



Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr.

Patent application No.

Demande de brevet n°

06120309.7 / EP06120309

The organization code and number of your priority application, to be used for filing abroad under the Paris Convention, is EP06120309

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

R.C. van Dijk



Anmeldung Nr:  
Application no.: 06120309.7  
Demande no:

Anmeldetag:  
Date of filing: 07.09.06  
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

BASF Plant Science GmbH  
Carl-Bosch-Strasse 64  
67117 Limburgerhof/DE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:  
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.  
If no title is shown please refer to the description.  
Si aucun titre n'est indiqué se référer à la description.)

### Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren

In anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)  
Staat/Tag/Aktenzeichen / State/Date/File no. / Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation / International Patent Classification / Classification internationale de brevets:

C12N1/00

Am Anmeldetag benannte Vertragsstaaten / Contracting states designated at date of filing / Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

## Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren

Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure in transgenen Pflanzen, indem in  
5 der Pflanze bereitgestellt werden mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität kodiert; mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert; mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität kodiert; und mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert,  
10 wobei die Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist.

15 In einer bevorzugten Ausführungsform werden zusätzlich weitere Nukleinsäuresequenzen, die für ein Polypeptid mit der Aktivität einer  $\omega 3$ -Desaturase und/oder einer  $\Delta 4$ -Desaturase kodieren, in der Pflanze bereitgestellt.

20 In einer weiteren bevorzugten Ausführungsform werden weitere Nukleinsäuresequenzen, die für Acyl-CoA-Dehydrogenase(n), Acyl-ACP(= acyl carrier protein)-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Acyl-CoA:Lysophospholipid-Acyltransferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) kodieren, in der Pflanze bereitgestellt.

Die Erfindung betrifft weiterhin rekombinante Nukleinsäuremoleküle, umfassend mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität  
30 kodiert; mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 5$ -

Desaturase-Aktivität kodiert; mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert; und mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert und die gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist.

Ein weiterer Teil der Erfindung betrifft Öle, Lipide und/oder Fettsäuren, die nach dem erfindungsgemäßen Verfahren hergestellt wurden, und deren Verwendung.

Schließlich betrifft die Erfindung auch transgene Pflanzen, die nach dem erfindungsgemäßen Verfahren hergestellt wurden oder die ein erfindungsgemäßes rekombinantes Nukleinsäuremolekül enthalten, und deren Verwendung als Nahrungs- oder Futtermittel.

Die Lipidsynthese lässt sich in zwei Abschnitte unterteilen: die Synthese von Fettsäuren und ihre Bindung an sn-Glycerin-3-Phosphat sowie die Addition oder Modifikation einer polaren Kopfgruppe. Übliche Lipide, die in Membranen verwendet werden, umfassen Phospholipide, Glycolipide, Sphingolipide und Phosphoglyceride. Die Fettsäuresynthese beginnt mit der Umwandlung von Acetyl-CoA in Malonyl-CoA durch die Acetyl-CoA-Carboxylase oder in Acetyl-ACP durch die Acetyltransacylase. Nach einer Kondensationsreaktion bilden diese beiden Produktmoleküle zusammen Acetoacetyl-ACP, das über eine Reihe von Kondensations-, Reduktions- und Dehydratisierungsreaktionen umgewandelt wird, so dass ein gesättigtes Fettsäuremolekül mit der gewünschten Kettenlänge erhalten wird. Die Produktion der ungesättigten Fettsäuren aus diesen Molekülen wird durch spezifische Desaturasen katalysiert, und zwar entweder aerob mittels molekularem Sauerstoff oder anaerob (bezüglich der Fettsäuresynthese in Mikroorganismen siehe F.C. Neidhardt et al. (1996) *E. coli* und *Salmonella*. ASM Press: Washington, D.C., S. 612-636 und darin enthaltene Literaturstellen; Lengeler et al. (Hrsgb.) (1999) *Biology of Prokaryotes*. Thieme: Stuttgart, New York, und die enthaltene Literaturstellen, sowie Magnuson, K., et al. (1993) *Microbiological Reviews* 57:522-542 und die enthaltenen Literaturstellen). Die so hergestellten an Phospholipide gebundenen Fettsäuren müssen anschließend für die weiteren Elongationen aus den Phospholipiden wieder in den Fett-



säureCoA-Ester-Pool überführt werden. Dies ermöglichen Acyl-CoA:Lysophospholipid-Acyltransferasen. Weiterhin können diese Enzyme die elongierten Fettsäuren wieder von den CoA-Estern auf die Phospholipide übertragen. Diese Reaktionsabfolge kann gegebenenfalls mehrfach durchlaufen werden.

5

Ferner müssen Fettsäuren anschließend an verschiedene Modifikationsorte transportiert und in das Triacylglycerin-Speicherlipid eingebaut werden. Ein weiterer wichtiger Schritt bei der Lipidsynthese ist der Transfer von Fettsäuren auf die polaren Kopfgruppen, beispielsweise durch Glycerin-Fettsäure-Acyltransferase (siehe Frentzen, 1998, Lipid, 100(4-5):161-166).

10

Eine Übersicht über die Pflanzen-Fettsäurebiosynthese, Desaturierung, den Lipidstoffwechsel und Membrantransport von fetthaltigen Verbindungen, die Betaoxidation, Fettsäuremodifikation und Cofaktoren, Triacylglycerin-Speicherung und –Assemblierung geben einschließlich der Literaturstellen die folgenden Artikel: Kinney (1997) Genetic Engineering, Hrsgb.: JK Setlow, 19:149-166; Ohlrogge und Browse (1995) Plant Cell 7:957-970; Shanklin und Cahoon (1998) Annu. Rev. Plant Physiol. Plant Mol. Biol. 49:611-641; Voelker (1996) Genetic Engineering, Hrsgb.: JK Setlow, 18:111-13; Gerhardt (1992) Prog. Lipid R. 31:397-417; Gühnemann-Schäfer & Kindl (1995) Biochim. Biophys Acta 1256:181-186; Kunau et al. (1995) Prog. Lipid Res. 34:267-342; Stymne et al. (1993) in: Biochemistry and Molecular Biology of Membrane and Storage Lipids of Plants, Hrsgb.: Murata und Somerville, Rockville, American Society of Plant Physiologists, 150-158; Murphy & Ross (1998) Plant Journal. 13(1):1-16.

15

20

25

Die mehrfach ungesättigten Fettsäuren können entsprechend ihrem Desaturierungsmuster in zwei große Klassen, die  $\omega$ -6- und die  $\omega$ -3-Fettsäuren, eingeteilt werden, die metabolisch und funktionell unterschiedliche Aktivitäten haben.

30

Im Folgenden werden mehrfach ungesättigte Fettsäuren als PUFA, PUFAs, LCPUFA oder LCPUFAs bezeichnet (poly unsaturated fatty acids, PUFA, mehrfach ungesättigte Fettsäuren; long chain poly unsaturated fatty acids, LCPUFA, langkettige mehrfach ungesättigte Fettsäuren).

Als Ausgangsstoff für den  $\omega$ -6-Stoffwechselweg fungiert die Fettsäure Linolsäure ( $18:2^{\Delta 9,12}$ ), während der  $\omega$ -3-Weg über Linolensäure ( $18:3^{\Delta 9,12,15}$ ) abläuft. Linolensäure wird dabei durch die Aktivität einer  $\omega$ -3-Desaturase aus Linolsäure gebildet (Tocher et al. (1998) Prog. Lipid Res. 37: 73-117 ; Domergue et al. (2002) Eur. J. Biochem. 269: 4105-4113).

Säugetiere und damit auch der Mensch verfügen über keine entsprechende Desaturaseaktivität ( $\Delta$ -12- und  $\omega$ -3-Desaturase) zur Bildung dieser Ausgangsstoffe und müssen daher diese Fettsäuren (essentielle Fettsäuren) über die Nahrung aufnehmen. Über eine Abfolge von Desaturase- und Elongase-Reaktionen werden dann aus diesen Vorstufen die physiologisch wichtigen mehrfach ungesättigten Fettsäuren Arachidonsäure (= ARA,  $20:4^{\Delta 5,8,11,14}$ ), eine  $\omega$ -6-Fettsäure und die beiden  $\omega$ -3-Fettsäuren Eicosapentaen- (= EPA,  $20:5^{\Delta 5,8,11,14,17}$ ) und Docosahexaensäure (DHA,  $22:6^{\Delta 4,7,10,13,17,19}$ ) synthetisiert.

Die Verlängerung von Fettsäuren durch Elongasen um 2 bzw. 4 C-Atome ist für die Produktion von  $C_{20}$ - bzw.  $C_{22}$ -PUFAs von entscheidender Bedeutung. Dieser Prozess verläuft über 4 Stufen. Den ersten Schritt stellt die Kondensation von Malonyl-CoA an das Fettsäure-Acyl-CoA durch die Ketoacyl-CoA-Synthase (KCS, im weiteren Text als Elongase bezeichnet) dar. Es folgt dann ein Reduktionsschritt (Ketoacyl-CoA-Reduktase, KCR), ein Dehydratationsschritt (Dehydratase) und ein abschließender Reduktionsschritt (Enoyl-CoA-Reduktase). Es wurde postuliert, dass die Aktivität der Elongase die Spezifität und Geschwindigkeit des gesamten Prozesses beeinflusst (Millar and Kunst (1997) Plant Journal 12:121-131).

Für Fettsäuren und Triacylglyceride besteht eine Vielzahl von Anwendungen in der Lebensmittelindustrie, der Tierernährung, der Kosmetik und im Pharmabereich.

Je nachdem, ob es sich um freie gesättigte und ungesättigte Fettsäuren oder um Triacylglyceride mit einem erhöhten Gehalt an gesättigten oder ungesättigten Fettsäuren handelt, sind sie für die unterschiedlichsten Anwendungen geeignet. So werden z.B. in der humanen Ernährung Lipide mit ungesättigten, speziell mehrfach ungesättigten, Fettsäuren bevorzugt. Den mehrfach ungesättigten  $\omega$ -3-Fettsäuren wird dabei ein positiver

Effekt auf den Cholesterinspiegel im Blut und damit auf die Prävention einer Herzerkrankung zugeschrieben. Durch Zugabe dieser  $\omega$ -3-Fettsäuren zur Nahrung kann das Risiko einer Herzerkrankung, eines Schlaganfalls oder von Bluthochdruck deutlich verringert werden (Shimikawa (2001) World Rev. Nutr. Diet. 88: 100-108).

5

Auch entzündliche, speziell chronisch entzündliche, Prozesse im Rahmen immunologischer Erkrankungen wie rheumatoider Arthritis lassen sich durch  $\omega$ -3-Fettsäuren positiv beeinflussen (Calder (2002) Proc. Nutr. Soc. 61: 345-358; Cleland und James (2000) J. Rheumatol. 27: 2305-2307). Sie werden deshalb Lebensmitteln, speziell diätetischen Lebensmitteln, zugegeben oder finden in Medikamenten Anwendung.

10

$\omega$ -6-Fettsäuren wie Arachidonsäure üben bei diesen rheumatischen Erkrankungen eher einen negativen Effekt aus.

$\omega$ -3- und  $\omega$ -6-Fettsäuren sind Vorläufer von Gewebshormonen, den sogenannten

15

Eicosanoiden wie den Prostaglandinen, die sich von der Dihomo- $\gamma$ -linolensäure, der Arachidonsäure und der Eicosapentaensäure ableiten, und den Thromboxanen und Leukotrienen, die sich von der Arachidonsäure und der Eicosapentaensäure ableiten.

Eicosanoide (sog. PG<sub>2</sub>-Serie), die aus  $\omega$ -6-Fettsäuren gebildet werden, fördern in der Regel Entzündungsreaktionen, während Eicosanoide (sog. PG<sub>3</sub>-Serie) aus  $\omega$ -3-

20

Fettsäuren geringe oder keine entzündungsfördernde Wirkung haben.

Mehrfach ungesättigte langkettige  $\omega$ -3-Fettsäuren wie Eicosapentaensäure (= EPA, C<sub>20</sub>:5 $\Delta$ <sup>5,8,11,14,17</sup>) oder Docosahexaensäure (= DHA, C<sub>22</sub>:6 $\Delta$ <sup>4,7,10,13,16,19</sup>) sind wichtige

25

Komponenten der menschlichen Ernährung aufgrund ihrer verschiedenen Rollen in der Gesundheit, die Aspekte wie die Entwicklung des kindlichen Gehirns, die Funktionalität des Auges, die Synthese von Hormonen und anderer Signalstoffe, sowie die Vorbeugung von Herz-Kreislauf-Beschwerden, Krebs und Diabetes umfassen (Poulos, A (1995) Lipids 30:1-14; Horrocks, LA und Yeo YK (1999) Pharmacol Res 40:211-225).

30

Aufgrund der heute üblichen Zusammensetzung der menschlichen Nahrung ist ein Zusatz von mehrfach ungesättigten  $\omega$ -3-Fettsäuren, die bevorzugt in Fischölen vorkommen, zur Nahrung besonders wichtig. So werden beispielsweise mehrfach ungesättigte

Fettsäuren wie Docosahexaensäure (= DHA, C22:6 $\Delta^{4,7,10,13,16,19}$ ) oder Eicosapentaensäure (= EPA, C20:5 $\Delta^{5,8,11,14,17}$ ) der Babynahrung zur Erhöhung des Nährwertes zugesetzt. Es besteht aus diesem Grund ein Bedarf an der Produktion mehrfach ungesättigter langkettiger Fettsäuren.

5

Hauptsächlich werden die verschiedenen Fettsäuren und Triglyceride aus Mikroorganismen wie *Mortierella* oder *Schizochytrium* oder aus Öl-produzierenden Pflanzen wie Soja, Raps und Algen wie *Cryptocodinium* oder *Phaeodactylum* und weiteren gewonnen, wobei sie in der Regel in Form ihrer Triacylglyceride (= Triglyceride = Tri-  
10 glycerole) anfallen. Sie können aber auch aus Tieren wie z.B. Fischen gewonnen werden. Die freien Fettsäuren werden vorteilhaft durch Verseifung der Triacylglyceride hergestellt. Sehr langkettige mehrfach ungesättigte Fettsäuren wie DHA, EPA, Arachidonsäure (ARA, C20:4 $\Delta^{5,8,11,14}$ ), Dihomo- $\gamma$ -linolensäure (DHGL, C20:3 $\Delta^{8,11,14}$ ) oder Docosapentaensäure (DPA, C22:5 $\Delta^{7,10,13,16,19}$ ) werden in Ölfruchtpflanzen wie Raps, Soja, Sonnenblume, Färbersaflor jedoch nicht synthetisiert. Übliche natürliche Quellen für diese  
15 Fettsäuren sind Fische wie Hering, Lachs, Sardine, Goldbarsch, Aal, Karpfen, Forelle, Heilbutt, Makrele, Zander oder Thunfisch oder Algen.

