatcacc	ctgtggagcc	cactcctccc	720
accgccgagg	attgctcctg	catcatgtac	acttctggat	cgaccggaaa	ccctaaggga	780
gtcatcctca	ctcacggaaa	cctcattgcc	gccattggcg	gagttaacaa	gatgctggaa	840
aagtacattc	gcgaaggcga	tgtcttgctt	gcctaccttc	ccttggctca	cgttctggaa	900
ttcatgggtg	agaacctctg	tctcttctgg	gggtgaaccc	ttggatatgg	tactgtccgc	960
acgctgacgg	atgcctctgt	gcgtgagtgc	cagggtgata	tcaaggaatt	gcggcctacg	1020
cttatgaccg	gcgttccagc	agtgtgggag	accatccgca	aagggtttct	cgcccaagta	1080
aaccagggtt	cacctctggg	tcaatccgtc	ttcaacgcgg	ctctgaacgc	caaggcctgg	1140
tgcattggacc	gcaaaactagg	cgttttgact	ggaattttcg	acactgtggg	gttcaacaag	1200
gtccgtcagc	aaactggagg	tcgtcttcgc	tacgcgctct	cgggcggcgc	gcctatctcc	1260
caggagaccc	agcgtttctt	gaccacagca	ctgtgcccta	tccttcaagc	ctatggcatg	1320
actgagtcgt	gcggcatgtg	ctcgatcatg	actccagagg	cgttcaacta	caaccgcgtc	1380
ggttccctctg	ttccctgcac	agaggtcaag	ctcgtggatg	tgcccgatgc	aggatacttt	1440
tcgactgatt	cgccccgccc	tcgtgggtgag	atttggtatc	gcggaccctc	catcacctct	1500
ggatacttca	agaacgctga	ggagacctcg	gcagccatca	cagaggaccg	ctggctcaag	1560
actggagata	ttggagagtg	gcatgctgat	ggcacactct	cggtcattga	tcgcaagaag	1620
aacttggctca	agttgtcgca	tggcgagtac	attgctctag	agaaacttga	gtcgggtgtac	1680
aagagcacgg	cttactgcaa	caacatctgc	gtttatgccg	attccatgca	aaacaagcct	1740
gtggcgctga	ttgttgcgag	tgaaccccg	atcctcgagc	tggccaaggc	caagggcctg	1800
gagagcccg	actttgcagt	gctctgccac	gataaggtga	tcattcaaggc	tgctcctcgac	1860
gcctgtctcg	cgactgccaa	aaaggctggc	ctcaagcccc	ccgagttgct	gcagggtgtg	1920
tacctggagt	ctgaggagtg	gaccgctcaa	ggcggtttgt	tgactgctgc	tcagaaattg	1980
aagcgcaagg	aatcaacca	ggcttatgct	gaccagatca	agcagatcta	tggtccaag	2040
taa						2043

<210> 54

<211> 2052

<212> DNA

<213> Mortierella alpina

<400> 54

catgccaaag	tgctttaccg	tcaacgtcgg	ccccaggagac	gtcaagggcg	agactcgcac	60
cogtcgctcc	atccaggccg	tcgacaaact	catggactca	ccctcaagcg	acatcaagac	120
cttgtacgat	gtcatccagt	actctgccaa	ggctccgcccc	aacctcaacg	ccatcggcta	180
ccgcaagatt	gtcaagatga	tcgaagagga	aaaggagatc	accaagatgg	tcagcggcga	240
gcctgtcaag	gagaaaaaga	cgtggaaata	cttcaagctc	tccggctacc	actatctgac	300
ctacaaggac	accaaggccg	tcattcgacag	cattggaagt	ggcctgcgca	agtgggggtgt	360
tgagcccaag	gagaggatca	ccgtcttttg	ttccacaagt	gccaaactggc	tgctggctcg	420
tcattggtgcc	ttcacgcagt	ccattgaccat	cgtcaccgtg	tatgacactt	tgggcgaggga	480
aggattgctg	cactcgatga	acgaggccga	gggtgggaacg	gcctacacga	acgctgattt	540
gatcaagaca	atgaccaacg	tttcaggacg	ctgccccacc	ctcaagagga	tcgtctatga	600
cggcgaagcc	aacgcagcag	acgtgatcgc	ccttcagacg	gccatcctc	accttcagct	660
tatcactctg	gaggagctga	agcagctcgg	tgtggatcac	cctgtggagc	ccactcctcc	720
caccgccgag	gattgctcct	gcatcatgta	cacttctgga	tcgaccggaa	accctaaggg	780

```

agtcatectc actcacggaa acctcattgc cgccattggc ggagttaaca agatgctgga 840
aaagtacatt cgcgaaggcg atgtcttgct tgcctacctt cccttggctc acgttctgga 900
attcatgggt gagaacctct gtctcttctg ggggtgtaacc cttggatatg gtactgtccg 960
cacgctgaag gatgcctctg tgcgtgagtg ccagggtgat atcaaggaat tgcggcctac 1020
gcttatgacc ggcgttccag cagtgtggga gaccatccgc aaagggtgtt tgcgccaaagt 1080
aaaccagggt tcacctctgg ttcaatccgt cttcaacgcg gctctgaacg ccaaggcctg 1140
gtgcatggac cgcaaactag gcgctttgac tgggaattttc gacactgtgg tgttcaacaa 1200
ggtcgctcag caaactggag gtctgtcttc ctacgcgctc tcgggcggcg cgcctatctc 1260
ccaggagacc cagcgcttct tgaccacagc actgtgccct atccttcaag cctatggcat 1320
gactgagtcg tgcggcatgt gctcgatcat gactccagag gcgttcaact acaaccgcgt 1380
cggttccctt gttccctgca cagagggtcaa gctcgtggat gtgcccgatg caggatactt 1440
ttcgactgat tcgccccgcc ctctgtgtga gatttggatt cgcggaccct ccatacctc 1500
tggatacttc aagaacgctg aggagacctc ggcagccatc acagaggacc gctggctcaa 1560
gactggagat attggagagt ggcattgctga tggcacactc tcggtcattg atcgcaagaa 1620
gaacttggtc aagttgtcgc atggcgagta cattgctcta gagaaacttg agtcggtgta 1680
caagagcacg gcttactgca acaacatctg cgtttatgcc gattccatgc aaaacaagcc 1740
tgtggcgctg attgttgca gtgaaccccg catctcgag ctggccaagg ccaagggcct 1800
ggagagccgc gactttgca tgccttgcca cgataagggt atcatcaagg ctgtcctcga 1860
cgcctgtctc gcgactgcca aaaaggctgg cctcaagccc gccgagttgc tgcagggtgt 1920
gtacctggag tctgaggagt ggaccgctca aggcggtttg ttgactgctg ctcagaaatt 1980
gaagcgcaag gaaatcaacc aggccttatgc tgaccagatc aagcagatct atggctccaa 2040
gtaaaaatga ac 2052

```

```

<210> 55
<211> 2569
<212> DNA
<213> Mortierella alpina

```

```

<220>
<221> misc_feature
<222> (557)..(767)
<223> n is a, c, g, or t

```

```

<400> 55
atgccaaagt gctttaccgt caacgtcggc cccgaggacg tcaagggcga gactcgcac 60
cgtcgctcca tccaggccgt cgacaaactc atggactcac cctcaagcga catcaagacc 120
ttgtacgatg tcatccagta ctctgccaaag gtccgccccca acctcaacgc catcggctac 180
cgcaagattg tcaagatgat cgaagaggaa aaggagatca ccaagatggt cagcggcgag 240
cctgtcaagg agaaaaagac gtggaaatac ttcaagctct ccggtacca ctatctgacc 300
tacaaggaca ccaaggccgt catcgacagc attggaagtg gcttgcgcaa gtgggggtgt 360
gagcccaagg agaggatcac cgtcttttgt tccacaaggt aatgtgtagc gccacgaaaa 420
tacgatcatt gcagcgtgaa ggggtgggaag aattaggggg gaaatgacat cgataacagg 480
aacgaaaaaa aaaaaaaaca agacgaagtc ggagatcctc gctattggcg tttaagcacc 540
gccagcgttc tttttttnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnctt gttaaccaa 780
ggctgaccgc tacctcgttt gcttgacttg acacgtatag tgccaactgg ctgctggtcg 840
ctcatggtgc cttcacgcag tccatgacca tcgtcaccgt gtatgacact ttgggcgagg 900
aaggattgct gcactcgatg aacgaggccg aggtgggaac ggcctacacg aacgctgatt 960
tgatcaagac aatgaccaac gtttcaggac gctgccccac cctcaagagg atcgtctatg 1020
acggcgaagc caacgcagca gacgtgatcg cccttcagac ggcccatcct caccttcagc 1080
ttatcactct ggaggagctg aagcagctcg gtgtggatca ccctgtggag cccactcctc 1140
ccaccgcca ggattgctcc tgcacatgt acacttctgg atcgaccgga aaccctaagg 1200
gagtcactct cactcacgga aacctcattg ccgccagtag gtgtttctct cactctcttt 1260
aaccctctct ctttcacttg caaatatgct gggaaatctct acttacctga atgttactgt 1320

```

tcttcggggtt	atcaacctag	ttggcggagt	taacaagatg	ctggaaaagt	acattcgcga	1380
aggcgatgtc	ttgcttgctt	accttcctct	ggctcacgtt	ctggaattca	tggttgagaa	1440
cctctgtctc	ttctgggggtg	taacccttgg	atatggtaact	gtccgcacgc	tgacggatgc	1500
ctctgtgcgt	gagtgccagg	gtgatataca	ggaattgcgg	cctacgctta	tgaccggcgt	1560
tccagcagtg	tgggagacca	tccgcaaagg	tgttctcgcc	caagtaaacc	agggttcacc	1620
tctggttcaa	tccgtcttca	acgcggctct	gaacgccaag	gcctggtgca	tggaccgcaa	1680
actaggcgct	ttgactggaa	ttttcgacac	tgtggtgttc	aacaagggtcc	gtcagcaaac	1740
tggaggtcgt	cttcgctacg	cgctctcggg	cggcgcgcct	atctcccagg	agaccagcgc	1800
cttcttgacc	acagcactgt	gccctatcct	tcaagcctat	ggcatgactg	agtcgtgcgg	1860
catgtgctcg	atcatgactc	cagaggcggt	caactacaac	cgcgtcgggt	cccctgttcc	1920
ctgcacagag	gtcaagctcg	tggatgtgcc	cgatgcagga	tacttttcga	ctgattcgcc	1980
ccgcctcgt	ggtgagattt	ggattegcgg	accctccatc	acctctggat	acttcaagaa	2040
cgctgaggag	acctcggcag	ccatcacaga	ggaccgctgg	ctcaagactg	gagatattgg	2100
agagtggcat	gctgatggca	cactctcggt	cattgatcgc	aagaagaact	tggtcaagtt	2160
gtcgcattgg	gagtacattg	ctctagagaa	acttgagtcg	gtgtacaaga	gcacggctta	2220
ctgcaacaac	atctgcgttt	atgccgatcc	catgcaaaaac	aagcctgtgg	cgctgattgt	2280
tgcgagtga	ccccgcaccc	tgcgactggc	caaggccaag	ggcctggaga	gccgcgactt	2340
tgcagtgttc	tgccacgata	aggtgatcat	caaggctgtc	ctcgacgcct	gtctcgcgac	2400
tgccaaaaag	gctggcctca	agcccgcgca	gttgctgcag	ggtgtgtacc	tggagtctga	2460
ggagtggacc	gctcaaggcg	gtttgttgac	tgctgctcag	aaattgaagc	gcaaggaaat	2520
caaccaggct	tatgctgacc	agatcaagca	gatctatggc	tccaagtaa		2569

<210> 56

<211> 2040

<212> DNA

<213> Mortierella alpina

<400> 56

atgacaaaagt	gcctcaccgt	cgaagtcgga	cccgccgacg	tccagggcga	gaccgcgcatc	60
cgccgctccg	tctctctctg	aaagcgcttc	atgtcctcgc	cctcgatga	catcaagacc	120
ctctacgacg	tcttcaacca	ctccgtcacc	gtccgccccca	acctcaacgc	gatcggatac	180
cgcaaggctg	tcaagattgt	cgaggaagaa	aaggaggctg	tcaaggttgt	caacggcgag	240
gaagtcaagg	aaaagaagac	ctggaagttc	ttcaagatgt	ccggctacca	ctggctcacc	300
tacaaggatg	cgaagcaggt	cgctcgacagc	atcgatgctg	gtcttcgcaa	gtttggcgctc	360
gagcccaagg	acaagctgac	cgtttttcggt	gccacaagtg	ccaactggct	cctgcttgcc	420
cacgggtgctt	tcaccagtc	catcaccatt	gttaccgcct	acgacaacct	gggcgaggac	480
ggtctttttg	actctatgaa	cgaggccgag	gtggccaccg	cttacacaaa	cgccgacttg	540
ctcaacacta	tcaagaacgt	tgccggcaaa	tgccccaccc	tgaagaagat	catctacgac	600
ggcgatgcca	agcccgcaga	tgtcattgcc	ctccaggagg	cccatcctca	cctccagctc	660
atcacccctg	aggagctgaa	gcagctcgga	gtggacaacc	ctgtcgcccc	aaccctcct	720
gctgccaaagg	actactgctg	catcatgtac	acttcgggat	cgactggcaa	ccccaaaggga	780
gtgttgctga	cccatggaaa	cctcgttgct	gccatcgag	gtgtgaacaa	gatgctgaca	840
aagtacgttc	acgagggaga	cgctcttgctc	gcgtacttgc	ctcttgctca	cgttctcgag	900
ttcctggctg	aaaacgtctg	tctcttctgg	ggtgtgactc	ttggctacgg	taccgtccgc	960
acattgactg	atgcctcagt	ccgtgagtgc	cagggtgata	tcaaggagtt	gcgccctaca	1020
ttgatgaccg	gtgttcctgc	tgtgtgggag	acgattcgta	aggagtggtt	ggctcagggtt	1080
tcccagggtc	cacctcttgt	tcaaaagatc	ttccatgctg	ctttgaacgc	caaggcctgg	1140
tgcctggacc	gcaagttggg	tgcgttgact	ggaatcttcg	atactgtcgt	cttcaacaag	1200
gtcaagcagc	agacaggagg	acgtcttcgc	ttcgcccttt	cgggagggtgc	acccatctct	1260
caggagaccc	agcgcttctt	gacgacagct	ttgtgcccta	tcctccaggg	ctacggtatg	1320
acagagtctt	gcggcatgtg	cgccattttg	acccccgatg	tcttcaacta	cagccgtgtc	1380
ggatccccag	ttccttgac	ggagggtcaag	ttggctcgatg	tgcccgatgc	aggataccac	1440
tcaacggact	tgcctctccc	ccgtggtgag	gtctgcatte	gtggaccctc	catcactgct	1500
ggatacttca	agaacccccga	ggagacctcc	gccacattga	ctgctgatcg	ctggctcaag	1560
actggagata	tccgagagtg	gcaccccgac	ggcactatct	cgatcattga	ccgcaagaag	1620
aacttggtca	agctgtcaca	cggagagtag	attgcttttg	agaagcttga	gtctgtctac	1680

aagagcacag cctactgcaa caacatttgc gtgtatgccg actcgatgca gaacaagccc 1740
gttgccatta ttgttgccag cgaaccccg atcctcgagt tggccaaggc caagggcatt 1800
gagagccgcg actttgctgc tctctgccac gacaaggta tcatcaaggc tgtccacgat 1860
gcctgcctcg ccactgcca gcggtgctgga ctcaagcccg ctgagatgct tcagggagt 1920
tacttggagt cagaagaatg gacggcccag gctggcatgt tgactgccgc tcagaagctc 1980
aagcgcaagg agatcaacca ggcctatgtc tcacagatca agcagcttta tggaacggcc 2040

<210> 57
<211> 680
<212> PRT
<213> Mortierella alpina

<400> 57
Met Thr Lys Cys Leu Thr Val Glu Val Gly Pro Ala Asp Val Gln Gly
1 5 10 15
Glu Thr Arg Ile Arg Arg Ser Val Leu Ser Ala Lys Arg Leu Met Ser
20 25 30
Ser Pro Ser Asp Asp Ile Lys Thr Leu Tyr Asp Val Phe Asn His Ser
35 40 45
Val Thr Val Arg Pro Asn Leu Asn Ala Ile Gly Tyr Arg Lys Val Val
50 55 60
Lys Ile Val Glu Glu Glu Lys Glu Val Val Lys Val Val Asn Gly Glu
65 70 75 80
Glu Val Lys Glu Lys Lys Thr Trp Lys Phe Phe Lys Met Ser Gly Tyr
85 90 95
His Trp Leu Thr Tyr Lys Asp Ala Lys Gln Val Val Asp Ser Ile Gly
100 105 110
Cys Gly Leu Arg Lys Phe Gly Val Glu Pro Lys Asp Lys Leu Thr Val
115 120 125
Phe Gly Ala Thr Ser Ala Asn Trp Leu Leu Leu Ala His Gly Ala Phe
130 135 140
Thr Gln Ser Ile Thr Ile Val Thr Ala Tyr Asp Thr Leu Gly Glu Asp
145 150 155 160
Gly Leu Leu His Ser Met Asn Glu Ala Glu Val Ala Thr Ala Tyr Thr
165 170 175
Asn Ala Asp Leu Leu Asn Thr Ile Lys Asn Val Ala Gly Lys Cys Pro
180 185 190
Thr Leu Lys Lys Ile Ile Tyr Asp Gly Asp Ala Lys Pro Ala Asp Val
195 200 205
Ile Ala Leu Gln Glu Ala His Pro His Leu Gln Leu Ile Thr Leu Glu
210 215 220
Glu Leu Lys Gln Leu Gly Val Asp Asn Pro Val Ala Pro Thr Pro Pro
225 230 235 240
Ala Ala Lys Asp Tyr Cys Cys Ile Met Tyr Thr Ser Gly Ser Thr Gly
245 250 255
Asn Pro Lys Gly Val Leu Leu Thr His Gly Asn Leu Val Ala Ala Ile
260 265 270
Gly Gly Val Asn Lys Met Leu Thr Lys Tyr Val His Glu Gly Asp Val
275 280 285
Leu Leu Ala Tyr Leu Pro Leu Ala His Val Leu Glu Phe Leu Val Glu
290 295 300
Asn Val Cys Leu Phe Trp Gly Val Thr Leu Gly Tyr Gly Thr Val Arg
305 310 315 320
Thr Leu Thr Asp Ala Ser Val Arg Glu Cys Gln Gly Asp Ile Lys Glu
325 330 335

Leu Arg Pro Thr Leu Met Thr Gly Val Pro Ala Val Trp Glu Thr Ile
 340 345 350
 Arg Lys Gly Val Leu Ala Gln Val Ser Gln Gly Ser Pro Leu Val Gln
 355 360 365
 Lys Ile Phe His Ala Ala Leu Asn Ala Lys Ala Trp Cys Leu Asp Arg
 370 375 380
 Lys Leu Gly Ala Leu Thr Gly Ile Phe Asp Thr Val Val Phe Asn Lys
 385 390 395 400
 Val Lys Gln Gln Thr Gly Gly Arg Leu Arg Phe Ala Leu Ser Gly Gly
 405 410 415
 Ala Pro Ile Ser Gln Glu Thr Gln Arg Phe Leu Thr Thr Ala Leu Cys
 420 425 430
 Pro Ile Leu Gln Gly Tyr Gly Met Thr Glu Ser Cys Gly Met Cys Ala
 435 440 445
 Ile Leu Thr Pro Asp Val Phe Asn Tyr Ser Arg Val Gly Ser Pro Val
 450 455 460
 Pro Cys Thr Glu Val Lys Leu Val Asp Val Pro Asp Ala Gly Tyr His
 465 470 475 480
 Ser Thr Asp Leu Pro Leu Pro Arg Gly Glu Val Cys Ile Arg Gly Pro
 485 490 495
 Ser Ile Thr Ala Gly Tyr Phe Lys Asn Pro Glu Glu Thr Ser Ala Thr
 500 505 510
 Leu Thr Ala Asp Arg Trp Leu Lys Thr Gly Asp Ile Gly Glu Trp His
 515 520 525
 Pro Asp Gly Thr Ile Ser Ile Ile Asp Arg Lys Lys Asn Leu Val Lys
 530 535 540
 Leu Ser His Gly Glu Tyr Ile Ala Leu Glu Lys Leu Glu Ser Val Tyr
 545 550 555 560
 Lys Ser Thr Ala Tyr Cys Asn Asn Ile Cys Val Tyr Ala Asp Ser Met
 565 570 575
 Gln Asn Lys Pro Val Ala Ile Ile Val Ala Ser Glu Pro Arg Ile Leu
 580 585 590
 Glu Leu Ala Lys Ala Lys Gly Ile Glu Ser Arg Asp Phe Ala Ala Leu
 595 600 605
 Cys His Asp Lys Val Ile Ile Lys Ala Val His Asp Ala Cys Leu Ala
 610 615 620
 Thr Ala Lys Arg Ala Gly Leu Lys Pro Ala Glu Met Leu Gln Gly Val
 625 630 635 640
 Tyr Leu Glu Ser Glu Glu Trp Thr Ala Gln Ala Gly Met Leu Thr Ala
 645 650 655
 Ala Gln Lys Leu Lys Arg Lys Glu Ile Asn Gln Ala Tyr Val Ser Gln
 660 665 670
 Ile Lys Gln Leu Tyr Gly Thr Ala
 675 680

<210> 58
 <211> 2043
 <212> DNA
 <213> Mortierella alpina

<400> 58
 atgacaaagt gcctcaccgt cgaagtcgga cccgccgacg tccagggcga gaccgcgcatc 60
 cgccgctccg tcctctctgc aaagcgctc atgtcctcgc cctcggatga catcaagacc 120
 ctctacgacg tcttcaacca ctccgtcacc gtccgccccca acctcaacgc gatcggatac 180
 cgcaaggctc tcaagattgt cgaggaagaa aaggaggtcg tcaaggttgt caacggcgag 240
 gaagtcaagg aaaagaagac ctggaagtgc ttcaagatgt ccggctacca ctggctcacc 300

tacaaggatg	cgaagcaggt	cgtcgacagc	atcggatgcg	gtcttcgcaa	gtttggcgtc	360
gagcccaagg	acaagctgac	cgttttcggg	gccacaagtg	ccaactggct	cctgcttgcc	420
cacggtgctt	tcacccagtc	catcaccatt	gttaccgcct	acgacaccct	gggcgaggac	480
ggtcttttgc	actctatgaa	cgaggccgag	gtggccaccg	cttacacaaa	cgccgacttg	540
ctcaacacta	tcaagaacgt	tgcgggcaaa	tgccccaccc	tgaagaagat	catctacgac	600
ggcgatgcc	agcccgcaga	tgtcattgcc	ctccaggagg	cccatcctca	cctccagctc	660
atcaccctcg	aggagctgaa	gcagctcgga	gtggacaacc	ctgtcgcccc	aaccctcct	720
gctgccaaagg	actactgctg	catcatgtac	acttcgggat	cgactggcaa	ccccaaagga	780
gtgttgctga	cccatggaaa	cctcgttgct	gccatcgag	gtgtgaacaa	gatgctgaca	840
aagtacgttc	acgagggaga	cgtcttgctc	gcgtacttgc	ctcttgctca	cgttctcgag	900
ttcctggctg	aaaacgtctg	tctcttctgg	ggtgtgactc	ttggctacgg	taccgtccgc	960
acattgactg	atgcctcagt	ccgtgagtgc	cagggtgata	tcaaggagtt	gcgccctaca	1020
ttgatgaccg	gtgttcctgc	tgtgtgggag	acgattcgta	agggagtgtt	ggctcaggtt	1080
tcccagggtc	cacctcttgt	tcaaaagatc	ttccatgctg	ctttgaacgc	caaggcctgg	1140
tgcctggacc	gcaagttggg	tgcgttgact	ggaatcttcg	atactgtcgt	cttcaacaag	1200
gtcaagcagc	agacaggagg	acgtcttcgc	ttcgcccttt	cgggaggtgc	acccatctct	1260
caggagaccc	agcgcttctt	gacgacagct	ttgtgcccta	tcctccaggg	ctacggtatg	1320
acagagtctt	gcggcatgtg	cgccattttg	acccccgatg	tcttcaacta	cagccgtgtc	1380
ggatccccag	ttccttgcac	ggaggtcaag	ttggctcgatg	tgcccgatgc	aggataccac	1440
tcaacggact	tgcctctccc	ccgtggtgag	gtctgcattc	gtggaccctc	catcactgct	1500
ggatacttca	agaaccccga	ggagacctcc	gccacattga	ctgctgatcg	ctggctcaag	1560
actggagata	tcggagagtg	gcaccccagc	ggcactatct	cgatcattga	ccgcaagaag	1620
aacttggtca	agctgtcaca	cggagagtac	attgctttgg	agaagcttga	gtctgtctac	1680
aagagcacag	cctactgcaa	caacatttgc	gtgtatgccg	actcgatgca	gaacaagccc	1740
gttgccatta	ttgttgccag	cgaaccccgc	atcctcgagt	tggccaaggc	caagggcatt	1800
gagagccgcg	actttgctgc	tctctgccac	gacaaggtta	tcacaaaggc	tgtccacgat	1860
gcctgcctcg	ccactgccaa	gcgtgctgga	ctcaagcccg	ctgagatgct	tcagggagtg	1920
tacttgagtg	cagaagaatg	gacggcccag	gctggcatgt	tgactgccgc	tcagaagctc	1980
aagcgcaagg	agatcaacca	ggcctatgtc	tcacagatca	agcagcttta	tggaacggcc	2040
taa						2043

<210> 59
 <211> 2392
 <212> DNA
 <213> Mortierella alpina

<400> 59						
ccttttatccc	cgacccgcc	tctctcgccg	ccaccatctc	gcattccttt	caatccacac	60
tcccacctgt	gccccctgct	tttcacgtcc	cgctctcctc	ccgccttctc	ctttcatcac	120
cccaattcaa	catgacaaag	tgcctcaccg	tcgaagtcgg	acccgccgac	gtccagggcg	180
agaccgcgat	ccgccgctcc	gtcctctctg	caaagcgccct	catgtcctcg	ccctcggatg	240
acatcaagac	cctctacgac	gtcttcaacc	actccgtcac	cgtcgcgccc	aacctcaacg	300
cgatcggata	ccgcaaggtc	gtcaagattg	tcgaggaaga	aaaggaggtc	gtcaaggttg	360
tcaacggcga	ggaagtcaag	gaaaagaaga	cctggaagtt	cttcaagatg	tccggctacc	420
actggctcac	ctacaaggat	gcgaagcagg	tcgtcgacag	catcggatgc	ggtcttcgca	480
agtttgccgt	cgagcccaag	gacaagctga	ccgttttcgg	tgccacaagt	gccaactggc	540
tcctgcttgc	ccacgggtgct	ttcaccagct	ccatcaccat	tgttaccgcc	tacgacaccc	600
tgggcgagga	cggctctttt	cactctatga	acgaggccga	ggtggccacc	gcttacacaa	660
acgccgactt	gctcaacact	atcaagaacg	ttgccggcaa	atgccccacc	ctgaagaaga	720
tcactctacga	cggcgatgcc	aagcccgcag	atgtcattgc	cctccaggag	gcccatactc	780
acctccagct	catcacctc	gaggagctga	agcagctcgg	agtggacaac	cctgtcgccc	840
caaccctcc	tgttgccaag	gactactgct	gcacatgta	cacttcggga	tcgactggca	900
accccaaggg	agtgttgctg	acccatggaa	acctcgttgc	tgccatcgga	ggtgtgaaca	960
agatgctgac	aaagtacgtt	cacgaggag	acgtcttgct	cgcgtacttg	cctcttgctc	1020
acgttctcga	gttcctggtc	gaaaacgtct	gtctcttctg	gggtgtgact	cttggctacg	1080
gtaccgtccg	cacattgact	gatgcctcag	tccgtgagtg	ccagggatgat	atcaaggagt	1140

tgcgccctac	attgatgacc	ggtgttcctg	ctgtgtggga	gacgattcgt	aagggagtgt	1200
tggetcaggt	ttcccagggc	tcacctcttg	ttcaaaagat	cttccatgct	gctttgaacg	1260
ccaaggcctg	gtgcctggac	cgcaagttgg	gtgcgttgac	tggaatcttc	gatactgtcg	1320
tcttcaacaa	ggtcaagcag	cagacaggag	gacgtcttcg	cttcgccctt	tcgggaggtg	1380
cacccatctc	tcaggagacc	cagcgcttct	tgacgacagc	tttgtgccct	atcctccagg	1440
gctacggtat	gacagagtct	tgcggcatgt	gcgccatttt	gacccccgat	gtcttcaact	1500
acagccgtgt	cggatcccca	gttccttgca	cggaggtcaa	gttggtcgat	gtgcccgatg	1560
caggatacca	ctcaacggac	ttgcctctcc	cccgtggtga	ggtctgcatt	cgtggaccct	1620
ccatcactgc	tggatacttc	aagaacccccg	aggagacctc	cgccacattg	actgctgac	1680
gctggctcaa	gactggagat	atcggagagt	ggcacccega	cggcactatc	tcgatcattg	1740
accgcaagaa	gaacttggtc	aagctgtcac	acggagagta	cattgctttg	gagaagcttg	1800
agtctgtcta	caagagcaca	gcctactgca	acaacatttg	cgtgtatgcc	gactcgatgc	1860
agaacaagcc	cgttgccatt	attgttgcca	gcgaacccccg	catcctcgag	ttggccaagg	1920
ccaagggcat	tgagagccgc	gactttgctg	ctctctgcca	cgacaagggt	atcatcaagg	1980
ctgtccacga	tgcttgctc	gccactgcca	agcgtgctgg	actcaagccc	gctgagatgc	2040
ttcagggagt	gtacttgag	tcagaagaat	ggacggccca	ggctggcatg	ttgactgccg	2100
ctcagaagct	caagcgcaag	gagatcaacc	aggcctatgt	ctcacagatc	aagcagcttt	2160
atggaacggc	ctaagtcgct	gaaagggtgtg	cctttgtccg	tctcttcaac	cccacaagtc	2220
ctatgtataa	tgacccgcgc	ggccctcctt	taatcctata	cccacccttt	tttacacggt	2280
aaagaagcca	catttttggg	tctttttttt	ctctcgcaca	cactacacac	tcccatcca	2340
ttccctccaa	acaggatggg	tgtctgcaaa	taaattgacg	aattttctct	tg	2392

<210> 60

<211> 2537

<212> DNA

<213> Mortierella alpina

<400> 60

atgacaaagt	gcctcaccgt	cgaagtcgga	cccgccgacg	tccagggcga	gacccgcac	60
cgccgctccg	tctctctcgc	aaagcgcttc	atgtcctcgc	cctcggatga	catcaagacc	120
ctctacgacg	tcttcaacca	ctccgtcacc	gtccgcccc	acctcaacgc	gatcggatac	180
cgcaaggteg	tcaagattgt	cgaggaagaa	aaggaggteg	tcaaggttgt	caacggcgag	240
gaagtcaagg	aaaagaagac	ctggaagttc	ttcaagatgt	ccggctacca	ctggctcacc	300
tacaaggatg	cgaagcaggt	cgtcgacagc	atcggatgcg	gtcttcgcaa	gtttggcgtc	360
gagcccaagg	acaagctgac	cgttttcggt	gccacaaggt	aagaaagagg	cataacaaga	420
aaatgcaaga	gaggcaaaaa	aaatggcttg	acgtgagagc	ataagggaac	caacagacag	480
gtgtttgtgt	gggttgcgga	tagtgggtga	gcatgcttcg	ttatcgaatg	tgggagaaga	540
gagcggacgc	gaatatggct	ctcgtctctg	gcgggatgcg	agtggccaag	tgtgggatac	600
atatcctcgc	ggtgggggtg	ccgggtcggc	ccttgaaatc	tgttgaaagc	tgataatgtg	660
aatgtggacc	gcaatcacgc	tcagattatg	cgtagcaagc	gtgttgctag	tctacatcat	720
gctcacacgt	attcacattt	attcattttc	actctatctc	gctcttagtg	ccaactggct	780
cctgcttgcc	cacgggtgctt	tcacccagtc	catcaccatt	gttaccgcct	acgacaccct	840
gggcgaggac	ggtcttttgc	actctatgaa	cgaggccgag	gtggccaccg	cttacacaaa	900
cgccgacttg	ctcaacacta	tcaagaacgt	tgccggcaaa	tgccccaccc	tgaagaagat	960
catctacgac	ggcgatgcca	agcccgcaga	tgtcattgcc	ctccaggagg	cccatcctca	1020
cctccagctc	atcacccctc	aggagetgaa	gcagctcgga	gtggacaacc	ctgtcgcccc	1080
aacccctcct	gctgccaaag	actactgctg	catcatgtac	acttcgggat	cgactggcaa	1140
ccccaaagga	gtgttgctga	cccatggaaa	cctcgttgct	gccagtacgt	atctttctcg	1200
tcatgatcgt	cctcccgcat	ttccactgcg	cttgttacca	tttgatggga	aatgtattta	1260
acccgaacca	cacatttttt	ctttttctca	cacttgccac	gtcactagtc	ggagggtgtg	1320
acaagatgct	gacaaagtac	gttcacgagg	gagacgtctt	gctcgcgtac	ttgcctcttg	1380
ctcacgttct	cgagttcctg	gtcgaaaacg	tctgtctctt	ctgggggtgtg	actcttggct	1440
acggtaccgt	ccgcacattg	actgatgcct	cagtcctgta	gtgccagggt	gatatacaagg	1500
agttgcgccc	tacattgatg	accggtgttc	ctgctgtgtg	ggagacgatt	cgtaagggag	1560
tgttggctca	ggtttcccag	ggctcacctc	ttgttcaaaa	gatcttccat	gctgctttga	1620
acgccaaggc	ctggtgcctg	gaccgcaagt	tgggtgcggt	gactggaatc	ttcgatactg	1680

```

tcgtcttcaa caaggtcaag cagcagacag gaggacgtct tcgcttcgcc ctttcgggag 1740
gtgcacccat ctctcaggag acccagcget tcttgacgac agctttgtgc cctatcctcc 1800
agggctacgg tatgacagag tcttgcgga tgtgcgccat tttgaccccc gatgtcttca 1860
actacagccg tgcggatcc ccagttcctt gcacggaggt caagttggtc gatgtgcccg 1920
atgcaggata ccactcaacg gacttgcctc tcccccgtag tgaggtctgc attcgtggac 1980
cctccatcac tgctggatac ttcaagaacc ccgaggagac ctccgccaca ttgactgctg 2040
atcgctggct caagactgga gatatcggag agtggcacc cgcaggcact atctcgatca 2100
ttgaccgcaa gaagaacttg gtcaagctgt cacacggaga gtacattgct ttggagaagc 2160
ttgagtctgt ctacaagagc acagcctact gcaacaacat ttgctgtat gccgactcga 2220
tgcagaacaa gcccgttgcc attattgttg ccagcgaacc ccgcctcctc gagttggcca 2280
aggccaaggg cattgagagc cgcgactttg ctgctctctg ccacgacaag gttatcatca 2340
aggctgtcca cgatgcctgc ctgcgccactg ccaagcgtgc tggactcaag cccgctgaga 2400
tgcttcaggg agtgtacttg gagtcagaag aatggacggc ccaggctggc atgttgactg 2460
ccgctcagaa gctcaagcgc aaggagatca accaggccta tgtctcacag atcaagcagc 2520
tttatggaac ggcctaa 2537