Aufgrund der positiven Eigenschaften der mehrfach ungesättigten Fettsäuren hat es in  
20 der Vergangenheit nicht an Ansätzen gefehlt, Gene, die an der Synthese dieser Fettsäuren bzw. Triglyceride beteiligt sind, für die Herstellung von Ölen in verschiedenen Organismen mit geändertem Gehalt an ungesättigten Fettsäuren verfügbar zu machen. So wird in WO 91/13972 und seinem US-Äquivalent eine  $\Delta$ -9-Desaturase beschrieben. In WO 93/11245 wird eine  $\Delta$ -15-Desaturase und in WO 94/11516 eine  $\Delta$ -12-Desaturase  
25 beansprucht. Weitere Desaturasen werden beispielsweise in EP-A-0 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP-A-0 794 250, Stukey et al. (1990) J. Biol. Chem., 265: 20144-20149, Wada et al. (1990) Nature 347: 200-203 oder Huang et al. (1999) Lipids 34: 649-659 beschrieben. Die biochemische Charakterisierung der verschiedenen Desaturasen ist jedoch bisher nur unzureichend erfolgt, da  
30 die Enzyme als membrangebundene Proteine nur sehr schwer zu isolieren und zu charakterisieren sind (McKeon et al. (1981) Methods in Enzymol. 71: 12141-12147, Wang et al. (1988) Plant Physiol. Biochem., 26: 777-792).

In der Regel erfolgt die Charakterisierung membrangebundener Desaturasen durch Einbringung in einen geeigneten Organismus, der anschließend auf Enzymaktivität mittels Edukt- und Produktanalyse untersucht wird.  $\Delta$ -6-Desaturasen werden in WO 93/06712, US 5,614,393, WO 96/21022, WO 00/21557 und WO 99/27111 beschrieben. Die Anwendung dieses Enzyms zur Produktion von Fettsäuren in transgenen Organismen wird in WO 98/46763, WO 98/46764 und WO 98/46765 beschrieben. Die Expression verschiedener Desaturasen und die Bildung mehrfach ungesättigter Fettsäuren wird auch in WO 99/64616 oder WO 98/46776 beschrieben und beansprucht. Bzgl. der Effektivität der Expression von Desaturasen und ihrem Einfluss auf die Bildung mehrfach ungesättigter Fettsäuren ist anzumerken, dass durch Expression einer einzelnen Desaturase wie bisher beschrieben lediglich geringe Gehalte an ungesättigten Fettsäuren/Lipiden wie z.B.  $\gamma$ -Linolensäure und Stearidonsäure erreicht wurden.

In der Vergangenheit wurden zahlreiche Versuche unternommen, Elongase-Gene zu erhalten. Millar and Kunst (1997) Plant Journal 12:121-131 und Millar et al. (1999) Plant Cell 11:825-838 beschreiben die Charakterisierung von pflanzlichen Elongasen zur Synthese von einfach ungesättigten langkettigen Fettsäuren (C22:1) bzw. zur Synthese von sehr langkettigen Fettsäuren für die Wachsbildung in Pflanzen (C<sub>28</sub>-C<sub>32</sub>). Beschreibungen zur Synthese von Arachidonsäure und EPA finden sich beispielsweise in WO 01/59128, WO 00/12720, WO 02/077213 und WO 02/08401. Die Synthese von mehrfach ungesättigter C24-Fettsäuren ist beispielsweise in Tvrdik et al. (2000) J. Cell Biol. 149: 707-718 oder in WO 02/44320 beschrieben.

Besonders geeignete Mikroorganismen zur Herstellung von PUFAs sind Mikroalgen wie Phaeodactylum tricornutum, Porphyridium-Arten, Thraustochytrien-Arten, Schizochytrien-Arten oder Crypthecodinium-Arten, Ciliaten, wie Stylonychia oder Colpidium, Pilze, wie Mortierella, Entomophthora oder Mucor und/oder Moose wie Physcomitrella, Ceratodon und Marchantia (R. Vazhappilly & F. Chen (1998) Botanica Marina 41: 553-558; K. Totani & K. Oba (1987) Lipids 22: 1060-1062; M. Akimoto et al. (1998) Appl. Biochemistry and Biotechnology 73: 269-278). Durch Stammselektion ist eine Anzahl von Mutantstämmen der entsprechenden Mikroorganismen entwickelt worden, die eine Reihe wün-

schenswerter Verbindungen, einschließlich PUFAs, produzieren. Die Mutation und Selektion von Stämmen mit verbesserter Produktion eines bestimmten Moleküls wie den mehrfach ungesättigten Fettsäuren ist jedoch ein zeitraubendes und schwieriges Verfahren. Mit Hilfe der vorgenannten Mikroorganismen lassen sich zudem nur begrenzte Mengen der gewünschten mehrfach ungesättigten Fettsäuren wie DPA, EPA oder ARA herstellen, die noch dazu in der Regel als Fettsäuregemische anfallen. Deshalb werden, wann immer möglich, gentechnologische Verfahren bevorzugt.

Höhere Pflanzen enthalten mehrfach ungesättigte Fettsäuren wie Linolsäure (C18:2) und Linolensäure (C18:3). ARA, EPA und DHA kommen im Samenöl höherer Pflanzen gar nicht oder nur in Spuren vor (E. Ucciani: Nouveau Dictionnaire des Huiles Végétales. Technique & Documentation – Lavoisier, 1995. ISBN: 2-7430-0009-0). Es wäre jedoch vorteilhaft, in höheren Pflanzen, bevorzugt in Ölsaaten wie Raps, Lein, Sonnenblume und Soja, LCPUFAs herzustellen, da auf diese Weise große Mengen qualitativ hochwertiger LCPUFAs für die Lebensmittelindustrie, die Tierernährung und für pharmazeutische Zwecke kostengünstig gewonnen werden können. Hierzu werden vorteilhafterweise über gentechnische Methoden Gene, die für Enzyme der Biosynthese von LCPUFAs kodieren, in Ölsaaten eingeführt und exprimiert. Dies sind Gene, die beispielsweise für  $\Delta$ -6-Desaturasen,  $\Delta$ -6-Elongasen,  $\Delta$ -5-Desaturasen oder  $\Delta$ -4-Desaturasen kodieren. Diese Gene können vorteilhaft aus Mikroorganismen und niederen Pflanzen isoliert werden, die LCPUFAs herstellen und in den Membranen oder Triacylglyceriden einbauen. So konnten bereits  $\Delta$ -6-Desaturase-Gene aus dem Moos *Physcomitrella patens* und  $\Delta$ -6-Elongase-Gene aus *P. patens* und dem Nematoden *C. elegans* isoliert werden.

Transgene Pflanzen, die für Enzyme der LCPUFA-Biosynthese kodierende Gene enthalten und exprimieren und als Folge dessen LCPUFAs produzieren, wurden beispielsweise in DE-A-102 19 203 (Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren in Pflanzen) beschrieben. Diese Pflanzen produzieren allerdings LCPUFAs in Mengen, die für eine Aufarbeitung der in den Pflanzen enthaltenen Öle noch weiter optimiert werden müssen. So beträgt der Gehalt von ARA in den in DE-A-102 19 203 beschriebenen Pflanzen lediglich 0,4 bis 2% und der Gehalt von EPA lediglich 0,5 bis 1%, jeweils bezogen auf den Gesamtlipidgehalt der Pflanze.

Um eine Anreicherung der Nahrung und des Futters mit mehrfach ungesättigten, langkettigen Fettsäuren zu ermöglichen, besteht daher ein großer Bedarf an einem einfachen, kostengünstigen Verfahren zur Herstellung von mehrfach ungesättigten, langkettigen Fettsäuren speziell in pflanzlichen Systemen.

Eine Aufgabe der Erfindung ist es daher, ein Verfahren bereitzustellen, mit dem langkettige, mehrfach ungesättigte Fettsäuren, insbesondere Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure, in transgenen Pflanzen in großer Menge preiswert hergestellt werden können.

Es wurde nun überraschenderweise herausgefunden, dass durch die Expression einer optimierten  $\Delta 5$ -Elongase-Sequenz in transgenen Pflanzen die Ausbeute an langkettigen, mehrfach ungesättigten Fettsäuren, insbesondere Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure, gesteigert werden kann.

Die durch das erfindungsgemäße Verfahren hergestellten PUFAs umfassen eine Gruppe von Molekülen, die höhere Tiere nicht mehr synthetisieren können und somit aufnehmen müssen oder die höhere Tiere nicht mehr ausreichend selbst herstellen können und somit zusätzlich aufnehmen müssen, obwohl sie leicht von anderen Organismen, wie Bakterien, synthetisiert werden können.

Entsprechend wird die Aufgabe der Erfindung gelöst durch das erfindungsgemäße Verfahren zur Herstellung von Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure in einer transgenen Pflanze, umfassend das Bereitstellen in der Pflanze von mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität kodiert; mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert; mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität kodiert; und mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, wobei die Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die



Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist. Für die Produktion von DHA muss desweiteren mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 4$ -Desaturase-Aktivität kodiert, in der Pflanze bereitgestellt werden.

5

Das "Bereitstellen in der Pflanze" bedeutet im Sinne der vorliegenden Erfindung, dass Maßnahmen getroffen werden, so dass die Nukleinsäuresequenzen kodierend für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität, ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität, ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität und ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität zusammen in einer Pflanze vorliegen. Das "Bereitstellen in der Pflanze" umfasst somit das Einbringen der Nukleinsäuresequenzen in die Pflanze sowohl durch Transformation einer Pflanze mit einem oder mehreren rekombinanten Nukleinsäuremolekülen, die die genannten Nukleinsäuresequenzen enthalten, als auch durch Verkreuzung von geeigneten Elternpflanzen, die eine oder mehrere der genannten Nukleinsäuresequenzen enthalten.

10

15

Die Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, ist erfindungsgemäß gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist. Dies bedeutet, dass die Nukleinsäuresequenz gezielt für die Zwecke der Erfindung optimiert wurde, ohne dass dadurch die von der Nukleinsäuresequenz kodierte Aminosäuresequenz verändert wurde.

20

Der genetische Code ist redundant, da er 61 Kodons verwendet, um 20 Aminosäuren zu spezifizieren. Daher werden die meisten der 20 proteinogenen Aminosäuren von mehreren Tripletts (Kodons) kodiert. Die synonymen Kodons, die eine einzelne Aminosäure spezifizieren, werden in einem bestimmten Organismus jedoch nicht mit gleicher Häufigkeit verwendet, sondern es gibt bevorzugte Kodons, die häufig verwendet werden und Kodons, die seltener verwendet werden. Diese Unterschiede in der Kodonverwendung werden zurückgeführt auf selektive evolutionäre Drücke und vor allem die Effizienz der Translation. Ein Grund für die geringere Translationseffizienz von selten auftretenden

25

30



Kodons könnte darin liegen, dass die entsprechenden Aminoacyl-tRNA-Pools erschöpft werden und damit nicht mehr zur Proteinsynthese zur Verfügung stehen.

Außerdem bevorzugen unterschiedliche Organismen unterschiedliche Kodons. Daher  
5 läuft beispielsweise die Expression einer rekombinanten DNA, die aus einer Säugerzelle  
stammt, in E. coli-Zellen häufig nur suboptimal ab. Deshalb kann der Austausch selten  
verwendeter Kodons gegen häufig verwendete Kodons in manchen Fällen die Expressi-  
on erhöhen. Ohne an eine Hypothese gebunden sein zu wollen, wird angenommen, dass  
die kodonoptimierten DNA-Sequenzen eine effizientere Translation ermöglichen und die  
10 daraus gebildeten mRNAs möglicherweise eine höhere Halbwertszeit in der Zelle besit-  
zen und daher häufiger für die Translation zur Verfügung stehen. Aus dem vorstehend  
gesagten folgt, dass eine Kodonoptimierung nur dann nötig ist, wenn der Organismus, in  
dem die Nukleinsäuresequenz exprimiert werden soll, ein anderer ist als der Organis-  
mus, aus dem die Nukleinsäuresequenz ursprünglich stammt.

15

Für viele Organismen, von denen die DNA-Sequenz einer größeren Zahl von Genen be-  
kannt ist, gibt es Tabellen, denen man die Häufigkeit der Verwendung bestimmter Ko-  
dons in dem jeweiligen Organismus entnehmen kann. Mit Hilfe dieser Tabellen lassen  
sich Proteinsequenzen mit relativ großer Genauigkeit in eine DNA-Sequenz zurück-  
20 übersetzen, die die im jeweiligen Organismus bevorzugten Kodons für die verschiedenen  
Aminosäuren des Proteins enthält. Tabellen zur Kodonverwendung können u.a. unter  
der folgenden Internet-Adresse aufgefunden werden:

<http://www.kazusa.or.jp/Kodon/E.html>. Darüber hinaus bieten mehrere Firmen Software  
für die Genoptimierung an, wie z.B. die Firma Entelechon (Software Leto) oder die Firma  
25 Geneart (Software GeneOptimizer).

Die Anpassung der Sequenzen an die Kodonverwendung in einem bestimmten Orga-  
nismus kann unter Zuhilfenahme verschiedener Kriterien erfolgen. Zum einen kann für  
eine bestimmte Aminosäure immer das am häufigsten im ausgewählten Organismus  
30 vorkommende Kodon verwendet werden, zum anderen kann aber auch die natürliche  
Frequenz der verschiedenen Kodons berücksichtigt werden, so dass alle Kodons für ei-  
ne bestimmte Aminosäure entsprechend ihrer natürlichen Häufigkeit in die optimierte

Sequenz eingebaut werden. Dabei kann die Auswahl, an welcher Position welches Basen-Triplett verwendet wird, zufällig stattfinden. Erfindungsgemäß wurde die DNA-Sequenz unter Berücksichtigung der natürlichen Häufigkeit einzelner Kodons angepasst, wobei die Verwendung des am häufigsten im ausgewählten Organismus vorkommenden  
5 Kodons ebenfalls geeignet ist.

Besonders bevorzugt ist eine Nukleinsäuresequenz aus *Ostreococcus tauri*, die für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, wie beispielsweise das in Seq ID No. 110 dargestellte Polypeptid, zumindest an die Kodonverwendung in Raps, Soja  
10 und/oder Lein angepasst. Bei der ursprünglich aus *Ostreococcus tauri* stammenden Nukleinsäuresequenz handelt es sich bevorzugt um die in Seq ID No. 109 )dargestellte Sequenz. Die für die  $\Delta 5$ -Elongase kodierende DNA-Sequenz ist an mindestens 20% der Positionen, bevorzugt an mindestens 30% der Positionen, besonders bevorzugt an mindestens 40% Positionen und am meisten bevorzugt an mindestens 50% der Positionen  
15 an die Kodonverwendung in Raps, Soja und/oder Lein angepasst.

Am meisten bevorzugt handelt es sich bei der verwendeten Nukleinsäuresequenz um die in SEQ ID No. 64 angegebene Sequenz.

20 Es versteht sich, dass auch solche kodonoptimierten DNA-Sequenzen von der Erfindung erfasst sind, die für ein Polypeptid mit der Aktivität einer  $\Delta 5$ -Elongase kodieren, dessen Aminosäuresequenz an einer oder mehreren Positionen gegenüber der Wildtyp-Sequenz verändert ist, das aber noch im wesentlichen die gleiche Aktivität aufweist wie das Wildtyp-Protein.

25 Bevorzugt ist die Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität kodiert, ausgewählt aus der Gruppe bestehend aus:

a) Nukleinsäuresequenzen mit der in Seq ID NO.

1,3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37,39 oder 41, bevorzugt mit der in

30 Seq ID No. 1 dargestellten Sequenz,

b) Nukleinsäuresequenzen, die für die in Seq ID No.

2,4,6,8,10,12,14,16,18,20,22,24,26,28,30,32,34,36,38,40 oder 42, bevorzugt in Seq ID No. 2 angegebene Aminosäuresequenz kodieren,

c) Nukleinsäuresequenzen, die mit dem komplementären Strang der a) oder b) oberhalb  
5 angegebenen Nukleinsäuresequenzen, insbesondere der in Seq ID No. 1 angegebenen Nukleinsäuresequenz unter stringenten Bedingungen hybridisieren,

d) Nukleinsäuresequenzen, die zu den in a) oder b) oberhalb angegebenen Nukleinsäuresequenzen, insbesondere zur der in Seq ID No. 1 angegebenen Sequenz zu mindestens 60%, 65%, 70%, 75% oder 80%, bevorzugt zu mindestens 81%, 82%, 83%, 84%,  
10 85%, 86%, 87%, 88%, 89% oder 90%, besonders bevorzugt zu mindestens 91%, 92%, 93%, 94% oder 95% und insbesondere zu mindestens 96%, 97%, 98% oder 99% identisch sind, und

e) Nukleinsäuresequenzen, die für eine Aminosäuresequenz kodieren, die mindestens eines, beispielsweise 2, 3, 4, 5, 6, 7 oder 8, vorzugsweise alle der in Seq ID No.

15 43,44,45,46,47,48,49 oder 50 angegebenen Aminosäurepattern aufweisen.

Unter Aminosäurepattern sind kurze Aminosäuresequenzen zu verstehen, die vorzugsweise weniger als 50, besonders bevorzugt weniger als 40 und insbesondere von 10 bis 40 und noch weiter bevorzugt von 10 bis 30 Aminosäuren umfassen.

20 Für die vorliegende Erfindung wird die Identität vorzugsweise über die Volllänge der erfindungsgemäßen Nukleotid- oder Aminosäuresequenzen ermittelt, beispielsweise für die in SEQ ID NO: 64 angegebene Nukleinsäuresequenz über die Volllänge von 903 Nukleotiden.

25 Bevorzugt ist die Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert, ausgewählt aus der Gruppe bestehend aus:

a) Nukleinsäuresequenzen mit der in Seq ID No. 171, 173,175,177,179,181 oder 183, insbesondere mit der in Seq ID No. 171 dargestellten Sequenz,

30 b) Nukleinsäuresequenzen, die für die in Seq ID No. 172, 174,176,178,180,182 oder 184, insbesondere für die in Seq ID No. 172 angegebene Aminosäuresequenz kodieren,

- c) Nukleinsäuresequenzen, die mit dem komplementären Strang der a) oder b) oberhalb angegebenen Nukleinsäuresequenzen, insbesondere der in Seq ID No. 1 angegebenen Nukleinsäuresequenz unter stringenten Bedingungen hybridisieren,
- d) Nukleinsäuresequenzen, die zu den in a) oder b) oberhalb angegebenen Nukleinsäuresequenzen, insbesondere zu der in Seq ID No. 171 angegebenen Sequenz zu mindestens 60%, 65%, 70%, 75% oder 80%, bevorzugt zu mindestens 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% oder 90%, besonders bevorzugt zu mindestens 91%, 92%, 93%, 94% oder 95% und insbesondere zu mindestens 96%, 97%, 98% oder 99% identisch sind, und
- 5 e) Nukleinsäuresequenzen, die für eine Aminosäuresequenz kodieren, die mindestens eines, beispielsweise 2, 3, 4, 5, 6, 7 oder 8, vorzugsweise alle der in Seq ID No. 185, 186, 187, 188, 189, 190, 191 oder 192 angegebenen Aminosäurepattern aufweisen.
- 10
- 15 Insbesondere handelt es sich bei der Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert ebenfalls um eine nach der vorliegenden Erfindung kodonoptimierten Sequenz, bevorzugt um die in der SEQ ID NO: 122 dargestellte Nukleinsäuresequenz.
- 20 Bevorzugt ist die Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität kodiert, ausgewählt aus der Gruppe bestehend aus:
- a) Nukleinsäuresequenzen mit der in Seq ID No. 51, 53 oder 55, bevorzugt mit der in Seq ID No. 51 dargestellten Sequenz,
- b) Nukleinsäuresequenzen, die für die in Seq ID No. 52, 54 oder 56, bevorzugt mit der in Seq ID No. 52 angegebenen Aminosäuresequenz kodieren,
- 25 c) Nukleinsäuresequenzen, die mit dem komplementären Strang der in a) oder b) oberhalb angegebenen Nukleinsäuresequenzen, insbesondere der in Seq ID No. 51 angegebenen Nukleinsäuresequenz unter stringenten Bedingungen hybridisieren, und
- 30 d) Nukleinsäuresequenzen, die zu den in a) oder b) oberhalb angegebenen Nukleinsäuresequenzen, insbesondere zur der in Seq ID No. 51 angegebenen Sequenz zu mindestens 60%, 65%, 70%, 75% oder 80%, bevorzugt zu mindestens 81%, 82%,

83%, 84%, 85%, 86%, 87%, 88%, 89% oder 90%, besonders bevorzugt zu mindestens 91%, 92%, 93%, 94% oder 95% und insbesondere zu mindestens 96%, 97%, 98% oder 99% identisch sind, und

- 5 e) Nukleinsäuresequenzen, die für eine Aminosäuresequenz kodieren, die mindestens eines, beispielsweise 2, 3, 4, 5, 6 oder 7, vorzugsweise alle der in Seq ID No. 57,58,59,60,61,62 oder 63 angegebenen Aminosäurepattern aufweisen.

Weitere geeignete Nukleinsäuresequenzen kann der Fachmann der Literatur bzw. den bekannten Genbanken wie z.B. <http://www.ncbi.nlm.nih.gov> entnehmen.

10

In einer weiteren bevorzugten Ausführungsform des Verfahrens werden zusätzlich eine oder mehrere Nukleinsäuresequenzen, die für ein Polypeptid mit der Aktivität einer  $\omega$ -3-Desaturase und/oder einer  $\Delta$ 4-Desaturase kodieren, in die Pflanze eingebracht.

- 15 Bevorzugt ist die Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\omega$ -3-Desaturase-Aktivität kodiert, ausgewählt aus der Gruppe bestehend aus:

- a) Nukleinsäuresequenzen mit der in Seq ID No. 193 oder 195, vorzugsweise der in Seq ID No. 193 dargestellten Sequenz,
- 20 b) Nukleinsäuresequenzen, die für die in Seq ID No. 194 angegebene Aminosäuresequenz kodieren,
- c) Nukleinsäuresequenzen, die mit dem komplementären Strang der in Seq ID No. 193 oder 195 angegebenen Nukleinsäuresequenz unter stringenten Bedingungen hybridisieren, und
- 25 d) Nukleinsäuresequenzen, die zu der in Seq ID No. 193 oder 195 angegebenen Sequenz zu mindestens 60%, 65%, 70%, 75% oder 80%, bevorzugt zu mindestens 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% oder 90%, besonders bevorzugt zu mindestens 91%, 92%, 93%, 94% oder 95% und insbesondere zu mindestens 96%, 97%, 98% oder 99% identisch sind.

- 30 Die im erfindungsgemäßen Verfahren vorteilhaft verwendete  $\omega$ -3-Desaturase ermöglicht eine Verschiebung vom  $\omega$ -6-Biosyntheseweg zum  $\omega$ -3-Biosyntheseweg, was zu einer Verschiebung von C<sub>18:2</sub>- zu C<sub>18:3</sub>-Fettsäuren führt. Weiterhin ist vorteilhaft, dass die  $\omega$ -3-

Desaturase eine breite Palette von Phospholipiden wie Phosphatidylcholin (= PC), Phosphatidylinositol (= PIS) oder Phosphatidylethanolamin (= PE) umsetzt. Schließlich lassen sich auch Desaturierungsprodukte in den Neutrallipiden (= NL), das heißt in den Triglyceriden finden.

5

Bevorzugt ist die Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 4$ -Desaturase-Aktivität kodiert, ausgewählt aus der Gruppe bestehend aus:

- a) Nukleinsäuresequenzen mit der in Seq ID No. 77, 79,81,83,85,87,89,91 oder 93, bevorzugt mit der in der Seq ID No. 77 dargestellten Sequenz,
- 10 b) Nukleinsäuresequenzen, die für die in Seq ID No. 78, 80,82,84,86,88,90,92 oder 94, bevorzugt für die in Seq ID No. 78 angegebene Aminosäuresequenz kodieren,
- c) Nukleinsäuresequenzen, die mit dem komplementären Strang der in a) oder b) oberhalb angegebenen Nukleinsäuresequenzen, insbesondere der in Seq ID No. 77 angegebenen Nukleinsäuresequenz unter stringenten Bedingungen hybridisieren,
- 15 und
- d) Nukleinsäuresequenzen, die zu der in Seq ID No. 77 angegebenen Sequenz zu mindestens 60%, 65%, 70%, 75% oder 80%, bevorzugt zu mindestens 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% oder 90%, besonders bevorzugt zu mindestens 91%, 92%, 93%, 94% oder 95% und insbesondere zu mindestens 96%, 97%,
- 20 98% oder 99% identisch sind, und
- e) Nukleinsäuresequenzen, die für eine Aminosäuresequenz kodieren, die mindestens eines, beispielsweise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 oder 14, vorzugsweise alle der in Seq ID No. 95,96,97,98,99,100,101,102,103,104,105,106,107 oder 108 angegebenen Aminosäurepattern aufweisen.

25

Die  $\Delta 4$ -Desaturase, die im erfindungsgemäßen Verfahren vorteilhaft verwendet wird, katalysiert die Einführung einer Doppelbindung in die Fettsäure Docosapentaensäure, was zur Bildung von Docosahexaensäure führt.

30

Für das erfindungsgemäße beschriebene Verfahren ist es vorteilhaft, in die Pflanzen zusätzlich zu den Nukleinsäuresequenzen, die für Polypeptide mit einer  $\Delta 6$ -Desaturase-Aktivität, einer  $\Delta 6$ -Elongase-Aktivität, einer  $\Delta 5$ -Desaturase-Aktivität und einer  $\Delta 5$ -

Elongase-Aktivität kodieren, sowie den ggf. eingebrachten Nukleinsäuresequenzen, die für ein Polypeptid mit einer  $\omega$ -3-Desaturase-Aktivität und/oder einer  $\Delta$ 4-Desaturase-Aktivität kodieren, zusätzlich weitere Nukleinsäuren einzubringen, die für Enzyme des Fettsäure- oder Lipidstoffwechsels kodieren.

5

Im Prinzip können alle Gene des Fettsäure- oder Lipidstoffwechsels in Kombination mit den im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen verwendet werden; bevorzugt werden Gene des Fettsäure- oder Lipidstoffwechsels ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP(= acyl carrier protein)-

10

Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Acyl-CoA:Lysophospholipid-Acyltransferasen, Fettsäure-Synthase(n), Fettsäure-

Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) in Kombination

15

mit der  $\Delta$ -6-Elongase,  $\Delta$ -6-Desaturase,  $\Delta$ -5-Desaturase und der  $\Delta$ -5-Elongase sowie ggf. der  $\omega$ -3-Desaturase und/oder der  $\Delta$ -4-Desaturase verwendet, wobei einzelne Gene oder mehrere Gene in Kombination verwendet werden können.

Vorteilhaft werden die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren im vegetativen Gewebe (= somatischem Gewebe) exprimiert. Unter vegetativem Gewebe ist im Sinne dieser Erfindung ein Gewebe zu verstehen, das sich durch mitotische Teilungen vermehrt. Derartiges Gewebe entsteht auch durch asexuelle Fortpflanzung (= Apomixis) und Vermehrung. Von Vermehrung spricht man dann, wenn sich die Zahl der Individuen in aufeinander folgenden Generationen erhöht. Diese durch asexuelle Vermehrung

25

entstandenen Individuen sind mit ihren Eltern weitestgehend identisch. Beispiele für derartige Gewebe sind Blatt, Blüte, Wurzel, Stengel, oberirdische oder unterirdische Ausläufer (Seitensprosse, Stolonen), Rhizome, Knospen, Knollen wie Wurzelknollen oder Ausläuferknollen, Zwiebel, Brutkörper, Brutknospen, Bulbillen oder Turione. Derartige Gewebe können auch durch unechte, echte oder durch den Mensch verursachte Viviparie

30

entstehen. Aber auch Samen, die durch Agamospermie, wie sie für Asteraceae, Poaceae oder Rosaceae typisch sind, entstanden sind, gehören zu den vegetativen Geweben, in denen vorteilhaft die Expression stattfindet. Zu einem geringeren Teil oder gar

nicht werden die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren im generativen Gewebe (Keimbahngewebe) exprimiert. Beispiele für derartige Gewebe sind Gewebe, die durch geschlechtliche Fortpflanzung, d.h. meiotische Zellteilungen entstehen, wie z.B. Samen, die durch geschlechtliche Prozesse entstanden sind. Unter zu einem geringen Teil ist zu verstehen, dass im Vergleich zum vegetativen Gewebe die Expression gemessen auf RNA- und/oder Proteinebene weniger als 5 %, vorteilhaft weniger als 3 %, besonders vorteilhaft weniger als 2 %, am meisten bevorzugt weniger als 1; 0,5; 0,25 oder 0,125 % beträgt.