```

```

<210> 61
<211> 22
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> Synthetic DNA

```

```

<400> 61
gtcggctcca agcttgcaat cc 22

```

```

<210> 62
<211> 25
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> Synthetic DNA

```

```

<400> 62
ggacagctcc agcactgtgg taaag 25

```

```

<210> 63
<211> 24
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> Synthetic DNA

```

```

<400> 63
gaccacggga ttccccaagg ctgc 24

```

```

<210> 64
<211> 25
<212> DNA
<213> Artificial sequence

```

<220>
<223> Synthetic DNA

<400> 64
cttggtcgcg cttgttcctg gccac 25

<210> 65
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 65
tacagctttg ttgctgtccc catc 24

<210> 66
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 66
gatgatgggt gtgcttgcaa agatc 25

<210> 67
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 67
aacccaaagc tgcgccaggc tgtcc 25

<210> 68
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 68
ttacagcttg gattcctttt gatgg 25

<210> 69
<211> 22

<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 69
gtcgtgccccg atgcggagac gc

22

<210> 70
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 70
tcagtggatc ccgttatata tcag

24

<210> 71
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 71
gcgtccccct ctatgataca ttg

23

<210> 72
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 72
gtgggatgca ggacggcaac atcg

24

<210> 73
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 73
ggatgccgaa caacagcgcg tgg

23

<210> 74
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 74
gcaccctcct cagaaacagc cctc

24

<210> 75
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 75
cagtcgagta cattgtcaac cacg

24

<210> 76
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 76
gcggttcaag aggcgaggca cagc

24

<210> 77
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 77
gttcatcttc tgctggctgg gtctc

25

<210> 78
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 78
gttgcgttgt tcacgcggca atcc

24

<210> 79
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 79
atggaaacct tggttaacgg aaag

24

<210> 80
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 80
tcagcaaaga tggccttggg ctgg

24

<210> 81
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 81
gtcaagggcg agactcgcat cc

22

<210> 82
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 82
cggtgacgat ggtcatggac tgc

23

<210> 83
<211> 23
<212> DNA
<213> Artificial sequence

<220>
 <223> Synthetic DNA

 <400> 83
 gcgagacccg catccgccgc tcc 23

<210> 84
 <211> 23
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

 <400> 84
 gaccgtcctc gccaggggtg tcg 23

<210> 85
 <211> 31
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

 <400> 85
 ggatccatgc cttccttcaa aaagtacaac c 31

<210> 86
 <211> 30
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

 <400> 86
 cccgggcaaa gagttttcta tctacagctt 30

<210> 87
 <211> 25
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

 <400> 87
 gaattcatgg ttgctctccc actcg 25

<210> 88
 <211> 27

<212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 88
 ggatccctac tatagcttgg ccttgcc 27

<210> 89
 <211> 29
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 89
 ggatccatgt atgtcggctc caagcttgc 29

<210> 90
 <211> 31
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 90
 gtcgactcaa agcctggctt tgccgctgac g 31

<210> 91
 <211> 30
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 91
 ggatccatgg aaaccttggt taacggaaag 30

<210> 92
 <211> 29
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 92
 ggtacctaga acttcttcca catctcctc 29

<210> 93
 <211> 31
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

<400> 93
 gagctcatgc caaagtgctt taccgtcaac g

31

<210> 94
 <211> 31
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

<400> 94
 ggatccttac ttggagccat agatctgctt g

31

<210> 95
 <211> 28
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

<400> 95
 tctagaatgg cacctcccaa cactattg

28

<210> 96
 <211> 30
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

<400> 96
 aagcttttac ttcttgaaaa agaccacgtc

30

<210> 97
 <211> 29
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthetic DNA

<400> 97
tctagaatgg ctgctgctcc cagtgtgag 29

<210> 98
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 98
aagcttttac tgtgccttgc ccatcttgg 29

<210> 99
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 99
tctagaatgg agtcgattgc gcaattcc 28

<210> 100
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 100
gagctcttac tgcaacttcc ttgccttctc 30

<210> 101
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 101
tctagaatgg gtgcggacac aggaaaaacc 30

<210> 102
<211> 30
<212> DNA
<213> Artificial sequence

<220>
 <223> Synthetic DNA

 <400> 102
 aagctttttac tcttccttgg gacgaagacc 30

 <210> 103
 <211> 28
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 103
 gaattcatga caaagtgcct caccgtcg 28

 <210> 104
 <211> 29
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 104
 cccgggactt aggccgttcc ataaagctg 29

 <210> 105
 <211> 54
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 105
 aattcataag aatgcggccg ctaaactatt ctagactagg tcgacggcgc gcc 54

 <210> 106
 <211> 54
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 106
 agcttggcgc gccgtcgacc tagtctagaa tagtttagcg gccgcattct tatg 54

 <210> 107
 <211> 26

<212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 107
 agcggccgca taggggagat cgaacc 26

<210> 108
 <211> 33
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 108
 agaattcggc gcgccatgca cgggtccttc tca 33

<210> 109
 <211> 20
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 109
 gtcgaccatg acaagtttgc 20

<210> 110
 <211> 21
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 110
 gtcgactgga agacgagcac g 21

<210> 111
 <211> 43
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthetic DNA

 <400> 111
 ggcaaacttg tcatgaagcg aaagagagat tatgaaaaca agc 43

<210> 112
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 112
cactcccttt tottaattgt tgagagagtg ttgggtgaga gt

42

<210> 113
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 113
taagaaaagg gagtgaatcg cataggg

27

<210> 114
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 114
catgacaagt ttgccaagat gcg

23

<210> 115
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 115
attgttgaga gagtgttggg tgagagtg

28

<210> 116
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 116
cactctctca acaatatgga aaccttggtt aacggaaagt 40

<210> 117
<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 117
cactcccttt tcttactaga acttcttcca catctcctca atatc 45

<210> 118
<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 118
cactcccttt tcttattact tggagccata gatctgcttg a 41

<210> 119
<211> 39
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 119
cactctctca acaatatgcc aaagtgcctt accgtcaac 39

<210> 120
<211> 16
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 120
gtcccgaatg gttcct 16

<210> 121
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 121
agcggttttc tacttgc

17

<210> 122
<211> 16
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 122
aactacaacc gcgtcg

16

<210> 123
<211> 16
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic DNA

<400> 123
cggcataaac gcagat

16

30179-221

CLAIMS:

1. A polynucleotide as defined in (a) or (b) below:
 - (a) a polynucleotide encoding a protein consisting of an amino acid sequence, wherein 1 to 10 amino acids are deleted, substituted, inserted and/or added in the amino acid
5 sequence set forth as SEQ ID NO: 47, wherein the protein has acyl-CoA synthetase activity or, when expressed in a host cell, functions to increase the amount of, or change the composition of, the fatty acids produced in the host cell; or
 - (b) a polynucleotide encoding a protein having an amino acid sequence having at least 90% identity to the amino acid sequence set forth as SEQ ID NO: 47, wherein the
10 protein has acyl-CoA synthetase activity or, when expressed in a host cell, functions to increase the amount of, or change the composition of, the fatty acids produced in the host cell.
2. The polynucleotide according to claim 1, comprising the nucleotide sequence set forth as SEQ ID NO: 46.
3. The polynucleotide according to claim 1, encoding a protein consisting of the
15 amino acid sequence set forth as SEQ ID NO: 47.
4. The polynucleotide according to any one of claims 1 to 3, which is a DNA.
5. A protein encoded by the polynucleotide according to any one of claims 1 to 4.
6. A vector comprising the polynucleotide according to any one of claims 1 to 4.
7. A non-human cell which is transformed with the polynucleotide according to
20 any one of claims 1 to 4, or into which the vector according to claim 6 is introduced.
8. A method for producing a lipid or fatty acid composition, which comprises collecting the lipid or fatty acid composition from the culture of the transformant according to claim 7, wherein the lipid is a triacylglycerol and, wherein the fatty acid is a polyunsaturated fatty acid having at least 18 carbon atoms.

Figure 1

1 GCTCTTTTTTGTCTGTTCTTTGCCACCCCACTCTGCTGTGCCTCTCCACTCCCCCTGCCCGTGAACGTTCTTCTGTCACTTTGCACAGCAGTATCTCC
 101 TCTGATCTCGCTTGRITATATTCCTCTAATCTCGTTTGGTTATATTCCTCTGATCTCGCTCGGTTATATTCTTCAGATATGGATGCTGTCCCTGCAGT
 M D A V P A V
 201 TGCTGCTGCGGCCATCCCGCAGCCATGTATGTCGGCTCCAAGCTTGCAATCCCCGTGATGTCAAGTTAGCTAAAGGCCTAGTCAGTGCCAAGCTAGGT
 A A A A I P A A M Y V G S K L A I P R D V K L A K G L V S A K L G
 301 TACAGGTCTACGAGAAGAACGACTCGATCAATATCTTATCGTTTTGAAGAGACCTGTAAGAAGCACCTCATCGGAAGCTTTGGTGTGTTGAAGGCA
 Y R S Y E K N D S I N I S Y R F E E T C K K H P H R E A L V F E G K
 401 AATCGTACACCTTCCAGGACATCCAGCGAGAATCGAATAGGGTGGGACACTGGCTGTTGTCCAAAGGCGTCAAGCGAGGAGAGATCGTGTGCTCTTCAT
 S Y T F Q D I Q R E S N R V G H W L L S K G V K R G E I V S L F M
 501 GCAAAATAAGCCAGAGTTTCTCTTCTTCTGGCTTGGACTCAACAAGATCGGCGCTACGGGAGCATTCAACACGAACCTCTCGGGCAAACCTCTGACG
 Q N K P E F L F F W L G L N K I G A T G A F I N T N L S G K P L T
 601 CACTCATTGCGTACCGGACAGCATCCATTCTGATTATGGATGCGGAAGTCCGACGCCCATTTATAGTGTCTCGATGAAGTCCTTGAGATGGGATATC
 H S L R T A T A S I L I M D A E L P T P I Y S V L D E V L E M G Y Q
 701 AGATATATTCCTACGGAGGATCCAGCAACACGCCTTTGCTACACAAGTTGAACCTTCTCAAATCTCGGATGCGGCCTTGCCCAAGAGTCTGCGAAGGAA
 I Y S Y G G S Q Q H A F A T Q V E L S Q I S D A A L P K S L R R K
 801 AACCCTGCAATGATATTGCCATGTTGATTTACACCTCCGGAACGACGGGTTTGCCCAAAGCTGGACGGTTCTCCCATGCTCGAGCCAACGTTGCCGCA
 T T A N D I A M L I Y T S G T T G L P K A G R F S H A R A N V A A
 901 CTTTTCTGGACGTCTTTCTACCACTTCAGCGAAAAAGACCGCCTGTACATCGCCTTGCTCTTTACCACAGTGTGGAGCTGTCTTGAATATGTGTGG
 L F W T S F Y H F S E K D R L Y I A L P L Y H S A G A V L G I C V A
 1001 CCTGGGTACCGGTGCTACGGTGGTCTGGCGCGCAAGTTTTCAACTACTTCTTCTGGGACGAATGCAGGGCCAACAAGGTACCGTGATCCAGTATAT
 W V T G A T V V L A R K F S T T S F W D E C R A N K V T V I Q Y I
 1101 TGGAGAAATCTGCCGATACTTACTGAATGCTCCTCTCTCCCTTGGACAAGACACACAGATCCGAATGGCGCATGGCAACGGCATGCGTCCGGATGTA
 G E I C R Y L L N A P P S P L D K T H T I R M A H G N G M R P D V
 1201 TGGAACAGATTCAGAGATCGTTTCGGCATCCCTTTGATCGGAGAATGGTATGCAAGCACTGAGGGCACCGGAATCTTGACAACTATAACACAGGACCCA
 W N R F R D R F G I P L I G E W Y A S T E G T G I L T N Y N T G P N
 1301 ATGGCGCTGGTGCGATAGGATACAGAGGCTCCTTGGCCAGAAGTGTGATAAGGGTCTGAAGATTGCGAAGTTGACATCCAAACCGAGGAACCTTATTCG
 G A G A I G Y R G S L A R T V D K G L K I A K F D I Q T E E L I R
 1401 TGACAAAATGGTCGATGCATTGAGTGTGTGCGAGATGAGCCCGGCGAGCTCTTGACAATGATTGATTCAAGTGATCCCACTCGCGCTTTCCAAGGGTAC
 D K N G R C I E C V A D E P G E L L T M I D S S D P T R A F Q G Y
 1501 CATAAAAATGCAGGTGCAAACTCCAAGAAAGTCCGTCAGGATGCATTCAAGTGTGGCGACCAATACTTTCTGACTGGTGACATCCTTCGTGCGGACGCTG
 H K N A G A N S K K V V Q D A F S V G D Q Y F R T G D I L R R D A D
 1601 ATGGCTATTTCTATTTTGGCGATCGTGTGGAGATACTTTCCGCTGGAAATCTGAAAACGTGTCAACTGCGGAGGTTTCTGAGGTGCTCTCAGCATACCC
 G Y F Y F G D R V G D T F R W K S E N V S T A E V S E V L S A Y P
 1701 GGACTGCATCGAGGTCAACGTTTATGGCGTTCAAGTCCCTGGACACGACGGCCGCGCAGGCATGGCTGCCATTGTCTCAAGGACACCATGAAGTGGGAT
 D C I E V N V Y G V Q V P G H D G R A G M A A I V S K D T M N W D
 1801 AGTTTCGCCAAGTTTGCACCTCAAAAATCTGCCGAAGTACTCTGTGCCGATTTTCATCCGCAAGGTCCAGAGATGGAGATTACGGGAACGTTCAAGCAAC
 S F A K F A L K N L P K Y S V P I F I R K V P E M E I T G T F K Q R
 1901 GAAAGGTTGAACTGGTGAACGAGGGCATGGACCGAGCAAGATCAAAGACGAAATGCTGTGGTTGGATGGGCACTCCTACCGGCCCTTCAAAGAGGGCGGA
 K V E L V N E G M D P S K I K D E M L W L D G H S Y R P F K E A E
 2001 GCATACTAGAGTCGTGAGCGGCAAGCCAGGCTTTGACGAATAAAATTATTTCTGTTTTGTCCGTTGAAAAA
 H T R V V S G K A R L

Figure 2A

	1	100
genome	ATGGATGCTGTCCCTGCAGTTGCTGCTGCGGCCATCCCCGCAGCCATGTATGTCGGCTCCAAGCTTGCAATCCCCCGTGATGTCAAGTTAGCTAAAGGCC	
CDS	ATGGATGCTGTCCCTGCAGTTGCTGCTGCGGCCATCCCCGCAGCCATGTATGTCGGCTCCAAGCTTGCAATCCCCCGTGATGTCAAGTTAGCTAAAGGCC	
	101	200
genome	TAGTCAGTGCCAAGCTAGGTTACAGGTCCTACGAGAAGAACGACTCGATCAATATCTCTTATCGTTTTGAAGAGACCTGTAAGAAGCACCCCTCATCGCGA	
CDS	TAGTCAGTGCCAAGCTAGGTTACAGGTCCTACGAGAAGAACGACTCGATCAATATCTCTTATCGTTTTGAAGAGACCTGTAAGAAGCACCCCTCATCGCGA	
	201	300
genome	AGCTTTGGTGTTTGAAGGCAAATCGTACACCTTCCAGGACATCCAGCGAGGTAACAAGAAAAACAATTGTCCAAAGTGACAGTCGAACGCATCCGAAATT	
CDS	AGCTTTGGTGTTTGAAGGCAAATCGTACACCTTCCAGGACATCCAGCGAG-----	
	301	400
genome	TTTACTCAAGATAAATTGGATCATCGCACCACACCCCTCCGTCGATTTCACTAACCCTTGACTTGGAATGTAGAATCGAATAGGGTGGGACACTGGGTGT	
CDS	-----AATCGAATAGGGTGGGACACTGGGTGT	
	401	500
genome	TGTCCAAAGGCGTCAAGCGAGGAGAGATCGTGTGCTCTTCATGCAAAATAAGCCAGAGTTTCTCTTCTTCTGGCTTGGACTCAACAAGATCGGCGCTAC	
CDS	TGTCCAAAGGCGTCAAGCGAGGAGAGATCGTGTGCTCTTCATGCAAAATAAGCCAGAGTTTCTCTTCTTCTGGCTTGGACTCAACAAGATCGGCGCTAC	
	501	600
genome	GGGAGCATTTCATCAACACGAACCTCTCGGGCAAACCTCTGACGCACTCATTGCGTACCGCGACAGCATCCATTCTGATTATGGATGCGGAAGTCCCGACG	
CDS	GGGAGCATTTCATCAACACGAACCTCTCGGGCAAACCTCTGACGCACTCATTGCGTACCGCGACAGCATCCATTCTGATTATGGATGCGGAAGTCCCGACG	
	601	700
genome	CCCATTATAGTGTCTCGATGAAGTCCTTGAGATGGGATATCAGATATATTCTACGGAGGATCCCAGCAACACGCCTTTGCTACACAAGTTGAACTTT	
CDS	CCCATTATAGTGTCTCGATGAAGTCCTTGAGATGGGATATCAGATATATTCTACGGAGGATCCCAGCAACACGCCTTTGCTACACAAGTTGAACTTT	
	701	800
genome	CTCAAATCTCGGATGCGGCCTTGCCCAAGAGTCTGCGAAGGAAAACCACTGCAAATGATATTGCCATGTTGATTTACACCTCCGGAACGACGGGTTTGCC	
CDS	CTCAAATCTCGGATGCGGCCTTGCCCAAGAGTCTGCGAAGGAAAACCACTGCAAATGATATTGCCATGTTGATTTACACCTCCGGAACGACGGGTTTGCC	
	801	900
genome	CAAAGCTGGACGGTTCTCCATGCTCGAGCCAACGGTAGGATTATACCCCCCTCCTCCCCCCCCCTTTTTTTCATTTGCTGTGAAGTTATTAGCTG	
CDS	CAAAGCTGGACGGTTCTCCATGCTCGAGCCAACG-----	
	901	1000
genome	TTCCACTAGCATATTGACTCATATTCACGTTTCCTTTTACACGTGGGATCCAGTTGCCGCACTTTTCTGGACGTCTTTCTACCACTTCAGCGAAAAAGAC	
CDS	-----TTGCCGCACTTTTCTGGACGTCTTTCTACCACTTCAGCGAAAAAGAC	
	1001	1100
genome	CGCCTGTACATCGCCTTGCCCTCTTTACCACAGTGCTGGAGCTGTCCTTGGAATATGTGTGGCCTGGGTACCCGGTGCTACGGTGGTCTGGCGCGCAAGT	
CDS	CGCCTGTACATCGCCTTGCCCTCTTTACCACAGTGCTGGAGCTGTCCTTGGAATATGTGTGGCCTGGGTACCCGGTGCTACGGTGGTCTGGCGCGCAAGT	
	1101	1200
genome	TTTCAACTACTTCCTTCTGGGACGAATGCAGGGCCAACAAGGTCAACCGTGATCCAGTATATTGGAGAAATCTGCCGATACTTACTGAATGCTCCTCCTTC	
CDS	TTTCAACTACTTCCTTCTGGGACGAATGCAGGGCCAACAAGGTCAACCGTGATCCAGTATATTGGAGAAATCTGCCGATACTTACTGAATGCTCCTCCTTC	
	1201	1300
genome	TCCCTTGGACAAGACACACAGATCCGAATGGCGCATGGCAACGGCATGCGTCCGGATGTATGGAACAGATTGAGAGATCGTTTCGGCATCCCTTTGATC	
CDS	TCCCTTGGACAAGACACACAGATCCGAATGGCGCATGGCAACGGCATGCGTCCGGATGTATGGAACAGATTGAGAGATCGTTTCGGCATCCCTTTGATC	
	1301	1400
genome	GGAGAATGGTATGCAAGCACTGAGGGCACCGGAATCTTGACAACTATAACACAGGACCCAATGGCGCTGGTGCGATAGGATACAGAGGCTCCTTGGCCA	
CDS	GGAGAATGGTATGCAAGCACTGAGGGCACCGGAATCTTGACAACTATAACACAGGACCCAATGGCGCTGGTGCGATAGGATACAGAGGCTCCTTGGCCA	

Figure 2B

	1401		1500
genome	GAACTGTCGATAAGGGTCTGAAGATTGCGAAGTTCGACATCCAAACCGAGGAACCTTATTCGTGACAAAAATGGTCGATGCATTGAGGTAAAGTTGACAGT		
CDS	GAACTGTCGATAAGGGTCTGAAGATTGCGAAGTTCGACATCCAAACCGAGGAACCTTATTCGTGACAAAAATGGTCGATGCATTGAG-----		
	1501		1600
genome	ATTAAGTTGAACATATTCCACAGTACCCTTTGTCTTCGGTGTCCAAAATACTGACTACTTGCTTGATGCCCTTCAAGTGTGTCGCAGATGAGCCCGGCG		
CDS	-----TGTGTCGCAGATGAGCCCGGCG		
	1601		1700
genome	AGCTCTTGACAATGATTGATTCAAGTGATCCCACTCGCGCTTCCAAAGGTACCATAAAAAATGCAGGTGCAAACTCCAAGAAAGTCGTCCAGGATGCATT		
CDS	AGCTCTTGACAATGATTGATTCAAGTGATCCCACTCGCGCTTCCAAAGGTACCATAAAAAATGCAGGTGCAAACTCCAAGAAAGTCGTCCAGGATGCATT		
	1701		1800
genome	CAGTGTTGGTAGGTGTAATCTTCACCTTGGTGCCTTCAACAGCAAACTTTGTGATCCGCGCTGACTCGATGACCCTGTCTGCCTACATTTCTAGGCGA		
CDS	CAGTGTTGG-----CGA		
	1801		1900
genome	CCAATACTTTCTGACTGGTGACATCCTTCGTGCGGACGCTGATGGCTATTTCTATTTTGGCGATCGTGTTGGAGATACTTCCGCTGGAAATCTGAAAAC		
CDS	CCAATACTTTCTGACTGGTGACATCCTTCGTGCGGACGCTGATGGCTATTTCTATTTTGGCGATCGTGTTGGAGATACTTCCGCTGGAAATCTGAAAAC		
	1901		2000
genome	GTGTCAACTGCGGAGGTTTCTGAGGTGCTCTCAGCATACCCGGACTGCATCGAGGTCAACGTTTATGGCGTTCAAGTCCCTGGACACGACGGCCGCGCAG		
CDS	GTGTCAACTGCGGAGGTTTCTGAGGTGCTCTCAGCATACCCGGACTGCATCGAGGTCAACGTTTATGGCGTTCAAGTCCCTGGACACGACGGCCGCGCAG		
	2001		2100
genome	GCATGGCTGCCATTGTCTCCAAGGACACCATGAACTGGGATAGTTTCGCCAAGTTTGCACTCAAAAATCTGCCGAAGTACTCTGTGCCGATTTTCATCCG		
CDS	GCATGGCTGCCATTGTCTCCAAGGACACCATGAACTGGGATAGTTTCGCCAAGTTTGCACTCAAAAATCTGCCGAAGTACTCTGTGCCGATTTTCATCCG		
	2101		2200
genome	CAAGGTCCCAGAGATGGAGATTACGGGAACGTTCAAGCAACGAAAGGTTGAACTGGTGAACGAGGGCATGGACCCGAGCAAGATCAAAGACGAAATGCTG		
CDS	CAAGGTCCCAGAGATGGAGATTACGGGAACGTTCAAGCAACGAAAGGTTGAACTGGTGAACGAGGGCATGGACCCGAGCAAGATCAAAGACGAAATGCTG		
	2201	2278	
genome	TGGTTGGATGGGCACTCCTACCGGCCCTTCAAAGAGGCGGAGCATACTAGAGTCGTGAGCGGCAAAGCCAGGCTTTGA		
CDS	TGGTTGGATGGGCACTCCTACCGGCCCTTCAAAGAGGCGGAGCATACTAGAGTCGTGAGCGGCAAAGCCAGGCTTTGA		

Figure 3A

1 ATCCGCCCATCCGCTCTCTTGCCACTGATCTCAAAGCGTGATCCAAAGGTCATTCTTAGGCAGCACTCACGCAGCTACTTAGAACTCTACCCACATATCC
 101 CTTATTGATACAATGGCTCCCGTCGCTGCACTCGCCGCCGCTCTGGCGGCAGGATACTATCTCAATGGCAAGTACCAAATCACTAAGGACTTGGCGCTTG
 M A S T K S L R T W R L
 201 TTCGCGTTGGTCTCCATGCACGCAAAAGACTTGAGGCCTTGGTCGAGAATCGGGACTGCAGTCTTTACAACAGGTTGAGGAACAATGCCAGATCCGGCC
 F A L V S M H A K D L R P W S R I G T A V F T T G S R N N A R S G L
 301 TTTCTCTGTTGCCCTTGTTTTGAGAACACGTCTTACACCTGGAGAGACTTGGAGCTGGCGTCCAACAGGATGGCCATTGGTTTGTGCTCAAGGAATC
 S L L P L F L R T R L T P G E T W S W R P T G W P I G L L L K E S
 401 CAAAAAAGGAGCGTGTGGCGATGATGATGCATAACTCGCCTCTGTTTATTATCACCTGGCTGGCAATGCTCAAGATCATGGTTGTACCTGCTTTTATC
 K K K E R V A M M M H N S P L F I I T W L A M L K I M V V P A F I
 501 AATAACCAGATTGCAGGACCTGTTCTGGTTCATTCTCTTAAAGTGGCCGACGCCAAGTTTCTTGTTCGATTACGAGTTGGCACCTGTCATCCAAAAGT
 N N Q I A G P V L V H S L K V A D A K F L L F D Y E L A P V I Q K S
 601 CGCTCAATGAGATCAAGGACATGGGTTACAATCTCTACACTGTACACCCAAGGATCAAGTTCTAGGTCAACTTTACGCCAATCTGCCGAGGCTGCTCG
 L N E I K D M G Y N L Y T V T P K D Q V L G Q L Y A N L P E A A R
 701 TCAGGTGTTGGATGAGGCTCCTTCATTCTTTGGTTATGTGAATGGCAGAACCTCAGTACCGAAGGTTTCTCGAACGAGAGTCGTCAGGAGGTGGTGATC
 Q V L D E A P S F F G Y V E W Q N L S T E G F S N E S R Q E V V I
 801 TCCGACCCCGCAGCCTTGATTACACCAGCGGGACACGGGATTCCCCAAGGCTGCTATCATGGACCATGGACGTTGCAACTTGGCTTCGATCTCTTATG
 S D P A A L I Y T S G T T G F P K A A I M D H G R C N L A S I S Y G
 901 GCACTCTATGCGGCATCAAACCAGAGAACAAGGTTTACATCACATTGCCGCTCTATCATTCTGCTGGAGCCATCATTGGTCTGGGCCAGAGCTTCACCAG
 T L C G I K P E N K V Y I T L P L Y H S A G A I I G L G Q S F T S
 1001 CGGATGCACCATTGTGCTGGCGCGAAAGTTCTCCGTGACAAAGTTTTGGCGTGATTGCGTTGAGTACGACGTAACCTATTTTCAGTACATTGGCGAACTC
 G C T I V L A R K F S V T K F W R D C V E Y D V T H F Q Y I G E L
 1101 TGCCGCTACCTTCTAAATGCCCGGAAAGTCCACTGGACAAAAGGCATAAGGTTGCGATGGCGTTTGGCAACGGAATGCGCCCGGATGTTTGGGCAAAGT
 C R Y L L N A P E S P L D K R H K V R M A F G N G M R P D V W A K F
 1201 TTCAGGAACGATTCAATATCCCCATTATTGTTGAGTACTACGCCATGAGCGAAGGAACATCGTCGCTTTTGAATGTGGCCAGGAACAAGCGCGACCAAGG
 Q E R F N I P I I V E Y Y A M S E G T S S L L N V A R N K R D Q G
 1301 TGGGTGGGATTCCGTGGCCCCGTCGTGAGGGCCTTGACGCCTCCCGTTCAACTGGTCAAGGTGGACTTTGACACGGAGGAGCTGATCCGCGATAAGAAG
 A V G F R G P V V R A L T P P V Q L V K V D F D T E E L I R D K K
 1401 ACGGGACTTTGCGTCCTATGCCAGCCTGGTGAGATTGGAGAAGTGGTCACGCTAGCCGACAACAAGACGACTGGCGCACGCTATGCTGGGTATTTCAATC
 T G L C V L C Q P G E I G E L V T L A D N K T T G A R Y A G Y F N Q
 1501 AGCCAGAGGTTTGAAGGCAAGGCTGGTCCAGAACGTGGTAGTGAAGGACGACATCTACTCCGGACGGGTGACCTCTTGTACTCCAAGGACCACTACTG
 P E V S K A R L V Q N V V V K D D I Y F R T G D L L Y S K D Q Y W
 1601 GTACTTTGCTGATCGCGCAGGAGACACGTACCGGTGAAAGGAGAGAAGCTGTGACAGCCGAGATTGCAGACACTATCGGCCGTGTTGAGGGCGTGGCT
 Y F A D R A G D T Y R W K G E N V S T A E I A D T I G R V E G V A
 1701 AGTTGTAAGTTTATGGCGTATCGGTCCCGGGCATGGATGGACGCGCGGGCATGGCTGCTTTGGTGCTCAAGAACTCGATTGTGCAGATGGCAGGTGGAA
 S C T V Y G V S V P G M D G R A G M A A L V L K N S I V Q M A G G S
 1801 GCCAGGCAAAGTTCCATGTGGATGAGGCTGCGCTGAACGCGTTTTTGGTGACTTGAGCAAGGATGTGGTCAAAAACTGCCGGCGTATGCGATTCTCG
 Q A K F H V D E A A L N A F L R D L S K D V V K K L P A Y A I P R
 1901 GTTCTTGCGCATTGCAGAGCAGGAACTGGAGACGACGGGCACGTTCAAGAACAAGAAGGTGGAGCTGAAGAAGGAAGGGTTTCGACCTCGGTAAGGTCAAG
 F L R I A E Q E L E T T G T F K N K K V E L K K E G F D L G K V K

Figure 3B

```
2001  GAGCGGCTGTACTGGTGGACACCCAAGGGTGAATATGCCCTTTTGGCGTGGGGGAGAACGAGCAGATCCTCGCAGGACGCGCTCGTCTTTGAGCGATGT
      E R L Y W W T P K G E Y A P F G V A E N E Q I L A G R A R L

2101  TTGTCAATGAAGTCATCGGCATCATCATCATCAAAAAAAAAAAAAAAAAAAAA
```

[illegible]