10 Besonders bevorzugt werden die Nukleinsäuresequenzen in den Blättern der transgenen Pflanzen exprimiert. Dies hat den Vorteil, dass die erfindungsgemäß hergestellten LCPUFAs von Tieren und Menschen direkt durch den Verzehr der Blätter aufgenommen werden können und keine vorherige Aufarbeitung des Pflanzenmaterials erforderlich ist.

15 Die Expression der erfindungsgemäßen Nukleinsäuresequenzen im Blatt kann durch die Verwendung von konstitutiven oder blattspezifischen Promotoren erreicht werden.

"Konstitutive Promotoren" sind Promotoren, die die Expression in einer Vielzahl, vorzugsweise in allen, Geweben über einen wesentlichen Zeitraum während der pflanzlichen Entwicklung, bevorzugt während der gesamten pflanzlichen Entwicklung, ermöglichen. Bevorzugt wird ein Promotor aus einer Pflanze oder aus einem Pflanzenvirus verwendet. Bevorzugt sind der Promotor des CaMV (cauliflower mosaic virus) 35S-Transkripts (Franck et al. (1980) Cell 21: 285-294), der 19S CaMV-Promotor (US 5,352,605), der Aktinpromotor aus Reis (McElroy et al. (1990) Plant Cell 2: 163-171), der Legumin B-Promotor (GenBank Acc. No. X03677), der Promotor der Nopalinsynthase aus Agrobacterium, der TR duale Promotor, der Octopinsynthase-Promotor aus Agrobacterium, der Ubiquitin-Promotor (Holtorf et al. (1995) Plant Mol. Biol. 29: 637-649), der Smas-Promotor, der Cinnamoylalkoholdehydrogenase-Promotor (US 5,683,439), die Promotoren der vacuolaren ATPase-Untereinheiten, der pEMU-Promotor (Last et al. (1991) Theor. Appl. Genet. 81: 581-588), der MAS-Promotor (Velten et al. (1984) EMBO J. 3(12): 2723-2730), der Histon-H3-Promotor aus Mais (Lepetit et al. (1992) Mol. Gen. Genet. 231: 276-285), der Promotor des Nitrilase 1-Gens aus Arabidopsis (GenBank



Acc. No. U38846, Nukleotide 3862-5325) und der Promotor eines Prolin-reichen Proteins aus Weizen (WO 91/13991) und weitere Promotoren, die konstitutive Genexpression vermitteln. Besonders bevorzugt ist der Promotor des CaMV 35S-Transkripts.

- 5 Es ist im Prinzip möglich, alle natürlich auftretenden konstitutiven Promotoren mit ihren Regulationssequenzen, wie die oben genannten, für das neue Verfahren zu verwenden. Es ist aber ebenfalls möglich, zusätzlich oder alleine synthetische Promotoren zu verwenden.
- 10 "Blattspezifische Promotoren" sind Promotoren, die eine hohe Aktivität im Blatt und keine oder nur eine geringe Aktivität in anderen Geweben zeigen. Unter "geringer Aktivität" wird im Rahmen der Erfindung verstanden, dass die Aktivität in anderen Geweben weniger als 20%, bevorzugt weniger als 10%, besonders bevorzugt weniger als 5% und am meisten bevorzugt weniger als 3, 2 oder 1% der Aktivität im Blatt beträgt. Geeignete
- 15 blattspezifische Promotoren sind z.B. die Promotoren der kleinen Untereinheit von Rubisco (Timko et al. (1985) Nature 318: 579-582) und des Chlorophyll a/b-bindenden Proteins (Simpson et al. (1985) EMBO J. 4: 2723-2729).

- Dem Fachmann sind weitere blattspezifische Promotoren bekannt bzw. er kann mit be-
- 20 kannten Methoden weitere geeignete Promotoren zu isolieren. So kann der Fachmann mit Hilfe gängiger molekularbiologischer Methoden, z. B. Hybridisierungsexperimenten oder DNA-Protein-Bindungsstudien, Blatt-spezifische regulatorische Nukleinsäureelemente identifizieren. Dabei wird z. B. in einem ersten Schritt aus Blattgewebe des gewünschten Organismus, aus dem die regulatorischen Sequenzen isoliert werden sollen, die gesamte poly(A)<sup>+</sup>-RNA isoliert und eine cDNA-Bank angelegt. In einem zweiten Schritt
- 25 werden mit Hilfe von cDNA-Klonen, die auf poly(A)<sup>+</sup>-RNA-Molekülen aus einem Nicht-Blattgewebe basieren, aus der ersten Bank mittels Hybridisierung diejenigen Klone identifiziert, deren korrespondierende poly(A)<sup>+</sup>-RNA-Moleküle lediglich im Blattgewebe akkumulieren. Anschließend werden mit Hilfe dieser so identifizierten cDNAs Promotoren
- 30 isoliert, die über Blatt-spezifische regulatorische Elemente verfügen. Dem Fachmann stehen darüber hinaus weitere auf PCR basierende Methoden für die Isolierung geeigneter Blatt-spezifischer Promotoren zur Verfügung.

Selbstverständlich können die Nukleinsäuresequenzen der vorliegenden Erfindung auch in den Samen der transgenen Pflanzen exprimiert werden, indem samen-spezifische Promotoren verwendet werden, die im Embryo und/oder im Endosperm aktiv sind. Samen-spezifische Promotoren können prinzipiell sowohl aus dikotyledonen als auch aus monokotyledonen Pflanzen isoliert werden. Im Folgenden sind bevorzugte Promotoren aufgeführt: USP (= unknown seed protein) und Vicilin (*Vicia faba*) (Bäumlein et al. (1991) Mol. Gen Genet. 225(3): 459-467), Napin (Raps) (US 5,608,152), Conlinin (Lein) (WO 02/102970), Acyl-Carrier Protein (Raps) (US 5,315,001 und WO 92/18634), Oleosin (*Arabidopsis thaliana*) (WO 98/45461 und WO 93/20216), Phaseolin (*Phaseolus vulgaris*) (US 5,504,200), Bce4 (WO 91/13980), Leguminosen B4 (LegB4-Promotor) (Bäumlein et al. (1992) Plant J. 2(2): 233-239), Lpt2 und lpt1(Gerste) (WO 95/15389 und WO95/23230), Samen-spezifische Promotoren aus Reis, Mais und Weizen (WO 99/16890), Amy32b, Amy 6-6 und Aleurain (US 5,677,474), Bce4 (Raps) (US 5,530,149), Glycinin (Soja) (EP 571 741), Phosphoenol-Pyruvatcarboxylase (Soja) (JP 06/62870), ADR12-2 (Soja) (WO 98/08962), Isocitratlyase (Raps) (US 5,689,040) oder  $\alpha$ -Amylase (Gerste) (EP 781 849).

In einer besonders bevorzugten Ausführungsform der vorliegenden Erfindung werden die verwendeten Nukleinsäuresequenzen, insbesondere die für eine  $\Delta$ -5 Elongase kodierende Nukleinsäuresequenz, welche gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist, bevorzugt die in SEQ ID NO: 64 beschriebene Nukleinsäuresequenz im generativen Gewebe, insbesondere im Samen exprimiert. Die spezifische Expression im Samen erfolgt vorteilhaft unter Verwendung eines der oben erwähnten samenspezifischen Promotoren, insbesondere unter Verwendung des Napin Promotors. In dieser besonders bevorzugten Ausführungsform beträgt der Gehalt an den hergestellten LCPUFAs, insbesondere an den C22 Fettsäuren im Samenöl mindestens 5 Gew.-%, vorteilhaft mindestens 6, 7, 8, 9 oder 10 Gew.-%, bevorzugt von mindestens 11, 12, 13, 14 oder 15 Gew.-%, besonders bevorzugt von mindestens 16, 17, 18, 19, oder 20 Gew.-%, ganz besonders bevorzugt von mindestens 25, 30, 35 oder 40 Gew.-% des Samenölgehalt. In einer weiteren besonders bevorzug-

ten Ausführungsform mit der in SEQ ID NO: 63 beschriebenen Nukleinsäuresequenz beträgt der Gehalt an C22-Fettsäuren im Samenöl mindestens 8 Gew.-% des Samenölgehalts.

5 In einer weiteren besonders bevorzugten Ausführungsform der vorliegenden Erfindung werden die verwendeten Nukleinsäuresequenzen, insbesondere die für eine  $\Delta$ -5 Elongase kodierende Nukleinsäuresequenz, welche gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist, bevorzugt die in  
10 SEQ ID NO: 64 beschriebene Nukleinsäuresequenz im generativen Gewebe, insbesondere im Samen exprimiert. Die spezifische Expression im Samen erfolgt vorteilhaft unter Verwendung eines der oben erwähnten samenspezifischen Promotoren, insbesondere unter Verwendung des Napin-Promotors. In dieser besonders bevorzugten Ausführungsform beträgt der Gehalt an Docosahexaensäure im Samenöl mindestens 1 Gew.-%, be-  
15 vorzugt mindestens 1,1, 1,2, 1,3, 1,4 oder 1,5 Gew.-%, besonders bevorzugt mindestens 1,6, 1,7, 1,8 oder 1,9 Gew.-%, insbesondere mindestens 2, 2,1, 2,2, 2,3, 2,4, 2,5, 2,6, 2,7, 2,8 oder 2,9 Gew.-%, weiterhin bevorzugt mindestens 3, 3,5 oder 4 Gew.-% des Samenölgehalts. In einer weiteren besonders bevorzugten Ausführungsform mit der in  
20 SEQ ID NO: 63 beschriebenen Nukleinsäuresequenz beträgt der Gehalt an Docosahexaensäure im Samenöl mindestens 1,9 Gew.-% des Samenölgehalts. Es ist dem Fachmann dabei bekannt, dass zur Herstellung von Docosahexaensäure zusätzlich eine oder mehrere Nukleinsäuresequenzen, die für ein Polypeptid mit der Aktivität einer  $\Delta$ 4-Desaturase-Aktivität kodiert, benötigt werden. Vorteilhaft wird eine Nukleinsäuresequenz, die für ein Polypeptid mit der Aktivität einer  $\Delta$ 4-Desaturase-Aktivität kodiert, ausgewählt  
25 aus der Gruppe bestehend aus Nukleinsäuresequenzen mit der in Seq ID No. 77, 79,81,83,85,87,89,91 oder 93, bevorzugt mit der in der Seq ID No. 77 dargestellten Sequenz.

In einer weiteren besonders bevorzugten Ausführungsform der vorliegenden Erfindung  
30 werden die verwendeten Nukleinsäuresequenzen, insbesondere die für eine  $\Delta$ -5 Elongase kodierende Nukleinsäuresequenz, welche gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die

Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist, bevorzugt die in SEQ ID NO: 64 beschriebene Nukleinsäuresequenz im generativen Gewebe, insbesondere im Samen exprimiert. Die spezifische Expression im Samen erfolgt vorteilhaft unter Verwendung eines der oben erwähnten samenspezifischen Promotoren, insbesondere unter Verwendung des NapinPromotors. In dieser besonders bevorzugten Ausführungsform beträgt der Gehalt an Docosahexaensäure im Samenöl mindestens 1 Gew.-%, bevorzugt mindestens 1,1, 1,2, 1,3, 1,4 oder 1,5 Gew.-%, besonders bevorzugt mindestens 1,6, 1,7, 1,8 oder 1,9 Gew.-%, insbesondere mindestens 2, 2,1, 2,2, 2,5, 2,6, 2,7, 2,8 oder 2,9 Gew.-%, weiterhin bevorzugt mindestens 3, 3,5, oder 4 Gew.-% des Samenölgehalts. Dabei beträgt der Gehalt an den hergestellten LCPUFAs, insbesondere an den C22 Fettsäuren im Samenöl mindestens 5 Gew.-%, vorteilhaft mindestens 6, 7, 8, 9 oder 10 Gew.-%, bevorzugt von mindestens 11, 12, 13, 14 oder 15 Gew.-%, besonders bevorzugt von mindestens 16, 17, 18, 19, oder 20 Gew.-%, ganz besonders bevorzugt von mindestens 25, 30, 35 oder 40 Gew.-% des Samenölgehalt. In einer weiteren besonders bevorzugten Ausführungsform mit der in SEQ ID NO: 63 beschriebenen Nukleinsäuresequenz beträgt der Gehalt an Docosahexaensäure im Samenöl mindestens 1,9 Gew.-% des Samenölgehalts, wobei der Gehalt an C22-Fettsäuren im Samenöl mindestens 8 Gew.-% des Samenölgehalts beträgt.

Die Pflanzengenexpression lässt sich auch über einen chemisch induzierbaren Promotor erleichtern (siehe eine Übersicht in Gatz (1997) Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:89-108). Chemisch induzierbare Promotoren eignen sich besonders, wenn gewünscht wird, dass die Genexpression auf zeitspezifische Weise erfolgt. Beispiele für solche Promotoren sind ein Salicylsäure-induzierbarer Promotor (WO 95/19443), ein Tetracyclin-induzierbarer Promotor (Gatz et al. (1992) Plant J. 2, 397-404) und ein Ethanol-induzierbarer Promotor.

Auch Promotoren, die auf biotische oder abiotische Stressbedingungen reagieren, sind geeignete Promotoren, wie beispielsweise der pathogeninduzierte PRP1-Gen-Promotor (Ward et al. (1993) Plant. Mol. Biol. 22: 361-366), der hitzeinduzierbare hsp80-Promotor aus Tomate (US 5,187,267), der kälteinduzierbare Alpha-Amylase-Promotor aus Kartoffel-

fel (WO 96/12814) oder der durch Wunden induzierbare pinII-Promotor (EP-A-0 375 091).

5 Ebenfalls besonders geeignet sind Promotoren, welche die plastidenspezifische Expression herbeiführen, da Plastiden das Kompartiment sind, in dem die Vorläufer sowie einige Endprodukte der Lipidbiosynthese synthetisiert werden. Geeignete Promotoren, wie der virale RNA-Polymerase-Promotor, sind beschrieben in WO 95/16783 und WO 97/06250, und der clpP-Promotor aus Arabidopsis, beschrieben in WO 99/46394.

10 Es versteht sich, dass die erfindungsgemäß hergestellten mehrfach ungesättigten Fettsäuren nicht nur in intakten transgenen Pflanzen produziert werden können, sondern auch in pflanzlichen Zellkulturen oder in Kalluskulturen.

15 Die im Verfahren hergestellten mehrfach ungesättigten Fettsäuren sind vorteilhaft in Phospholipiden und/oder Triacylglyceriden gebunden, können aber auch als freie Fettsäuren oder aber gebunden in Form anderer Fettsäureester in den Organismen vorkommen. Dabei können sie als "Reinprodukte" oder aber vorteilhaft in Form von Mischungen verschiedener Fettsäuren oder Mischungen unterschiedlicher Phospholipide wie Phosphatidylglycerol, Phosphatidylcholin, Phosphatidylethanolamin und/oder  
20 Phosphatidylserin und/oder Triacylglyceriden, Monoacylglyceriden und/oder Diacylglyceriden vorliegen. Vorteilhaft liegen die im Verfahren hergestellten LCPUFAS EPA, DPA und DHA im Phosphatidylcholin und/oder Phosphatidylethanolamin und/oder in den Triacylglyceriden vor. Die Triacylglyceride können außerdem noch weitere Fettsäuren enthalten wie kurzkettige Fettsäuren mit 4 bis 6 C-Atomen, mittelkettige Fettsäuren mit 8  
25 bis 12 C-Atomen oder langkettige Fettsäuren mit 14 bis 24 C-Atomen. Bevorzugt enthalten sie langkettige Fettsäuren, besonders bevorzugt C<sub>20</sub>- oder C<sub>22</sub>-Fettsäuren.

30 Unter dem Begriff "Glycerid" wird ein mit ein, zwei oder drei Carbonsäureresten verestertes Glycerin verstanden (Mono-, Di- oder Triglycerid). Unter "Glycerid" wird auch ein Gemisch aus verschiedenen Glyceriden verstanden. Vorteilhaft handelt es sich bei dem Glycerid um ein Triglycerid. Das Glycerid oder das Glyceridgemisch kann weitere Zusät-

ze, z.B. freie Fettsäuren, Antioxidantien, Proteine, Kohlenhydrate, Vitamine und/oder andere Substanzen enthalten.

Unter einem "Glycerid" im Sinne des erfindungsgemäßen Verfahrens werden ferner vom Glycerin abgeleitete Derivate verstanden. Dazu zählen neben den oben beschriebenen Fettsäureglyceriden auch Glycerophospholipide und Glyceroglycolipide. Bevorzugt seien hier die Glycerophospholipide wie Lecithin (Phosphatidylcholin), Cardiolipin, Phosphatidylglycerin, Phosphatidylserin und Alkylacylglycerophospholipide beispielhaft genannt.

Unter Phospholipiden im Sinne der Erfindung sind zu verstehen Phosphatidylcholin, Phosphatidylethanolamin, Phosphatidylserin, Phosphatidylglycerin und/oder Phosphatidylinositol.

Die Fettsäureester mit mehrfach ungesättigten C<sub>18</sub>-, C<sub>20</sub>- und/oder C<sub>22</sub>-Fettsäuremolekülen können aus den Nutzpflanzen, die für die Herstellung der Fettsäureester verwendet wurden, in Form eines Öls oder Lipids beispielsweise in Form von Verbindungen wie Sphingolipiden, Phosphoglyceriden, Lipiden, Glycolipiden wie Glycosphingolipiden, Phospholipiden wie Phosphatidylethanolamin, Phosphatidylcholin, Phosphatidylserin, Phosphatidylglycerol, Phosphatidylinositol oder Diphosphatidylglycerol, Monoacylglyceriden, Diacylglyceriden, Triacylglyceriden oder sonstigen Fettsäureestern wie den Acetyl-CoenzymA-Estern, die die mehrfach ungesättigten Fettsäuren mit mindestens zwei, drei oder vier, bevorzugt vier, fünf oder sechs Doppelbindungen enthalten, isoliert werden, vorteilhaft werden sie in der Form ihrer Diacylglyceride, Triacylglyceride und/oder in Form des Phosphatidylesters isoliert, besonders bevorzugt in der Form der Triacylglyceride, Phosphatidylcholin und/oder Phosphatidylethanolamin. Neben diesen Estern sind die mehrfach ungesättigten Fettsäuren auch als freie Fettsäuren oder gebunden in anderen Verbindungen in den Pflanzen enthalten. In der Regel liegen die verschiedenen vorgenannten Verbindungen (Fettsäureester und freie Fettsäuren) in den Organismen in einer ungefähren Verteilung von 80 bis 90 Gew.-% Triglyceride, 2 bis 5 Gew.-% Diglyceride, 5 bis 10 Gew.-% Monoglyceride, 1 bis 5 Gew.-% freie Fettsäuren, 2 bis 8 Gew.-% Phospholipide vor, wobei sich die Summe der verschiedenen Verbindungen zu 100 Gew.-% ergänzt.

Im erfindungsgemäßen Verfahren werden die hergestellten LCPUFAs mit einem Gehalt von mindestens 4 Gew.-%, vorteilhaft von mindestens 5, 6, 7, 8, 9 oder 10 Gew.-%, bevorzugt von mindestens 11, 12, 13, 14 oder 15 Gew.-%, besonders bevorzugt von mindestens 16, 17, 18, 19, oder 20 Gew.-%, ganz besonders bevorzugt von mindestens 25, 30, 35 oder 40 Gew.-% bezogen auf die gesamten Fettsäuren in der transgenen Pflanze hergestellt. Dabei sind die im erfindungsgemäßen Verfahren hergestellten Fettsäuren EPA, DPA und/oder DHA mit einem Gehalt von jeweils mindestens 5 Gew.-%, bevorzugt von jeweils mindestens 6, 7, 8 oder 9 Gew.-%, besonders bevorzugt von jeweils mindestens 10, 11 oder 12 Gew.-%, am meisten bevorzugt von jeweils mindestens 13, 14, 15, 16, 17, 18, 19 oder 20 Gew.-%, bezogen auf die gesamten Fettsäuren in der transgenen Pflanze enthalten.

Vorteilhaft werden die Fettsäuren in gebundener Form hergestellt. Mit Hilfe der im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren lassen sich diese ungesättigten Fettsäuren an sn1-, sn2- und/oder sn3-Position der vorteilhaft hergestellten Triacylglyceride bringen. Vorteilhaft sind mindestens 11% der Triacylglyceride doppelt (das heißt an sn1- und sn2- oder sn2- und sn3-Position) substituiert. Auch dreifach substituierte Triacylglyceride sind nachweisbar. Da im erfindungsgemäßen Verfahren von den Ausgangsverbindungen Linolsäure (C18:2) bzw. Linolensäure (C18:3) mehrere Reaktionsschritte durchlaufen werden, fallen die Endprodukte des Verfahrens wie beispielsweise Arachidonsäure (ARA) oder Eicosapentaensäure (EPA) nicht als absolute Reinprodukte an, es sind immer auch Spuren oder größere Mengen der Vorstufen im Endprodukt enthalten. Sind in der Ausgangspflanze beispielsweise sowohl Linolsäure als auch Linolensäure vorhanden, so liegen die Endprodukte wie ARA oder EPA und/oder DPA und/oder DHA als Mischungen vor. Die Vorstufen sollten vorteilhaft nicht mehr als 20 Gew.-%, bevorzugt nicht mehr als 15 Gew.-%, besonders bevorzugt nicht als 10 Gew.-%, ganz besonders bevorzugt nicht mehr als 5 Gew.-% bezogen auf die Menge des jeweiligen Endprodukts betragen. Vorteilhaft werden in einer transgenen Pflanze als Endprodukte nur ARA oder EPA und/oder DPA und/oder DHA im erfindungsgemäßen Verfahren gebunden oder als freie Säuren hergestellt.



Fettsäureester bzw. Fettsäuregemische, die nach dem erfindungsgemäßen Verfahren hergestellt wurden, enthalten vorteilhaft 6 bis 15 % Palmitinsäure, 1 bis 6 % Stearinsäure; 7 – 85 % Ölsäure; 0,5 bis 8 % Vaccensäure, 0,1 bis 1 % Arachinsäure, 7 bis 25 % gesättigte Fettsäuren, 8 bis 85 % einfach ungesättigte Fettsäuren und 60 bis 85 % mehrfach ungesättigte Fettsäuren, jeweils bezogen auf 100 % und auf den Gesamtfettsäuregehalt der Organismen. Als vorteilhafte mehrfach ungesättigte Fettsäure sind in den Fettsäureester bzw. Fettsäuregemische bevorzugt mindestens 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9 oder 1 % bezogen auf den Gesamtfettsäuregehalt an Arachidonsäure enthalten. Weiterhin enthalten die Fettsäureestern bzw. Fettsäuregemischen, die nach dem erfindungsgemäßen Verfahren hergestellt wurden, vorteilhaft Fettsäuren ausgewählt aus der Gruppe der Fettsäuren Erucasäure (13-Docosaensäure), Sterculinsäure (9,10-Methylene octadec-9-enonsäure), Malvalinsäure (8,9-Methylen Heptadec-8-enonsäure), Chaulmoogrinsäure (Cyclopenten-dodecansäure), Furan-Fettsäure (9,12-Epoxyoctadeca-9,11-dienonsäure), Vernonsäure (9,10-Epoxyoctadec-12-enonsäure), Tarinsäure (6-Octadecynonsäure), 6-Nonadecynonsäure, Santalbinsäure (t11-Octadecen-9-ynoic acid), 6,9-Octadecenynonsäure, Pyrulinsäure (t10-Heptadecen-8-ynonsäure), Crepenyninsäure (9-Octadecen-12-ynonsäure), 13,14-Dihydrooropheinsäure, Octadecen-13-ene-9,11-diynonsäure, Petroselensäure (cis-6-Octadecenonsäure), 9c,12t-Octadecadiensäure, Calendulasäure (8t10t12c-Octadecatriensäure), Catalpinsäure (9t11t13c-Octadecatriensäure), Eleosterinsäure (9c11t13t-Octadecatriensäure), Jacarinsäure (8c10t12c-Octadecatriensäure), Punicinsäure (9c11t13c-Octadecatriensäure), Parinarinsäure (9c11t13t15c-Octadecatetraensäure), Pinolensäure (all-cis-5,9,12-Octadecatriensäure), Laballensäure (5,6-Octadecadienallensäure), Ricinolsäure (12-Hydroxyölsäure) und/oder Coriolinsäure (13-Hydroxy-9c,11t-Octadecadienonsäure). Die vorgenannten Fettsäuren kommen in den nach dem erfindungsgemäßen Verfahren hergestellten Fettsäureester bzw. Fettsäuregemischen in der Regel vorteilhaft nur in Spuren vor, das heißt sie kommen bezogen auf den Gesamtfettsäuregehalt zu weniger als 30 %, bevorzugt zu weniger als 25 %, 24 %, 23 %, 22 % oder 21 %, besonders bevorzugt zu weniger als 20 %, 15 %, 10 %, 9 %, 8 %, 7%, 6 % oder 5%, ganz besonders bevorzugt zu weniger als 4 %, 3 %, 2 % oder 1 % vor. In einer weiteren bevorzugten Form der Erfindung kommen diese vorgenannten Fettsäuren bezogen auf die Gesamtfettsäuren zu weniger als 0,9%; 0,8%; 0,7%; 0,6%; oder 0,5%, besonders bevorzugt zu weniger als



0,4%; 0,3%; 0,2%; 0,1% vor. Vorteilhaft enthalten die nach dem erfindungsgemäßen Verfahren hergestellten Fettsäureester bzw. Fettsäuregemische weniger als 0,1 % bezogen auf die Gesamtfettsäuren und/oder keine Buttersäure, kein Cholesterin sowie keine Nisinsäure (Tetracosahexaensäure, C23:6<sup>Δ3,8,12,15,18,21</sup>).

5

Durch die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen kann in den transgenen Pflanzen eine Steigerung der Ausbeute an LCPUFAs von mindestens 50%, vorteilhaft von mindestens 80%, besonders vorteilhaft von mindestens 100%, ganz besonders vorteilhaft von mindestens 150% gegenüber den nicht transgenen Pflanzen erreicht werden.

10

Auch chemisch reine mehrfach ungesättigte Fettsäuren oder Fettsäurezusammensetzungen sind nach den vorbeschriebenen Verfahren darstellbar. Dazu werden die Fettsäuren oder die Fettsäurezusammensetzungen aus den Pflanzen in bekannter Weise beispielsweise über Extraktion, Destillation, Kristallisation, Chromatographie oder Kombinationen dieser Methoden isoliert. Diese chemisch reinen Fettsäuren oder Fettsäurezusammensetzungen sind für Anwendungen im Bereich der Lebensmittelindustrie, der Kosmetikindustrie und besonders der Pharmaindustrie vorteilhaft.

15

20

Für das erfindungsgemäße Verfahren sind prinzipiell alle dicotylen oder monokotylen Nutzpflanzen geeignet. Unter Nutzpflanzen sind Pflanzen zu verstehen, die der Nahrungsproduktion für Mensch und Tier, der Produktion von Genussmitteln, Fasern und Pharmazeutika dienen, wie Getreide, z.B. Mais, Reis, Weizen, Gerste, Hirse, Hafer, Roggen, Buchweizen; wie Knollen, z.B. Kartoffel, Maniok, Batate, Yams etc.; wie Zuckerpflanzen, z.B. Zuckerrohr oder Zuckerrübe; wie Hülsenfrüchte, z.B. Bohnen, Erbsen, Saubohne etc.; wie Öl- und Fettfrüchte, z.B. Sojabohne, Raps, Sonnenblume, Färberdiestel, Lein, Camelina etc., um nur einige zu nennen. Vorteilhafte Pflanzen sind ausgewählt aus der Gruppe der Pflanzenfamilien bestehend aus den Familien der Aceraceae, Actinidiaceae, Anacardiaceae, Apiaceae, Arecaceae, Asteraceae, Arecaceae, Betulaceae, Boraginaceae, Brassicaceae, Bromeliaceae, Cannabaceae, Cannaceae, Caprifoliaceae, Chenopodiaceae, Convolvulaceae, Cucurbitaceae, Dioscoreaceae, E-laeagnaceae, Ericaceae, Euphorbiaceae, Fabaceae, Fagaceae, Grossulariaceae, Juglandaceae, Lauraceae, Liliaceae, Linaceae, Malvaceae, Moraceae, Musaceae, Oleaceae, Oxalida-

25

30

ceae, Papaveraceae, Poaceae, Polygonaceae, Punicaceae, Rosaceae, Rubiaceae, Rutaceae, Scrophulariaceae, Solanaceae, Sterculiaceae und Valerianaceae.