Figure 4B

	1401		1500
genome	CGGGACCACGGGATTCCCAAGGCTGCTATCATGGACCATGGACGTTGCAACTGTAAGCAATCGCATAGGATCGATAGCGCTGAATGGCTGGCGAGTGGA		
CDS	CGGGACCACGGGATTCCCAAGGCTGCTATCATGGACCATGGACGTTGCAACT-----		
	1501		1600
genome	TGCAAATGGTCGAGATGCTTACCATTATCGTGGTGTGCCTTTTATAGTGGCTTCGATCTCTTATGGCACTCTATGCGGCATCAAACCAGAGACAAGGT		
CDS	-----TGGCTTCGATCTCTTATGGCACTCTATGCGGCATCAAACCAGAGACAAGGT		
	1601		1700
genome	TTACATCACATTGCCGCTCTATCATTCTGCTGGAGGTACGTGCTTCCATCTCACCTCAACATCTCTACGACGGTTTGATCCTGTTCTTACACTCATT		
CDS	TTACATCACATTGCCGCTCTATCATTCTGCTGGAG-----		
	1701		1800
genome	CTTCTGGGCATGGGAACAAAGCCATCATTGGTCTGGGCCAGAGCTTCACCAGCGGATGCACCATTGTGCTGGCGGAAAGTTCTCCGTGACAAAGTTTTG		
CDS	-----CCATCATTGGTCTGGGCCAGAGCTTCACCAGCGGATGCACCATTGTGCTGGCGGAAAGTTCTCCGTGACAAAGTTTTG		
	1801		1900
genome	GCGTGATTGCGTTGAGTACGACGTAACCTCATTTTCAGGTACAAGTCCTATCCAATGGTCTACATACCGTCCTTGTGTATTTCAACGGCGACCGCCACTA		
CDS	GCGTGATTGCGTTGAGTACGACGTAACCTCATTTTCAG-----		
	1901		2000
genome	ACCCGCTTTTATGTATACCGCAGTACATTGGCGAACTCTGCCGCTACCTTCTAAATGCCCCGAAAGTCCACTGGACAAAAGGCATAAGGTTCCGAT		
CDS	-----TACATTGGCGAACTCTGCCGCTACCTTCTAAATGCCCCGAAAGTCCACTGGACAAAAGGCATAAGGTTCCGAT		
	2001		2100
genome	GGCGTTTGGCAACGGAATGCGCCCGGATGTTTGGGCAAAGTTTCAGGAACGATTCAATATCCCCATTATTGTTGAGTACTACGCCATGAGCGAAGGAACA		
CDS	GGCGTTTGGCAACGGAATGCGCCCGGATGTTTGGGCAAAGTTTCAGGAACGATTCAATATCCCCATTATTGTTGAGTACTACGCCATGAGCGAAGGAACA		
	2101		2200
genome	TCGTCGCTTTTGAATGTGGCCAGGAACAAGCGCGACCAAGGTGCGGTGGGATTCCGTGGCCCCGTGCTGAGGTATGCGGCATCTGGGCGCTTTAGTTCTT		
CDS	TCGTCGCTTTTGAATGTGGCCAGGAACAAGCGCGACCAAGGTGCGGTGGGATTCCGTGGCCCCGTGCTGAGG-----		
	2201		2300
genome	CGTGTTCAATGGTTTCATTATAACATCTTCAGCTCAACTTTTGCCCGTGCTTTTCCTTTCAATTTGTTTCCACTAGGGCCTTGACGCCTCCCGTTCAA		
CDS	-----GCCTTGACGCCTCCCGTTCAA		
	2301		2400
genome	CTGGTCAAGGTGGACTTTGACACGGAGGAGCTGATCCGCGATAAGAAGACGGGACTTTGCGTCCTATGCCAGCCTGGTGAGATTGGAGAACTGGTCACGC		
CDS	CTGGTCAAGGTGGACTTTGACACGGAGGAGCTGATCCGCGATAAGAAGACGGGACTTTGCGTCCTATGCCAGCCTGGTGAGATTGGAGAACTGGTCACGC		
	2401		2500
genome	TAGCCGACAACAAGACGACTGGCGCACGCTATGCTGGGTATTTCAATCAGCCAGAGGTTTGAAGGCAAGGCTGGTCCAGAACGTGGTAGTGAAGGACGA		
CDS	TAGCCGACAACAAGACGACTGGCGCACGCTATGCTGGGTATTTCAATCAGCCAGAGGTTTGAAGGCAAGGCTGGTCCAGAACGTGGTAGTGAAGGACGA		
	2501		2600
genome	CATCTACTTCCGGACGGGTGACCTCTTGTACTCCAAGGACCAGTACTGGTACTTTGCTGATCGCGCAGGAGACACGTACCGGTGGAAAGGAGAGAACGTG		
CDS	CATCTACTTCCGGACGGGTGACCTCTTGTACTCCAAGGACCAGTACTGGTACTTTGCTGATCGCGCAGGAGACACGTACCGGTGGAAAGGAGAGAACGTG		
	2601		2700
genome	TCGACAGCCGAGATTGCAGACACTATCGGCCGTGTTGAGGGCGTGGCTAGTTGTACTGTTTATGGCGTATCGGTCCCGGGCATGGATGGACGCGCGGGCA		
CDS	TCGACAGCCGAGATTGCAGACACTATCGGCCGTGTTGAGGGCGTGGCTAGTTGTACTGTTTATGGCGTATCGGTCCCGGGCATGGATGGACGCGCGGGCA		
	2701		2800
genome	TGGCTGCTTTGGTGCTCAAGAACTCGATTGTGCAGATGGCAGGTGGAAGCCAGGCAAAGTTCCATGTGGATGAGGCTGCGCTGAACGCGTTTTTGCCTGA		
CDS	TGGCTGCTTTGGTGCTCAAGAACTCGATTGTGCAGATGGCAGGTGGAAGCCAGGCAAAGTTCCATGTGGATGAGGCTGCGCTGAACGCGTTTTTGCCTGA		

Figure 4C

	2801		2900
genome	CTTGAGCAAGGATGTGGTCAAAAACTGCCGGCGTATGCGATTCTCGGTTCTTGCGCATTGCAGAGCAGGAACTGGAGACGACGGGCACGTTCAAGAAC		
CDS	CTTGAGCAAGGATGTGGTCAAAAACTGCCGGCGTATGCGATTCTCGGTTCTTGCGCATTGCAGAGCAGGAACTGGAGACGACGGGCACGTTCAAGAAC		
	2901		3000
genome	AAGAAGGTGGAGCTGAAGAAGGAAGGGTTCGACCTCGGTAAAGTCAAGGAGCGGCTGTACTGGTGGACACCCAAGGGTGAATATGCCCTTTTGGCGTGG		
CDS	AAGAAGGTGGAGCTGAAGAAGGAAGGGTTCGACCTCGGTAAAGTCAAGGAGCGGCTGTACTGGTGGACACCCAAGGGTGAATATGCCCTTTTGGCGTGG		
	3001		3041
genome	CGGAGAACGAGCAGATCCTCGCAGGACGCGCTCGTCTTTGA		
CDS	CGGAGAACGAGCAGATCCTCGCAGGACGCGCTCGTCTTTGA		

1 GCCTACTTTGCGCTCGCCTCATCGACCCAAAGGCAGCAATGGAAACGGATGCTCTTACCATCGCTTTGACCATCGCCATCGCCATCGTGCTGGCTTTGGT
M E T D A L T I A L T I A I A I V L A L V

101 CAAATTCACGAAAAAGAGCCTGACCTGCATCCGCTCCTGCTCGGGCAGCAATCGTCTGTACGCCCATTCGGAACGAGGGCGAGTCCGTTATCCATAGA
K F N E K E P D L H P L L L G Q Q S S V T P I R N E G E S V I H R

201 TCCAAAACGGTGCCACACGGGACACTGCTGACGAAGCGCCGAGCGAGAAAATCAAGACTCTGCACGATGTCTGGCAGACTGGAGCAGCTGTCAACCCAG
S K T V P H G T L L T K R P S E K I K T L H D V W Q T G A A V N P A

301 CCGGCCGATCGTTGATGTTTATGCTGCAGAACCAGTTTGCCTTTATCGAGGCCACGTATGAGCAAGTCAATAGGAGGATTGGCGGCTTCGGAACAGGTTT
G R S L M F M L Q N Q F A F I E A T Y E Q V N R R I G G F G T G F

401 CGTGAAGGCAACAGGGCTAAAGCCCAAGACGGACACACCAGTAGGAATCTTTATGCCCTACTCTCAAGAATCGTTTCGTTGCCCAGCAGGCATTCTATCGA
V K A T G L K P K T D T P V G I F M P Y S Q E S F V A Q Q A F Y R

501 TACAGCTTTGTTGCTGTCCCATCCATGATCTGAGGAACAACGACCTCTTGGTGGAGGTAGTAGACCAGACCAAGCTCAAGGCCATCATAGTCTCACAAA
Y S F V A V P I H D L R N N D L L V E V V D Q T K L K A I I V S Q K

601 AGGTGCTCCCGTTATTGCTGCAATCTCTGAAGGAGTGTCCAACCATCAAGACAATCATCATGGCAGGAATCTACATCTCACAGGAGCAGCTGGAATGGC
V L P L L L Q S L K E C P T I K T I I M A G I Y I S Q E Q L E M A

701 AGCACAGCATGGAGTAAAGCTGCTCAAATTCGCGGCAGTGGAAATATGAGGGATCCTCGACTCTGATGGAGCCTGTTACGCTGATCCGGAGGATGTTGCC
A Q H G V K L L K F A A V E Y E G S S T L M E P V Q P D P E D V A

801 ATGATCAACTATAACACAAAGTCGTCTTCGCTCTCGAAAGGCGTCATGCTTACCCATGCCAACCTGATCGCGGCATGACTGCCTTCACGGAGTCACTTC
M I N Y N T K S S S L S K G V M L T H A N L I A A M T A F T E S L P

901 CGGCAAAAAAGCGTTTCTCCAGCAAAGATCGTCTTCTCTCATTTTTTCCAATGGAGATGTCATCTCTGTCTTCATGTGAGCGCCATCATCCTGATGGG
A K K R F S S K D R L L S H F S N G D V I S V F M S S A I I L M G

1001 AGGTTCTTTGGTCTTTCCATCTGGTTTGATGAAGAAGCTTTTGCATGATCCCAAGCTTCTGCACCAACGATCTTTGCAAGCACACCCATCATCCTGGAA
G S L V F P S G L M K N V L H D S Q A S A P T I F A S T P I I L E

1101 AAGATTCACGAAGCACTTCAAGTGACGTATGGCCAAGGCTCCATGTTACGGCGCGGCTTTGCTGCCAAATTGGCCATACTTCAAGCTGGACGAATCACTA
K I H E A L Q L T Y G Q G S M F R R G F A A K L A I L Q A G R I T T

1201 CAACAAGCCTATGGGACTTGATTGGACTGGGCGAGGTCCGCAGCAAACCTTGGTGAAAGGTTGGAATGGTTGTAACAACACATCCTACCAAACCTGAGAC
T S L W D L I G L G E V R S K L G G K V R M V V T T H P T K P E T

1301 GCTGGATTATATCAGAGCCGCGATGGGCATCCATGTCAATTACCACTTACGGCAGGACAGAGACGTCGGGCATTGTGACAGCCCGCAACATGCTGGATTAT
L D Y I R A A M G I H V I T T Y G R T E T S G I V T A R N M L D Y

1401 GCCAACGCACCTCATTTAGGACCACAGTGGGTTGCAACGAGGTTAAGCTTGTGGACGATGTTGCAGCTGGCTTTACAAGTGACAGCAGGCCAACCCAC
A N A P H L G P P V G C N E V K L V D D V A A G F T S A D E P N P R

1501 GAGGCGAGATCCTTATCCGAGGCCCAATGTGATGAAAGGTTATTACAAGAAGCGGGTGCCACTTCAACGGCTATCGATGAGGAAGGGTGGTTCCATTC
G E I L I R G P N V M K G Y Y K K P G A T S T A I D E E G W F H S

1601 AGGAGAGCTGGGCACATTCCACTCCAACGGCACTTTAGACGTGTTGGGCAAGAAGAAGACGAAGTCTGCAGTTGGATCACCGTCATGAAAGGAGATG
G E L G T F H S N G T L D V L G K K K K T K S A V G S P S

1701 CTGCATGTGCTACAGAATATAAAAAGGGAGAAGATACGTTCCGTTAACCACATCAAAAAAAAAAAAAAAAAAAAA

Figure 6A

	1	100
genome	ATGGAAACGGATGCTCTTACCATCGCTTTGACCATCGCCATCGCCATCGTGCTGGCTTTGGTCAAATTCACGAAAAAGAGCCTGACCTGCATCCGCTCC	
CDS	ATGGAAACGGATGCTCTTACCATCGCTTTGACCATCGCCATCGCCATCGTGCTGGCTTTGGTCAAATTCACGAAAAAGAGCCTGACCTGCATCCGCTCC	
	101	200
genome	TGCTCGGGCAGCAATCGTCTGTCACGCCATTTCGGAACGAGGGCGAGTCCGTTATCCATAGATCCAAAACGGTGCCACACGGGACACTGCTGACGAAGCG	
CDS	TGCTCGGGCAGCAATCGTCTGTCACGCCATTTCGGAACGAGGGCGAGTCCGTTATCCATAGATCCAAAACGGTGCCACACGGGACACTGCTGACGAAGCG	
	201	300
genome	CCCGAGCGAGAAAATCAAGACTCTGCACGATGTCTGGCAGACTGGAGCAGCTGTCAACCCAGCCGGCCGATCGTTGATGTTTATGCTGCAGAACCAGTTT	
CDS	CCCGAGCGAGAAAATCAAGACTCTGCACGATGTCTGGCAGACTGGAGCAGCTGTCAACCCAGCCGGCCGATCGTTGATGTTTATGCTGCAGAACCAGTTT	
	301	400
genome	GCGTTTATCGAGGTACGATGGACCCGCTGTAGTAACCCCGCTGTCTCTTGAGCAATATCGCAGGAGTCTCACCATTAGAGGATTCATTCTCCTTCGCATA	
CDS	GCGTTTATCGAGG-----	
	401	500
genome	GGCCACGTATGGTAACGTGTGATTCCGGTAGCCTTTCTGTCTGTTGAAGATGCGGTGATGTGGATCTCTAACAGATCTTGTTCAATGGTGACACAGAGCA	
CDS	--CCACGTATG-----AGCA	
	501	600
genome	AGTCAATAGGAGGATTGGCGGCTTCGGAACAGGTTTCGTGAAGGCAACAGGGCTAAAGCCCAAGACGGACACACCAGTAGGAATCTTTATGCCCTACTCT	
CDS	AGTCAATAGGAGGATTGGCGGCTTCGGAACAGGTTTCGTGAAGGCAACAGGGCTAAAGCCCAAGACGGACACACCAGTAGGAATCTTTATGCCCTACTCT	
	601	700
genome	CAAGGTACGCGAACAAGCGCGTGAGTACTGATCCAGCAACGCATAGGGACTGACGCGAGTGAGCCACGTGTTTGACCTTGACCGCGCCTCTTTACTGTA	
CDS	CAAG-----	
	701	800
genome	GAATCGTTCGTTGCCAGCAGGCATTCTATCGATACAGCTTTGTTGCTGTCCCCATCCATGATCTGAGGAACAACGACCTCTTGTTGGAGGTAGTAGACC	
CDS	-AATCGTTCGTTGCCAGCAGGCATTCTATCGATACAGCTTTGTTGCTGTCCCCATCCATGATCTGAGGAACAACGACCTCTTGTTGGAGGTAGTAGACC	
	801	900
genome	AGACCAAGCTCAAGGCCATCATAGTCTCACAAAAGGTGCTCCGTTATTGCTGCAATCTCTGAAGGAGTGTCCAACCATCAAGACAATCATATGGCAGG	
CDS	AGACCAAGCTCAAGGCCATCATAGTCTCACAAAAGGTGCTCCGTTATTGCTGCAATCTCTGAAGGAGTGTCCAACCATCAAGACAATCATATGGCAGG	
	901	1000
genome	AATCTACATCTCACAGGAGCAGCTGGAAATGGCAGCACAGCATGGAGTAAAGCTGCTCAAATTCGCGGCAGTGGAATATGAGGGATCCTCGACTCTGATG	
CDS	AATCTACATCTCACAGGAGCAGCTGGAAATGGCAGCACAGCATGGAGTAAAGCTGCTCAAATTCGCGGCAGTGGAATATGAGGGATCCTCGACTCTGATG	
	1001	1100
genome	GAGCCTGTTCAGCCTGGTATGTGAAGCAAAAGTCAAGGAAATGCGGTTGCTTGATGTTGCTGCGATGTTTTGACCACCAGACCTCTTTAATAGATCC	
CDS	GAGCCTGTTCAGCCTG-----ATCC	
	1101	1200
genome	GGAGGATGTTGCCATGATCAACTATAACACAAAAGTCGTCTTCGGTATGGATGCTTGTTTTTCAGTGGTAGTTTTGTATAAGCGGGCATGGAGATGATTGA	
CDS	GGAGGATGTTGCCATGATCAACTATAACACAAAAGTCGTCTTCG-----	
	1201	1300
genome	GCTTACTTGAAGTACTTTCCGGCTGCATCATTGATCCGATAGCTCTCGAAAGGCGTCATGCTTACCCATGCCAACCTGATCGCGGCGATGACTGCCTTCA	
CDS	-----CTCTCGAAAGGCGTCATGCTTACCCATGCCAACCTGATCGCGGCGATGACTGCCTTCA	
	1301	1400
genome	CGGAGTCAC TTCGGCAAAAAAGCGTTTCTCCAGCAAAGATCGTCTTCTCTCATTTTTCCAATGGAGATGTCATCTCTGTCTTCATGTCGAGCGCCAT	
CDS	CGGAGTCAC TTCGGCAAAAAAGCGTTTCTCCAGCAAAGATCGTCTTCTCTCATTTTTCCAATGGAGATGTCATCTCTGTCTTCATGTCGAGCGCCAT	

Figure 6B

	1401	1500
genome	CATCCTGATGGGAGGTTCTTTGGTCTTTCCATCTGGTTTGATGAAGAACGTTTTGCATGATCCCAAGCTTCTGCACCAACGATCTTTGCAAGGTACAAG	
CDS	CATCCTGATGGGAGGTTCTTTGGTCTTTCCATCTGGTTTGATGAAGAACGTTTTGCATGATCCCAAGCTTCTGCACCAACGATCTTTGCAAG-----	
	1501	1600
genome	ATGTTCCAGTGTGATCTGTGGATTGTCTCGTTATTCATGTGCAAGATACTTACGGCGCTGTTCCATTATTTTTGACCCCTCCTAGCACACCCATCATCC	
CDS	-----CACACCCATCATCC	
	1601	1700
genome	TGGAAAAGATTACGAAGCACTTCAGTTGACGTATGGCCAAGGCTCCATGTTCAGGCGGGCTTTGCTGCCAAATTGGCCATACTTCAAGCTGGACGAAT	
CDS	TGGAAAAGATTACGAAGCACTTCAGTTGACGTATGGCCAAGGCTCCATGTTCAGGCGGGCTTTGCTGCCAAATTGGCCATACTTCAAGCTGGACGAAT	
	1701	1800
genome	CACTACAACAAGCCTATGGGACTTGATTGGACTGGGCGAGGTCCGCAGCAAACCTTGGTGGAAAGGTTGGAATGGTTGTAACAACACGTAAGTCTCCTTTT	
CDS	CACTACAACAAGCCTATGGGACTTGATTGGACTGGGCGAGGTCCGCAGCAAACCTTGGTGGAAAGGTTGGAATGGTTGTAACAACAC-----	
	1801	1900
genome	TTACCATGCGGGCTTACGTACTTGCAGCATGATATTCGAAGATACTAACAGTCTTTTCATGACGCTCCTGTAGATCCTACCAAACCTGAGACGCTGGATT	
CDS	-----ATCCTACCAAACCTGAGACGCTGGATT	
	1901	2000
genome	ATATCAGAGCCGCGATGGGCATCCATGTCATTACCACTTACGGCAGGACAGAGACGTCGGGCATTGTGACAGCCCGCAACATGCTGGATTATGCCAACGC	
CDS	ATATCAGAGCCGCGATGGGCATCCATGTCATTACCACTTACGGCAGGACAGAGACGTCGGGCATTGTGACAGCCCGCAACATGCTGGATTATGCCAACGC	
	2001	2100
genome	ACCTCATTTAGGACCACCAGTGGGTTGCAACGAGGTTAAGCTTGTGGACGATGTTGCAGCTGGCTTTACAAGTGCAGACGAGCCCAACCCACGAGGGCGAG	
CDS	ACCTCATTTAGGACCACCAGTGGGTTGCAACGAGGTTAAGCTTGTGGACGATGTTGCAGCTGGCTTTACAAGTGCAGACGAGCCCAACCCACGAGGGCGAG	
	2101	2200
genome	GTAAGGAGTGCATATCTTTGCTGCTTCATGGGCGATAAAAAGGCTCAAATGCCGTTGATTCTACTCTAACAGCGTTATTTTATTGTTCAACAATATTTT	
CDS	-----	
	2201	2300
genome	GTGCCTAGATCCTTATCCGAGGCCCAATGTGATGAAAGGTTATTACAAGAAGCCGGGTGCCACTTCAACGGCTATCGATGAGGAAGGGTGGTTCCATTTC	
CDS	-----ATCCTTATCCGAGGCCCAATGTGATGAAAGGTTATTACAAGAAGCCGGGTGCCACTTCAACGGCTATCGATGAGGAAGGGTGGTTCCATTTC	
	2301	2391
genome	AGGAGAGCTGGGCACATTCCACTCCAACGGCACTTTAGACGTGTTGGGCAAGAAGAAGACGAAGTCTGCAGTTGGATCACCGTCATGA	
CDS	AGGAGAGCTGGGCACATTCCACTCCAACGGCACTTTAGACGTGTTGGGCAAGAAGAAGACGAAGTCTGCAGTTGGATCACCGTCATGA	

Figure 7A

```

1   CACACGCTCACGTTGCTCTCACCCGACCCACTCCCACTCTCGCTCTCATTCTCTCCCTTGCTCCCTTTCAAGGTCTAACAGCATCAAC
101  ATCAGCATCAGCATCAAGCTTCTCATTCCCTCCCTCGTCTAAATCTCTGAAAGAGTTGCTTTGCAATTCAGCAATGTCCTCGACCAGAACGCCAGTCC
      M S L D Q N A Q S
201  GTTGAGCTCCCAGGCACCCGGCAACCAGGCCAGACAGGCATCTATCGCCGCAAAGGCTTCGAGAATGCCCTTCTCGCCGTCCCACCCAGCAGACCGCACA
      V E L P G T R Q P G Q T G I Y R R K G F E N A L L A V P P S R P H I
301  TCAAGACCATCTACGATGCCTTCCAGCAGGACTGAAGCTTAATCCCAACGGAGCTGCCCTGGGCAGCCGAGTGTACGACCCGGTGACGGACACCTTTGG
      K T I Y D A F Q H G L K L N P N G A A L G S R V Y D P V T D T F G
401  AGGCTATGTCTGGCAGACGTATGCACAGGTGAACGACCGCATCACTCGCTTCGGCAGTGGATTGGTCAAAATTCATAAGGACGTCCATGGTCTTGCCACC
      G Y V W Q T Y A Q V N D R I T R F G S G L V K I H K D V H G L A T
501  GTGGGCCAGAAGTGGTCTCTCGGAATCTGGGCCATCAACCGACCCGAGTGGACCATCGCGTCCGAGGCTTGCTCGGCCTACAACCTGGTCTCCGTGGGTC
      V G Q K W S L G I W A I N R P E W T I A S E A C S A Y N L V S V G L
601  TTTACGATACTTTGGGACCCGAGGCTGTGACTTATGGCATTAAATCACGCTGAGTGGCTCTATTGTCGTAACAAGTGTGGATCATATCGCGACGCTGCTGAA
      Y D T L G P E A V T Y G I N H A E C S I V V T S V D H I A T L L N
701  CGAATCTTCCAAGATGCCTGGGCTCAAAATCATCATCAGCATGGATGACCTCGATACTGGGAGAGCAGGCCAGGACTGGCTCCACCGGCACCATCCTC
      E S S K M P G L K I I I S M D D L D T G R A G P G L A P T G T I L
801  AGGACTTACGCTCAGGACAAAGGGTACTACTTTATGATTGGTCTGAGGTTGAAGCCGTGGTATTGAGCATGGACGAAAGCATACGCCACCAACCTCCT
      R T Y A Q D K G V L L Y D W S E V E A V G I Q H G R K H T P P T S S
901  CCGACGCATATACGATCTGCTATACCAGCGGGACAACAGGCTTGCCAAAAGGTGCCATTTTGACCCATGGAAACTTGATCGCCCTTTTGGCCTCCAGTGA
      D A Y T I C Y T S G T T G L P K G A I L T H G N L I A L L A S S D
1001 TGTGGCCACACCAAGTGGTGGCTGACGATTGCCTCATCAGTTTCTTGCCCTGCCTCAGCTCTTTGGTCGGGTGATGGAGCTCTTCGCGATGGCCGAGGA
      V A T P V L A D D C L I S F L P L P H V F G R V M E L F A M A A G
1101 GGAAAGATTGGCTACAGCACGGGAGATCCTTTGCGTCTCTTGAGGACGTCTCGCACCTAAAGCCCTCCATCTTCCCGCTGTGCCAGACTGCTGAACC
      G K I G Y S T G D P L R L L E D V S H L K P S I F P A V P R L L N R
1201 GCGTGATGCCAAGGTGTATGCGGCAACTGTTGGAGCGCCTGGACTCACAGGGGCACTGGCGGACGAGGATTGGCCACCAAGCTCACCAATTTGAGAGA
      V Y A K V Y A A T V G A P G L T G A L A R R G L A T K L T N L R E
1301 GGGCAAAGGTTTCCACCACCCATTGTGGGACCGAATCCTCTTCTCAAAGGTCAAGCAAGCGCTCGGGCGCAATGTGAGACTGATGTTGACTGCCTCCGCT
      G K G F H H P L W D R I L F S K V K Q A L G G N V R L M L T A S A
1401 CCCATCTCGGCCGAGATCTTGAATTGTCGCTGTCGCTTTCTGCTGCGAGGTCGTGGAGGCATATGGACAGACTGAGGGCGGTGGAGCGGCCACAAACA
      P I S A E I L E F V R V A F C C E V V E A Y G Q T E G G G A A T N T
1501 CCGTGATTGGCGAGACCGAGGCTGGACACGTGGTCCTCCTCAAGCTTGTTGCGAGATCAAACCTGGTGGATGTACCGAGCTGAACTACTTTGCGACCGA
      V I G E T E A G H V G P P Q A C C E I K L V D V P E L N Y F A T D
1601 TAAACCATTCCTCGTGGTGAGATTTGTGTCCGTGGACCCGGTGTCTTCTGTTATCTCAAGGATGAGGCCAAGACCAAGGAGACCATTTGATGAGGAG
      K P F P R G E I C V R G P G V I P G Y L K D E A K T K E T I D E E
1701 GGCTGGCTGCACTCGGGCGATATCGCCATCATGAGTGGCAAAGGCACCGTTACCATCATTGACAGGAAGAAGACGTGTTCAAGCTGAGCCAAGGAGAAT
      G W L H S G D I A I M S G K G T V T I I D R K K N V F K L S Q G E Y
1801 ACATCGCGGCAGAGAACATTGAAGGGGTTTCTCTCCAAGGTTCCATTTCATCCAACAAATTCTGGTGCACGGCGACTCGACCGAGAGCTGTTTGGTGGC
      I A A E N I E G R F L S K V P F I Q Q I L V H G D S T E S C L V A
1901 CATCTTGATCCCAGAGCCTGAGGCCTTCATCCCTTTGTGAACAAAGTGCTCGAGAACGTCAATCTTCAACCTGGAGATCTTGACGCCTACAGGAAGATC
      I L I P E P E A F I P F V N K V L E N V N L Q P G D L A A Y R K I

```

Figure 7B

2001 GTTAAACAACCCAAAGCTGCGCCAGGCTGTCCTCAAAGAGCTGATCAAGGCTGGCAAGGATGCTGGATTGAAAGGCTTTGAGATTCCAAAGGCGATCCTCC
V N N P K L R Q A V L K E L I K A G K D A G L K G F E I P K A I L L

2101 TCGAATCTGAGGCATTACGGTCGAAAACGACAAGATGACCCGACTTTCAAGATCAAAGACACCCCTGTCGTCCAGGCTTACCGCGAGCAACTGACAGC
E S E A F T V E N D K M T P T F K I K R H P V V Q A Y R E Q L T A

2201 CCTCTACAACGAAATCCATCAAAGGAATCCAAGCTGTAAAAAGAAACCCCTTAGAACCTGCGGTGCTCGCAGCAATTAAAAAAAAGAGAGATATTACT
L Y N E I H Q K E S K L *

2301 CTCACAGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 8A

	1	100
genome	ATGTCCCTCGACCAGAACGCCAGTCCGTTGAGCTCCCAGGCACCGGCAACCAGGCCAGACAGGTAACACAGGCCAGCAGTACTGCACCCGTTTTCCAT	
CDS	ATGTCCCTCGACCAGAACGCCAGTCCGTTGAGCTCCCAGGCACCGGCAACCAGGCCAGAC-----	
	101	200
genome	TAAGGTTCCAACCCAGTCTTGGACATGGGGCGAGACCTGCTTGTGTTACTGTGGATTACACCCCGCGCGCCCTTCCCTTCTGGTGCCTTGACCCGT	
CDS	-----	
	201	300
genome	GTGACACCCCGCAAACCTGCTCAGCTCTTCTCGGCTGACCAACCTTCATTACATCCACGTAGGCATCTATCGCCGCAAAGGCTTCGAGAATGCCCTTCTC	
CDS	-----AGGCATCTATCGCCGCAAAGGCTTCGAGAATGCCCTTCTC	
	301	400
genome	GCCGTCCCACCCAGCAGACCGCACATCAAGACCATCTACGATGCCCTTCAGCACGGACTGAAGCTTAATCCCAACGGAGCTGCCCTGGGCAGCCGAGTGT	
CDS	GCCGTCCCACCCAGCAGACCGCACATCAAGACCATCTACGATGCCCTTCAGCACGGACTGAAGCTTAATCCCAACGGAGCTGCCCTGGGCAGCCGAGTGT	
	401	500
genome	ACGACCCGGTGACGGACACCTTTGGAGGCTATGTCTGGCAGACGTATGCACAGGTGAACGACCGCATCACTCGCTTCGGCAGTGGATTGGTCAAAATTCA	
CDS	ACGACCCGGTGACGGACACCTTTGGAGGCTATGTCTGGCAGACGTATGCACAGGTGAACGACCGCATCACTCGCTTCGGCAGTGGATTGGTCAAAATTCA	
	501	600
genome	TAAGGACGTCCATGGTCTTGCCACCGTGGGCCAGAAGTGGTCTCTCGGAATCTGGGCCATCAACCGACCCGAGTGGACCATCGCGTCCGAGGCTTGCTCG	
CDS	TAAGGACGTCCATGGTCTTGCCACCGTGGGCCAGAAGTGGTCTCTCGGAATCTGGGCCATCAACCGACCCGAGTGGACCATCGCGTCCGAGGCTTGCTCG	
	601	700
genome	GCCTACAACCTGGTCTCCGTGGGTCTTTACGATACTTTGGGACCCGAGGCTGTGACTTATGGCATTAAATCAGCTGAGTGCTCTATTGTCGTAACAAGTG	
CDS	GCCTACAACCTGGTCTCCGTGGGTCTTTACGATACTTTGGGACCCGAGGCTGTGACTTATGGCATTAAATCAGCTGAGTGCTCTATTGTCGTAACAAGTG	
	701	800
genome	GTAAGGACATGAAGCCATAACGACAACGGCTAAAAAAAACATGGTTCTCATGACAAGTATCGATAGTAACGTTATTCTTGGCGCTTGATGTGTTCTA	
CDS	-----	
	801	900
genome	GTGGATCATATCGCGACGCTGCTGAACGAATCTTCCAAGATGCCTGGGCTCAAAATCATCATCAGCATGGATGACCTCGATACTGGGAGAGCAGGCCAG	
CDS	-TGGATCATATCGCGACGCTGCTGAACGAATCTTCCAAGATGCCTGGGCTCAAAATCATCATCAGCATGGATGACCTCGATACTGGGAGAGCAGGCCAG	
	901	1000
genome	GACTGGCTCCACCGGCACCATCCTCAGGACTTACGCTCAGGACAAAGGGTACTACTTTATGATTGGTCTGAGGTTGAAGCCGTCGGTATTCAGCATGG	
CDS	GACTGGCTCCACCGGCACCATCCTCAGGACTTACGCTCAGGACAAAGGGTACTACTTTATGATTGGTCTGAGGTTGAAGCCGTCGGTATTCAGCATGG	
	1001	1100
genome	ACGAAAGCATACGCCACCAACCTCCTCCGACGCATATACGATCTGCTATACCAGCGGGACAACAGGCTTGCCAGTAATATGTTGCTTTTATCCACGCA	
CDS	ACGAAAGCATACGCCACCAACCTCCTCCGACGCATATACGATCTGCTATACCAGCGGGACAACAGGCTTGCCA-----	
	1101	1200
genome	TACAGTGTGCCGATATTTTCAATGTTCAATTGCTCTCATTAGATGCATGACACTTATCATTACTTAAGCGACTTTCCTTTGGCGTTCATAGAAAGGTGCC	
CDS	-----AAAGGTGCC	
	1201	1300
genome	ATTTTGACCCATGGAACTTGATCGCCCTTTTGGCCTCCAGTGATGTGGCCACACCAAGTGCTGGCTGACGATTGCCTCATCAGTTTCTTGCCCTGCCTC	
CDS	ATTTTGACCCATGGAACTTGATCGCCCTTTTGGCCTCCAGTGATGTGGCCACACCAAGTGCTGGCTGACGATTGCCTCATCAGTTTCTTGCCCTGCCTC	
	1301	1400
genome	ACGTCTTTGGTCGGGTTCATGGAGCTCTTCGCGATGGCCGAGGAGGAAGATTGGCTACAGCACGGGAGATCCTTTGCGTCTCTTGGAGGACGTCTCGCA	
CDS	ACGTCTTTGGTCGGGTTCATGGAGCTCTTCGCGATGGCCGAGGAGGAAGATTGGCTACAGCACGGGAGATCCTTTGCGTCTCTTGGAGGACGTCTCGCA	