Beispielhaft seien die folgenden Pflanzen genannt: Anacardiaceae wie die Gattungen  
5 Pistacia, Mangifera, Anacardium z.B. die Gattung und Arten Pistacia vera (Pistazie),  
Mangifer indica (Mango) oder Anacardium occidentale (Cashew), Asteraceae wie die  
Gattungen Calendula, Carthamus, Centaurea, Cichorium, Cynara, Helianthus, Lactuca,  
Locusta, Tagetes, Valeriana z.B. die Gattung und Arten Calendula officinalis (Garten-  
Ringelblume), Carthamus tinctorius (Färberdistel, safflower), Centaurea cyanus (Korn-  
10 blume), Cichorium intybus (Wegwarte), Cynara scolymus (Artischocke), Helianthus an-  
nus (Sonnenblume), Lactuca sativa, Lactuca crispa, Lactuca esculenta, Lactuca scariola  
L. ssp. sativa, Lactuca scariola L. var. integrata, Lactuca scariola L. var. integrifolia, Lac-  
tuca sativa subsp. romana, Locusta communis, Valeriana locusta (Salat), Tagetes lucida,  
Tagetes erecta oder Tagetes tenuifolia (Studentenblume), Apiaceae wie die Gattung  
15 Daucus z.B. die Gattung und Art Daucus carota (Karotte), Betulaceae wie die Gattung  
Corylus z.B. die Gattungen und Arten Corylus avellana oder Corylus columna (Hasel-  
nuss), Boraginaceae wie die Gattung Borago z.B. die Gattung und Art Borago officinalis  
(Borretsch), Brassicaceae wie die Gattungen Brassica, Camelina, Melanosinapis, Sina-  
pis, Arabidopsis z.B. die Gattungen und Arten Brassica napus, Brassica rapa ssp.  
20 (Raps), Sinapis arvensis Brassica juncea, Brassica juncea var. juncea, Brassica juncea  
var. crispifolia, Brassica juncea var. foliosa, Brassica nigra, Brassica sinapioides, Came-  
lina sativa, Melanosinapis communis (Senf), Brassica oleracea (Futterrübe) oder Arabi-  
dopsis thaliana, Bromeliaceae wie die Gattungen Anana, Bromelia (Ananas) z.B. die  
Gattungen und Arten Anana comosus, Ananas ananas oder Bromelia comosa (Ananas),  
25 Caricaceae wie die Gattung Carica wie die Gattung und Art Carica papaya (Papaya),  
Cannabaceae wie die Gattung Cannabis wie die Gattung und Art Cannabis sativa (Hanf),  
Convolvulaceae wie die Gattungen Ipomoea, Convolvulus z.B. die Gattungen und Arten  
Ipomoea batatas, Ipomoea pandurata, Convolvulus batatas, Convolvulus tiliaceus, Ipo-  
moea fastigiata, Ipomoea tiliacea, Ipomoea triloba oder Convolvulus panduratus (Süß-  
30 kartoffel, Batate), Chenopodiaceae wie die Gattung Beta wie die Gattungen und Arten  
Beta vulgaris, Beta vulgaris var. altissima, Beta vulgaris var. vulgaris, Beta maritima, Be-  
ta vulgaris var. perennis, Beta vulgaris var. conditiva oder Beta vulgaris var. esculenta

(Zuckerrübe), Cucurbitaceae wie die Gattung Cucurbita z.B. die Gattungen und Arten Cucurbita maxima, Cucurbita mixta, Cucurbita pepo oder Cucurbita moschata (Kürbis), Elaeagnaceae wie die Gattung Elaeagnus z.B. die Gattung und Art Olea europaea (Olive), Ericaceae wie die Gattung Kalmia z.B. die Gattungen und Arten Kalmia latifolia, Kalmia angustifolia, Kalmia microphylla, Kalmia polifolia, Kalmia occidentalis, Cistus chamaerhodendros oder Kalmia lucida (Berglorbeer), Euphorbiaceae wie die Gattungen Manihot, Janipha, Jatropha, Ricinus z.B. die Gattungen und Arten Manihot utilissima, Janipha manihot, Jatropha manihot, Manihot aipil, Manihot dulcis, Manihot manihot, Manihot melanobasis, Manihot esculenta (Manihot) oder Ricinus communis (Rizinus), Fabaceae wie die Gattungen Pisum, Albizia, Cathormion, Feuillea, Inga, Pithecolobium, Acacia, Mimosa, Medicago, Glycine, Dolichos, Phaseolus, Soja z.B. die Gattungen und Arten Pisum sativum, Pisum arvense, Pisum humile (Erbse), Albizia berteriana, Albizia julibrissin, Albizia lebbeck, Acacia berteriana, Acacia littoralis, Albizia berteriana, Albizzia berteriana, Cathormion berteriana, Feuillea berteriana, Inga fragrans, Pithecellobium berterianum, Pithecellobium fragrans, Pithecolobium berterianum, Pseudalbizia berteriana, Acacia julibrissin, Acacia nemu, Albizia nemu, Feuillea julibrissin, Mimosa julibrissin, Mimosa speciosa, Sericanrda julibrissin, Acacia lebbeck, Acacia macrophylla, Albizia lebbeck, Feuillea lebbeck, Mimosa lebbeck, Mimosa speciosa (Seidenbaum), Medicago sativa, Medicago falcata, Medicago varia (Alfalfa) Glycine max Dolichos soja, Glycine gracilis, Glycine hispida, Phaseolus max, Soja hispida oder Soja max (Sojabohne), Geraniaceae wie die Gattungen Pelargonium, Cocos, Oleum z.B. die Gattungen und Arten Cocos nucifera, Pelargonium grossularioides oder Oleum cocois (Kokusnuss), Gramineae wie die Gattung Saccharum z.B. die Gattung und Art Saccharum officinarum, Juglandaceae wie die Gattungen Juglans, Wallia z.B. die Gattungen und Arten Juglans regia, Juglans ailanthifolia, Juglans sieboldiana, Juglans cinerea, Wallia cinerea, Juglans bixbyi, Juglans californica, Juglans hindsii, Juglans intermedia, Juglans jamaicensis, Juglans major, Juglans microcarpa, Juglans nigra oder Wallia nigra (Walnuss), Lauraceae wie die Gattungen Persea, Laurus z.B. die Gattungen und Arten Laurus nobilis (Lorbeer), Persea americana, Persea gratissima oder Persea persea (Avocado), Leguminosae wie die Gattung Arachis z.B. die Gattung und Art Arachis hypogaea (Erdnuss), Linaceae wie die Gattungen Linum, Adenolinum z.B. die Gattungen und Arten Linum usitatissimum, Linum humile, Linum austriacum, Linum bienne, Linum angustifolium, Li-

num catharticum, *Linum flavum*, *Linum grandiflorum*, *Adenolinum grandiflorum*, *Linum lewisii*, *Linum narbonense*, *Linum perenne*, *Linum perenne* var. *lewisii*, *Linum pratense* oder *Linum trigynum* (Lein), *Lythraeae* wie die Gattung *Punica* z.B. die Gattung und Art *Punica granatum* (Granatapfel), *Malvaceae* wie die Gattung *Gossypium* z.B. die Gattungen und Arten *Gossypium hirsutum*, *Gossypium arboreum*, *Gossypium barbadense*, *Gossypium herbaceum* oder *Gossypium thurberi* (Baumwolle), *Musaceae* wie die Gattung *Musa* z.B. die Gattungen und Arten *Musa nana*, *Musa acuminata*, *Musa paradisiaca*, *Musa* spp. (Banane), *Onagraceae* wie die Gattungen *Camissonia*, *Oenothera* z.B. die Gattungen und Arten *Oenothera biennis* oder *Camissonia brevipes* (Nachtkerze), *Palmae* wie die Gattung *Elaeis* z.B. die Gattung und Art *Elaeis guineensis* (Ölpalme), *Papaveraceae* wie die Gattung *Papaver* z.B. die Gattungen und Arten *Papaver orientale*, *Papaver rhoeas*, *Papaver dubium* (Mohn), *Pedaliaceae* wie die Gattung *Sesamum* z.B. die Gattung und Art *Sesamum indicum* (Sesam), *Piperaceae* wie die Gattungen *Piper*, *Artanthe*, *Peperomia*, *Steffensia* z.B. die Gattungen und Arten *Piper aduncum*, *Piper amalago*, *Piper angustifolium*, *Piper auritum*, *Piper betel*, *Piper cubeba*, *Piper longum*, *Piper nigrum*, *Piper retrofractum*, *Artanthe adunca*, *Artanthe elongata*, *Peperomia elongata*, *Piper elongatum*, *Steffensia elongata* (Cayennepfeffer), *Poaceae* wie die Gattungen *Hordeum*, *Secale*, *Avena*, *Sorghum*, *Andropogon*, *Holcus*, *Panicum*, *Oryza*, *Zea* (Mais), *Triticum*, z.B. die Gattungen und Arten *Hordeum vulgare*, *Hordeum jubatum*, *Hordeum murinum*, *Hordeum secalinum*, *Hordeum distichon*, *Hordeum aegiceras*, *Hordeum hexastichon*., *Hordeum hexastichum*, *Hordeum irregulare*, *Hordeum sativum*, *Hordeum secalinum* (Gerste), *Secale cereale* (Roggen), *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida* (Hafer), *Sorghum bicolor*, *Sorghum halepense*, *Sorghum saccharatum*, *Sorghum vulgare*, *Andropogon drummondii*, *Holcus bicolor*, *Holcus sorghum*, *Sorghum aethiopicum*, *Sorghum arundinaceum*, *Sorghum caffrorum*, *Sorghum cernuum*, *Sorghum dochna*, *Sorghum drummondii*, *Sorghum durra*, *Sorghum guineense*, *Sorghum lanceolatum*, *Sorghum nervosum*, *Sorghum saccharatum*, *Sorghum subglabrescens*, *Sorghum verticilliflorum*, *Sorghum vulgare*, *Holcus halepensis*, *Sorghum miliaceum*, *Panicum militaceum* (Hirse), *Oryza sativa*, *Oryza latifolia* (Reis), *Zea mays* (Mais) *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum* oder *Triticum vulgare* (Weizen), *Porphyridiaceae* wie die Gattungen *Chroothece*, *Flintiella*, *Petrovanella*, *Porphyridium*, *Rhodella*, *Rhodorus*, Van-

- hoeffenia z.B. die Gattung und Art *Porphyridium cruentum*, Proteaceae wie die Gattung *Macadamia* z.B. die Gattung und Art *Macadamia intergrifolia* (*Macadamia*), Rubiaceae wie die Gattung *Coffea* z.B. die Gattungen und Arten *Coffea* spp., *Coffea arabica*, *Coffea canephora* oder *Coffea liberica* (Kaffee), Scrophulariaceae wie die Gattung *Verbascum* z.B. die Gattungen und Arten *Verbascum blattaria*, *Verbascum chaixii*, *Verbascum densiflorum*, *Verbascum lagurus*, *Verbascum longifolium*, *Verbascum lychnitis*, *Verbascum nigrum*, *Verbascum olympicum*, *Verbascum phlomoides*, *Verbascum phoenicum*, *Verbascum pulverulentum* oder *Verbascum thapsus* (Königskerze), Solanaceae wie die Gattungen *Capsicum*, *Nicotiana*, *Solanum*, *Lycopersicon* z.B. die Gattungen und Arten *Capsicum annuum*, *Capsicum annuum* var. *glabriusculum*, *Capsicum frutescens* (Pfeffer), *Capsicum annuum* (Paprika), *Nicotiana tabacum*, *Nicotiana alata*, *Nicotiana attenuata*, *Nicotiana glauca*, *Nicotiana langsdorffii*, *Nicotiana obtusifolia*, *Nicotiana quadrivalvis*, *Nicotiana repanda*, *Nicotiana rustica*, *Nicotiana sylvestris* (Tabak), *Solanum tuberosum* (Kartoffel), *Solanum melongena* (Aubergine) *Lycopersicon esculentum*, *Lycopersicon lycopersicum*., *Lycopersicon pyriforme*, *Solanum integrifolium* oder *Solanum lycopersicum* (Tomate), Sterculiaceae wie die Gattung *Theobroma* z.B. die Gattung und Art *Theobroma cacao* (Kakao) oder Theaceae wie die Gattung *Camellia* z.B. die Gattung und Art *Camellia sinensis* (Tee).
- 20 In einer vorteilhaften Ausführungsform des Verfahren werden als Nutzpflanzen Ölfruchtpflanzen verwendet, die große Mengen an Lipidverbindungen enthalten, wie Erdnuss, Raps, Canola, Sonnenblume, Saflor (*Carthamus tinctoria*), Mohn, Senf, Hanf, Rizinus, Olive, Sesam, Calendula, Punica, Nachtkerze, Königskerze, Distel, Wildrosen, Haselnuss, Mandel, *Macadamia*, Avocado, Lorbeer, Kürbis, Lein, Soja, Pistazien, Borretsch, Bäume (Ölpalme, Kokosnuss oder Walnuss) oder Feldfrüchte, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa oder Buschpflanzen (Kaffee, Kakao, Tee), *Salix*-Arten sowie ausdauernde Gräser und Futterfeldfrüchte. Vorteilhafte erfindungsgemäße Pflanzen sind Ölfruchtpflanzen, wie Erdnuss, Raps, Canola, Sonnenblume, Saflor, Mohn, Senf, Hanf, Rhizinus, Olive, Calendula, Punica, Nachtkerze, Kürbis, Lein, Soja, Borretsch, Bäume (Ölpalme, Kokosnuss). Besonders bevorzugt sind Pflanzen, die reich an C18:2- und/oder C18:3-Fettsäuren sind, wie

Sonnenblume, Färberdistel, Tabak, Königskerze, Sesam, Baumwolle, Kürbis, Mohn, Nachtkerze, Walnuss, Lein, Hanf oder Distel. Ganz besonders bevorzugt sind Pflanzen wie Färberdistel, Sonnenblume, Mohn, Nachtkerze, Walnuss, Lein oder Hanf.

5 Es ist auch vorteilhaft, die erfindungsgemäßen Nukleinsäuresequenzen in den Blättern von Futter- oder Nahrungspflanzen zu exprimieren und dadurch den Gehalt der Blätter an Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure zu steigern. Bevorzugte Futterpflanzen sind z.B. Kleearten wie Rotklee (*Trifolium pratense*), Weißklee (*Trifolium repens*), Bastardklee (*Trifolium hybridum*), Esparsette (*Onobrychis viciifolia*), Alexandrinerklee (*Trifolium alexandrinum*) und Perserklee (*Trifolium resupinatum*). Bevorzugte Nahrungspflanzen sind etwa Salatarten wie *Lactuca sativa*, *Lactuca crisper*, *Lactuca esculenta*, *Lactuca scariola* L. ssp. *sativa*, *Lactuca scariola* L. var. *integrata*, *Lactuca scariola* L. var. *integrifolia*, *Lactuca sativa* subsp. *romana*, *Locusta communis* und *Valeriana locusta*.