Figure 8B

	1401		1500
genome	CCTAAAGCCCTCCATCTTCCCGCTGTGCCAGACTGCTGAACCGCGTGTATGCCAAGGTGTATGCGGCAACTGTTGGAGCGCCTGGACTCACAGGGGCA		
CDS	CCTAAAGCCCTCCATCTTCCCGCTGTGCCAGACTGCTGAACCGCGTGTATGCCAAGGTGTATGCGGCAACTGTTGGAGCGCCTGGACTCACAGGGGCA		
	1501		1600
genome	CTGGCGGACGAGGATTGGCCACCAAGCTCACCAATTTGAGAGAGGGCAAAGGTTTCCACCACCCATTGTGGGACCGAATCCTCTTCTCAAAGGTCAAGC		
CDS	CTGGCGGACGAGGATTGGCCACCAAGCTCACCAATTTGAGAGAGGGCAAAGGTTTCCACCACCCATTGTGGGACCGAATCCTCTTCTCAAAGGTCAAGC		
	1601		1700
genome	AAGCGCTCGGCGGCAATGTGAGACTGATGTTGACTGGTAAGTGTGCTTTTGGAAGATGAAATCACGTTTATGTAACCCCCCCCCCCCCCTTGTTTATA		
CDS	AAGCGCTCGGCGGCAATGTGAGACTGATGTTGACTG-----		
	1701		1800
genome	AACATTAATCGTTGTACTGTCGTCGTCTAGCCTCCGCTCCCATCTCGGCCGAGATCTTGAATTGTCGCGTGTGCGTTTCTGCTGCGAGGTCGTGGAGGC		
CDS	-----CCTCCGCTCCCATCTCGGCCGAGATCTTGAATTGTCGCGTGTGCGTTTCTGCTGCGAGGTCGTGGAGGC		
	1801		1900
genome	ATATGGACAGACTGAGGGCGGTGGAGCGGCCACAAACACCGTGATTGGCGAGACCGAGGCTGGACACGTCGGTCCTCCTCAAGCTTGTTGCGAGATCAAA		
CDS	ATATGGACAGACTGAGGGCGGTGGAGCGGCCACAAACACCGTGATTGGCGAGACCGAGGCTGGACACGTCGGTCCTCCTCAAGCTTGTTGCGAGATCAAA		
	1901		2000
genome	CTGGTGGATGTACCCGAGCTGAACTACTTTGCGACCGATAAACCATTCCCTCGTGGTGAGATTGTGTCCGTGGACCCGGTGTCAATTCCTGGTTATCTCA		
CDS	CTGGTGGATGTACCCGAGCTGAACTACTTTGCGACCGATAAACCATTCCCTCGTGGTGAGATTGTGTCCGTGGACCCGGTGTCAATTCCTGGTTATCTCA		
	2001		2100
genome	AGGATGAGGCCAAGACCAAGGAGACCATTGATGAGGAGGGCTGGCTGCACTCGGGCGATATCGCCATCATGAGTGGCAAAGGCACCGTTACCATCATTGA		
CDS	AGGATGAGGCCAAGACCAAGGAGACCATTGATGAGGAGGGCTGGCTGCACTCGGGCGATATCGCCATCATGAGTGGCAAAGGCACCGTTACCATCATTGA		
	2101		2200
genome	CAGGAAGAAGAACGTGTTCAAGGTAACAAAACAGTGGCTCTCCACAGATCTGTTAGCGCGTTTTTCTGCGCACAGTACACTGAAACCACCCTGTTTT		
CDS	CAGGAAGAAGAACGTGTTCAAG-----		
	2201		2300
genome	GCTTTGTTCTGAACTAGCTGAGCCAAGGAGAATACATCGGGCAGAGAACATTGAAGGGCGTTTCCTCTCCAAGGTTCCATTATCCAACAAATTCTGG		
CDS	-----CTGAGCCAAGGAGAATACATCGGGCAGAGAACATTGAAGGGCGTTTCCTCTCCAAGGTTCCATTATCCAACAAATTCTGG		
	2301		2400
genome	TGCACGGCGACTCGACCGAGAGCTGTTTGGTGGCCATCTTGATCCAGAGCCTGAGGCCTTCATCCCTTTGTGAACAAAGTGCTCGAGAACGTCAATCT		
CDS	TGCACGGCGACTCGACCGAGAGCTGTTTGGTGGCCATCTTGATCCAGAGCCTGAGGCCTTCATCCCTTTGTGAACAAAGTGCTCGAGAACGTCAATCT		
	2401		2500
genome	TCAACCTGGAGATCTTGACGCCTACAGGAAGATCGTTAACAACCCAAAGCTGCGCCAGGCTGTCTCAAAGAGCTGATCAAGGCTGGCAAGGATGCTGGA		
CDS	TCAACCTGGAGATCTTGACGCCTACAGGAAGATCGTTAACAACCCAAAGCTGCGCCAGGCTGTCTCAAAGAGCTGATCAAGGCTGGCAAGGATGCTGGA		
	2501		2600
genome	TTGAAAGGGTACGTACAGAACTCTCTTTGCTACCGTAGCGGAGGCCACTAGAGTTGAGGTGATACAGATCGACAGAAAAAAAAAAAAAACTAAACAATC		
CDS	TTGAAAG-----		
	2601		2700
genome	TCTCTTCAAAAAC TTGGTGTCTTGTACCACAGCTTTGAGATTCAAAGGCGATCCTCCTCGAATCTGAGGCATTACGGTCAAAAACGACAAGATGA		
CDS	-----GCTTTGAGATTCAAAGGCGATCCTCCTCGAATCTGAGGCATTACGGTCAAAAACGACAAGATGA		
	2701		2800
genome	CCCGGACTTTCAAGATCAAAAGACACCCTGTGTCAGGCTTACCGCGAGCAACTGACAGCCCTCTACAACGAAATCCATCAAAAGGAATCCAAGCTGTA		
CDS	CCCGGACTTTCAAGATCAAAAGACACCCTGTGTCAGGCTTACCGCGAGCAACTGACAGCCCTCTACAACGAAATCCATCAAAAGGAATCCAAGCTGTA		

Figure 8C

	2801
genome	A
CDS	A

Figure 9A

```

1   TTTTTTTTTTTCTTTCTCTCCAACCCCTTTCACCCCCACGGCTCGGCTCGTACTCAAGCCTCAGTCCACACTCTCGTCCTCTAGCCTGCTGCATTCA
101  CGATTCACATTCTCTCTCGACTCCAGCATCGCTACTCCCTCGTGCTACTTTACCATGACCAACCAATTGTAATCCATCGAAGTGGCAGGCAGCCAGAG
      M T T Q L Y S I E V A G S P E
201  ATTCGGGGCGAGGGCAAACCTCGGCGCAGCGTTCTCAGCCCAGACAACTCGTCCAGAGCTATCAGTCTTTCAAGGGCGACGGCTCCATCACCCTCTAT
      I P G E G K P R R S V L S P D K L V Q S Y Q S F K G D G S I T T L Y
301  ATGAGAACTTTTGGAGGGCATCCAGCGCTCAGAGGGAGAGAGTTTCTCGGACACCGCCCATCGTCGATAATGTAGCTCAGCCGTACGAATGGCTAAG
      E N F L E G I Q R S E G G E F L G H R P I V D N V A Q P Y E W L S
401  CTACACGCGCGTTTCAGGAACGTGTGCCAACTTTGGCGTGGTCTCATCCAGCTGGGCGTGAAGTGGACTCGAACTTTGGCATCTTTCCATCAACAGG
      Y T R V Q E R V A N F G A G L I Q L G L K V D S N F G I F S I N R
501  CCCGAATGGACAATGAGTGAGCTGGCAGGCTACATGTACAACCTTACATCTGTGCCGCTTTACGACACTCTGGGCGTCTGGCCATCGAATACATCGTTA
      P E W T M S E L A G Y M Y N F T S V P L Y D T L G V S A I E Y I V N
601  ATCAGACCGAGATGGAGACCATCATCGCGTCGGGTGATAAAGCCTCGATCCTGTAAACATGAAATCAACTCTGCCGACACTCAAGAACATTGTCGTTAT
      Q T E M E T I I A S A D K A S I L L N M K S T L P T L K N I V V M
701  GGGCTCGCTCGAAGACGCGCTCGTTGTGAGGGTAGGGAAATGGATATCCACATCGTTGCGTGGAGTGACGTGGAACGGATGGCTTCAACAACCCCGCG
      G S L E D A L V V E G R E M D I H I V A W S D V E R D G F N N P A
801  CCAGCCAACCCCTCCAACACCGGACGACGTGCCACCATCTGCTACACGTGAGGAACAACCGGGACACCAAAGGGCGCAATCCTGACCCACAAAACTTTG
      P A N P P T P D D V A T I C Y T S G T T G T P K G A I L T H K N F V
901  TGGCTGGCCTTGCTCGTTCCATATGATGGCAAAGCACCAAAAGTTTTTCATCCCCTCGAGCGTTGACACTCACATATCTTACCTGCCCTGGCACATGT
      A G L A S F H M M A K H Q K F F I P S S V D T H I S Y L P L A H V
1001 GTTCGAGCGTTTGTCTCAGGCTGTTATGATTTCTGGCGCAGCTCGGATTGGGTATTACCAAGGAGACACTTTGAAGCTACTCGATGATGTGGCGATCTTG
      F E R L S Q A V M I S G A A R I G Y Y Q G D T L K L L D D V A I L
1101 CAGCCCACCATCTTTGTGTCCGTTCCAGCACTCTTTAACAGGATTTACGACAAGTTCTAGCAGGTGTGAAAGCCAAGGGCGGTCTCGCAGCTTTCTTAT
      Q P T I F V S V P R L F N R I Y D K V L A G V K A K G G L A A F L F
1201 TCAACGCGCTTTTGAACCAAGAAGGCTAATTTGAAACGCGGTATCCTGGAGCAGGCCATCTGGGATCGACTGGTATTTGGTGCAATTCGTGCGCGACT
      N R A F E T K K A N L K R G I L E H A I W D R L V F G A I R A R L
1301 CGGTGGCAAAGTTAAGCATATTGTCTCAGGATCAGCCCCTATAGCCCCGGAGGTCATGGATTTCCTTCGCATTTGCTTCAGTGCCGACGTTTATGAAGGG
      G G K V K H I V S G S A P I A P D V M D F L R I C F S A D V Y E G
1401 TATGGACAGACGGAGCAGGCTGCTGGTTTGTGTATGAGCTACAGAGGTGACTTGACCTCGGGTCAAGTGGGACCCCTCAGCTGTGGGTGGAAGTGAAGC
      Y G Q T E Q A A G L C M S Y R G D L T S G Q V G P P Q L C V E V K L
1501 TCAGAGACGTTCCGGACATGCACTACACAAGCCAGGACAAGCCTCGCCCTCGCGGGGAGATCATGCTTCGAGGCCATTGAGTTTCAAAGGCTATTACAA
      R D V P D M H Y T S Q D K P R P R G E I M L R G H S V F K G Y Y K
1601 GGCTCCAAGCAAACAGAGGAGACTGGACGCACAGGGATGGGCAAGCACTGGAGACGTTGGTGAATGGGACGAGCGTGGCCGCTTGGTGGTGATCGAC
      A P K Q T E E T L D A Q G W A S T G D V G E W D E R G R L V V I D
1701 CGTGTCAAAAACATTTTCAAGTTGGCTCAAGGCGAATACATTGCACCTGAAAAGATCGAAGCCGCTCTGGCCAAACACTACCTTGTGCCCCAGGTCTTTG
      R V K N I F K L A Q G E Y I A P E K I E A V L A K H Y L V A Q V F V
1801 TCTACGGAGACTCCTTCCAAGCGACATTGGTGGGAGTTGTGTCGCCGATGCGGAGACGCTAAAGCCTTGGGCGGATGACCATGGCCTTGGAGGCAAGAG
      Y G D S F Q A T L V G V V V P D A E T L K P W A D D H G L G G K S

```

Figure 9B

```
1901  CTATGAAGAACTATGCGCTCATCCGCTGTCAAAGAACTTTGCTGAAGGAGCTCAAAGAGTTTGGTCGTGAAAATGATCTGAAGGGCTTTGAGATATTG
      Y E E L C A H P A V K E T L L K E L K E F G R E N D L K G F E I L

2001  AAGAACATTCATGTAACGGCGGAGCAATTCTCAATTGAGAATGATCTTTTGACACCCACATTCAAGCTGAAGAGACACACCGCGAAAGAGAAGTACATCG
      K N I H V T A E Q F S I E N D L L T P T F K L K R H T A K E K Y I A

2101  CCGAGATTGAGCTGATGTATAACGGGATCCACTGAAAGAGTCTAGCCAAAGCAGATCTTTTATTACTGTCGTTAAAAAACTACTCGTAACCATCAAAA
      E I E L M Y N G I H *

2201  AAAAAAAAAAAAAA
```

Figure 10A

	1	100
genome	ATGACCACCCAATTGTACTCCATCGAAGTGGCAGGCAGCCAGAGATTCGGGGCGAGGGCAAACCTCGGCGCAGCGTTCTCAGCCCAGACAAACTCGTCC	
CDS	ATGACCACCCAATTGTACTCCATCGAAGTGGCAGGCAGCCAGAGATTCGGGGCGAGGGCAAACCTCGGCGCAGCGTTCTCAGCCCAGACAAACTCGTCC	
	101	200
genome	AGAGCTATCAGTCTTTCAAGGGCGACGGCTCCATCACCCTCTATATGAGAACTTTTTGGAGGGCATCCAGCGCTCAGGTA	
CDS	AGAGCTATCAGTCTTTCAAGGGCGACGGCTCCATCACCCTCTATATGAGAACTTTTTGGAGGGCATCCAGCGCTCAG-----	
	201	300
genome	GCGAGACTTCTCCACATGACACCAGCGTTTCTCTATTAACGTTATCCCTTGGCGCCTTTGGCTTTGTTGTGGTTAACGATGTGCCTAATCATAGAGGGAG	
CDS	-----AGGGAG	
	301	400
genome	GAGAGTTTCTCGGACACCGCCCATCGTCGATAATGTAGCTCAGCCGTACGAATGGCTAAGCTACACGCGGTTCAGGAACGTGTGCGCAACTTTGGCGC	
CDS	GAGAGTTTCTCGGACACCGCCCATCGTCGATAATGTAGCTCAGCCGTACGAATGGCTAAGCTACACGCGGTTCAGGAACGTGTGCGCAACTTTGGCGC	
	401	500
genome	TGGTCTCATCCAGCTGGGCCTGAAAGTCGACTCGAACTTTGGCATCTTTCCATCAACAGGCCCGAATGGGTGAGTGCAGTGGGAGGATCCTTTTTCTTT	
CDS	TGGTCTCATCCAGCTGGGCCTGAAAGTCGACTCGAACTTTGGCATCTTTCCATCAACAGGCCCGAATGG-----	
	501	600
genome	TATGAAATGGGCCGTCGAGTCTAGATTCGATGATGGCTGCGGTTGGACCCGGCCTCGGCCCCCAACCAAACAGGTGTGGGAAACCCCAACCGACTCTA	
CDS	-----	
	601	700
genome	ATAGAACTCCGTTCTCTGATTTGGCGACGAACATACTATTACCTGTAGACAATGAGTGAGCTGGCAGGCTACATGTACAACTTTACATCTGTGCCGCTT	
CDS	-----ACAATGAGTGAGCTGGCAGGCTACATGTACAACTTTACATCTGTGCCGCTT	
	701	800
genome	TACGACACTCTGGGCGTCTCGGCCATCGAATACATCGTTAATCAGACCGAGATGGAGACCATCATCGCGTCGGCTGATAAAGCCTCGATCCTGTTAAACA	
CDS	TACGACACTCTGGGCGTCTCGGCCATCGAATACATCGTTAATCAGACCGAGATGGAGACCATCATCGCGTCGGCTGATAAAGCCTCGATCCTGTTAAACA	
	801	900
genome	TGAAATCAACTCTGCCGACACTCAAGAACATTGTCGTTATGGGCTCGCTCGAAGACGCGCTCGTTGTCGAGGGTAGGGAAATGGATATCCACATCGTTGC	
CDS	TGAAATCAACTCTGCCGACACTCAAGAACATTGTCGTTATGGGCTCGCTCGAAGACGCGCTCGTTGTCGAGGGTAGGGAAATGGATATCCACATCGTTGC	
	901	1000
genome	GTGGAGTGACGTGGAACGCGATGGCTTCAACAACCCCGCGCCAGCCAAACCTCCAACACCGGAGGACGTGCGCACCATCTGCTACACGTGAGGAACAACC	
CDS	GTGGAGTGACGTGGAACGCGATGGCTTCAACAACCCCGCGCCAGCCAAACCTCCAACACCGGAGGACGTGCGCACCATCTGCTACACGTGAGGAACAACC	
	1001	1100
genome	GGGACACCAAAGTAACGTTAAGATGTCGATATGAGAATATCTCTAAAATGTGGCATGCAGCTTGAATGAGAATTCACCTCCTTCATAGACACTTGTGTCC	
CDS	GGGACACCAAAG-----	
	1101	1200
genome	TTTCTCAAGTCACTCACCTCTCTTCATTTATCCTACGTGTATGTGTATGTAGGGGCGCAATCCTGACCCACAAAACTTTGTGGCTGGCCTTGCCCTCGTT	
CDS	-----GGCGCAATCCTGACCCACAAAACTTTGTGGCTGGCCTTGCCCTCGTT	
	1201	1300
genome	CCATATGATGGCAAAGCACAAAAGTTTTTCATCCCCTCGAGCGTTGACACTCACATATCTTACCTGCCCTGGCACATGTGTTGAGCGTTTGTCTCAG	
CDS	CCATATGATGGCAAAGCACAAAAGTTTTTCATCCCCTCGAGCGTTGACACTCACATATCTTACCTGCCCTGGCACATGTGTTGAGCGTTTGTCTCAG	
	1301	1400
genome	GCTGTTATGATTTCTGGCGCAGCTCGGATTGGGTATTACCAAGGAGACACTTTGAAGCTACTCGATGATGTGGCGATCTTGACGCCACCATCTTTGTGT	
CDS	GCTGTTATGATTTCTGGCGCAGCTCGGATTGGGTATTACCAAGGAGACACTTTGAAGCTACTCGATGATGTGGCGATCTTGACGCCACCATCTTTGTGT	

Figure 10B

	1401		1500
genome	CCGTTCCACGACTCTTTAACAGGATTTACGACAAGGTTCTAGCAGGTGTGAAAGCCAAGGGCGGTCTCGCAGCTTTCTTATTCAACCGCGCTTTTGAAAC		
CDS	CCGTTCCACGACTCTTTAACAGGATTTACGACAAGGTTCTAGCAGGTGTGAAAGCCAAGGGCGGTCTCGCAGCTTTCTTATTCAACCGCGCTTTTGAAAC		
	1501		1600
genome	CAAGAAGGCTAATTTGAAACGCGGTATCCTGGAGCACGCCATCTGGGATCGACTGGTATTTGGTGCAATTCGTGCGCGACTCGGTGGCAAAGTTAAGCAT		
CDS	CAAGAAGGCTAATTTGAAACGCGGTATCCTGGAGCACGCCATCTGGGATCGACTGGTATTTGGTGCAATTCGTGCGCGACTCGGTGGCAAAGTTAAGCAT		
	1601		1700
genome	ATTGTCTCAGGATCAGCCCTATAGCCCCGGACGTCATGGATTTCTTCGCATTTGCTTCAGTGCCGACGTTTATGAAGGGTATGGACAGACGGAGCAGG		
CDS	ATTGTCTCAGGATCAGCCCTATAGCCCCGGACGTCATGGATTTCTTCGCATTTGCTTCAGTGCCGACGTTTATGAAGGGTATGGACAGACGGAGCAGG		
	1701		1800
genome	CTGCTGGTTTGTGTATGAGCTACAGAGGTGACTTGACCTCGGGTCAAGTGGGACCCCTCAGCTGTGCGTCGAAGTGAAGCTCAGAGACGTTCCGGACAT		
CDS	CTGCTGGTTTGTGTATGAGCTACAGAGGTGACTTGACCTCGGGTCAAGTGGGACCCCTCAGCTGTGCGTCGAAGTGAAGCTCAGAGACGTTCCGGACAT		
	1801		1900
genome	GCACTACACAAGCCAGGACAAGCCTCGCCCTCGCGGGGAGATCATGCTTCGAGGCCAATCAGTTTTCAAAGGCTATTACAAGGCTCCAAAGCAAACAGAG		
CDS	GCACTACACAAGCCAGGACAAGCCTCGCCCTCGCGGGGAGATCATGCTTCGAGGCCAATCAGTTTTCAAAGGCTATTACAAGGCTCCAAAGCAAACAGAG		
	1901		2000
genome	GAGACACTGGACGCACAGGGATGGGCAAGGTATGGTTTTGTGCAACCACTATTTTGTCAATTACGCTTAACACTGGTTGTCATCAGCCCGCATCTGACT		
CDS	GAGACACTGGACGCACAGGGATGGGCAAG-----		
	2001		2100
genome	AAGGTCACATTGTACGCGACATAGCACTGGAGACGTTGGTGAATGGGACGAGCGTGGCGGCTTGGTGGTGATCGACCGTGTCAAAAACATTTTCAAGTTG		
CDS	-----CACTGGAGACGTTGGTGAATGGGACGAGCGTGGCGGCTTGGTGGTGATCGACCGTGTCAAAAACATTTTCAAGTTG		
	2101		2200
genome	GCTCAAGGCGAATACATTGCACCTGAAAAGATCGAAGCCGTCTGGCCAAACACTACCTTGTGCCCCAGGTCTTTGTCTACGGAGACTCCTTCCAAGCGA		
CDS	GCTCAAGGCGAATACATTGCACCTGAAAAGATCGAAGCCGTCTGGCCAAACACTACCTTGTGCCCCAGGTCTTTGTCTACGGAGACTCCTTCCAAGCGA		
	2201		2300
genome	CATTGGTGGGAGTTGTGCGTGCCCGATGCGGAGACGCTAAAGCCTTGGGCCGATGACCA/TGGCCTTGGAGGCAAGAGCTATGAAGAACTATGCGCTCATCC		
CDS	CATTGGTGGGAGTTGTGCGTGCCCGATGCGGAGACGCTAAAGCCTTGGGCCGATGACCA/TGGCCTTGGAGGCAAGAGCTATGAAGAACTATGCGCTCATCC		
	2301		2400
genome	CGCTGTCAAAGAAACCTTGCTGAAGGAGCTCAAAGAGTTTGGTCGTGAAAATGATCTCAAGGGCTTTGAGATATTGAAGAACATTCATGTAACGGCGGGAG		
CDS	CGCTGTCAAAGAACTTTGCTGAAGGAGCTCAAAGAGTTTGGTCGTGAAAATGATCTCAAGGGCTTTGAGATATTGAAGAACATTCATGTAACGGCGGGAG		
	2401		2500
genome	CAATTCTCAATTGAGAATGATCTTTTGGTGAGTCGATCGGACTTGCGCACCGAACAGTTTGAGAGTCAAGGATCGTATATTAACCTAAGCCTTGCTTTGGT		
CDS	CAATTCTCAATTGAGAATGATCTTTG-----		
	2501		2586
genome	AGACACCCACATTCAAGCTGAAGAGACACACCGCGAAAGAGAAGTACATCGCCGAGATTGAGCTGATGTATAACGGGATCCACTGA		
CDS	--ACACCCACATTCAAGCTGAAGAGACACACCGCGAAAGAGAAGTACATCGCCGAGATTGAGCTGATGTATAACGGGATCCACTGA		

Figure 11A

```

1   TTTCTCACCTTCCCTCCGCTGCCCTCTGCTGCACACTCCTCTGGCTTATACCATCCACCCCTCTAGCCCCGCCACTTCGCCGCCAACCTCATCCGACTC
101  ACACCGCAATGGCTACTCAAATGTAAGTGGTGGTCCCAACAGCCCGACATTCCCGGCGAAGGCAAGCCCGCGTAGTGTGCTTTGTCCAGACAA
    M A T Q M Y S V V V P N S P D I P G E G K P R R S V L C P D K
201  GCTCCTGGAGAACTACCCCTCAGTGAAGCAGGCTCAACGATCAGGACCTGTACGAGAACTTCCAAGAAGGTGTTCTCCGTTCAGGCGGCGCCATTTT
    L L E N Y P S V K A G S T I T T L Y E N F Q E G V L R S G G A H F
301  TTGGGCCATCGTCCATTGTGAATGGCCAGCCTCAGGCTTACAAGTGGCAGTCGTATGTCGATGTCAGCAAGCGTGTACGCACTTCGGCGCTGGCCTGG
    L G H R P I V N G Q P Q A Y K W Q S Y V D V S K R V T H F G A G L A
401  CTCATCTCGGCTTGTCTCCAAAGCAAACTTTGGAATTTTCTCTATCAACCGGCTGAGTGGACCATGAGTGAGCTTGTGGCTATATGCACAACTACAC
    H L G L S P K Q N F G I F S I N R P E W T M S E L A G Y M H N Y T
501  CAGCGTCCCCCTCTATGATACATTGGGAGTCGCCGCGATCGAGTATATCGTTAACCAGACTGAGATGCAGATCATCATTGCTTCGTCGACAAAGCTTCT
    S V P L Y D T L G V A A I E Y I V N Q T E M Q I I I A S S D K A S
601  ATCATCCTCCACATGAAATCAGCACTTCCAACCGTTCAGACGATTGTGCTCATGGGGGAATTTACTGACGCTCTCGTCGAGAGGGTAAGGAGCTCAACA
    I I L H M K S A L P T V Q T I V V M G E F T D A L V A E G K E L N I
701  TCAACATTGTATCCTGGACCGATGTCGAAAAGAGCGGTCTTGAGCGGCCTGTGGAAGCCGTGCACCCACAGCCGAGGATATCGCTACCATCTGTTACAC
    N I V S W T D V E K S G L E R P V E A V H P T A E D I A T I C Y T
801  ATCTGGAACCACTGGAACGCCAAAAGGTGCTATCTTGACCCACAAGAACTTTGTTGCCACTATCGCTTCATTCCACATGATGGCAAAGCATGGCAGGTTG
    S G T T G T P K G A I L T H K N F V A T I A S F H M M A K H G R F
901  TTCATTCCCTCGCCTGCCGACACACATGTATCCTACCTGCCCTTGCCGACGTCTTTGAGCGCCTTTGCCAGGCTGTTATGATCTCGGGCGCTGCGCGTA
    F I P S P A D T H V S Y L P L A H V F E R L C Q A V M I S G A A R I
1001 TTGGTTACTACCAAGGAGATACGCTGAAGCTGCTGGACGATGTTGCCGCTCTGCATCCCACCATTTTGCCTCCGTCCTCTCTTTAACCGTATCTA
    G Y Y Q G D T L K L L D D V A V L H P T I F A S V P R L F N R I Y
1101 CGACAAGGTGCTTGTGGCGTCAAGGCCAAGGGTGGTATCGCCGCTTCTTGTGTTAACCGCGCATATAATTCCAAGAAGGCCAACTTGCGAAAGGGCGTA
    D K V L A G V K A K G G I A A F L F N R A Y N S K K A N L R K G V
1201 CTTGAGCATCCGCTCTGGGACAAGCTGGTCTTTGGAGCGATTGCGCGCGCTTGGGTGGCAAGGTTAAGCACATCGTGTGAGGATCTGCCCCATCTCTC
    L E H P L W D K L V F G A I R A R L G G K V K H I V S G S A P I S P
1301 CTGATGTGATGGATTTCTCCGCATCTGCTTCAGCGCTGATGTGTATGAGGATATGGCCAGACGGAACAGGCAGCCGGATTAAGTATGAGCTATCGCGG
    D V M D F L R I C F S A D V Y E G Y G Q T E Q A A G L S M S Y R G
1401 TGATTTGACTCCAGGACAGGTTGGCCACCTCAACTGTGCACAGAGGTCAAGTTGAAGGACATCCCTAGTATGAACTATAGCAGCGCGGACAAGCCTTTC
    D L T P G Q V G P P Q L C T E V K L K D I P S M N Y S S A D K P F
1501 CCCCCTGGAGAAATCATGCTTCGCGGAAACTCTGTGTTCAAGGGCTATTACAAAGCACAAAGCAGACTGAAGAAACATTGGATGCTGACGGTTGGTCCA
    P R G E I M L R G N S V F K G Y Y K A P K Q T E E T L D A D G W S S
1601 GTACCGGAGACGTTGGACAGTGGGATGCCAAGGCCGTCTGGTGGTCATTGATCGCGTCAAGAACATCTTCAAGTTGGCGCAAGGAGAATATATTGCGCC
    T G D V G Q W D A Q G R L V V I D R V K N I F K L A Q G E Y I A P
1701 TGAAAAGATCGAGGCTGTCTCGCCAAGCACTTCTCGTTGCCAGATTTTGTCTATGGGCACTCGCTCCAGGCCACCATTGTGCGGGTGGTTGTCCCT
    E K I E A V L A K H F L V A Q I F V Y G H S L Q A T I V A V V V P
1801 GATGCTGAGACGCTCAAGTTGTGGGCTAAAGAAAACAAGCTGGGTGACAAGTCTTACGAGGAGCTGTGCGCTCTCCCTCAGCTTCGCACAACCCTCCAAA
    D A E T L K L W A K E N K L G D K S Y E E L C A L P Q L R T T L Q K
1901 AGGAGTTGGCTACTTTTGGCAAAGAATCGGATCTGAAGGGCTTTGAGATTCTAAGAACATTCTGTTATCTCCGAGCAGTTTTCAATTGAGAACGATCT
    E L A T F G K E S D L K G F E I P K N I H V I S E Q F S I E N D L

```

Figure 11B

```
2001   TTGACCCCCACCTTCAAGCTGAAGAGACATGCTGCCAAAGAGAAGTATAACGCCGAAATCGACCGCATGTATGCAGAAATCGCTTAATATAAATAATGG
      L T P T F K L K R H A A K E K Y N A E I D R M Y A E I A

2101   TTGTACTCAATATAAAAAAAAAA
```

Figure 12A

	1	100
genome	ATGGCTACTCAAATGTACTCGGTGGTCGTCCCAACAGCCCCGACATTCCCGGCGAAGGCAAGCCCCGCCGTAGTGTGCTTTGTCCAGACAAGCTCCTGG	
CDS	ATGGCTACTCAAATGTACTCGGTGGTCGTCCCAACAGCCCCGACATTCCCGGCGAAGGCAAGCCCCGCCGTAGTGTGCTTTGTCCAGACAAGCTCCTGG	
	101	200
genome	AGAACTACCCCTCAGTGAAAGCAGGCTCAACGATCACGACCCTGTACGAGAACTTCCAAGAAGGTGTTCTCCGTTACAGTAACAACGCTTACCACGTCGA	
CDS	AGAACTACCCCTCAGTGAAAGCAGGCTCAACGATCACGACCCTGTACGAGAACTTCCAAGAAGGTGTTCTCCGTTACAG-----	
	201	300
genome	CGTTCGCCTCGCAATGGCACCTTTCCTTCGGTCAAACCTCTTAAATGTTTCCTTTACAATCGTAGGCGGCGCCCATTTTTTGGGCCATCGTCCCATTGT	
CDS	-----CGGCGCCCATTTTTTGGGCCATCGTCCCATTGT	
	301	400
genome	GAATGGCCAGCCTCAGGCTTACAAGTGGCAGTCGTATGTCGATGTCAGCAAGCGTGTTACGCACTTCGGCGCTGGCCTGGCTCATCTCGGCTTGTCTCCA	
CDS	GAATGGCCAGCCTCAGGCTTACAAGTGGCAGTCGTATGTCGATGTCAGCAAGCGTGTTACGCACTTCGGCGCTGGCCTGGCTCATCTCGGCTTGTCTCCA	
	401	500
genome	AAGCAAACTTTGGAATTTTCTCTATCAACCGGCCTGAGTGGCTAGGTGATGCCTTGCTTTCTTTGCCCACTCGTCGTCAAGGTAACGCAGGGCCTGCG	
CDS	AAGCAAACTTTGGAATTTTCTCTATCAACCGGCCTGAGTGG-----	
	501	600
genome	CCGATTTTAAACAGTACATTCTATGCACTGTCGTTACATAGACCATGAGTGAGCTTGCTGGCTATATGCACAACACACCAGCGTCCCCCTCTATGATAC	
CDS	-----ACCATGAGTGAGCTTGCTGGCTATATGCACAACACACCAGCGTCCCCCTCTATGATAC	
	601	700
genome	ATTGGGAGTCGCCGCGATCGAGTATATCGTTAACCAGACTGAGATGCAGATCATCATTGCTTCGTCCGACAAAGCTTCTATCATCCTCCACATGAAATCA	
CDS	ATTGGGAGTCGCCGCGATCGAGTATATCGTTAACCAGACTGAGATGCAGATCATCATTGCTTCGTCCGACAAAGCTTCTATCATCCTCCACATGAAATCA	
	701	800
genome	GCACITCCAACCGTTACAGACGATTGTCGTATGGGGGAATTTACTGACGCTCTCGTCGCAGAGGGTAAGGAGCTCAACATCAACATTGTATCCTGGACCG	
CDS	GCACITCCAACCGTTACAGACGATTGTCGTATGGGGGAATTTACTGACGCTCTCGTCGCAGAGGGTAAGGAGCTCAACATCAACATTGTATCCTGGACCG	
	801	900
genome	ATGTCGAAAAGAGCGGTCTTGAGCGGCCTGTGCAAGCGGTGCACCCACAGCCGAGGATATCGCTACCATCTGTTACACATCTGGAACCACTGGAACGCC	
CDS	ATGTCGAAAAGAGCGGTCTTGAGCGGCCTGTGCAAGCGGTGCACCCACAGCCGAGGATATCGCTACCATCTGTTACACATCTGGAACCACTGGAACGCC	
	901	1000
genome	AAAGTAAGTCAAGATCATTACATGGTGAGCCTCCATTGCTTGGACTGAACAGTCTACTCAGCAGGTTCTTCGTTTACTTTGACATGCGCAGAGGTGCTA	
CDS	AAA-----AGGTGCTA	
	1001	1100
genome	TCTTGACCCACAAGAACTTTGTTGCCACTATCGCTTCATTCCACATGATGGCAAAGCATGGCAGGTTCTTCATTCCCTCGCCTGCCGACACACATGTATC	
CDS	TCTTGACCCACAAGAACTTTGTTGCCACTATCGCTTCATTCCACATGATGGCAAAGCATGGCAGGTTCTTCATTCCCTCGCCTGCCGACACACATGTATC	
	1101	1200
genome	CTACCTGCCCTTGCCACGCTTTGAGCGCCTTTGCCAGGCTGTTATGATCTCGGGCGCTGCGCGTATTGGTTACTACCAAGGAGATACGCTGAAGCTG	
CDS	CTACCTGCCCTTGCCACGCTTTGAGCGCCTTTGCCAGGCTGTTATGATCTCGGGCGCTGCGCGTATTGGTTACTACCAAGGAGATACGCTGAAGCTG	
	1201	1300
genome	CTGGACGATGTTGCCGTCTGCATCCCACCATTTTTGCCTCCGTCCCTCGTCTCTTTAACCGTATCTACGACAAGGTGCTTGCTGGCGTCAAGGCCAAGG	
CDS	CTGGACGATGTTGCCGTCTGCATCCCACCATTTTTGCCTCCGTCCCTCGTCTCTTTAACCGTATCTACGACAAGGTGCTTGCTGGCGTCAAGGCCAAGG	
	1301	1400
genome	GTGGTATCGCCGCCTTCTTGTTTAACCGCGCATATAATTCCAAGAAGGCCAACTTGCGAAAGGGCGTACTTGAGCATCCGCTCTGGGACAAGCTGGTCTT	
CDS	GTGGTATCGCCGCCTTCTTGTTTAACCGCGCATATAATTCCAAGAAGGCCAACTTGCGAAAGGGCGTACTTGAGCATCCGCTCTGGGACAAGCTGGTCTT	

Figure 12B

Figure 13

1 GGCACGAGGCTCTACTCTCCATTGCCCACTCACTCATTGCCCTCTGTCCATCACCGGCATTGCTCGTTCGCGCCTTCGCGCACTCCACTCTTTCTTTCA
 101 TTCCTTCTTTACAACGGCCATCTCCCCCTCGCTCTGCGCTTCTCCCATCCAGCTAACAAATGCACATTCTGAATGCCACAAGACCATTCTCCAGGCTGTC
 M H I L N A T R P F S R L S
 201 TCCAACCGTAAGGAGACCTTGGCTAGGACTCGGCCAGACGCGCCCTTATGCTATCGCGCAGACCGAGGCCAGTCCTAGGCTGTCATATGTCCGAGGCACC
 P T V R R P W L G L G Q T R P Y A I A Q T E A S P R L S Y V R G T
 301 ACCGTCGGCAGCCAGCTATGCGAGGATCCCATCGGTGCGTACTGGGACAGGGTCGTCAATCGTCACGGTGACCGCCTCGGACTTGTGTCGCAAGCAGAGC
 T V G T Q L C E D P I G A Y W D R V V N R H G D R L G L V V K H E Q
 401 AGGACCTGCACTGGACCTTCCGTCACTTTGGCGGGCAGGTTGATAGCCTCTGCCGTGGGCTCTATGAGTCTGGCTTGCAGAAAGGGTGATCGACTCGCTGT
 D L H W T F R Q F G G Q V D S L C R G L Y E S G L R K G D R L A V
 501 CTGGATGCCGAACAACAGCGCGTGGGCCAGCTCCAGTATGCTACTGCCAAGTCTGGCATCATTCTGGTAACTCTCAACCCTGCGTACCGGAGGCAAGAG
 W M P N N S A W A T L Q Y A T A K S G I I L V T L N P A Y R R Q E
 601 CTAATGCAGACATTGTCTTTGGTCGAGTGCAAGTCACTGGTCTATGTGCCAAGTCTAAAGACTTCGAATTATAGCGAGATGTTGCTCGACCTCCTACCAG
 L L Q T L S L V E C K S L V Y V P S L K T S N Y S E M L L D L L P E
 701 AACTCCAGTACCAGTCGCCAAATCAGCTCTTGACCGAGAAGCTACCCTCACTTCGTCAAGTCATCGTGTGTTGACAATGGCTCGCAAGTCCAGAGACAGC
 L Q Y Q S P N Q L L T E K L P S L R Q V I V F D N G S Q V P E T A
 801 AAAATTGAAGGGATTGACAAAGTATCAGGATTTGTTGATCAAGAATCCCTCGACCGCTGTGACGGAGCTCTTGAAAAGGAACGGCTCGCTATCGACAAC
 K L K G L T K Y Q D L L I K N P S T A V D G A L E K E R L A I D N
 901 AGGGATATCATCAATCTCCAGTTTACTAGCGGAACCTACAGGCCTTCCCAAGGGCGTCTCGCTGTGCGCATCGAAACATCTTGAATAACGGCATTCAATTG
 R D I I N L Q F T S G T T G L P K G V S L S H R N I L N N G I H I G
 1001 GAGATAACATGCGACTGACGGAAAAGGATTTGCTTTGCTGCCCGGTCCCGCTCTTTCACTGCTTTGGACTGGTGGTGGCAAGCTTGGCTGCAATGACCCA
 D N M R L T E K D L L C C P V P L F H C F G L V L A S L A A M T H
 1101 TGGCGCAGGAATTATTTACCTTCGCGAGTCCTTTGATGCTGAGGCCACACTGAGGGCTGTTTCTGAGGAGGGTGCTACAGCGCTGCATGGCGTGCCGACT
 G A G I I Y P S Q S F D A E A T L R A V S E E G A T A L H G V P T
 1201 ATGCTGTTGGAAGAGATGAACCACCCCACTTTGCAAAGTACAACCTTTGACACTTCGGACAGGAATTGCAGCTGGATCCCCTGTGCCATTGAGGTCA
 M L L E E M N H P N F A K Y N L S T L R T G I A A G S P V P I E V M
 1301 TGAAGACGTGCAGACAAAGATGAACCTGAAGGAGCTGACTATCTGTTACGGCATGACCGAGACCTCGCCCGTGTCTTCATGACACTCACAACGGATGA
 K N V Q T K M N L K E L T I C Y G M T E T S P V S F M T L T T D E
 1401 ATTACGGGATCGATGTGAGACTGTTGGACGAATTATGCCACATCTCGAGGCCAAAGTCGTCAACCCTGAGACGGGAGAGACTTTGCCAGTGAATTCATCA
 L R D R C E T V G R I M P H L E A K V V N P E T G E T L P V N S S
 1501 GGAGAGTTGTGCACGCGCGGGTATGCTGTGATGGAGGGTGGTTACTGGCGATCCCAGGAGCAGACAGATGCAGTGGTGGACAAGGATGGCTGGATGCACA
 G E L C T R G Y A V M E G G Y W R S Q E Q T D A V V D K D G W M H T
 1601 CTGGCGACACTGCCGTGCTCGATGACCGTGGCTTTTGCAGGATCGACGGACGCATCAAGGACATGGTGATCCGAGGAGGCGAAAAATCCATCCTGTAGA
 G D T A V L D D R G F C R I D G R I K D M V I R G G E K I H P V E
 1701 GGTGAGAACTGTCTCTTTGAGATGGACGGCGTCAAGAACGTGTCTGTGATTGGCGTTCCGACAAAGCGGTATGGCGAGCAGGTGTGTGCGTGGATCTCG
 V E N C L F E M D G V K N V S V I G V P D K R Y G E Q V C A W I S
 1801 ACCAAGGACGGGAAGACGGTCAGTCTGGAGGCAGTGCAAAAGTTCTGTGAGGGCAAGATTGCGCACTACAAGGTGCCGCGGTATGTGGTTGTGGTGGAGT
 T K D G K T V S L E A V Q K F C E G K I A H Y K V P R Y V V V V E S
 1901 CCAATGAGTTCCCGACTACCCCTCGGGCAAGATCCAAAAGAATGTGATGCGCGAGCTGACCAAGGCGAAGCTGCAGCTGCCTTGATGGTACTAGGATAT
 N E F P T T P S G K I Q K N V M R E L T K A K L Q L P
 2001 GGAGCCGACGAAAGTAATAAAGGCGTATGCTGGCATGGCGCAAGATCTGAGCCCTGCGGTGAGGTGCATTCACTGACGCCATTAG

Figure 14A

	1	100
genome	ATGCACATTCTGAATGCCACAAGACCATTCTCCAGGCTGTCTCCAACCGTAAGGAGACCTTGGCTAGGACTCGGCCAGACGCGCCCTTATGCTATCGCGC	
CDS	ATGCACATTCTGAATGCCACAAGACCATTCTCCAGGCTGTCTCCAACCGTAAGGAGACCTTGGCTAGGACTCGGCCAGACGCGCCCTTATGCTATCGCGC	
	101	200
genome	AGACCGAGGCCAGTCCTAGGCTGTCATATGTCCGAGGGACCACCGTCGGCACCCAGCTATGCGAGGATCCCATCGGTGCGTACTGGGACAGGGTCGTCAA	
CDS	AGACCGAGGCCAGTCCTAGGCTGTCATATGTCCGAGGGACCACCGTCGGCACCCAGCTATGCGAGGATCCCATCGGTGCGTACTGGGACAGGGTCGTCAA	
	201	300
genome	TCGTCACGGTGACCGCCTCGGACTTGTGTCGAAGCACGAGCAGGACCTGCACTGGACCTTCCGTCAGTTTGGCGGGCAGGTTGATAGCCTCTGCCGTGGG	
CDS	TCGTCACGGTGACCGCCTCGGACTTGTGTCGAAGCACGAGCAGGACCTGCACTGGACCTTCCGTCAGTTTGGCGGGCAGGTTGATAGCCTCTGCCGTGGG	
	301	400
genome	CTCTATGAGTCTGGCTTGCGAAAGGGTGATCGACTCGCGTATGTACACTCTCGTCGTGCCTGAAGCGCTTGAAGGGAAAACTGTATTGAACCCCGGCC	
CDS	CTCTATGAGTCTGGCTTGCGAAAGGGTGATCGACTCGC-----	
	401	500
genome	GCTGCATTAGCGTCGTTTCAACATTTGGCTAATTTTTCTTTCCCTCGTCCTTTGCCGATGCAAACAGTGTCTGGATGCCGAACAACAGCGCGTGGGCCA	
CDS	-----TGTCTGGATGCCGAACAACAGCGCGTGGGCCA	
	501	600
genome	CGCTCCAGTATGCTACTGCCAAGTCTGGCATCATTCTGGTAACTCTCAACCCTCGGTATGTAAGTGCATGTCATAATCATACTCGCGTCGCCTCAAAAATA	
CDS	CGCTCCAGTATGCTACTGCCAAGTCTGGCATCATTCTGGTAACTCTCAACCCTCGGT-----	
	601	700
genome	TGCTTTTTTGACGTGTTTACTTATCCTGCGGTTCTCATAACAGGTACCGGAGGCAAGAGCTACTGCAGACATTGTCTTTGGTCGAGTGCAAGTCATTGG	
CDS	-----CCGGAGGCAAGAGCTACTGCAGACATTGTCTTTGGTCGAGTGCAAGTCATTGG	
	701	800
genome	TCTATGTGCCAAGTCTAAAGACTTCGAATTATAGCGAGATGTTGCTCGACCTCCTACCAGAACTCCAGTACCAGTCGCCAAATCAGCTCTTGACCGAGAA	
CDS	TCTATGTGCCAAGTCTAAAGACTTCGAATTATAGCGAGATGTTGCTCGACCTCCTACCAGAACTCCAGTACCAGTCGCCAAATCAGCTCTTGACCGAGAA	
	801	900
genome	GCTACCCTCACTTCGTCAAGTCATCGTGTGTTGACAATGGCTCGCAAGTCCCAGAGACAGCAAAATTGAAGGGATTGACAAAGTATCAGGATTTGTTGATC	
CDS	GCTACCCTCACTTCGTCAAGTCATCGTGTGTTGACAATGGCTCGCAAGTCCCAGAGACAGCAAAATTGAAGGGATTGACAAAGTATCAGGATTTGTTGATC	
	901	1000
genome	AAGAATCCCTCGACCGCTGTGACGGAGCTCTTGAAAAGGAACGGCTCGCTATCGACAACAGGGATATCATCAATCTCCAGTTTACTAGCGGAACCTACAG	
CDS	AAGAATCCCTCGACCGCTGTGACGGAGCTCTTGAAAAGGAACGGCTCGCTATCGACAACAGGGATATCATCAATCTCCAGTTTACTAGCGGAACCTACAG	
	1001	1100
genome	GCCTTCCCAAGTATGTGTCAGAAATTGAACCTTGACACGGGATGCATAATCGACGTTCAACCACCTTGTCTGACACCGTCTTGATCTACAATACTATG	
CDS	GCCTTCCCAAG-----	
	1101	1200
genome	GTCCAAATTTTAGGGGCGTCTCGCTGTGCGATCGAAACATCTTGAATAACGGCATTGATTTGGAGATAACATGCGACTGACGGAAAAGGATTTGCTTTG	
CDS	-----GGCGTCTCGCTGTGCGATCGAAACATCTTGAATAACGGCATTGATTTGGAGATAACATGCGACTGACGGAAAAGGATTTGCTTTG	
	1201	1300
genome	CTGCCCGGTCCCGCTCTTTCACTGCTTTGGACTGGTGTGGCAAGCTTGGCTGCAATGACCCATGGCGCAGGAATTATTTACCCTTCGCAGTCCTTTGAT	
CDS	CTGCCCGGTCCCGCTCTTTCACTGCTTTGGACTGGTGTGGCAAGCTTGGCTGCAATGACCCATGGCGCAGGAATTATTTACCCTTCGCAGTCCTTTGAT	
	1301	1400
genome	GCTGAGGCCACACTGAGGGCTGTTTCTGAGGAGGGTGCTACAGCGCTGCATGGCGTGCCGACTATGCTGTTGGAAGAGATGAACCAACCCCACTTTGCAA	
CDS	GCTGAGGCCACACTGAGGGCTGTTTCTGAGGAGGGTGCTACAGCGCTGCATGGCGTGCCGACTATGCTGTTGGAAGAGATGAACCAACCCCACTTTGCAA	

Figure 14B

	1401		1500
genome	AGTACAACCTTTGACACTTCGGACAGGAATTGCAGCTGGATCCCCTGTGCCATTGAGGTCATGAAGAACGTGCAGACAAAGATGAACCTGAAGGAGCT		
CDS	AGTACAACCTTTGACACTTCGGACAGGAATTGCAGCTGGATCCCCTGTGCCATTGAGGTCATGAAGAACGTGCAGACAAAGATGAACCTGAAGGAGCT		
	1501		1600
genome	GACTATCTGTTACGGCATGACCGAGACCTCGCCGTGTCCTTCATGACACTCACAACGGATGAATTACGGGATCGATGTGAGACTGTTGGACGAATTATG		
CDS	GACTATCTGTTACGGCATGACCGAGACCTCGCCGTGTCCTTCATGACACTCACAACGGATGAATTACGGGATCGATGTGAGACTGTTGGACGAATTATG		
	1601		1700
genome	CCACATCTCGAGGCCAAAGGTAAAACGGCTGTCAGCCTGGTGTTTCATAATTTGTCTTGAATCTCGGTGGTCTGTTGCTAATTTGGGTACATTGCTTGC		
CDS	CCACATCTCGAGGCCAAAG-----		
	1701		1800
genome	ACAATACGGTAGTCGTCAACCCTGAGACGGGAGAGACTTTGCCAGTGAATTCATCAGGAGAGTTGTGCACGCGCGGGTATGCTGTGATGGAGGGTGGTTA		
CDS	-----TCGTCAACCCTGAGACGGGAGAGACTTTGCCAGTGAATTCATCAGGAGAGTTGTGCACGCGCGGGTATGCTGTGATGGAGGGTGGTTA		
	1801		1900
genome	CTGGCGATCCAGGAGCAGACAGATGCAGTGGTGGACAAGGATGGCTGGATGCACACTGGCGACACTGCCGTGCTCGATGACCGTGGCTTTTGCAGGATC		
CDS	CTGGCGATCCAGGAGCAGACAGATGCAGTGGTGGACAAGGATGGCTGGATGCACACTGGCGACACTGCCGTGCTCGATGACCGTGGCTTTTGCAGGATC		
	1901		2000
genome	GACGGACGCATCAAGGACATGGTGATCCGAGGAGGCGAAAAATCCATCCTGTAGAGGTCGAGAACTGTCTCTTTGAGATGGACGGCGTCAAGAACGTGT		
CDS	GACGGACGCATCAAGGACATGGTGATCCGAGGAGGCGAAAAATCCATCCTGTAGAGGTCGAGAACTGTCTCTTTGAGATGGACGGCGTCAAGAACGTGT		
	2001		2100
genome	CTGTGATTGGCGTTCCGACAAGCGGTATGGCGAGCAGGTGTGTGCGTGGATCTCGACCAAGGACGGGAAGACGGTCAGTCTGGAGGCAGTGCAAAAGTT		
CDS	CTGTGATTGGCGTTCCGACAAGCGGTATGGCGAGCAGGTGTGTGCGTGGATCTCGACCAAGGACGGGAAGACGGTCAGTCTGGAGGCAGTGCAAAAGTT		
	2101		2200
genome	CTGTGAGGGCAAGATTGCGCACTACAAGGTGCCGCGGTATGTGGTTGTGGTGGAGTCCAATGAGTTCCCGACTACCCCTCGGGCAAGATCCAAAAGAAT		
CDS	CTGTGAGGGCAAGATTGCGCACTACAAGGTGCCGCGGTATGTGGTTGTGGTGGAGTCCAATGAGTTCCCGACTACCCCTCGGGCAAGATCCAAAAGAAT		
	2201	2242	
genome	GTGATGCGCGAGCTGACCAAGGCGAAGCTGCAGCTGCCTTGA		
CDS	GTGATGCGCGAGCTGACCAAGGCGAAGCTGCAGCTGCCTTGA		

Figure 15A

```

1   TGCTTTTCTCTTCTCGTCACCCCTCCTTCTTCCCATTCCTCGGGTCCTCCTCGGTTCTTAATCAGTTTCTCAGACCCTGTCCATTCTCTGGCCTCCACA
101  CACACCCCACTCTCCCTTGAACAAATACCTTATCCAGATCAAAGACATGCCTTCCCTCAAAAAGTACAACCTCGACAAGCAGAGTGTGAGGTCCTGGC
      M P S F K K Y N L D K Q S V E V P G
201  ACTCGGAAGCCTGGCGCTTCAGGCCACTACAGACATGCCGCTACGGCGATGCTCTTGTCACCAACATCCGTGAGGCCCTCATATCGAACTCTTTACG
      T R K P G A S G H Y R H A A Y G D A L V T N I R E A P H I E T L Y D
301  ACATGTGGCAGAACTCTGTGACAAAGTATGGCGGCAATGACTTTTTGGGTACCGTCCCTTCAACACTGTTGCCAGACCTATGGTGGCTACAGTTGGGA
      M W Q N S V T K Y G G N D F L G H R P F N T V A Q T Y G G Y S W E
401  GACGTACCGCCAGATTAACCAGCGGTTAATGCGTTCCGGCAGCGGTATCATGCACCTGAACGAGGTGATCCTCGGCAACCGCCAGCTTAACGCTGGGCG
      T Y R Q I N Q R V N A F G S G I M H L N E V I L G N R Q L N R W A
501  TTGGGCATCTGGTCCACGGTCGCCCTGAGTGGTTTATTACGGAGATGAGGTGCAACTGCTACAACCTCATTTCTGTTGCATTGTACGACACCCTTGGAC
      L G I W S H G R P E W F I T E M S C N C Y N L I S V A L Y D T L G P
601  CTGATGCAGTCGAGTACATTGTCAACCACGCCGAGATTGAGATTGTTGTCTCAAGTGCCAACCATATCGCCTCTTTGCTCGAGAAGCCGAGAAGCTCCC
      D A V E Y I V N H A E I E I V V S S A N H I A S L L E N A E K L P
701  CAAGCTCAAGGCCATTGTGAGCATGGATGCTCTTACGATACCGTCCCGTCCCGGGCGCCACCTCTGCCGCACAGGTTCTTCTGTCCTGGGGTGACAA
      K L K A I V S M D A L H D T V P V P G A T S A A Q V L R A W G A Q
801  AAGGGCATCAAGGTCTATGACTTTAACGAGATTGAGTCCCTCGGTGCCGAGTTCCCTCGCAAGCACCTGCCTCCACCGCTGATGAGGTGCGCTCCATCT
      K G I K V Y D F N E I E S L G A E F P R K H L P P T A D E V A S I C
901  GCTACACTTCGGGACCAACCGGTGAGCCTAAAGGAGCCATGCTCACCCACAGAACTTTGTTGCTACTGTTGGTACCAACCGCGAGGGCATGCTTCTCAC
      Y T S G T T G Q P K G A M L T H R N F V A T V G T N R E G M L L T
1001 CGAGGACGACGTTTTGATCAGTTTCTTGCCCTTGGGTCACATTATGGGACGCGTCATTGACACTTGCTCGATGTACAGCGGTGGCAAGATTGGTTACTTC
      E D D V L I S F L P L A H I M G R V I D T C S M Y S G G K I G Y F
1101 CGTGGAGATATTCTTTGCTTCTCGAGGACGTTGCTGAGCTCCGTCACATTCTTCCAGCTGTGCCTCGCCTCTTGAACCGCATTATGCCAAGCTCG
      R G D I L L L L E D V A E L R P T F F P A V P R L L N R I Y A K L V
1201 TTGCCTCTACCATTGAGGCCCGCGTTTGGTGGGTGCCTTGGCCGTCGCGGTGTCGCCGCCAAGATGGCCAACCTTGCTGCCGGAAGGGTGTAACCA
      A S T I E A P G L V G A L A R R G V A A K M A N L A A G K G V N H
1301 CGCTCTCTGGGACAGACTGCTGTTCAACAAGGTCAAGATGGCCCTGGGTGGTGGCTTCAAGTTCAGGTCATCCTGACTGGATCTGCGCCATTGCCAAGGAGTT
      A L W D R L L F N K V K M A L G G R V Q V I L T G S A P I A K E V
1401 CTCAGCTTCTTGAGAATTGCTTTGCGATGCGTGGTTTTGGAGGGATACGGCTCCACTGAGGGCATGGCTACCGCCACCATCACAATGGCTGATGAGTACA
      L S F L R I A F G C V V L E G Y G S T E G M A T A T I T M A D E Y I
1501 TTCCTGGTCACATTGGCTGCCCTCGTGGTGGATGCGAGCTCAAGCTGGTGGATGTGCCCGCGATGAACTACCTCTCTACCGACCAGCCCTACCCCGTGG
      P G H I G C P R A G C E L K L V D V P A M N Y L S T D Q P Y P R G
1601 AGAGATCTGGATCCGTGGTGACACTGTTTTCAAAGGATACTTCAAGGACGAGAAGAACAAGTGTGAGACTATCGACTCTGAAGGCTGGCTCGCTACCGGT
      E I W I R G D T V F K G Y F K D E K N T S E T I D S E G W L A T G
1701 GATATTGGATTTGTGGATAAGCGTGGATGCTTTACGATCATTGACCGCAAGAAGAACATCTTCAAGTTGGCACAAGGTGAATACATTGCTCCTGAAAAGA
      D I G F V D K R G C F T I I D R K K N I F K L A Q G E Y I A P E K I
1801 TTGAGAACGTCTTGGGCGCACGCTGCAATCTTGTCCAGCAGATCTATGTTTCATGGTGATTCGCTTGAGTCCACCTTGGTGGCAGTTCTTATCCCGAGCC
      E N V L G A R C N L V Q Q I Y V H G D S L E S T L V A V L I P E P
1901 CGAGACCTTCTGCCCTTCGCGAATGCCATTGCTGGTGCCTCCGTCACTGCTGGAGATGTTGAGGGTTTGAACAAGCTGTGCCAAGATCCCAAGGTCAAG
      E T F L P F A N A I A G A S V T A G D V E G L N K L C Q D P K V K

```

Figure 15B

2001 ATCGCGGTTCTGAAGGAGTTGGAGAAGGCCGGAAGGCCGGTGCGATGCGCGGATTCGAGTTCGTGAAGCGTGTCCTTGACCACGGATGCATTCTCGG
I A V L K E L E K A G K A G A M R G F E F V K R V H L T T D A F S V

2101 TCGACAACGGCATGATGACACCTACCTTCAAGGTCGGTCGCCCACAAGTAGCCGAGCATTTCAAGGAGCAAATCACGGCCATGTATAAGGAGATCAATGC
D N G M M T P T F K V R R P Q V A E H F R E Q I T A M Y K E I N A

2201 CTCGACCCCTGTTGCCAAGCTGTAGATAGAAAACCTTTGCCCTTATTACCCCTTGAATAGAAGGTGACACGTTGTTTGATTACACAAAAAAAAAAAA
S T P V A K L

2301 AAAAAA

Figure 16A

	1	100
genome	ATGCCTTCCTTCAAAAAGTACAACCTCGACAAGCAGAGTGTGAGGTCCCTGGCACTCGGAAGCCTGGCGCTTCAGGCAAGTTGGATAGTGGCTCATGAG	
CDS	ATGCCTTCCTTCAAAAAGTACAACCTCGACAAGCAGAGTGTGAGGTCCCTGGCACTCGGAAGCCTGGCGCTTCAG-----	
	101	200
genome	ATCCAAGTGTGGGCCACGCTACCATGAACTAGTCGCTAATGCAGATTCTCTATTGCCITTCACCTTCTCAGAGGCCACTACAGACATGCCGCCTACGG	
CDS	-----GCCACTACAGACATGCCGCCTACGG	
	201	300
genome	CGATGCTCTTGTCAACCAACATCCGTGAGGCCCTCATATCGAAACTCTTTACGACATGTGGCAGAACTGTAAGTTGATTCCCGAGGACCTGGGAAATCTT	
CDS	CGATGCTCTTGTCAACCAACATCCGTGAGGCCCTCATATCGAAACTCTTTACGACATGTGGCAGAACT-----	
	301	400
genome	TACTCGGATCTGTCTACAAGAAGACTCAGACTCTGCTCTCCACAATTACCTTGTAACCTTCAGCTGTGACAAAGTATGGCGGCAATGACTTTTGGGTC	
CDS	-----CTGTGACAAAGTATGGCGGCAATGACTTTTGGGTC	
	401	500
genome	ACCGTCCCTTCAACACTGTTGCCAGACCTATGGTGGCTACAGTTGGGAGACGTACCGCCAGATTAACCAGCGGTTAATGCGTTCGGCAGCGGTATCAT	
CDS	ACCGTCCCTTCAACACTGTTGCCAGACCTATGGTGGCTACAGTTGGGAGACGTACCGCCAGATTAACCAGCGGTTAATGCGTTCGGCAGCGGTATCAT	
	501	600
genome	GCACCTGAACGAGGTGATCCTCGGCAACCGCCAGCTTAACCGCTGGGCGTTGGGCATCTGGTCCACGGTCGCCCTGAGTGGTTTCATTACGGAGATGAGC	
CDS	GCACCTGAACGAGGTGATCCTCGGCAACCGCCAGCTTAACCGCTGGGCGTTGGGCATCTGGTCCACGGTCGCCCTGAGTGGTTTCATTACGGAGATGAGC	
	601	700
genome	TGCAACTGCTACAACCTCATTTCTGTTGCATTGTACGACACCCCTGGACCTGATGCAGTCGAGTACATTGTCAACCACGCCGAGATTGAGATTGTTGTCT	
CDS	TGCAACTGCTACAACCTCATTTCTGTTGCATTGTACGACACCCCTGGACCTGATGCAGTCGAGTACATTGTCAACCACGCCGAGATTGAGATTGTTGTCT	
	701	800
genome	CAAGTGGTAAGTCGCTTTTTTTTTTTTGTATTGGCGCCTGCTTGCTTGCTGCACGTGAAACTATAATGGTTTCTGATCTCCCTTTTGTACCTCCT	
CDS	CAAGTG-----	
	801	900
genome	CGTTTCGATCAACTTCTTTGTAATTTTATCCCGGGCATTGACTCTATATGCAGCCAAACCATATCGCCTCTTTGCTCGAGAAGCCGAGAAGCTCCCCAA	
CDS	-----CCAACCATATCGCCTCTTTGCTCGAGAAGCCGAGAAGCTCCCCAA	
	901	1000
genome	GCTCAAGGCCATTGTGAGCATGGATGCTCTTCAGGATACCGTCCCGTCCCGGGCGCCACCTCTGCCGCACAGGTTCTTCGTGCCTGGGGTGACAAAAAG	
CDS	GCTCAAGGCCATTGTGAGCATGGATGCTCTTCAGGATACCGTCCCGTCCCGGGCGCCACCTCTGCCGCACAGGTTCTTCGTGCCTGGGGTGACAAAAAG	
	1001	1100
genome	GGCATCAAGGTCTATGACTTTAACGAGATTGAGTCCCTCGGTGCCGAGTTCCCTCGCAAGCACCTGCCTCCACCGCTGATGAGGTGCCTCCATCTGCT	
CDS	GGCATCAAGGTCTATGACTTTAACGAGATTGAGTCCCTCGGTGCCGAGTTCCCTCGCAAGCACCTGCCTCCACCGCTGATGAGGTGCCTCCATCTGCT	
	1101	1200
genome	ACACTTCGGGCACACCGGTGAGCCTGTAAGTGTGTTTTCCCTTCACTGACGATGTGGCTGAGGATGCATTTTGCTGATTGAACTTCCCATCTAACAT	
CDS	ACACTTCGGGCACACCGGTGAGCCT-----	
	1201	1300
genome	CGCTTGATCTATAGAAAGGAGCCATGCTCACCACAGAACTTTGTTGCTACTGTTGGTACCAACCGCGAGGGCATGCTTCTCACCAGGACGACGTTTTT	
CDS	-----AAAGGAGCCATGCTCACCACAGAACTTTGTTGCTACTGTTGGTACCAACCGCGAGGGCATGCTTCTCACCAGGACGACGTTTTT	
	1301	1400
genome	GATCAGGTATAACTCATCTTGAAGCTGATCAATGACACGGATGTACAACGGACGATGGAGCGAGATCTACAATGCGGAATGGCTCACCTCCGGTTTAC	
CDS	GATCAG-----	