15

Durch die enzymatische Aktivität der im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen, die für Polypeptide mit  $\Delta$ -6-Elongase-,  $\Delta$ -6-Desaturase-,  $\Delta$ -5-Desaturase- und/oder  $\Delta$ -5-Elongase-Aktivität kodieren, vorteilhaft in Kombination mit Nukleinsäuresequenzen, die für Polypeptide mit  $\omega$ -3-Desaturase- und/oder  $\Delta$ -4-Desaturase-Aktivität kodieren, sowie weiteren Nukleinsäuresequenzen, die für Polypeptide des Fettsäure- oder Lipidstoffwechsels wie weiteren Polypeptiden mit  $\Delta$ -5-,  $\Delta$ -6-,  $\Delta$ -8-,  $\Delta$ -12-Desaturase- oder  $\Delta$ -5-,  $\Delta$ -6- oder  $\Delta$ -9-Elongaseaktivität kodieren, können unterschiedlichste mehrfach ungesättigte Fettsäuren im erfindungsgemäßen Verfahren hergestellt werden. Je nach Auswahl der für das erfindungsgemäße Verfahren verwendeten Nutzpflanzen lassen sich Mischungen der verschiedenen mehrfach ungesättigten Fettsäuren oder einzelne mehrfach ungesättigte Fettsäuren wie EPA, DPA oder DHA in freier oder gebundener Form herstellen. Je nachdem welche Fettsäurezusammensetzung in der Ausgangspflanze vorherrscht (C18:2- oder C18:3-Fettsäuren) entstehen so Fettsäuren, die sich von C18:2-Fettsäuren ableiten, wie GLA, DGLA oder ARA oder Fettsäuren, die sich von C18:3-Fettsäuren ableiten, wie EPA, DPA oder DHA. Liegt in der für das Verfahren verwendeten Pflanze als ungesättigte Fettsäure nur Linolsäure (= LA, C18:2 <sup>$\Delta$ 9,12</sup>) vor, so können als Produkte des Verfahrens nur GLA, DGLA und ARA ent-

20

25

30

stehen, die als freie Fettsäuren oder gebunden vorliegen können. Ist in der im Verfahren verwendeten Pflanze als ungesättigte Fettsäure nur  $\alpha$ -Linolensäure (= ALA, C18:3 <sup>$\Delta$ 9,12,15</sup>) vorhanden, beispielsweise wie in Lein, so können als Produkte des Verfahrens nur SDA, ETA, EPA, DPA und/oder DHA entstehen, die wie oben beschrieben als freie Fettsäuren oder gebunden vorliegen können. Durch Modifikation der Aktivität der im Verfahren verwendeten und an der Synthese beteiligten Enzyme  $\Delta$ -6-Elongase,  $\Delta$ -6-Desaturase,  $\Delta$ -5-Desaturase und/oder  $\Delta$ -6-Elongase vorteilhaft in Kombination mit weiteren Genen des Lipid- oder Fettsäurestoffwechsels lassen sich gezielt in den Pflanzen nur einzelne Produkte herstellen. Vorteilhaft werden nur EPA, DPA oder DHA oder deren Mischungen synthetisiert. Da die Fettsäuren in Biosyntheseketten synthetisiert werden, liegen die jeweiligen Endprodukte nicht als Reinsubstanzen in den Organismen vor. Es sind immer auch geringe Mengen der Vorläuferverbindungen im Endprodukt enthalten. Diese geringen Mengen betragen weniger als 20 Gew.-%, vorteilhaft weniger als 15 Gew.-%, besonders vorteilhaft weniger als 10 Gew.-%, ganz besonders vorteilhaft weniger als 5, 4, 3, 2 oder 1 Gew.-% bezogen auf die Endprodukte EPA, DPA oder DHA oder deren Mischungen.

Zur Steigerung der Ausbeute im beschriebenen Verfahren zur Herstellung von Ölen und/oder Triglyceriden mit einem vorteilhaft erhöhten Gehalt an mehrfach ungesättigten Fettsäuren ist es vorteilhaft, die Menge an Ausgangsprodukt für die Fettsäuresynthese zu steigern. Dies kann beispielsweise durch das Einbringen einer Nukleinsäure, die für ein Polypeptid mit  $\Delta$ -12-Desaturase kodiert, in den Organismus erreicht werden. Dies ist besonders vorteilhaft in Nutzpflanzen, wie Öl-produzierenden Pflanzen wie Pflanzen der Familie der Brassicaceae wie der Gattung Brassica z.B. Raps; der Familie der E-laeagnaceae wie die Gattung Elaeagnus z.B. die Gattung und Art Olea europaea oder der Familie Fabaceae wie der Gattung Glycine z.B. die Gattung und Art Glycine max, die einen hohen Ölsäuregehalt aufweisen. Da diese Organismen nur einen geringen Gehalt an Linolsäure aufweisen (Mikoklajczak et al. (1961) Journal of the American Oil Chemical Society 38: 678 - 681) ist die Verwendung der genannten  $\Delta$ -12-Desaturasen zur Herstellung des Ausgangsprodukts Linolsäure aus Ölsäure vorteilhaft. Daneben können die Ausgangsfettsäuren auch von außen zugefüttert werden, was aber aus Kostengründen weniger bevorzugt ist.



Moose und Algen sind die einzigen bekannten Pflanzensysteme, die erhebliche Mengen an mehrfach ungesättigten Fettsäuren, wie Arachidonsäure (ARA) und/oder Eicosapentaensäure (EPA) und/oder Docosahexaensäure (DHA) herstellen. Moose enthalten PUFAs in Membranlipiden, während Algen, algenverwandte Organismen und einige Pilze auch nennenswerte Mengen an PUFAs in der Triacylglycerolfraktion akkumulieren. Daher eignen sich Nukleinsäuremoleküle, die aus solchen Stämmen isoliert werden, die PUFAs auch in der Triacylglycerolfraktion akkumulieren, besonders vorteilhaft für das erfindungsgemäße Verfahren und damit zur Modifikation des Lipid- und PUFA-Produktionssystems in einer Pflanze, wie einer Nutzpflanze wie einer Ölfruchtpflanze, beispielsweise Raps, Canola, Lein, Hanf, Soja, Sonnenblumen, Borretsch. Sie sind deshalb vorteilhaft im erfindungsgemäßen Verfahren verwendbar.

Im erfindungsgemäßen Verfahren verwendete Nukleinsäuren stammen vorteilhaft aus Pflanzen wie Algen, beispielsweise Algen der Familie der Prasinophyceae wie aus den Gattungen *Heteromastix*, *Mammella*, *Mantoniella*, *Micromonas*, *Nephroselmis*, *Ostreococcus*, *Prasinocladus*, *Prasinococcus*, *Pseudoscourfielda*, *Pycnococcus*, *Pyramimonas*, *Scherffelia* oder *Tetraselmis* wie den Gattungen und Arten *Heteromastix longifillis*, *Mamiella gilva*, *Mantoniella squamata*, *Micromonas pusilla*, *Nephroselmis olivacea*, *Nephroselmis pyriformis*, *Nephroselmis rotunda*, *Ostreococcus tauri*, *Ostreococcus* sp., *Prasinocladus ascus*, *Prasinocladus lubricus*, *Pycnococcus provasolii*, *Pyramimonas amyliifera*, *Pyramimonas disomata*, *Pyramimonas obovata*, *Pyramimonas orientalis*, *Pyramimonas parkeae*, *Pyramimonas spinifera*, *Pyramimonas* sp., *Tetraselmis apiculata*, *Tetraselmis carteriaformis*, *Tetraselmis chui*, *Tetraselmis convolutae*, *Tetraselmis desikacharyi*, *Tetraselmis gracilis*, *Tetraselmis hazeni*, *Tetraselmis impellucida*, *Tetraselmis inconspicua*, *Tetraselmis levis*, *Tetraselmis maculata*, *Tetraselmis marina*, *Tetraselmis striata*, *Tetraselmis subcordiformis*, *Tetraselmis suecica*, *Tetraselmis tetrabrachia*, *Tetraselmis tetrathele*, *Tetraselmis verrucosa*, *Tetraselmis verrucosa* fo. *rubens* oder *Tetraselmis* sp. oder aus Algen der Familie Euglenaceae wie aus den Gattungen *Ascoglena*, *Astasia*, *Colacium*, *Cyclidiopsis*, *Euglena*, *Euglenopsis*, *Hyalophacus*, *Khawkinea*, *Lepocinclis*, *Phacus*, *Strombomonas* oder *Trachelomonas* wie die Gattungen und Art *Euglena acus*, *Euglena geniculata*, *Euglena gracilis*, *Euglena mixocylindracea*, *Euglena rostrifera*,



*Euglena viridis*, *Colacium stentorium*, *Trachelomonas cylindrica* oder *Trachelomonas volvocina*.

- Weitere vorteilhafte Pflanzen sind Algen wie *Isochrysis* oder *Cryptothecodinium*, Algen/  
5 Diatomeen wie *Thalassiosira* oder *Phaeodactylum*, Moose wie *Physcomitrella* o-  
der *Ceratodon* oder höheren Pflanzen wie den *Primulaceae* wie *Aleuritia*, *Calendula stel-*  
*lata*, *Osteospermum spinescens* oder *Osteospermum hyoseroides*, Mikroorganismen wie  
Pilzen wie *Aspergillus*, *Thraustochytrium*, *Phytophthora*, *Entomophthora*, *Mucor* oder  
10 *Mortierella*, Bakterien wie *Shewanella*, Hefen oder Tiere wie Nematoden wie *Ca-*  
*enorhabditis*, Insekten, Frösche, Seegurken oder Fische. Vorteilhaft stammen die erfin-  
dungsgemäßen isolierten Nukleinsäuresequenzen aus einem Tier aus der Ordnung der  
Vertebraten. Bevorzugt stammen die Nukleinsäuresequenzen aus der Klasse der Ver-  
tebrata; Euteleostomi, Actinopterygii; Neopterygii; Teleostei; Euteleostei, Protacanthopte-  
rygii, Salmoniformes; Salmonidae bzw. *Oncorhynchus* oder Vertebrata, Amphibia, Anura,  
15 *Pipidae*, *Xenopus* oder Evertebrata wie Protochordata, Tunicata, Holothuroidea, Cioni-  
dae wie *Amaroucium constellatum*, *Botryllus schlosseri*, *Ciona intestinalis*, *Molgula citri-*  
*na*, *Molgula manhattensis*, *Perophora viridis* oder *Styela partita*. Besonders vorteil-  
haft stammen die Nukleinsäuren aus Pilzen, Tieren oder aus Pflanzen wie Algen o-  
der Moosen, bevorzugt aus der Ordnung der Salmoniformes wie der Familie der Salmo-  
20 nidae wie der Gattung *Salmo* beispielsweise aus den Gattungen und Arten *Oncorhyn-*  
*chus mykiss*, *Trutta trutta* oder *Salmo trutta fario*, aus Algen wie den Gattungen *Manto-*  
*niella* oder *Ostreococcus* oder aus den Diatomeen wie den Gattungen *Thalassiosira* oder  
*Phaeodactylum* oder aus Algen wie *Cryptothecodinium*.
- 25 Bei einer bevorzugten Ausführungsform umfasst das Verfahren ferner den Schritt des  
Gewinnens einer Zelle oder einer ganzen Pflanze, die die im Verfahren verwendeten  
Nukleinsäuresequenzen, die für eine  $\Delta$ -6-Desaturase,  $\Delta$ -6-Elongase,  $\Delta$ -5-Desaturase  
und/oder  $\Delta$ -5-Elongase sowie ggf. Nukleinsäuresequenzen, die für eine  $\omega$ -3-Desaturase  
und/oder eine  $\Delta$ -4-Desaturase kodieren, enthält, wobei die Zelle und/oder die Nutzpflan-  
30 ze noch weitere Nukleinsäuresequenzen des Lipid- oder Fettsäurestoffwechsels enthal-  
ten kann. Die im Verfahren bevorzugt verwendeten Nukleinsäuresequenzen werden zur  
Expression vorteilhaft in mindestens ein Genkonstrukt und/oder einen Vektor wie nach-

folgend beschrieben, allein oder in Kombination mit weiteren Nukleinsäuresequenzen, die für Proteine des Fettsäure- oder Lipidstoffwechsels kodieren, eingebaut und schließlich in die Zelle oder Pflanze transformiert. Bei einer weiteren bevorzugten Ausführungsform umfasst dieses Verfahren ferner den Schritt des Gewinnens der Öle, Lipide oder freien Fettsäuren aus den Nutzpflanzen. Die so hergestellte Zelle oder die so hergestellte Nutzpflanze ist vorteilhaft eine Zelle einer Öl-produzierenden Pflanze, Gemüse-, Salat-, oder Zierpflanze oder die Pflanze selbst wie oben ausgeführt.

Unter Anzucht ist beispielsweise die Kultivierung im Falle von Pflanzenzellen, -gewebe oder -organe auf oder in einem Nährmedium oder der ganzen Pflanze auf bzw. in einem Substrat beispielsweise in Hydrokultur, Blumentopferde oder auf einem Ackerboden zu verstehen.

"Transgen" bzw. "rekombinant" im Sinne der Erfindung bedeutet bezüglich zum Beispiel einer Nukleinsäuresequenz, einer Expressionskassette (= Genkonstrukt) oder einem Vektor enthaltend die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen oder einer mit den im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen, Expressionskassette oder Vektor transformierten Pflanze alle solche durch gentechnische Methoden zustande gekommenen Konstruktionen, in denen sich entweder

- a) die Nukleinsäuresequenz, oder
- b) eine mit der Nukleinsäuresequenz funktionell verknüpfte genetische Kontrollsequenz, zum Beispiel ein Promotor, oder
- c) (a) und (b)

nicht in ihrer natürlichen, genetischen Umgebung befinden oder durch gentechnische Methoden modifiziert wurden, wobei die Modifikation beispielhaft eine Substitution, Addition, Deletion, Inversion oder Insertion eines oder mehrerer Nukleotidreste sein kann. Natürliche genetische Umgebung meint den natürlichen genomischen bzw. chromosomalen Locus in dem Herkunftsorganismus oder das Vorliegen in einer genomischen Bibliothek. Im Fall einer genomischen Bibliothek ist die natürliche, genetische Umgebung

der Nukleinsäuresequenz bevorzugt zumindest noch teilweise erhalten. Die Umgebung flankiert die Nukleinsäuresequenz zumindest an einer Seite und hat eine Sequenzlänge von mindestens 50 bp, bevorzugt mindestens 500 bp, besonders bevorzugt mindestens 1000 bp, ganz besonders bevorzugt mindestens 5000 bp. Eine natürlich vorkommende

5 Expressionskassette - beispielsweise die natürlich vorkommende Kombination des natürlichen Promotors der im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenz mit der Nukleinsäuresequenz, die für Proteine mit entsprechender  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase-,  $\Delta$ -5-Desaturase- und  $\Delta$ -5-Elongase-Aktivität kodiert, vorteilhaft in Kombination mit Nukleinsäuresequenzen, die für Proteine mit  $\omega$ -3-Desaturase- und/oder  $\Delta$ -4-Desaturase-Aktivität kodieren - wird zu einer transgenen Expressionskassette, wenn diese durch nicht-natürliche, synthetische ("künstliche") Verfahren wie beispielsweise einer Mutagenisierung geändert wird. Entsprechende Verfahren sind beispielsweise beschrieben in US 5,565,350 oder WO 00/15815.