Figure 16B

	1401	1500
genome	AATTACCACACTTCTAGTTTCTTGCCCTTGGCTCACATTATGGGACGCGTCATTGACACTTGCTCGATGTACAGCGGTGGCAAGATTGTTACTTCCGTG	
CDS	-----TTTCTTGCCCTTGGCTCACATTATGGGACGCGTCATTGACACTTGCTCGATGTACAGCGGTGGCAAGATTGTTACTTCCGTG	
	1501	1600
genome	GAGATATTCTTTTCTTCTCGAGGACGTTGCTGAGCTCCGTCCCACATTCTTCCAGCTGTGCCTCGCCTCTTGAACCGCATTATGCCAAGCTCGTTGC	
CDS	GAGATATTCTTTTCTTCTCGAGGACGTTGCTGAGCTCCGTCCCACATTCTTCCAGCTGTGCCTCGCCTCTTGAACCGCATTATGCCAAGCTCGTTGC	
	1601	1700
genome	CTCTACCATTGAGGCCCGCGTTTGGTCGGTGCCTTGGCCCGTCGCGGTGTCGCCGCCAAGATGGCCAACCTTGCTGCCGAAAGGGTGTCAACCACGCT	
CDS	CTCTACCATTGAGGCCCGCGTTTGGTCGGTGCCTTGGCCCGTCGCGGTGTCGCCGCCAAGATGGCCAACCTTGCTGCCGAAAGGGTGTCAACCACGCT	
	1701	1800
genome	CTCTGGGACAGACTGCTGTTCAACAAGGTCAAGATGGCCCTGGGTGGTCGCGTTCAGGTCATCCTGACTGGATCTGCGCCCATTGCCAAGGAGGTTCTCA	
CDS	CTCTGGGACAGACTGCTGTTCAACAAGGTCAAGATGGCCCTGGGTGGTCGCGTTCAGGTCATCCTGACTGGATCTGCGCCCATTGCCAAGGAGGTTCTCA	
	1801	1900
genome	GCTTCTTGAGAATTGCTTTGGATGCGTGGTTTTGGAGGGATACGGCTCCACTGAGGGCATGGCTACCGCCACCATCACAAATGGCTGAGTACGTGAACCT	
CDS	GCTTCTTGAGAATTGCTTTGGATGCGTGGTTTTGGAGGGATACGGCTCCACTGAGGGCATGGCTACCGCCACCATCACAAATGGCTGA-----	
	1901	2000
genome	AGTTATTTTATTGAAATGTCGTGGAGCCTGTCGACTGTAGCATTTCATCTAACCATTAGTAATCTTTTCAAAAATAGTGAGTACATTCTGCTGACATT	
CDS	-----TGAGTACATTCTGCTGACATT	
	2001	2100
genome	GGCTGCCCTCGTGCTGGATGCGAGCTCAAGCTGGTGGATGTGCCCGCATGAACCTACCTCTCTACCGACCAGCCCTACCCCGTGGAGAGATCTGGATCC	
CDS	GGCTGCCCTCGTGCTGGATGCGAGCTCAAGCTGGTGGATGTGCCCGCATGAACCTACCTCTCTACCGACCAGCCCTACCCCGTGGAGAGATCTGGATCC	
	2101	2200
genome	GTGGTGACACTGTTTTCAAAGGATACTTCAAGGACGAGAAGAACAAGTAGTGAGACTATCGACTCTGAAGGCTGGCTCGCTACCGGTGATATTGGATTTGT	
CDS	GTGGTGACACTGTTTTCAAAGGATACTTCAAGGACGAGAAGAACAAGTAGTGAGACTATCGACTCTGAAGGCTGGCTCGCTACCGGTGATATTGGATTTGT	
	2201	2300
genome	GGATAAGCGTGGATGCTTTACGATCATTGACCGCAAGAAGAACATCTTCAAGGTATGAGCAAAAAGTGGGATTGATCTGATCGTTTCTTCTCGTTTTC	
CDS	GGATAAGCGTGGATGCTTTACGATCATTGACCGCAAGAAGAACATCTTCAAG-----	
	2301	2400
genome	CCGTTAAGGAATCCGCTCATACTAACGTTCTCGTCAATGGTTGCATGTATTATAGTTGGCACAAGGTGAATACATTGCTCCTGAAAAGATTGAGAACG	
CDS	-----TTGGCACAAGGTGAATACATTGCTCCTGAAAAGATTGAGAACG	
	2401	2500
genome	TCTTGGGCGCACGCTGCAATCTTGTCCAGCAGATCTATGTTTCATGGTGATTGCTTGGTCCACCTTGGTCGCAGTTCTTATTCCCGAGCCCGAGACCTT	
CDS	TCTTGGGCGCACGCTGCAATCTTGTCCAGCAGATCTATGTTTCATGGTGATTGCTTGGTCCACCTTGGTCGCAGTTCTTATTCCCGAGCCCGAGACCTT	
	2501	2600
genome	CCTGCCCTTCGCGAATGCCATTGCTGGTGCCTCCGTCACTGCTGGAGATGTTGAGGGTTTGAACAAGCTGTGCCAAGATCCCAAGGTCAAGATCGCGGTT	
CDS	CCTGCCCTTCGCGAATGCCATTGCTGGTGCCTCCGTCACTGCTGGAGATGTTGAGGGTTTGAACAAGCTGTGCCAAGATCCCAAGGTCAAGATCGCGGTT	
	2601	2700
genome	CTGAAGGAGTTGGAGAAGGCCGGAAGGCCGGTGCATGCGCGGATTGAGTTCGTGAAGCGTGTCCACTTGACCACGGATGCATTCTCGGTGACAACG	
CDS	CTGAAGGAGTTGGAGAAGGCCGGAAGGCCGGTGCATGCGCGGATTGAGTTCGTGAAGCGTGTCCACTTGACCACGGATGCATTCTCGGTGACAACG	
	2701	2800
genome	GCATGATGACACCTACCTTCAAGGTCCGTGCCCCACAAGTAGCCGAGCATTTAGGGAGCAAATCAGGCCATGTATAAGGAGATCAATGCCTCGACCCC	
CDS	GCATGATGACACCTACCTTCAAGGTCCGTGCCCCACAAGTAGCCGAGCATTTAGGGAGCAAATCAGGCCATGTATAAGGAGATCAATGCCTCGACCCC	

Figure 16C

	2801	2816
genome	TGTTGCCAAGCTGTAG	
CDS	TGTTGCCAAGCTGTAG	

CA 02787832 2012-07-20

Figure 18A

	1	100
genome	ATGGTTGGTGAGTAAACAGCGATCGCCCTGGGTCCAATCGACTCATCTGCTAATCTGATCTCACACGTCCTTATCAATTACAAAAAGAAAAAAGAGA	
CDS	ATGGTT-----	
	101	200
genome	CAAGAGAGAGAATTACTAACATTGCTCTCTCCCTCGTGTGATGCAGCTCTCCCACTCGTGGCAGCAGCTGTCCAGCTGCCATGTATGTGAGCTCAAAG	
CDS	-----GCTCTCCCACTCGTGGCAGCAGCTGTCCAGCTGCCATGTATGTGAGCTCAAAG	
	201	300
genome	CTGGCACTTCCTCGGGATATGAAGTTGATTAAGAGCTTGATCGGAGCCAAGATGGCCTACAGTGCCATGGAAAAGAACGACGCCCTCAACCTGACACTCC	
CDS	CTGGCACTTCCTCGGGATATGAAGTTGATTAAGAGCTTGATCGGAGCCAAGATGGCCTACAGTGCCATGGAAAAGAACGACGCCCTCAACCTGACACTCC	
	301	400
genome	GCTTCGACGAGTGCTACCGCAAGTATCCTGACCGTGAAGCCCTGGTCTTTGAGGGCAAATCCTATTCATTCCGTGATATTCAGCTTGGTAAGCCATTGTT	
CDS	GCTTCGACGAGTGCTACCGCAAGTATCCTGACCGTGAAGCCCTGGTCTTTGAGGGCAAATCCTATTCATTCCGTGATATTCAGCTTG-----	
	401	500
genome	ACACGGCACTCACCCACGCTTCCTGCTTTCAGAGCTCAGACTGCTCGAACCTTCAATTTTTGTATCGATCGATACCGATCGATCATGCGCCACAT	
CDS	-----	
	501	600
genome	TGCCTAGAGCTATGGCACAGCGCTAATGCATGCCTTCTTGTATGAAGCCTCCAACAGGTGCGGCAACTGGTTGCTGGCCAAAGGGATCAAGCGAGGAG	
CDS	-----CCTCCAACAGGTGCGGCAACTGGTTGCTGGCCAAAGGGATCAAGCGAGGAG	
	601	700
genome	ATATCGTCTCGCTTTTCATGTTGAACAGGCCAGAGTTCATCTTCTGCTGGGTGGGTCTCAACAAGATTGGAGCCACTGGTGCCTTCATCAATACCAACCT	
CDS	ATATCGTCTCGCTTTTCATGTTGAACAGGCCAGAGTTCATCTTCTGCTGGGTGGGTCTCAACAAGATTGGAGCCACTGGTGCCTTCATCAATACCAACCT	
	701	800
genome	TACGGGCAAACCCCTGACACATTCCTCCGGACAGCCACGTCGTCAATGTTGATCATGGACACGGAGTTGACAGACGGATCGCCAACTCCCTGGATGAG	
CDS	TACGGGCAAACCCCTGACACATTCCTCCGGACAGCCACGTCGTCAATGTTGATCATGGACACGGAGTTGACAGACGGATCGCCAACTCCCTGGATGAG	
	801	900
genome	ATTGAGGAGATGGGCTATTCAATTTACTCTTACGGACCCGAAGCCGTGGACTTTGCTACCCCGATGGATATCTCGCAGGTCCAGACACCGATACACCCG	
CDS	ATTGAGGAGATGGGCTATTCAATTTACTCTTACGGACCCGAAGCCGTGGACTTTGCTACCCCGATGGATATCTCGCAGGTCCAGACACCGATACACCCG	
	901	1000
genome	AACACCTGCGCCGGAACAGACCGCGGATGACATTGCGATGCTCATCTACACCTCTGGAAGTCTGGTCTTCCCAAGGCCGGTCGTGTCTCTCATGCGCG	
CDS	AACACCTGCGCCGGAACAGACCGCGGATGACATTGCGATGCTCATCTACACCTCTGGAAGTCTGGTCTTCCCAAGGCCGGTCGTGTCTCTCATGCGCG	
	1001	1100
genome	TGCCTCTAGTAAGTTGAGAGTCTTCAGCCTTTGACATACGATATTTTTGAGCGTGCTACTAACAGTTCTCGTTGCCGTTATCTGCATATTTTTAGTGGGA	
CDS	TGCCTCTA-----TGGGA	
	1101	1200
genome	CCTCAGTTTTGGAACCGATTCTATCACTTCAGTGAGAGCGACAGGGTCTATCTGTCTTGCCCTTGTAACACAGTGCTGGCGCCATCTTGGGAGTGATTG	
CDS	CCTCAGTTTTGGAACCGATTCTATCACTTCAGTGAGAGCGACAGGGTCTATCTGTCTTGCCCTTGTAACACAGTGCTGGCGCCATCTTGGGAGTGATTG	
	1201	1300
genome	CTTGTTGGACCTCGGGAGCAACCTTGATCCTGGCCCGCAAGTTCTCCGCGACACATTTCTGGGAGGATTGCCGCGTGAACAACGCAACTGTGATTCAATA	
CDS	CTTGTTGGACCTCGGGAGCAACCTTGATCCTGGCCCGCAAGTTCTCCGCGACACATTTCTGGGAGGATTGCCGCGTGAACAACGCAACTGTGATTCAATA	
	1301	1400
genome	CATTGGAGAAATTTGCAGATATCTGCTCAACACGCCAGAATCACCCCTGGACAAGGCACACTCGATACGACTGGCACATGGTAATGGAATGCGACCCGAT	
CDS	CATTGGAGAAATTTGCAGATATCTGCTCAACACGCCAGAATCACCCCTGGACAAGGCACACTCGATACGACTGGCACATGGTAATGGAATGCGACCCGAT	

Figure 18B

	1401	1500
genome	GTCTGGACTCGCTTCAGAGATCGGTTCCGCATCCCGTTGATTGGCGAGTGGTATGCATCGACTGAGGGAAGTGGAGCCTTGTCGAATTATAACACAGGCC	
CDS	GTCTGGACTCGCTTCAGAGATCGGTTCCGCATCCCGTTGATTGGCGAGTGGTATGCATCGACTGAGGGAAGTGGAGCCTTGTCGAATTATAACACAGGCC	
	1501	1600
genome	CAGCGGGCGCTGGAGCGATTGGATACCGCGGTACCCCTGCCAGAGCATTGGATAAAGGACTCAGGATTGCGAGATTTGATGTCCAGACAGAGGAGTTGGT	
CDS	CAGCGGGCGCTGGAGCGATTGGATACCGCGGTACCCCTGCCAGAGCATTGGATAAAGGACTCAGGATTGCGAGATTTGATGTCCAGACAGAGGAGTTGGT	
	1601	1700
genome	TCGGGACAAAAACGGTTATTGCATTGAGGTAAACATAAGGCGCATTGTTGAAGTCTAAATCACTTCAGATGCTTTGTTCCGGCATGCTTACCAAACGCAC	
CDS	TCGGGACAAAAACGGTTATTGCATTGAG-----	
	1701	1800
genome	CGACCATCTTTTCACTGGTGCGGCATATAATAGTGCAAACCTGGCGAGCCCGGAGAATTGCTGACGCTTGTTGATGCTAAAGAGCCGAACAAAGACTTCA	
CDS	-----TGCAAACCTGGCGAGCCCGGAGAATTGCTGACGCTTGTTGATGCTAAAGAGCCGAACAAAGACTTCA	
	1801	1900
genome	AAGGATACCATCAAAACCAGGCAGCGACCAACAAAAAGATTGTCAAAGATGTTTTCAAAGCCGGCGACATGTACTTCGGTACCGGAGATATCCTTCGGCG	
CDS	AAGGATACCATCAAAACCAGGCAGCGACCAACAAAAAGATTGTCAAAGATGTTTTCAAAGCCGGCGACATGTACTTCGGTACCGGAGATATCCTTCGGCG	
	1901	2000
genome	CGATAGCGATGGGTACTTTTACTTTGGCGACCGTGTGGGCGATACATTCCGGTGGAAGTCCGAGAATGTGTCTACGGCCGAGGTGTCTGAAGTCCTCTCG	
CDS	CGATAGCGATGGGTACTTTTACTTTGGCGACCGTGTGGGCGATACATTCCGGTGGAAGTCCGAGAATGTGTCTACGGCCGAGGTGTCTGAAGTCCTCTCG	
	2001	2100
genome	CAGTATCCGGACTGTATCGAAGTCAATGTGTATGGAGTTCAGATCCAGGGCAGGACGGACGCGCCGGTATGGCAGCGATTGTGTCCAAGAGCACGATGG	
CDS	CAGTATCCGGACTGTATCGAAGTCAATGTGTATGGAGTTCAGATCCAGGGCAGGACGGACGCGCCGGTATGGCAGCGATTGTGTCCAAGAGCACGATGG	
	2101	2200
genome	ATTGGGAGAAATTTGCGGCGTATGCACTCAAGAACCTGCCGCGGTATTCTGTTCCGATCTTTATCCGCAAGATGCCGAGATGGAGATCACAGGGACGTT	
CDS	ATTGGGAGAAATTTGCGGCGTATGCACTCAAGAACCTGCCGCGGTATTCTGTTCCGATCTTTATCCGCAAGATGCCGAGATGGAGATCACAGGGACGTT	
	2201	2300
genome	CAAGCAGCGCAAAGTCGAGTTGGTGAATGAGGGAATCGACCCCAAGACGATTGCCAACGAGATGCTGTGGTTGGACGGACACCACTATAAGCCGTTCAAG	
CDS	CAAGCAGCGCAAAGTCGAGTTGGTGAATGAGGGAATCGACCCCAAGACGATTGCCAACGAGATGCTGTGGTTGGACGGACACCACTATAAGCCGTTCAAG	
	2301	2345
genome	GCGGCCGAGCACCAGCGGTCATCAGCGGCAAGGCCAAGCTATAG	
CDS	GCGGCCGAGCACCAGCGGTCATCAGCGGCAAGGCCAAGCTATAG	

Figure 19A

```

1  ATGGAAACCTTGGTTAACGGAAAGTATGCGGTGAGTACGACGAGGTCGATCACATCTATCGCAACGTCATGGCTACAGGCGGGCTCCTCGACAGGCCTA
   M E T L V N G K Y A V E Y D E V D H I Y R N V M A T G G L L D R P M
101  TGCCTCCATACTACGACATCAAGGAGCGCACCATGGCCACCTCTTTGAGTATATGGCCAACACCTACGAAGACAAAGACGCCATGGGCTGGCGAGACAT
   P P Y Y D I K E R T M A H L F E Y M A N T Y E D K D A M G W R D I
201  TATCAAGGTCCACAAGGTCGAGAAGCAGGCTGCCAATCCTGGCGAGAAGCCAAAGACCTGGATCACTTATGAGCTCTCGGACTACAACCTGGATGTCGTAC
   I K V H K V E K Q A A N P G E K P K T W I T Y E L S D Y N W M S Y
301  CGCCAAGCCAAGAAGTATGAGATCGAGTTGGCTTGGGCATCACACGCCCTTGGAGTTGAGAAGGGGAGACTTTGTCATGATCTTTGCTAGCACATGTCGG
   R Q A K N Y A D R V G L G I T R L G V E K G D F V M I F A S T C P E
401  AATGGTTCCTGACAGCGCATGGATGCTTCTCGCAGTCAGTGACTATCGTGACAGCCTACGACTCGATGGAGGAGAAGTCGATCCAGTTTATTGTTGACCA
   W F L T A H G C F S Q S V T I V T A Y D S M D E K S I Q F I V D Q
501  GTCCAGGCCAAGGCCATCTTTGCTGATGCGCACACGCTCCCTGTGGTGCCAAACTCATGCAGAAGGGCAACAGTGGTGTCAAGGCAGTCATTTACACA
   S Q P K A I F A D A H T L P V V S K L M Q K G N S G V K A V I Y T
601  GGCCAAGAGTGGGAAGTGACCGATGCAATCAAGAAGATGGAGCAAGTAGAAAACCGCTCATTTGAGCTGGTTCATATCGACGAAGTCAAGAAGACCAAGT
   G Q E W E V T D A I K K M E Q V E N R S F E L V H I D E L K K T K S
701  CAGCATCTAACGGCGAACAGTCTGCCGAAAGGGGAAGCAGAGATCATCTGAGGATGCCGAAGGGCGCTCAGGACGAGATCGAGGTATATACCCCTAAGGC
   A S N G E Q S A G K G K Q R S S E D A E G A Q D E I E V I Y P K A
801  GGATGATCTGGCCTGTATTATGTATACCTCTGGGTGACGGGTGAGCCCAAGGGCGCGCAATTGACACATGGCAACTTGATGGCGGCCATTGGAAGTGCT
   D D L A C I M Y T S G S T G Q P K G A Q L T H G N L M A A I G S A
901  GCGGCCATGGAGGGCGACAGCTGGACAAGGAACAGACATTGTTATTTTCATATCTGCCATTGGCCCATGTCTCGAGTTTGTCAATTTCCCACTTTGTGG
   A A M E G D Q L D K E T D I V I S Y L P L A H V L E F V I S H F V V
1001  TATCCATGGGCTGCCGCTTTGATTGCGGACGAGCAGCACTCTGATGGATGATGCAGTCGCTCCACCGCAGGAAGTGGCAGGTCCAAGGGCCTTGGTGA
   S M G C R L G F G R A R T L M D D A V A P T A G S G R S K G L G D
1101  TCTGAAGGCGCTCCAGCCAACATTGATGGCTGGTGTGCCAAGCATCTGGGAGCGTATCCGCAAGGGCATCCTGGCCGAGGTCAACAAGCAATCCTCCCT
   L K A L Q P T L M A G V P T I W E R I R K G I L A E V N K Q S F P
1201  ATCCGTACACTCTTCTTTGCTGCACTCAACACCAAGTGGGCTATCGTCCAGGCTACCGGATCTGAGAACTTTGTACCAAGACTATTGACTCGTTGGTCT
   I R T L F F A A L N T K W A I V Q A T G S E N F V T K T I D S L V F
1301  TTAGTAAGGCTAAGGAGCTCGTTGGAGGCAAGCTGCGCCTTACCTTGACTGGAGGGGCGGAATCAGTGATGAGACGCACCGGTTCTTGAGCATGGTAAT
   S K A K E L V G G K L R L T L T G G A G I S D E T H R F L S M V M
1401  GTGCTACGTTATCTCGGGATATGGTCTCACTGAAGTCTGTGGTGTGCCGCTGTCAACCTGCCACGTATGGGTACCGTCTCAGGACCGTTGGACCAACC
   C Y V I S G Y G L T E V C G V A A V T L P R M G H R L R T V G P P
1501  GCGCCAGTCTTGAGCTGAAGTTGGTGAATGTGCCGACACCGAGTACACAGGAGACAATGGATCGGGCGAAATCTGGTTCCGTGGACCTGCAGTGATGA
   A P S L E L K L V N V P D T E Y T G D N G S G E I W F R G P A V M K
1601  AGGGATACTTCAAACCTCGAGGAAGAGACCAAGAAGGTGATGACCGGGGATGGTTGGTTCAAGACAGGCGACATTGGCACGATGAACCCAGACGCGCACT
   G Y F K L E E E T K K V M T G D G W F K T G D I G T M N P D G T L
1701  GTCAATCAAGGACAGGGTCAAGAATCTGGTCAAGCTGTCTCATGGAGAATATGTGCCCTGGAGAAATGTGAAGCCGTTTATCGCGATTCCAAGGAGATC
   S I K D R V K N L V K L S H G E Y V A L E K C E A V Y R D S K E I
1801  AAGAGCATTTCATCGTTGCGGACAATGGGTGCCCTGTGTTGCTGGCCGTTGTGGAACCGAGCCACGAGGGGCGTCTGACAAGGAGATTTTGGATATCC
   K S I C I V A D N G C P V L L A V V E P S H A G A S D K E I L D I L

```

Figure 19B

```
1901 TGAAGAGCCAAGCCAAGGCGGCGGGCCTCTCCAAGTCCGAGACTGTGCAAGGCGTTATCATTGATGATTCGGACTGGATGACGAATGGGTTTCATGACCTC
    K S Q A K A A G L S K S E T V Q G V I I D D S D W M T N G F M T S

2001 GAGCAGCAAGGTCAAGAGACGCGAGGTCCGCAAGGCACACAACAAGGATATTGAGGAGATGTGGAAGAAGTTCTAG
    S S K V K R R E V R K A H N K D I E E M W K K F *
```

Figure 20A

	1	100
genome	ATGGAACCTTGGTTAACGGAAAGTATGCGGTCGAGTACGACGAGGTCGATCACATCTATCGCAACGTCATGGCTACAGGCGGGCTCCTCGACAGGCCTA	
CDS	ATGGAACCTTGGTTAACGGAAAGTATGCGGTCGAGTACGACGAGGTCGATCACATCTATCGCAACGTCATGGCTACAGGCGGGCTCCTCGACAGGCCTA	
	101	200
genome	TGCCTCCATACTACGACATCAAGGAGCGCACCATGGCCACCTCTTTGAGTATATGGCCAACACCTACGAAGACAAAGACGCCATGGGCTGGCGAGACAT	
CDS	TGCCTCCATACTACGACATCAAGGAGCGCACCATGGCCACCTCTTTGAGTATATGGCCAACACCTACGAAGACAAAGACGCCATGGGCTGGCGAGACAT	
	201	300
genome	TATCAAGGTATTGACTGCCCGGCCATGACTTTTTACCCACACGATCCTCCCTTTTTTCTCTCCATTCTTCTATCCTGACCGTATCGCTATCGAA	
CDS	TATCAAGGT-----	
	301	400
genome	CGAGTCAACGAGATCAGTTCCACGCTTACTTTACTCCCTCGTCTGATTCTGATTATTTCTCTCTGCTCTTCGTGATTTGTTCCGAGGTCCACAAG	
CDS	-----CCACAAG	
	401	500
genome	GTCGAGAAGCAGGCTGCCAATCCTGGCGAGAAGCCAAAGACCTGGATCACTTATGAGCTCTCGGACTACAACCTGGATGTCGTACCGCCAAGCCAAGAACT	
CDS	GTCGAGAAGCAGGCTGCCAATCCTGGCGAGAAGCCAAAGACCTGGATCACTTATGAGCTCTCGGACTACAACCTGGATGTCGTACCGCCAAGCCAAGAACT	
	501	600
genome	ATGCAGATCGAGTTGGCTTGGGCATCACACGCCCTTGGAGTTGAGAAGGGAGACTTTGTCATGATCTTTGCTAGCACATGGTATAACCTCATGCAGACAAC	
CDS	ATGCAGATCGAGTTGGCTTGGGCATCACACGCCCTTGGAGTTGAGAAGGGAGACTTTGTCATGATCTTTGCTAGCACATG-----	
	601	700
genome	TTTCGCATCCATGACAGCATCGCAAGAAAAAAAAAAGAGAAGGAAATTATGATAATTCGGGCCAAGCGACTAAACACCGCTATCGCACGCTTTTTTTTT	
CDS	-----	
	701	800
genome	TTTTTTCGTATCTTGCAATTTCTTATCGTTCAAATAGTCCCGAATGGTTCCTGACAGCGCATGGTAAGCTTTTTCTTTTGGGATGATGATTTCTTCTT	
CDS	-----TCCCGAATGGTTCCTGACAGCGCAT-----	
	801	900
genome	ATGACAGCATGATGAGATTCAACCGGGATTAGGTGACGGCTGTTGATTGTGCAAAGGGGGGCGGACAATTTGAAGAGACTTGGGGAGGTGTTTGTGAT	
CDS	-----	
	901	1000
genome	GACGACAAATCCAAGATTCAACAGAAGACTGGTGCCAGGGAAGAGTAGAGAGGGTGCCTCGTTCTTTGGCATATGAAAGTGGAATTACGATATGGCTGAGT	
CDS	-----	
	1001	1100
genome	TAGGAGTCTAATTCATCGCGGAAAAAGAGTCGCGACTGAAGTCCCAAGTCGGGGACTGTAGAAAGTATTCCACTCGTGTCTGGTGAAATGAGGAGGACTG	
CDS	-----	
	1101	1200
genome	GGTTGGGGTTTGGAGTCCGACGAGAAATCATGGAATACGCTTCTCGGTCTTCAGTCCACTGATCACTCATGGCGCAATTGACTCTACAATAATAGGATG	
CDS	-----GGATG	
	1201	1300
genome	CTTCTCGCAGTCAGTGACTATCGTGACAGCCTACGACTCGATGGACGAGAAGTCGATCCAGTTTATTGTTGACCAGTCCCAGCCCAAGGCCATCTTTGCT	
CDS	CTTCTCGCAGTCAGTGACTATCGTGACAGCCTACGACTCGATGGACGAGAAGTCGATCCAGTTTATTGTTGACCAGTCCCAGCCCAAGGCCATCTTTGCT	
	1301	1400
genome	GATGCGCACACGCTCCCTGTGGTGTCCAAACTCATGCAGAAGGGCAACAGTGGTGTCAAGGCAGTCATTTACACAGGCCAAGAGTGGGAAGTGACCGATG	
CDS	GATGCGCACACGCTCCCTGTGGTGTCCAAACTCATGCAGAAGGGCAACAGTGGTGTCAAGGCAGTCATTTACACAGGCCAAGAGTGGGAAGTGACCGATG	

Figure 20B

	1401	1500
genome	CAATCAAGAAGATGGAGCAAGTAGAAAACCGCTCATTTGAGCTGGTTCATATCGACGAACTCAAGAAGACCAAGTCAGCATCTAACGGCGAACAGTCTGC	
CDS	CAATCAAGAAGATGGAGCAAGTAGAAAACCGCTCATTTGAGCTGGTTCATATCGACGAACTCAAGAAGACCAAGTCAGCATCTAACGGCGAACAGTCTGC	
	1501	1600
genome	CGGAAAGGGGAAGCAGAGATCATCTGAGGATGCCGAAGGCGCTCAGGACGAGATCGAGGTCATATACCCTAAGGCGGATGATCTGGCCTGTATTATGTAT	
CDS	CGGAAAGGGGAAGCAGAGATCATCTGAGGATGCCGAAGGCGCTCAGGACGAGATCGAGGTCATATACCCTAAGGCGGATGATCTGGCCTGTATTATGTAT	
	1601	1700
genome	ACCTCTGGGTCGACGGGTCAGCCCAAGGGCGCGCAATTGACACATGGCAACTTGATGGCGGCCATTGGAAGTGCTGCGGCCATGGAGGGCGACCAAGCTGG	
CDS	ACCTCTGGGTCGACGGGTCAGCCCAAGGGCGCGCAATTGACACATGGCAACTTGATGGCGGCCATTGGAAGTGCTGCGGCCATGGAGGGCGACCAAGCTGG	
	1701	1800
genome	ACAAGGAAACAGACATTGTTATTTTCATATCTGCCATTGGCCCATGTCCTCGAGTTTGTCAATTTCCCACTTTGTGGTATCCATGGTAAGTCGAACATCCCT	
CDS	ACAAGGAAACAGACATTGTTATTTTCATATCTGCCATTGGCCCATGTCCTCGAGTTTGTCAATTTCCCACTTTGTGGTATCCATGG-----	
	1801	1900
genome	TTACATCTGCATCCCAAATGCGGCTAAAGTCAAGTTGTTGACCTGAACGTTTATATTCATCTTAGGGCTGCCGTCTTGGATTCCGACGAGCAGCGACTC	
CDS	-----GCTGCCGTCTTGGATTCCGACGAGCAGCGACTC	
	1901	2000
genome	TGATGGATGATGCAGTCGCTCCACCGCAGGAAGTGGCAGGTCCAAGGGCCTTGGTGATCTGAAGGCGCTCCAGCCAACATTGATGGGTATGATTGGCAT	
CDS	TGATGGATGATGCAGTCGCTCCACCGCAGGAAGTGGCAGGTCCAAGGGCCTTGGTGATCTGAAGGCGCTCCAGCCAACATTGATGG-----	
	2001	2100
genome	GAAGCCAGAGACTAATATGACGTAGAGTGGGACAAAACCTTATTTTGACGTATTGCATATTGTGTCGATTTTCAAAAACGATAGCTGGTGTGCCAACGAT	
CDS	-----CTGGTGTGCCAACGAT	
	2101	2200
genome	CTGGGAGCGTATCCGCAAGGGCATCCTGGCCGAGGTCAACAAGCAATCCTTCCCTATCCGTACACTCTTCTTTGCTGCACTCAACACCAAGTGGGCTATC	
CDS	CTGGGAGCGTATCCGCAAGGGCATCCTGGCCGAGGTCAACAAGCAATCCTTCCCTATCCGTACACTCTTCTTTGCTGCACTCAACACCAAGTGGGCTATC	
	2201	2300
genome	GTCCAGGCTACCGGATCTGAGAACTTTGTACCAAGACTATTGACTCGTTGGTCTTTAGTAAGGCTAAGGAGCTCGTTGGAGGCAAGCTGCGCCTTACCT	
CDS	GTCCAGGCTACCGGATCTGAGAACTTTGTACCAAGACTATTGACTCGTTGGTCTTTAGTAAGGCTAAGGAGCTCGTTGGAGGCAAGCTGCGCCTTACCT	
	2301	2400
genome	TGACTGGAGGGGCCGGAATCAGTGATGAGACGCACCGGTTCTTGAGCATGGTAATGTGCTACGTTATCTCGGGATATGGTCTCACTGAAGTCTGTGGTGT	
CDS	TGACTGGAGGGGCCGGAATCAGTGATGAGACGCACCGGTTCTTGAGCATGGTAATGTGCTACGTTATCTCGGGATATGGTCTCACTGAAGTCTGTGGTGT	
	2401	2500
genome	TGCCGCTGTACCCCTGCCACGTATGGGTACCGTCTCAGGACCGTTGGACCACCCGTAAGTCCGCTCGCATATCTTCTCGCGATCTGAATATGCGAGCTG	
CDS	TGCCGCTGTACCCCTGCCACGTATGGGTACCGTCTCAGGACCGTTGGACCACCCG-----	
	2501	2600
genome	TTTTTTTTTGTCTTTCAAAGCTAACACTTTGCTTTTTTCGCGACAACAGGCGCCAGTCTTGAGCTGAAGTTGGTGAATGTGCCGACACCGAGTACACA	
CDS	-----CGCCAGTCTTGAGCTGAAGTTGGTGAATGTGCCGACACCGAGTACACA	
	2601	2700
genome	GGAGACAATGGATCGGGCGAAATCTGGTTCCGTGGACCTGCAGTGATGAAGGGATACTTCAAACCTCGAGGAAGAGACCAAGAAGGTGATGACCGGGGATG	
CDS	GGAGACAATGGATCGGGCGAAATCTGGTTCCGTGGACCTGCAGTGATGAAGGGATACTTCAAACCTCGAGGAAGAGACCAAGAAGGTGATGACCGGGGATG	
	2701	2800
genome	GTTGGTTCAAGACAGGCGACATTGGCAGATGAACCCAGACGGCAGCTGTCAATCAAGGACAGGGTCAAGAATCTGGTCAAGCTGTCTCATGGAGAATA	
CDS	GTTGGTTCAAGACAGGCGACATTGGCAGATGAACCCAGACGGCAGCTGTCAATCAAGGACAGGGTCAAGAATCTGGTCAAGCTGTCTCATGGAGAATA	
	2801	2900
genome	TGTCGCCCTGGAGAAATGTGAAGCCGTTTATCGCGATTCCAAGGAGATCAAGAGCATTTGCATCGTTGCCGACAATGGGTGCCCTGTGTTGCTGGCCGTT	
CDS	TGTCGCCCTGGAGAAATGTGAAGCCGTTTATCGCGATTCCAAGGAGATCAAGAGCATTTGCATCGTTGCCGACAATGGGTGCCCTGTGTTGCTGGCCGTT	

Figure 20C

	2901		3000
genome	GTGGAACCGAGCCACGCAGGGGTGAGTGAGACGCTTTTGGCCTGACAAGGTCTCGTTCTTATGGGAATGTGGATCATTACTCTATCACTAACGAGAGAGG		
CDS	GTGGAACCGAGCCACGCAGGGG-----		
	3001		3100
genome	CTGTATTATTCTACTTGCCTAACGTGGCATAGGCGTCTGACAAGGAGATTTTGGATATCCTGAAGAGCCAAGCCAAGGCGGGGGCCTCTCCAAGTCCGA		
CDS	-----CGTCTGACAAGGAGATTTTGGATATCCTGAAGAGCCAAGCCAAGGCGGGGGCCTCTCCAAGTCCGA		
	3101		3200
genome	GACTGTGCAAGGCGTTATCATTGATGATTCGGACTGGATGACGAATGGGTTTCATGACCTCGAGCAGCAAGGTCAAGAGACGCGAGGTCCGCAAGGCACAC		
CDS	GACTGTGCAAGGCGTTATCATTGATGATTCGGACTGGATGACGAATGGGTTTCATGACCTCGAGCAGCAAGGTCAAGAGACGCGAGGTCCGCAAGGCACAC		
	3201	3236	
genome	AACAAGGATATTGAGGAGATGTGGAAGAAGTTCTAG		
CDS	AACAAGGATATTGAGGAGATGTGGAAGAAGTTCTAG		

Figure 21A

```

1   CATGCCAAAGTGCTTTACCGTCAACGTCGGCCCCGAGGACGTCAAGGGCGAGACTCGCATCCGTCGCTCCATCCAGGCCGTCGACAAACTCATGGACTCA
   M P K C F T V N V G P E D V K G E T R I R R S I Q A V D K L M D S
101  CCTCAAGCGACATCAAGACCTTGTACGATGTCATCCAGTACTCTGCCAAGTCCGCCCAACCTCAACGCCATCGGCTACCGCAAGATTGTCAAGATGA
   P S S D I K T L Y D V I Q Y S A K V R P N L N A I G Y R K I V K M I
201  TCGAAGAGGAAAAGGAGATCACCAAGATGGTCAGCGGCGAGCCTGTCAAGGAGAAAAAGACGTGGAATACTTCAAGCTCTCCGGCTACCACTATCTGAC
   E E E K E I T K M V S G E P V K E K K T W K Y F K L S G Y H Y L T
301  CTACAAGGACACCAAGGCCGTATCGACAGCATTGGAAGTGGCCTGCGCAAGTGGGGTGTGAGCCCAAGGAGAGGATCACCGTCTTTGGTTCCACAAGT
   Y K D T K A V I D S I G S G L R K W G V E P K E R I T V F G S T S
401  GCCAACTGGCTGCTGGTCGCTCATGGTGCCTTACGCGAGTCCATGACCATCGTCACCGTGTATGACACTTTGGGCGAGGAAGGATTGCTGCACTCGATGA
   A N W L L V A H G A F T Q S M T I V T V Y D T L G E E G L L H S M N
501  ACGAGGCCGAGGTGGGAACGGCCTACACGAACGCTGATTTGATCAAGACAATGACCAACGTTTCAGGACGCTGCCCCACCTCAAGAGGATCGTCTATGA
   E A E V G T A Y T N A D L I K T M T N V S G R C P T L K R I V Y D
601  CGGCGAAGCCAACGCAGCAGACGTGATCGCCCTTCAGACGGCCCATCCTCACCTTCAGCTTATCACTCTGGAGGAGCTGAAGCAGCTCGGTGTGGATCAC
   G E A N A A D V I A L Q T A H P H L Q L I T L E E L K Q L G V D H
701  CCTGTGGAGCCCACTCCTCCACCGCCGAGGATTGCTCCTGCATCATGTACACTTCTGGATCGACCGGAAACCTAAGGGAGTCATCCTCACTCACGGAA
   P V E P T P P T A E D C S C I M Y T S G S T G N P K G V I L T H G N
801  ACCTCATTGCCGCCATTGGCGGAGTTAACAAGATGCTGGAAAAGTACATTGCGGAAGGCGATGTCTTGCTTGCCTACCTTCCCTTGGCTCACGTTCTGGA
   L I A A I G G V N K M L E K Y I R E G D V L L A Y L P L A H V L E
901  ATTCATGGTTGAGAACCTCTGTCTCTTCTGGGGTGTAAACCTTGGATATGGTACTGTCCGCACGCTGACGGATGCCTCTGTGCGTGAGTGCCAGGGTGAT
   F M V E N L C L F W G V T L G Y G T V R T L T D A S V R E C Q G D
1001 ATCAAGGAATTGCGGCCCTACGCTTATGACCGGCGTTCCAGCAGTGTGGGAGACCATCCGCAAAGGTGTTCTCGCCCAAGTAAACCAGGGTTACCTCTGG
   I K E L R P T L M T G V P A V W E T I R K G V L A Q V N Q G S P L V
1101 TTCAATCCGCTTCAACGCGGCTCTGAACGCCAAGGCCTGGTGCATGGACCGCAAACCTAGGCGCTTTGACTGGAATTTTCGACACTGTGGTGTCAACAA
   Q S V F N A A L N A K A W C M D R K L G A L T G I F D T V V F N K
1201 GGTCCGTCAGCAAACCTGGAGGTCGTCTTCTGCTACGCGCTCTCGGGCGGCGGCGCTATCTCCAGGAGACCCAGCGCTTCTTGACCACAGCACTGTGCCCT
   V R Q Q T G G R L R Y A L S G G A P I S Q E T Q R F L T T A L C P
1301 ATCCTTCAAGCCTATGGCATGACTGAGTCGTGCGGCATGTGCTCGATCATGACTCCAGAGGCGTTCAACTACAACCGCGTCGGTTCCCTGTTCCCTGCA
   I L Q A Y G M T E S C G M C S I M T P E A F N Y N R V G S P V P C T
1401 CAGAGGTCAAGCTCGTGGATGTGCCGATGCAGGATACTTTTCGACTGATTGCCCCGCGCTCGTGGTGAGATTTGGATTGCGGACCCCTCCATCACCTC
   E V K L V D V P D A G Y F S T D S P R P R G E I W I R G P S I T S
1501 TGGATACTTCAAGAACGCTGAGGAGACCTCGGCAGCCATCACAGAGGACCGCTGGCTCAAGACTGGAGATATTGGAGAGTGGCATGCTGATGGCACACTC
   G Y F K N A E E T S A A I T E D R W L K T G D I G E W H A D G T L
1601 TCGGTCATTGATCGCAAGAAGAACTTGGTCAAGTTGTCGCATGGCGAGTACATTGCTCTAGAGAACTTGAGTCGGTGTACAAGAGCACGGCTTACTGCA
   S V I D R K K N L V K L S H G E Y I A L E K L E S V Y K S T A Y C N
1701 ACAACATCTGCGTTTATGCCGATTCCATGCAAAACAAGCCTGTGGCGCTGATTGTTGCGAGTGAACCCGCGATCCTCGAGCTGGCCAAGGCCAAGGGCCT
   N I C V Y A D S M Q N K P V A L I V A S E P R I L E L A K A K G L
1801 GGAGAGCCGCGACTTTGCAGTGCTCTGCCACGATAAGGTGATCATCAAGGCTGTCTCGACGCTGTCTCGCGACTGCCAAAAAGGCTGGCCTCAAGCCC
   E S R D F A V L C H D K V I I K A V L D A C L A T A K K A G L K P

```

Figure 21B

```
1901  GCCGAGTTGCTGCAGGGTGTGTACCTGGAGTCTGAGGAGTGGACCGCTCAAGGCGGTTTGTGACTGCTGCTCAGAAATTGAAGCGCAAGGAAATCAACC
      A E L L Q G V Y L E S E E W T A Q G G L L T A A Q K L K R K E I N Q

2001  AGGCTTATGCTGACCAGATCAAGCAGATCTATGGCTCCAAGTAAAAATGAAC
      A Y A D Q I K Q I Y G S K
```

Figure 22A

	1	100
genome	ATGCCAAAGTGCTTTACCGTCAACGTGGGCCCGAGGACGTCAAGGGCGAGACTCGCATCCGTCGCTCCATCCAGGCCGTCGACAAACTCATGGACTCAC	
CDS	ATGCCAAAGTGCTTTACCGTCAACGTGGGCCCGAGGACGTCAAGGGCGAGACTCGCATCCGTCGCTCCATCCAGGCCGTCGACAAACTCATGGACTCAC	
	101	200
genome	CCTCAAGCGACATCAAGACCTTGTACGATGTCATCCAGTACTCTGCCAAGGTCCGCCCAACCTCAACGCCATCGGCTACCGCAAGATTGTCAAGATGAT	
CDS	CCTCAAGCGACATCAAGACCTTGTACGATGTCATCCAGTACTCTGCCAAGGTCCGCCCAACCTCAACGCCATCGGCTACCGCAAGATTGTCAAGATGAT	
	201	300
genome	CGAAGAGGAAAAGGAGATCACCAAGATGGTCAGCGCGAGCCTGTCAAGGAGAAAAAGACGTGGAATACTTCAAGCTCTCCGGCTACCACTATCTGACC	
CDS	CGAAGAGGAAAAGGAGATCACCAAGATGGTCAGCGCGAGCCTGTCAAGGAGAAAAAGACGTGGAATACTTCAAGCTCTCCGGCTACCACTATCTGACC	
	301	400
genome	TACAAGGACACCAAGGCCGTATCGACAGCATTGGAAGTGGCCTGCGCAAGTGGGGTGTGAGCCCAAGGAGAGGATCACCGTCTTTGGTTCCACAAGGT	
CDS	TACAAGGACACCAAGGCCGTATCGACAGCATTGGAAGTGGCCTGCGCAAGTGGGGTGTGAGCCCAAGGAGAGGATCACCGTCTTTGGTTCCACAAG--	
	401	500
genome	AATGTGTAGCGCCACGAAAATACGATCATTGCAGCGTGAAGGGTGGGAAGAATTAGGGGGGAAATGACATCGATAACAGGAACGAAAAAAAAAAAAACA	
CDS	-----	
	501	600
genome	AGACGAAGTCGGAGATCCTCGCTATTGGCGTTTAAGCACCGCCAGCGTTCTTTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	
CDS	-----	
	601	700
genome	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	
CDS	-----	
	701	800
genome	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	
CDS	-----	
	801	900
genome	GCTTGACTTGACACGTATAGTGCCAAGTGGCTGCTGGTGGTGCCTTACGCGAGTCCATGACCATCGTCACCGTGTATGACACTTTGGGCGAGG	
CDS	-----TGCCAAGTGGCTGCTGGTGGTGCCTTACGCGAGTCCATGACCATCGTCACCGTGTATGACACTTTGGGCGAGG	
	901	1000
genome	AAGGATTGCTGCACTCGATGAACGAGGCCGAGGTGGGAACGGCTACACGAACGCTGATTTGATCAAGACAATGACCAACGTTTCAGGACGCTGCCCCAC	
CDS	AAGGATTGCTGCACTCGATGAACGAGGCCGAGGTGGGAACGGCTACACGAACGCTGATTTGATCAAGACAATGACCAACGTTTCAGGACGCTGCCCCAC	
	1001	1100
genome	CCTCAAGAGGATCGTCTATGACGGCGAAGCCAACGCAGCAGACGTGATCGCCCTTACAGCGGCCCATCCTCACCTTCAGCTTATCACTCTGGAGGAGCTG	
CDS	CCTCAAGAGGATCGTCTATGACGGCGAAGCCAACGCAGCAGACGTGATCGCCCTTACAGCGGCCCATCCTCACCTTCAGCTTATCACTCTGGAGGAGCTG	
	1101	1200
genome	AAGCAGCTCGGTGTGGATCACCTGTGGAGCCCACTCCTCCACGCCGAGGATTGCTCCTGCATCATGTACACTTCTGGATCGACCGGAAACCTAAGG	
CDS	AAGCAGCTCGGTGTGGATCACCTGTGGAGCCCACTCCTCCACGCCGAGGATTGCTCCTGCATCATGTACACTTCTGGATCGACCGGAAACCTAAGG	
	1201	1300
genome	GAGTCATCCTCACTCACGGAAACCTCATTGCCGCCAGTAGGTGTTTCTCTCACTCTCTTTAACCTCTCTTTTCACTTGCAAATATGCTGGGAATCTCT	
CDS	GAGTCATCCTCACTCACGGAAACCTCATTGCCGCCA-----	
	1301	1400
genome	ACTTACCTGAATGTTACTGTTCTTCGGGTATCAACCTAGTTGGCGGAGTTAACAAGATGCTGGAAGGTACATTGCGGAAGGCGATGTCTTGCTTGCTTGCCT	
CDS	-----TTGGCGGAGTTAACAAGATGCTGGAAGGTACATTGCGGAAGGCGATGTCTTGCTTGCTTGCCT	

Figure 22B

	1401	1500
genome	ACCTTCCCTTGGCTCACGTTCTGGAATTCATGGTTGAGAACCCTCTGTCTCTTCTGGGGTGTAAACCCTTGGATATGGTACTGTCCGCACGCTGACGGATGC	
CDS	ACCTTCCCTTGGCTCACGTTCTGGAATTCATGGTTGAGAACCCTCTGTCTCTTCTGGGGTGTAAACCCTTGGATATGGTACTGTCCGCACGCTGACGGATGC	
	1501	1600
genome	CTCTGTGCGTGAGTGCCAGGGTGATATCAAGGAATTGCCGCCCTACGCTTATGACCGGCGTTCCAGCAGTGTGGGAGACCATCCGCAAAGGTGTTCTCGCC	
CDS	CTCTGTGCGTGAGTGCCAGGGTGATATCAAGGAATTGCCGCCCTACGCTTATGACCGGCGTTCCAGCAGTGTGGGAGACCATCCGCAAAGGTGTTCTCGCC	
	1601	1700
genome	CAAGTAAACCAGGGTTCACCTCTGGTTCAATCCGTCTTCAACGCGGCTCTGAACGCCAAGGCCTGGTGCATGGACCGCAAACCTAGGCGCTTTGACTGGAA	
CDS	CAAGTAAACCAGGGTTCACCTCTGGTTCAATCCGTCTTCAACGCGGCTCTGAACGCCAAGGCCTGGTGCATGGACCGCAAACCTAGGCGCTTTGACTGGAA	
	1701	1800
genome	TTTTCGACACTGTGGTGTCAACAAGGTCCGTCAGCAAACCTGGAGGTGCTCTTCGCTACGCGCTCTCGGGCGGCGCGCCTATCTCCAGGAGACCCAGCG	
CDS	TTTTCGACACTGTGGTGTCAACAAGGTCCGTCAGCAAACCTGGAGGTGCTCTTCGCTACGCGCTCTCGGGCGGCGCGCCTATCTCCAGGAGACCCAGCG	
	1801	1900
genome	CTTCTTGACCACAGCACTGTGCCCTATCCTTCAAGCCTATGGCATGACTGAGTCGTGCGGCATGTGCTCGATCATGACTCCAGAGGCGTTCAACTACAAC	
CDS	CTTCTTGACCACAGCACTGTGCCCTATCCTTCAAGCCTATGGCATGACTGAGTCGTGCGGCATGTGCTCGATCATGACTCCAGAGGCGTTCAACTACAAC	
	1901	2000
genome	CGCGTCGGTTCOCCTGTTCCCTGCACAGAGGTCAAGCTCGTGGATGTGCCCGATGCAGGATACTTTTCGACTGATTGCCCCGCCCTCGTGGTGAGATTT	
CDS	CGCGTCGGTTCOCCTGTTCCCTGCACAGAGGTCAAGCTCGTGGATGTGCCCGATGCAGGATACTTTTCGACTGATTGCCCCGCCCTCGTGGTGAGATTT	
	2001	2100
genome	GGATTGCGGACCCCTCCATCACCTCTGGATACTTCAAGAACGCTGAGGAGACCTCGGCAGCCATCACAGAGGACCGCTGGCTCAAGACTGGAGATATTGG	
CDS	GGATTGCGGACCCCTCCATCACCTCTGGATACTTCAAGAACGCTGAGGAGACCTCGGCAGCCATCACAGAGGACCGCTGGCTCAAGACTGGAGATATTGG	
	2101	2200
genome	AGAGTGGCATGCTGATGGCACACTCTCGGTCATTGATCGCAAGAAGAACTTGGTCAAGTTGTGCGATGGCGAGTACATTGCTCTAGAGAAACTTGAGTCG	
CDS	AGAGTGGCATGCTGATGGCACACTCTCGGTCATTGATCGCAAGAAGAACTTGGTCAAGTTGTGCGATGGCGAGTACATTGCTCTAGAGAAACTTGAGTCG	
	2201	2300
genome	GTGTACAAGAGCACGGCTTACTGCAACAACATCTGCGTTTATGCCGATTCCATGCAAAACAAGCCTGTGGCGCTGATTGTTGCGAGTGAACCCCGCATCC	
CDS	GTGTACAAGAGCACGGCTTACTGCAACAACATCTGCGTTTATGCCGATTCCATGCAAAACAAGCCTGTGGCGCTGATTGTTGCGAGTGAACCCCGCATCC	
	2301	2400
genome	TCGAGCTGGCCAAGGCCAAGGGCCTGGAGAGCCGCGACTTTGCAGTGCTCTGCCACGATAAGGTGATCATCAAGGCTGTCTCGACGCCTGTCTCGCGAC	
CDS	TCGAGCTGGCCAAGGCCAAGGGCCTGGAGAGCCGCGACTTTGCAGTGCTCTGCCACGATAAGGTGATCATCAAGGCTGTCTCGACGCCTGTCTCGCGAC	
	2401	2500
genome	TGCCAAAAGGCTGGCCTCAAGCCCGCCGAGTTGCTGCAGGGTGTGTACCTGGAGTCTGAGGAGTGGACCGCTCAAGGCGGTTTGTGACTGCTGCTCAG	
CDS	TGCCAAAAGGCTGGCCTCAAGCCCGCCGAGTTGCTGCAGGGTGTGTACCTGGAGTCTGAGGAGTGGACCGCTCAAGGCGGTTTGTGACTGCTGCTCAG	
	2501	2569
genome	AAATTGAAGCGCAAGGAAATCAACCAGGCTTATGCTGACCAGATCAAGCAGATCTATGGCTCCAAGTAA	
CDS	AAATTGAAGCGCAAGGAAATCAACCAGGCTTATGCTGACCAGATCAAGCAGATCTATGGCTCCAAGTAA	

Figure 23A

```

1   CCTTTATCCCGCACCGCCATCTCTCGCGCCACCATCTCGCATTCTTTCAATCCACACTCCCACCTGTGCCCCCTGCTTTTCACGTCCCGCTCTCATC
101  CCGCCTTCTCCTTTTCATCACCCCAATTCAACATGACAAAGTGCCTCACCGTCGAAGTCGGACCCGCGGACGTCCAGGGCGAGACCCGCATCCGCGCTCC
      M T K C L T V E V G P A D V Q G E T R I R R S
201  GTCCTCTCTGCAAAGCGCCTCATGTCTCTCGCCCTCGGATGACATCAAGACCCTCTACGACGTCTTCAACCACTCCGTACCGTCCGCCCCAACCTCAACG
      V L S A K R L M S S P S D D I K T L Y D V F N H S V T V R P N L N A
301  CGATCGGATACCGCAAGGTCGTCAAGATTGTGAGGAAGAAAAGGAGGTGTCGAAGTTGTCAACGGCGAGGAAGTCAAGGAAAAGAAGACCTGGAAGTT
      I G Y R K V V K I V E E E K E V V K V V N G E E V K E K K T W K F
401  CTTCAAGATGTCCGGCTACCACTGGCTCACCTACAAGGATGCGAAGCAGGTGTCGACAGCATCGGATGCGGTCTTCGCAAGTTTGGCGTCGAGCCCAAG
      F K M S G Y H W L T Y K D A K Q V V D S I G C G L R K F G V E P K
501  GACAAGCTGACCGTTTTCGGTGCCACAAGTGCCAAGTGGCTCCTGCTTGCCACGGTGCTTTACCCAGTCCATCACCATTGTTACCGCTACGACACCC
      D K L T V F G A T S A N W L L L A H G A F T Q S I T I V T A Y D T L
601  TGGGCGAGGACGGTCTTTTGAAGTCTATGAACGAGGCGGAGGTGGCCACCGCTTACACAAACGCCGACTTGCTCAACACTATCAAGAACGTTGCCGGCAA
      G E D G L L H S M N E A E V A T A Y T N A D L L N T I K N V A G K
701  ATGCCCCACCCTGAAGAAGATCATCTACGACGGCGATGCCAAGCCGCGAGATGTCATTGCCCTCCAGGAGGCCCATCCTCACCTCCAGCTCATCACCTC
      C P T L K K I I Y D G D A K P A D V I A L Q E A H P H L Q L I T L
801  GAGGAGCTGAAGCAGCTCGGAGTGGACAACCCTGTGCCCCAACCCCTCCTGCTGCCAAGGACTACTGCTGCATCATGTACACTTCGGGATCGACTGGCA
      E E L K Q L G V D N P V A P T P P A A K D Y C C I M Y T S G S T G N
901  ACCCCAAGGGAGTGTGCTGACCCATGGAAACCTCGTTGCTGCCATCGGAGGTGTGAACAAGATGCTGACAAAGTACGTTACGAGGGAGACGTCTTGCT
      P K G V L L T H G N L V A A I G G V N K M L T K Y V H E G D V L L
1001 CGCGTACTTGCTCTTGCTCACGTTCTCGAGTTCCTGGTCGAAAACGTGTGTCTCTTCTGGGGTGTGACTCTTGCTACGGTACCGTCCGCACATTGACT
      A Y L P L A H V L E F L V E N V C L F W G V T L G Y G T V R T L T
1101 GATGCCTCAGTCCGTGAGTGCCAGGGTGATATCAAGGAGTTGCGCCCTACATTGATGACCGGTGTTCTGCTGTGTGGGAGACGATTGTAAGGGAGTGT
      D A S V R E C Q G D I K E L R P T L M T G V P A V W E T I R K G V L
1201 TGGCTCAGGTTTCCAGGGCTCACCTCTTGTTCAAAAGATCTTCCATGCTGCTTTGAACGCCAAGGCCTGGTGCCTGGACCGCAAGTTGGGTGCGTTGAC
      A Q V S Q G S P L V Q K I F H A A L N A K A W C L D R K L G A L T
1301 TGGAACTTCGATACTGTGCTTCAACAAGGTCAAGCAGCAGACAGGAGGACGTCTTCGCTTCGCCCTTTGGGAGGTGCACCCATCTCTCAGGAGACC
      G I F D T V V F N K V K Q Q T G G R L R F A L S G G A P I S Q E T
1401 CAGCGCTCTTGACGACAGCTTTGTGCCCTATCCTCCAGGGCTACGGTATGACAGAGTCTTGCGGCATGTGCGCCATTTTGACCCCGATGTCTTCAACT
      Q R F L T T A L C P I L Q G Y G M T E S C G M C A I L T P D V F N Y
1501 ACAGCCGTGTGCGATCCCGAGTTCCTTGACGGGAGGTCAAGTTGGTCGATGTGCCGATGCAGGATACCACTCAACGGACTTGCTCTCCCCGCTGGTGA
      S R V G S P V P C T E V K L V D V P D A G Y H S T D L P L P R G E
1601 GGTCTGCATTGCTGGACCCCTCCATCACTGCTGGATACTTCAAGAACCCCGAGGAGACCTCCGCCACATTGACTGCTGATCGCTGGCTCAAGACTGGAGAT
      V C I R G P S I T A G Y F K N P E E T S A T L T A D R W L K T G D
1701 ATCGGAGAGTGGCAGCCCGACGGCACTATCTCGATCATTGACCGCAAGAAGAACTTGGTCAAGCTGTCACACGGAGAGTACATTGCTTTGGAGAAGCTTG
      I G E W H P D G T I S I I D R K K N L V K L S H G E Y I A L E K L E
1801 AGTCTGTCTACAAGAGCACAGCCTACTGCAACAACATTTGCGTGTATGCCGACTCGATGCAGAACAAGCCCGTTGCCATTATTGTTGCCAGCGAACCCCG
      S V Y K S T A Y C N N I C V Y A D S M Q N K P V A I I V A S E P R
1901 CATCCTCGAGTTGGCCAAGGCCAAGGGCATTGAGAGCCGCGACTTTGCTGCTCTCTGCCACGACAAGGTTATCATCAAGGCTGTCCACGATGCCTGCCTC
      I L E L A K A K G I E S R D F A A L C H D K V I I K A V H D A C L

```

Figure 23B

2001 GCCACTGCCAAGCGTGGTGGACTCAAGCCCGCTGAGATGCTTCAGGGAGTGTACTTGGAGTCAGAAGAATGGACGGCCAGGCTGGCATGTTGACTGCCG
A T A K R A G L K P A E M L Q G V Y L E S E E W T A Q A G M L T A A

2101 CTCAGAAGCTCAAGCGCAAGGAGATCAACCAGGCCTATGTCTCAGATCAAGCAGCTTTATGGAACGGCCTAAGTCGCTGAAAGGTGTGCCTTTGTCCG
Q K L K R K E I N Q A Y V S Q I K Q L Y G T A

2201 TCTCTCAACCCACAAGTCCTATGTATAATGACCCGCGCGCCCTCCTTTAATCCTATACCCACCCTTTTTACACGTTAAAGAAGCCACATTTTGGT

2301 TCTTTTTTTCTCTCGCACACACTACACACTCCCATCCATTCCCTCCAAACAGGATGGTTGTCTGCAAATAAATTGACGAATTTCTCTTGAAAAAAA

2401 AAAAAAAAAAAAAA

Figure 24A

	1	100
genome	ATGACAAAGTGCCTCACCCTCGAAGTCGGACCCGCCGACGTCCAGGGCGAGACCCGCATCCGCCGCTCCGTCTCTCTGCAAAGCGCCTCATGTCTCGC	
CDS	ATGACAAAGTGCCTCACCCTCGAAGTCGGACCCGCCGACGTCCAGGGCGAGACCCGCATCCGCCGCTCCGTCTCTCTGCAAAGCGCCTCATGTCTCGC	
	101	200
genome	CCTCGGATGACATCAAGACCCCTCTACGACGTCTTCAACCACTCCGTACCGTCCGCCCAACCTCAACGCGATCGGATACCGCAAGGTCGTCAAGATTGT	
CDS	CCTCGGATGACATCAAGACCCCTCTACGACGTCTTCAACCACTCCGTACCGTCCGCCCAACCTCAACGCGATCGGATACCGCAAGGTCGTCAAGATTGT	
	201	300
genome	CGAGGAAGAAAAGGAGGTCGTCAAGGTTGTCAACGGCGAGGAAGTCAAGGAAAAGAAGACCTGGAAGTTCTTCAAGATGTCCGGCTACCACTGGCTCACC	
CDS	CGAGGAAGAAAAGGAGGTCGTCAAGGTTGTCAACGGCGAGGAAGTCAAGGAAAAGAAGACCTGGAAGTTCTTCAAGATGTCCGGCTACCACTGGCTCACC	
	301	400
genome	TACAAGGATGCGAAGCAGGTCGTGACAGCATCGGATGCGGTCTTCGCAAGTTTGGCGTCGAGCCCAAGGACAAGCTGACCGTTTTCGGTGCCACAAGGT	
CDS	TACAAGGATGCGAAGCAGGTCGTGACAGCATCGGATGCGGTCTTCGCAAGTTTGGCGTCGAGCCCAAGGACAAGCTGACCGTTTTCGGTGCCACAAG--	
	401	500
genome	AAGAAAGAGGCATAACAAGAAAATGCAAGAGAGGCCAAAAAAATGGCTTGACGTGAGAGCATAAGGGAACCAACAGACAGGTGTTTGTGTGGGTTGCGGA	
CDS	-----	
	501	600
genome	TAGTGGGTGAGCATGCTTCGTTATCGAATGTGGGAGAAGAGACGGACGCGAATATGGCTCTCGTCTCTGGCGGGATGCGAGTGGCCAAGTGTGGGATAC	
CDS	-----	
	601	700
genome	ATATCCTCGCGGTGGGGTGTCCGGGTCCGCCCTTGAATCTTGTTGAAGCATGATAATGTGAATGTGGACCGCAATCACGCTCAGATTATGCGTAGCAAGC	
CDS	-----	
	701	800
genome	GTGTTGCTAGTCTACATCATGCTCACACGTATTCACATTTATTCATTTTCACTCTATCTCGCTCTTAGTGCCAAGTGGCTCCTGCTTGCCACGGTGCTT	
CDS	-----TGCCAAGTGGCTCCTGCTTGCCACGGTGCTT	
	801	900
genome	TCACCCAGTCCATCACCATTGTTACCGCTACGACACCCTGGGCGAGGACGGTCTTTTGCACTCTATGAACGAGGCCGAGGTGGCCACCGCTTACACAAA	
CDS	TCACCCAGTCCATCACCATTGTTACCGCTACGACACCCTGGGCGAGGACGGTCTTTTGCACTCTATGAACGAGGCCGAGGTGGCCACCGCTTACACAAA	
	901	1000
genome	CGCCGACTTGCTCAACACTATCAAGAACGTTGCCGGCAAATGCCCCACCGTGAAGAAGATCATCTACGACGGCGATGCCAAGCCCGCAGATGTCATTGCC	
CDS	CGCCGACTTGCTCAACACTATCAAGAACGTTGCCGGCAAATGCCCCACCGTGAAGAAGATCATCTACGACGGCGATGCCAAGCCCGCAGATGTCATTGCC	
	1001	1100
genome	CTCCAGGAGGCCCATCCTCACCTCCAGCTCATCACCTCGAGGAGCTGAAGCAGCTCGGAGTGGACAACCCTGTGCCCCAACCCCTCCTGCTGCCAAGG	
CDS	CTCCAGGAGGCCCATCCTCACCTCCAGCTCATCACCTCGAGGAGCTGAAGCAGCTCGGAGTGGACAACCCTGTGCCCCAACCCCTCCTGCTGCCAAGG	
	1101	1200
genome	ACTACTGCTGCATCATGTACACTTCGGGATCGACTGGCAACCCCAAGGGAGTGTTGCTGACCCATGGAAACCTCGTTGCTGCCAGTACGTATCTTTCTCG	
CDS	ACTACTGCTGCATCATGTACACTTCGGGATCGACTGGCAACCCCAAGGGAGTGTTGCTGACCCATGGAAACCTCGTTGCTGCCA-----	
	1201	1300
genome	TCATGATCGTCCTCCCGCATTTCCACTGGGCTTGTTACCATTTGATGGGAAATGTATTTAACCCGAACCACACATTTTTCTTTTTCTCACACTTGCCAC	
CDS	-----	
	1301	1400
genome	GTCAC TAGTCGGAGGTGTGAACAAGATGCTGACAAAGTACGTTACGAGGGAGACGTCTTGCTCGCGTACTTGCCTCTTGCTCACGTTCTCGAGTTCCTG	
CDS	-----TCGGAGGTGTGAACAAGATGCTGACAAAGTACGTTACGAGGGAGACGTCTTGCTCGCGTACTTGCCTCTTGCTCACGTTCTCGAGTTCCTG	

Figure 24B

	1401	1500
genome	GTCGAAAACGTCTGTCTCTTCTGGGGTGTGACTCTTGGGTACGGTACCGTCCGCACATTGACTGATGCCTCAGTCCGTGAGTGCCAGGGTGATATCAAGG	
CDS	GTCGAAAACGTCTGTCTCTTCTGGGGTGTGACTCTTGGGTACGGTACCGTCCGCACATTGACTGATGCCTCAGTCCGTGAGTGCCAGGGTGATATCAAGG	
	1501	1600
genome	AGTTGCGCCCTACATTGATGACCGGTGTTCTGCTGTGTGGGAGACGATTTCGTAAGGGAGTGTTGGCTCAGGTTTCCAGGGCTCACCTCTTGTTCAAAA	
CDS	AGTTGCGCCCTACATTGATGACCGGTGTTCTGCTGTGTGGGAGACGATTTCGTAAGGGAGTGTTGGCTCAGGTTTCCAGGGCTCACCTCTTGTTCAAAA	
	1601	1700
genome	GATCTTCCATGCTGCTTTGAACGCCAAGGCCTGGTGCCTGGACCGCAAGTTGGGTGCGTTGACTGGAATCTTCGATACTGTCGTCTTCAACAAGGTCAAG	
CDS	GATCTTCCATGCTGCTTTGAACGCCAAGGCCTGGTGCCTGGACCGCAAGTTGGGTGCGTTGACTGGAATCTTCGATACTGTCGTCTTCAACAAGGTCAAG	
	1701	1800
genome	CAGCAGACAGGAGGACGTCTTCGCTTCGCCCTTTTCGGGAGGTGCACCCATCTCTCAGGAGACCCAGCGCTTCTTGACGACAGCTTTGTGCCCTATCCTCC	
CDS	CAGCAGACAGGAGGACGTCTTCGCTTCGCCCTTTTCGGGAGGTGCACCCATCTCTCAGGAGACCCAGCGCTTCTTGACGACAGCTTTGTGCCCTATCCTCC	
	1801	1900
genome	AGGGGTACGGTATGACAGAGTCTTGCGGCATGTGCGCCATTTTGACCCCGATGTCTTCAACTACAGCCGTGTCGGATCCCCAGTTCCCTTGACCGGAGGT	
CDS	AGGGGTACGGTATGACAGAGTCTTGCGGCATGTGCGCCATTTTGACCCCGATGTCTTCAACTACAGCCGTGTCGGATCCCCAGTTCCCTTGACCGGAGGT	
	1901	2000
genome	CAAGTTGGTCGATGTGCCCGATGCAGGATACCACTCAACGGACTTGCCCTCTCCCCCGTGGTGAGGTCTGCATTGCTGGACCCCTCCATCACTGCTGGATAC	
CDS	CAAGTTGGTCGATGTGCCCGATGCAGGATACCACTCAACGGACTTGCCCTCTCCCCCGTGGTGAGGTCTGCATTGCTGGACCCCTCCATCACTGCTGGATAC	
	2001	2100
genome	TTCAAGAACCCCGAGGAGACCTCCGCCACATTGACTGCTGATCGCTGGCTCAAGACTGGAGATATCGGAGAGTGGCACCCCGACGGCACTATCTCGATCA	
CDS	TTCAAGAACCCCGAGGAGACCTCCGCCACATTGACTGCTGATCGCTGGCTCAAGACTGGAGATATCGGAGAGTGGCACCCCGACGGCACTATCTCGATCA	
	2101	2200
genome	TTGACCGCAAGAAGAACTTGGTCAAGCTGTACACGGAGAGTACATTGCTTTGGAGAAGCTTGAGTCTGTCTACAAGAGCACAGCCTACTGCAACAACAT	
CDS	TTGACCGCAAGAAGAACTTGGTCAAGCTGTACACGGAGAGTACATTGCTTTGGAGAAGCTTGAGTCTGTCTACAAGAGCACAGCCTACTGCAACAACAT	
	2201	2300
genome	TTGCGTGTATGCCGACTCGATGCAGAACAAGCCGTTGCCATTATTGTTGCCAGCGAACCCCGCATCCTCGAGTTGGCCAAGGCCAAGGGCATTGAGAGC	
CDS	TTGCGTGTATGCCGACTCGATGCAGAACAAGCCGTTGCCATTATTGTTGCCAGCGAACCCCGCATCCTCGAGTTGGCCAAGGCCAAGGGCATTGAGAGC	
	2301	2400
genome	CGCGACTTTGCTGCTCTCTGCCACGACAAGGTTATCATCAAGGCTGTCCACGATGCCTGCCTCGCCACTGCCAAGCGTGCTGGACTCAAGCCCGCTGAGA	
CDS	CGCGACTTTGCTGCTCTCTGCCACGACAAGGTTATCATCAAGGCTGTCCACGATGCCTGCCTCGCCACTGCCAAGCGTGCTGGACTCAAGCCCGCTGAGA	
	2401	2500
genome	TGCTTCAGGGAGTGTACTTGGAGTCAGAAGAATGGACGGCCCGAGGCTGGCATGTTGACTGCCGCTCAGAAGCTCAAGCGCAAGGAGATCAACCAGGCCTA	
CDS	TGCTTCAGGGAGTGTACTTGGAGTCAGAAGAATGGACGGCCCGAGGCTGGCATGTTGACTGCCGCTCAGAAGCTCAAGCGCAAGGAGATCAACCAGGCCTA	
	2501	2537
genome	TGTCTCACAGATCAAGCAGCTTTATGGAACGGCCTAA	
CDS	TGTCTCACAGATCAAGCAGCTTTATGGAACGGCCTAA	

Figure 25A

	1	100
MaACS-3	-----METDALTIALTIAIAIVLALVKFNEKEPDHLPLLLGQSSVT---	PIRNE
MaACS-4	-----MSLDQNAQSVELPGTR-QPGQTGIYRRKGFENALLAVPPSRP---	HIKTI
MaACS-5	-----MTTQLYSIEVAGSPEIPGEGKPRRSVLSPKLVQSYQSFKGDGSIITL	
MaACS-6	-----MATQMYSVVVPSNDIPGEGKPRRSVLCPKLLENYPSVKAGSTIITL	
MaACS-8	-----MPSFKKYNLDKQSVEVPGR-KPGASGHYRHAAYGDALVTNIREAP---	HIETL
MaACS-10	-----METLVNGKYAVEYDEV DHIYRNVMATGGLDRMPPPYYDIKERTM	
MaACS-11	-----MPKCFTVNVGPELVKGETRIRRS-----IQAVDKLMDSP-SSDIKTL	
MaACS-12	-----MTKCLTVEVGPAVQGETRIRRS-----VLSAKRLMSSP-SSDIKTL	
ScFAA1	-----MVAQYTVPVGKAANEHETAPRRN-----YQCREKPLVRPPNTKCSVT	
ScFAA2	MAAPDYALTDLIESDPRFESLKTRLAGYTKGSDEYIEELYSQLPLTSYPRYKTF	LKKQAVATSNPDNEAGFSSIYRS-----SLSSENLVSCVDKNLRTA
ScFAA3	-----MSEQHSVAVGKAANEHETAPRRN-----VRVKKRPLIRPLNSSASTL	
ScFAA4	-----MTEQYSVAVGKAANEHETAPRRN-----IRVKDQPLIRPINSASTL	
	101	200
MaACS-3	GESVIHRSKTVPHGTLTKPSEKIKTLHDVWQTGAAVNPAGRSLMFLQNGFAFIEATYEQVNRRI	GGFTFVKATGLKPKTD-----TPVGI FMPYS
MaACS-4	YDAFQHGLKLNPNGAALGSVYDPVTD-----F-----GGYVWQTYAQVNDRI	TRFSLVKIHKDVHGLATVGQKWSLGIWAINR
MaACS-5	YENFLEGIQRSEGGEFLGHPIVDNVA-----QPYEWLSYTRVQERVANF	ALLIQLGLKVDSNF-----G-IFSINR
MaACS-6	YENFQEGVLRSGGAHFLGHPIVNGQP-----QAYKQWSYVDVSKRVTHF	ALLAHLGLSPKQNF-----G-IFSINR
MaACS-8	YDMWQNSVTKYGGNDFLGHPIFNTVAQT-----Y-----GGYSWETYRQINQRVNAF	SIMHLNEVILGNR-QLNRWALGIWSHGR
MaACS-10	AHLFEYMANTYEDKDAMGWDIKVKHVEKQAA-----PGEKPKTWITYELSDYNWMSYRQAKNYADRV	LITRLGVEKG-----DFVMIFASTC
MaACS-11	YDVIQYSAKVRPNLNAIGYKIVKMIIEEEKEITKMVSGEPVKEKTKWKYFKLSGYHYLYTKD	KAVIDSISLRKWGVPE--KE-----RITVFGSTS
MaACS-12	YDVFNHSVTVRPNLNAIGYKVVKIVEEEKEVVKVNGEEVKEKTKWKFFKMSGYHWLYTKDA	KQVVDSTGLRKFGVEP--KD-----KLTVFGATS
ScFAA1	YEFVLECFQKNKNSNAMGW DVKEIHEESKSYMKKVDGKETSEVKKWMYELSHYHNSFDQ	LTDIMHEIRLVKIGLKPND--KLHLYAATS
ScFAA2	YDHFMFSARRWPQRDCLGSPIDKATGT-----WEETFRFESYSTVSKRCHNI	SILSLVNTKRKRPLEANDFVVAIILSHNN
ScFAA3	YEFALCFNKGKGRDGMW DVIEIHETKKTIVRKVDGKDKSIEKTLWLYEMSPYKMMTYQEL	IWVMHDMRLAKIGIKPNGEH-----KFHIFASTS
ScFAA4	YEFALCFNKGKGRDGMW DIIDIHETKKTIVRKVDGKDKPIEKTWLYELTPYITMTYEEM	ICVMHDIRLIKIGVKPNGEN-----KFHIFASTS
	201	300
MaACS-3	QESFVAQQAFYRYSFVAVPIHDLRNNDLLVEVVDQTKLKAIIVSQVLPPLLQSLKE-CPTIKTI	IMAGIYISQEQLE-----M---AAQHG
MaACS-4	PEWTIASEACSAYNLVSGLYDTLGP EAVTYGINHAECSIVVTSVDHIATLLNESSK-MPGLK	IIISMDDLDTGRAGPG---LA-PTGTILRTYAQDKG
MaACS-5	PEWTMSELAGYMYNFTSVPL YDTLGVSAIEYIVNQTEMETIIASADKASILLNMKST-LPTL	KNIVVMGSLEDALVE-----GREMD
MaACS-6	PEWTMSELAGYMHNYTSVPL YDTLGVAAIEYIVNQTEMQIIASSDKASILLHMKSA-LPTV	QTIVVMGEFTDALVAE-----GKELN
MaACS-8	PEWFIEMSCNLYNLISVALYDTLGPDAVEYIVNHAEIEIVVSSANHISLLENAEK-LPKL	KAIVSMDALHDTVVPVPG---AT-SAAQVLRAWGAQKG
MaACS-10	PEWFLTAHGCFSGSVTIVTAYDSMDEKSIQFIVDQSQPKAIFADAHTLPVSKLMQKGN	SGVKAVIYTGQEWVTDATKKMEQVENRSFELVHIDELKKT
MaACS-11	ANWLLVAHGAFTQSMITIVTYDTLGE EGLLSHMNEAEVGTAYTNADLIKMTNVSGR-CPTL	KRIVYDGEANAAD-----VIALQTAHPH
MaACS-12	ANWLLLAHGAFTQSMITIVTAYDTLGE DGLLSHMNEAEVATAYTNADLLNTIKNVAGK-CPTL	KKIYDGDAPAD-----VIALQEAHPH
ScFAA1	HKWMKMF LGAQSQGI PVVTAYDTLGEKGLIHSLVQTGSKAIFTDNLPSLIPKVAQ-AQDV	KYIHFDSISSEDRRQSGKIYQS-AHDAINRIKEVRPD
ScFAA2	PEWILTDLACQAYSLTNTALYETLGPNTSEYILNLT EAPILIFAKSNMYHVLKMVPD-MKF	VNTLVGMDELTHDELRLN-----ESLLPVKCNLSNEK
ScFAA3	HKWMKIFLGCISQGI PVVTAYDTLGE SGLIHSMVETESAAIFTDNQLLAKMIVPLQS-AK	DIKFLIHNEPIDPNRRQNGKLYKA-AKDAINKIREVRPD
ScFAA4	HKWMKTF LGCMSQGI PVVTAYDTLGE SGLIHSMVETDSVAIFTDNQLLSKLAVPLKT-AK	NKFKVIHNEPIDPSDKRQNGKLYKA-AKDAVDKIREVRPD
	301	400
MaACS-3	VKLLKFAAVEYE-----SSTLMEPVQDPEDVAMNYNTKSSSLVMLTALIAAMTAFTESLPAK--	KRFSSKDRLLSHFSNGDVISVFM
MaACS-4	VLLYDWSEVEAV-----IQHGRKHTPTSSDAYTCYTSGTTGLPAILTGLIALLASSDVA-----	TPVLADDCLISFLPLPHVFGVRM
MaACS-5	IHIVAWS DVERD-----FNNPAPANPTDDVATCYTSGTTGTPAILTGFVAGLASFHMMAKH	QKFFIPSSVDTHISYPLAHVFERLS
MaACS-6	INIVSWTDVEKS-----LERPVEAVHTAEDIATCYTSGTTGTPAILTGFVATIASFHMMAKH	GRFFIPSPADTHVSYLPLAHVFERLC
MaACS-8	IKVYDFNEIESL-----AEFPRKHLPTADEVASCYTSGTTGTPAMLTRFVATVGTNREG-----	MLLTEDDVLISFLPLAHIMGRVI
MaACS-10	KSASNGEQSAGK KORSSEDAEGAQDEIEVIYKADDLACMYTSGSTGTPAOLTGLMAAIGSAAAMEG--	DQLDKETDIVISYPLAHVLEFVI
MaACS-11	LQLITLEELKQL-----VDHPVEPTPTAEDCSCMYTSGSTGNPVILTGLIAAIGGVNKMLEK----	YIREGDVLLAYLPLAHVLEFMV
MaACS-12	LQLITLEELKQL-----VDNPVAPTPTAAKDYCCMYTSGSTGNPVLLTGLVAAIGGVNKMLTK----	YVHEGDVLLAYLPLAHVLEFLV
ScFAA1	IKTFSFDDILKLKE-----SCNEIDVHPGKDDLCCMYTSGSTGTPVVLKSVVAGVGGASLNVLK----	FVGNTDRVICFLPLAHIFELVF
ScFAA2	ITFFSLEQVEQV-----CFNKIPAIPTPDSLYTSFTSGTTGLPVEMSR IASGIAFAFSTFRIP-PD	KRNQQLYDMCFLPLAHIFERMV
ScFAA3	IKIYSFE EVVKIKK-----SKDEVKLHPEPKDLACMYTSGSISAPVVLTYIVSGIAGVGHNVFG--	WIGSTDRVLSFLPLAHIFELVF
ScFAA4	IKIYSFDEIEIKK-----AKDEVELHFKPEDPACMYTSGSTGTPVVLTYIVAGIGGVGHNVIG--	WIGPTDRIIAFLPLAHIFELTF

Figure 25B

	401	500
MaACS-3	SSAIIIMGGSIVFPSGLMKN-----VLHISQASATIFASTPIILEKIHEALQLTYGGG-----SMFRRGFAAIATLQAGR-----ITTT	
MaACS-4	ELFAMAAGGKIGYSTGDPLR-----LLEIVSHLKSIIFPAVPRLLNRVYAKVYAATVGAPGLTGALARRGLATLTNLREGK-----GFHH	
MaACS-5	QAVMISGAARIGYYQGDTLK-----LLDVAILQITIFVSVPRLFNRIYDKVLAVGKAKGGAAFLNRAFETKANLRKG-----ILEH	
MaACS-6	QAVMISGAARIGYYQGDTLK-----LLDVAVLHTIFASVPRLFNRIYDKVLAVGKAKGGIAAFLNRAYNSKANLRKG-----VLEH	
MaACS-8	DTCMSYSGKIGYFRGDILL-----LLEIVAELRITFFPAVPRLLNRIYAKLVASTIEAPGLVGALARRGVAAIMANLAAGK-----GVNH	
MaACS-10	SHFVVSMGCRLLGFRARTLMDDAVAPTAGSGRSKGLGLKALQITLMAGVPTIWERIRKGILAEVNGQSFPITLFFAALNTWAI VQATGS---ENFVT	
MaACS-11	ENLCLFWGVTLGYGTVRTLTDASVR-----ECQGIKELRTLMTGVPVAVWETIRKGVLAQVNGQSPLVQSVFNAALNAWCMDRKL-----GALT	
MaACS-12	ENVCLFWGVTLGYGTVRTLTDASVR-----ECQGIKELRTLMTGVPVAVWETIRKGVLAQVNGQSPLVQKIFHAALNAWCLDRKL-----GALT	
ScFAA1	ELLSFYWGACIGYATVKTLTSSSVR-----NCQGLQEFKITIMVGVAAVWETVRKGI LNQIDNLPFLTCKIFWTAYNTLNMQR-LH-----IPGG	
ScFAA2	IAYDLAIGFGIGFLHKPDPT-----VLVELKILKYAVALVPRI LTRFEAGIKNALDKS-TVQRNVANTILDSARFTARGGPDKSIMNF	
ScFAA3	EFEAFYWNGILGYGSVKTLTNTSTR-----NCKGLVEFKITIMIGVAAVWETVRKAILEKISDLTPVLQKIFWSAYSM E-----KS-----VPCT	
ScFAA4	EFEAFYWNGILGYANVKTLTPTSTR-----NCQGLMEFKITVMVGVAAVWETVRKGI LAKINELPGWSQTLFWTVYAL E-----RN-----IPCS	
	501	600
MaACS-3	SLWDLIGLGEVRSKLIG-KVRMVVTH-PTKPETLDYIRAAMGIHVITTRITSGIVTARNMLDYANAPHLPPVGCNEVKVDDVAAGFTSADEPNP	
MaACS-4	PLWDRILFSKVQALIG-NVRLMLTASAPISAEILEFVRVAFCEVVEAQQGGGAATNTVI-GETEAGHVPPQACCEIKVDVPELNYFATDKPFP	
MaACS-5	AIWDRLVFGAIRARLIG-KVKHIVSGSAPIAPDVMDFLRICFSADVYEGQQAAGLCMSYR-GDLTSGQVPPQLCCEVKRDVPMHYTSQDKPRP	
MaACS-6	PLWDKLVFGAIRARLIG-KVKHIVSGSAPISPDVMDFLRICFSADVYEGQQAAGLSMSYR-GDLTPGQVPPQLCTEVK KDIPSMNYSSADKPPF	
MaACS-8	ALWDRLLFNKVKMALIG-RVQVILTGSAPIAKEVLSFLRIAFGCVVLEGSGMATATITMA-DEYIPGHI CPRAGCELKVDVPAMNYLSTDQYP	
MaACS-10	KTIDSLVFSKAKELVIG-KLRLTLTGAGISDETHRFLSMVM-CYVISGLVCGVAAVTLPRMGHRLRTVPPAPSLELVNVVPDEYTGDNQ---	
MaACS-11	GIFDTVFNKVRQQTIG-RLRYALSGGAPISQETORFLTTAL-CPILQAMSCGMCSIMTP-EAFNYNRVSPVPCTEVKVDVPDAGYFSTDSPRP	
MaACS-12	GIFDTVFNKVRQQTIG-RLRFALSGGAPISQETORFLTTAL-CPILQAMSCGMCAILTP-DVFNYSRVSPVPCTEVKVDVPDAGYHSTDPLP	
ScFAA1	GALGNLVFKKIRTATIG-QLRYLLNGGSPISRDAQEFITNLI-CPMLIGLITCASTTILDP-ANFELGVALDTCGVTVKVDVEELGYFAKNN---	
ScFAA2	LVYHRVLIDKIRDSLISNNSFIITGSAPISKDTLLFLRSALDIGIRQGLITFAGVGLSEP-FEKDVGSCAIGISAECRKSVPEMGYHADKD--L	
ScFAA3	GFLSRMVFKKVRQATIG-HLKYIMNGGSAISIDAQKFFSIVL-CPMIIGLITVANACVLEP-DHFEYGIVDLVGSVTAKVDVKDLGYAKNN---	
ScFAA4	GLLSGLIFKRIREATIG-NLRFILNGGSAISIDAQKFLSNLL-CPMLIGLITGVANACVLEP-EHFDYGIADLVGTITAKVDVEDLGYFAKNN---	
	601	700
MaACS-3	RLILIRPNVMKGYYKPGASTAIDEEGFHSIELGTFHSNITLDVLGKKIKTKSAVGSPS-----	
MaACS-4	RLICVRPGVIPGILDEAKKETIDEEGLHSDIAIMSGKTVTIIDRKLVFKLSQGEYIAAENIEGRFLSKVPFIQQILVHGDSTESCLVAILIP	
MaACS-5	RLIMLRHSVFKGYAPKQTEETLDAQGASTDVGEWDERRLVVIDRVNIFKLAQGEYIAPEKIEAVLAKHY-LVAQVFVYGDSDFOATLVGVVVP	
MaACS-6	RLIMLRNSVFKGYAPKQTEETLDADGASTDVGEWDAQRLVVIDRVNIFKLAQGEYIAPEKIEAVLAKHF-LVAQIFVYGHSLQATIVAVVVP	
MaACS-8	RLIWRIDTVFKGIFDEKNSETIDSEGLATDIGFVDKRCTTIIDRKLVFKLSQGEYIAPEKIEAVLAKHF-LVAQIFVYGHSLQATIVAVVVP	
MaACS-10	SLIWRIPAVMKGFLEEEKKVMTGDGFKTDIGTMNPDTLISKDRVNLVKLSHGGEYIALEKLEAVYRDSKE-IKSICIVADNGCPVLLAVVEP	
MaACS-11	RLIWRIPPSITSGIFNAEE SAAITEDRLKTDIGEWHADTLSDIDRKLVVKLSHGGEYIALEKLESVYKSTAY-CNNICVYADSMQNKPVAILVA	
MaACS-12	RLVCIRPSITAGFNPEESATLTADRLKTDIGEWHPDTISIIDRKLVVKLSHGGEYIALEKLESVYKSTAY-CNNICVYADSMQNKPVAILVA	
ScFAA1	QVWITANVTPEYNEEE SQALTSDDGFKTDIGEWANHLKIIDRKLVKTMNGEYIALEKLESVYRSNEY-VANICVYADQSKTKPVGIIVP	
ScFAA2	KQLQIRPQVFERFNPNESKAVDQDGFSTDVAFIDGKRISVIDRVNFFKLHGEYIAPEKIEAVLAKHY-LVAQVFVYGDSDFOATLVGVVVP	
ScFAA3	QALLKAPVCSEYNPIEAVSFTYDGFRTDIVETPKQLKIIDRRNLVKLTNGEYIALEKLESVYRSNSY-VKNICVYADESRVKPVGIIVP	
ScFAA4	QALLFKAPICSEYNPEEAAAFTDDGFRTDIAEWTPKQVKIIDRKLVKLTNGEYIALEKLESVYRSNPY-VQNICVYADENKVKPVGIIVP	
	701	800
MaACS-3	-----	
MaACS-4	EPEAFIPFVNKVLENVNLQPGDLAAYRKIVNNPKLRQAVLKELIKAGKDAGLKGFEIPKAILLESEAFVENDKMTPTFKIKRHPVVQAYREQLTALYNE	
MaACS-5	DAETLKPWADDHGL-----GGKSYEELCAHPAVKETLLKELKEFGRENDLKGFEIPKNIHVISEQFSIENDLLTPTFKLRHRTAKEKYIAEIELMYNG	
MaACS-6	DAETLKLWAKENKL-----GDKSYEELCALPQLRTTLQELATFGKESDLKGFEIPKNIHVISEQFSIENDLLTPTFKLRHAAKEKYNAEIDRMAYE	
MaACS-8	EPETFLPFANA IAG-ASVTAGDVEGLNKLQDPKVKI AVLKELEKAGKAGAMRGFEFVKRVHLTDAFSVDNGMMTPTFKVRPQVAEHFREQITAMYKE	
MaACS-10	S-----HAGASDKEILDILKSQAKAAGLSKSETVOGVIIDDSQWMT-NGFMTSSSKVKRREVRKAHNKDIEEMWKK	
MaACS-11	SEPRILELAKAKGLESR-----DFAVLCHDKVLIKAVLDACLATAKKAGLKPAELLOGVYLESEEWTAQGGLLTAAQKLKRKEINQAYADQIKQIYGS	
MaACS-12	SEPRILELAKAKGIESR-----DFAALCHDKVLIKAVHDACLATAKRAGLKPAEMLOGVYLESEEWTAQAGMLTAAQKLKRKEINQAYYSQIKQLYGT	
ScFAA1	NHAPLTKLAKKLGI MEQKDS--SINTENYLEDAKLKAVYSDLLKTGKDQGLVGIELLAGIVFFDGEWTPQNGFVTSQAQKLKRKIDILNAVKDKVDVYSS	
ScFAA2	DVDAAQPI LAAKHPEVKTWT-KEVLVENLNRNKKLRKEFLNKINKCT--DGLOGFEKLHNKVGLEPLTLEDVVPTPTFKIKRAKASKFFKDTLDQLYAE	
ScFAA3	NPGLSKFAVKLRIMKKGE-----DIENYIHDKALRNAVFKEMIATAKSOGLVGIELLCGIVFFDEEWTPENGFTVSAQKLKRREILA AVKSEVERVYKE	
ScFAA4	NLGHLSKLAIELGIMVPG-----DVESYIHEKKLODAVCKDMLSTAKSOGLNGIELLCGIVFFEEWTPENGLVTSQAQKLKRREILA AVKPDVERVYKE	

Figure 25C

801
MaACS-3 -----
MaACS-4 IHQKESKL--
MaACS-5 IH-----
MaACS-6 IA-----
MaACS-8 INASTPVAKL
MaACS-10 F-----
MaACS-11 K-----
MaACS-12 A-----
ScFAA1 S-----
ScFAA2 GSLVKTEKL-
ScFAA3 NS-----
ScFAA4 NT-----

Figure 26A

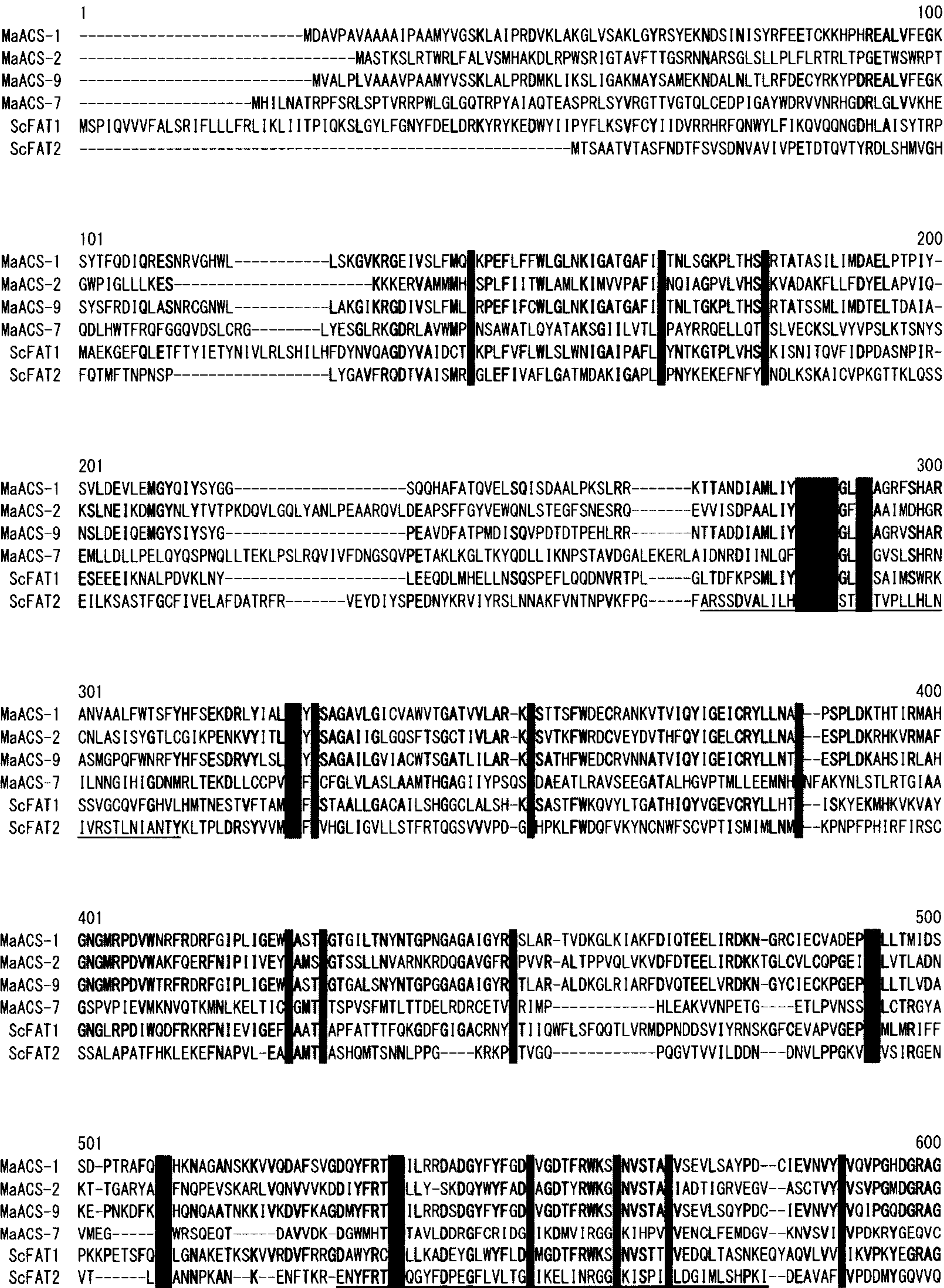


Figure 26B

	601		700
MaACS-1	MAAIVSKDTMWNDS-----FAKFALKNLPKYSVIFIRKVP---EMEITGTFKQRKELVNEGMDPSKIKDEMLWLDG-HSYRPF		
MaACS-2	MAALVLKNSIVQMAGGSQAKFHVDEAALNAFLRDLSKDVVKKLPAYAIRFLRIAEQ--ELETTGTFKNKKELKKEGFDLGKVKERLYWWTPKGEYAPF		
MaACS-9	MAAIVSKSTMDWEK-----FAAYALKNLPKYSVIFIRKMP---EMEITGTFKQRKELVNEGIDPKTIANEMLWLDG-HHYKPF		
MaACS-7	AWISTKDGTKVSLEAVQK-----FCEGKIAH--YKVRYVVVVESNEFPTTPSGKIQKNMRELTKAKLQLP-----		
ScFAT1	FAVIKLTDSLDTAKTKLLN-----DSLRLNLPYAMLFVKFVDE--IKMTDNHKILKKYREQLPKGLDGNDTIFWLKNYKRYEVL		
ScFAT2	AAIVLKKGEKMTYEELVN-----FLKKHLAS--FKITKVYFVDK--LPKTATGKIQRRIAETFAKSSRNKSKL-----		
	701	726	
MaACS-1	KEAEHTRVVSGKARL-RIKLFRFVR-		
MaACS-2	GVAENEQILAGRARL-----		
MaACS-9	KAAEHQRVISGKAKL-----		
MaACS-7	-----		
ScFAT1	TAADWEAIDAQTIKL-----		
ScFAT2	-----		

Figure 27

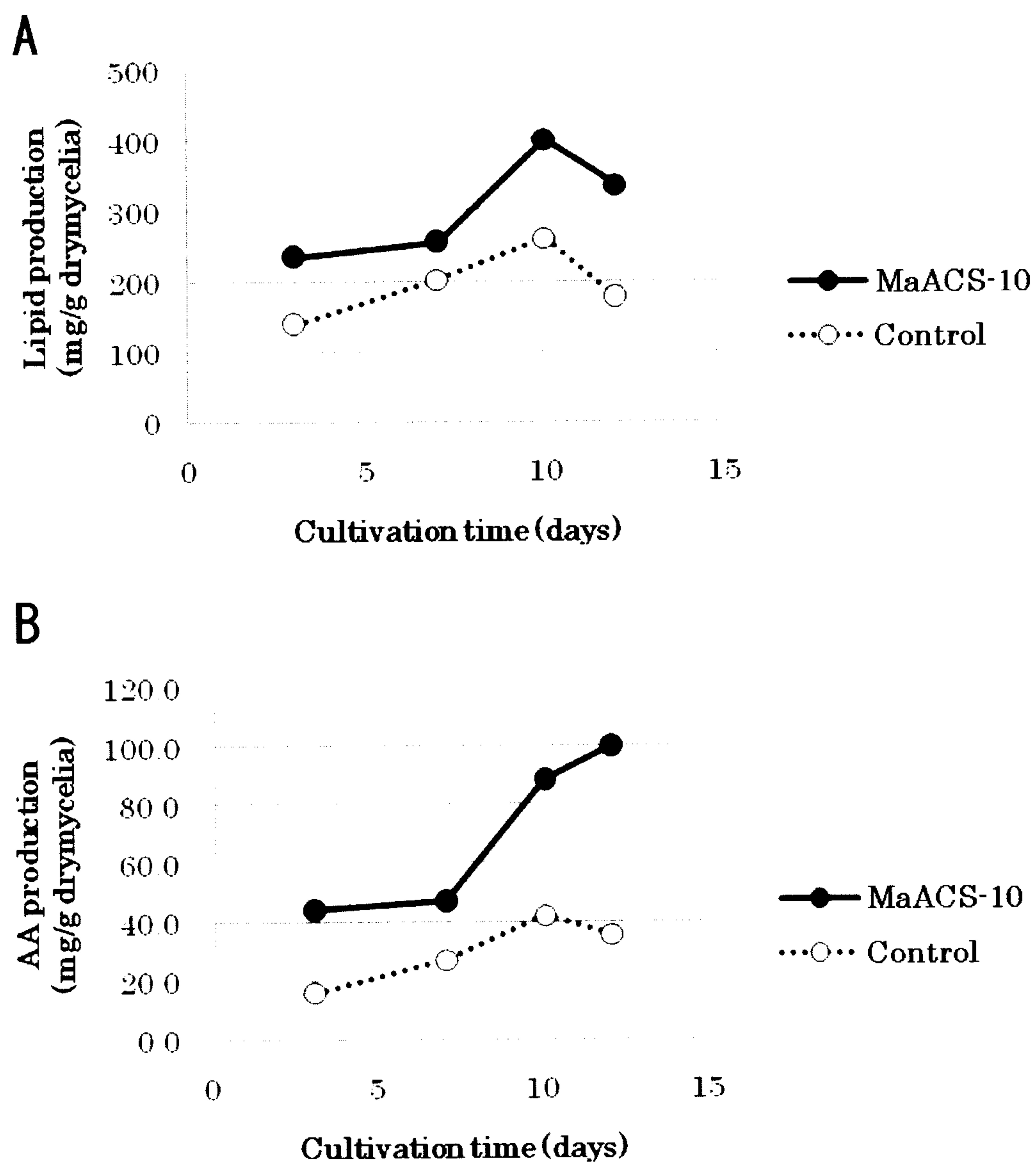


Figure 28