15 Unter "transgener Pflanze" im Sinne der Erfindung ist wie vorgenannt zu verstehen, dass die im Verfahren verwendeten Nukleinsäuren nicht an ihrer natürlichen Stelle im Genom der Pflanze sind. Dabei können die Nukleinsäuresequenzen homolog oder heterolog exprimiert werden. Transgen bedeutet aber auch, dass die erfindungsgemäßen Nukleinsäuren an ihrem natürlichen Platz im Genom der Pflanze sind, dass jedoch die Sequenz gegenüber der natürlichen Sequenz verändert wurde und/oder dass die Regulationssequenzen der natürlichen Sequenzen verändert wurden. Bevorzugt ist unter transgen die Expression der im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren an nicht-natürlicher Stelle im Genom zu verstehen, das heißt eine homologe oder bevorzugt heterologe Expression der Nukleinsäuresequenzen liegt vor.

25

Bevorzugte transgene Organismen sind Nutzpflanzen wie Öl-produzierende Pflanzen, Gemüse-, Salat- oder Zierpflanzen, die vorteilhaft ausgewählt sind aus der Gruppe der Pflanzenfamilien bestehend aus den Familien der Aceraceae, Actinidiaceae, Anacardiaceae, Apiaceae, Arecaceae, Asteraceae, Arecaceae, Betulaceae, Boraginaceae, Brassicaceae, Bromeliaceae, Cannabaceae, Cannaceae, Caprifoliaceae, Chenopodiaceae, Convolvulaceae, Cucurbitaceae, Dioscoreaceae, Elaeagnaceae, Ericaceae, Euphorbiaceae, Fabaceae, Fagaceae, Grossulariaceae, Juglandaceae, Lauraceae, Liliaceae, Li-

30

naceae, Malvaceae, Moraceae, Musaceae, Oleaceae, Oxalidaceae, Papaveraceae, Poaceae, Polygonaceae, Punicaceae, Rosaceae, Rubiaceae, Rutaceae, Scrophulariaceae, Solanaceae, Sterculiaceae und Valerianaceae.

- 5 Als Wirtspflanzen für die im erfindungsgemäßen Verfahren verwendeten Nucleinsäuren, die Expressionskassette oder den Vektor eignen sich prinzipiell vorteilhaft alle Nutzpflanzen, die in der Lage sind Fettsäuren, speziell ungesättigte Fettsäuren, zu synthetisieren bzw. die für die Expression rekombinanter Gene geeignet sind. Beispielhaft seien an dieser Stelle Pflanzen wie Arabidopsis, Asteraceae wie Calendula oder Nutzpflanzen wie  
10 Soja, Erdnuss, Rizinus, Sonnenblume, Mais, Baumwolle, Flachs, Raps, Kokosnuss, Ölpalme, FärberSaflor (*Carthamus tinctorius*) oder Kakaobohne genannt. Weitere vorteilhafte Pflanzen sind an anderer Stelle dieser Anmeldung aufgeführt.

- Für die Herstellung der transgenen Nutzpflanze werden in der Regel als Zwischenwirte  
15 Mikroorganismen verwendet. Derartige nutzbare Zwischenwirtszellen werden in: Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990) genannt.

- Vorteilhaft verwendbare Expressionsstämme für diesen Zweck sind z.B. solche, die eine geringe Proteaseaktivität aufweisen. Sie werden z.B. in: Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128 beschrieben.  
20

- Transgene Pflanzen, die die im erfindungsgemäßen Verfahren synthetisierten mehrfach ungesättigten, lange-kettigen Fettsäuren enthalten, können vorteilhaft direkt vermarktet werden, ohne dass die synthetisierten Öle, Lipide oder Fettsäuren isoliert werden müssen. Diese Form der Vermarktung ist besonders vorteilhaft.  
25

- Bei den "Pflanzen" im Sinne der vorliegenden Erfindung handelt es sich um ganze Pflanzen sowie alle Pflanzenteile, Pflanzenorgane oder Pflanzenteile wie Blatt, Stiel, Samen, Wurzel, Knollen, Antheren, Fasern, Wurzelhaare, Stängel, Embryos, Kalli, Kotelydonen, Petiolen, Erntematerial, pflanzliches Gewebe, reproduktives Gewebe, Zellkulturen, die  
30

sich von der transgenen Pflanze ableiten und/oder dazu verwendet werden können, die transgene Pflanze hervorzubringen. Der Samen umfasst dabei alle Samenteile wie die Samenhüllen, Epidermis- und Samenzellen, Endosperm oder Embryogewebe.

- 5 Die im erfindungsgemäßen Verfahren hergestellten Verbindungen können aber auch aus den Pflanzen in Form ihrer Öle, Fett, Lipide und/oder freien Fettsäuren isoliert werden. Durch das erfindungsgemäße Verfahren hergestellte mehrfach ungesättigte Fettsäuren lassen sich durch Ernten der Pflanzen oder Pflanzenzellen entweder aus der Kultur, in der sie wachsen, oder vom Feld gewinnen. Dies kann über Pressen oder Extraktion
- 10 der Pflanzenteile, bevorzugt der Pflanzensamen, erfolgen. Dabei können die Öle, Fette, Lipide und/oder freien Fettsäuren durch so genanntes Kaltschlagen oder Kaltpressen ohne Zuführung von Wärme durch Pressen gewonnen werden. Damit sich die Pflanzenteile, speziell die Samen, leichter aufschließen lassen, werden sie vorher zerkleinert, gedämpft oder geröstet. Die so vorbehandelten Samen können anschließend gepresst
- 15 oder mit Lösungsmittel wie warmem Hexan extrahiert werden. Anschließend wird das Lösungsmittel wieder entfernt. Auf diese Weise können mehr als 96 % der im erfindungsgemäßen Verfahren hergestellten Verbindungen isoliert werden. Anschließend werden die so erhaltenen Produkte weiter bearbeitet, das heißt raffiniert. Dabei werden zunächst beispielsweise die Pflanzenschleime und Trübstoffe entfernt. Die sogenannte
- 20 Entschleimung kann enzymatisch oder beispielsweise chemisch/physikalisch durch Zugabe von Säure wie Phosphorsäure erfolgen. Anschließend werden die freien Fettsäuren durch Behandlung mit einer Base beispielsweise Natronlauge entfernt. Das erhaltene Produkt wird zur Entfernung der im Produkt verbliebenen Lauge mit Wasser gründlich gewaschen und getrocknet. Um die noch im Produkt enthaltenen Farbstoffe zu entfer-
- 25 nen, werden die Produkte einer Bleichung mit beispielsweise Bleicherde oder Aktivkohle unterzogen. Zum Schluss wird das Produkt noch beispielsweise mit Wasserdampf desodoriert.

- Vorzugsweise sind die durch dieses Verfahren produzierten PUFAs bzw. LCPUFAs C<sub>20</sub>-
- 30 und/oder C<sub>22</sub>-Fettsäuremoleküle mit mindestens vier Doppelbindungen im Fettsäuremolekül, vorzugsweise fünf oder sechs Doppelbindungen. Diese C<sub>20</sub>- und/oder C<sub>22</sub>-Fettsäuremoleküle lassen sich aus der Pflanze in Form eines Öls, Lipids oder einer frei-

en Fettsäure isolieren. Geeignete transgene Pflanzen sind beispielsweise die vorstehend erwähnten.

Diese erfindungsgemäßen Öle, Lipide oder Fettsäuren enthalten wie oben beschrieben vorteilhaft 6 bis 15 % Palmitinsäure, 1 bis 6 % Stearinsäure; 7 – 85 % Ölsäure; 0,5 bis 8 % Vaccensäure, 0,1 bis 1 % Arachinsäure, 7 bis 25 % gesättigte Fettsäuren, 8 bis 85 % einfach ungesättigte Fettsäuren und 60 bis 85 % mehrfach ungesättigte Fettsäuren jeweils bezogen auf 100 % und auf den Gesamtfettsäuregehalt der Pflanzen.

Als vorteilhafte mehrfach ungesättigte, langkettige Fettsäuren sind in den Fettsäureestern bzw. Fettsäuregemischen wie Phosphatidylfettsäureestern oder Triacylglyceridestern bevorzugt mindestens 10; 11; 12; 13; 14; 15; 16; 17; 18; 19 oder 20 Gew.-% bezogen auf den Gesamtfettsäuregehalt an Eicosapentaensäure und/oder mindestens 1; 2; 3; 4; 5 oder 6 Gew.-% bezogen auf den Gesamtfettsäuregehalt an Docosapentaensäure und/oder mindestens 1; 2; 3; bevorzugt mindestens 4; 5; 6; besonders bevorzugt mindestens 7 oder 8 und am meisten bevorzugt mindestens 9 oder 10 Gew.-% bezogen auf den Gesamtfettsäuregehalt an Docosahexaensäure enthalten.

Weiterhin enthalten die Fettsäureester bzw. Fettsäuregemische, die nach dem erfindungsgemäßen Verfahren hergestellt wurden, Fettsäuren ausgewählt aus der Gruppe der Fettsäuren Erucasäure (13-Docosaensäure), Sterculinsäure (9,10-Methylene octadec-9-enonsäure), Malvalinsäure (8,9-Methylen Heptadec-8-enonsäure), Chaulmoogrinsäure (Cyclopentendodecansäure), Furan-Fettsäure (9,12-Epoxy-octadeca-9,11-dienonsäure), Vernonsäure (9,10-Epoxyoctadec-12-enonsäure), Tarinsäure (6-Octadecynonsäure), 6-Nonadecynonsäure, Santalbinsäure (t11-Octadecen-9-ynoic acid), 6,9-Octadecenynonsäure, Pyrulinsäure (t10-Heptadecen-8-ynonsäure), Crepenyninsäure (9-Octadecen-12-ynonsäure), 13,14-Dihydrooropheinsäure, Octadecen-13-ene-9,11-diyonsäure, Petroselensäure (cis-6-Octadecenonsäure), 9c,12t-Octadecadiensäure, Calendulasäure (8t10t12c-Octadecatriensäure), Catalpinsäure (9t11t13c-Octadecatriensäure), Eleosterinsäure (9c11t13t-Octadecatriensäure), Jacarinsäure (8c10t12c-Octadecatriensäure), Punicinsäure (9c11t13c-Octadecatriensäure), Parinarinsäure (9c11t13t15c-Octadecatetraensäure), Pinolensäure (all-cis-5,9,12-

Octadecatriensäure), Laballensäure (5,6-Octadecadienallensäure), Ricinolsäure (12-Hydroxyölsäure) und/oder Coriolinsäure (13-Hydroxy-9c,11t-Octadecadienonsäure). Die vorgenannten Fettsäuren kommen in den nach dem erfindungsgemäßen Verfahren hergestellten Fettsäureester bzw. Fettsäuregemischen in der Regel vorteilhaft nur in Spuren vor, das heißt sie kommen bezogen auf die Gesamtfettsäuren zu weniger als 30 %, bevorzugt zu weniger als 25 %, 24 %, 23 %, 22 % oder 21 %, besonders bevorzugt zu weniger als 20 %, 15 %, 10 %, 9 %, 8 %, 7%, 6 % oder 5%, ganz besonders bevorzugt zu weniger als 4 %, 3 %, 2 % oder 1 % vor. In einer weiteren bevorzugten Form der Erfindung kommen diese vorgenannten Fettsäuren bezogen auf die Gesamtfettsäuren zu weniger als 0,9%; 0,8%; 0,7%; 0,6%; oder 0,5%, besonders bevorzugt zu weniger als 0,4%; 0,3%; 0,2%; 0,1% vor. Vorteilhaft enthalten die nach dem erfindungsgemäßen Verfahren hergestellten Fettsäureester bzw. Fettsäuregemische weniger als 0,1 % bezogen auf die Gesamtfettsäuren und/oder keine Buttersäure, kein Cholesterin sowie keine Nisinsäure.

Eine weitere erfindungsgemäße Ausführungsform ist die Verwendung der Öle, Lipide, der Fettsäuren und/oder der Fettsäurezusammensetzung, die nach dem erfindungsgemäßen Verfahren hergestellt werden, in Futtermitteln, Nahrungsmitteln, Kosmetika oder Pharmazeutika. Die im erfindungsgemäßen Verfahren gewonnenen Öle, Lipide, Fettsäuren oder Fettsäuregemische können in der dem Fachmann bekannten Weise zur Abmischung mit anderen Ölen, Lipiden, Fettsäuren oder Fettsäuregemischen tierischen Ursprungs wie z.B. Fischölen verwendet werden. Auch diese so hergestellten Öle, Lipide, Fettsäuren oder Fettsäuregemische, die aus pflanzlichen und tierischen Bestandteilen bestehen, können zur Herstellung von Futtermitteln, Nahrungsmitteln, Kosmetika oder Pharmazeutika verwendet werden.

Unter dem Begriff "Öl", "Lipid" oder "Fett" wird ein Fettsäuregemisch verstanden, das ungesättigte und/oder gesättigte, vorzugsweise veresterte Fettsäure(n) enthält. Bevorzugt ist, dass das Öl, Lipid oder Fett einen hohen Anteil an mehrfach ungesättigten freien oder vorteilhaft veresterten Fettsäure(n), insbesondere Linolsäure,  $\gamma$ -Linolensäure, Dihomo- $\gamma$ -linolensäure, Arachidonsäure,  $\alpha$ -Linolensäure, Stearidonsäure, Eicosatetraensäure, Eicosapentaensäure, Docosapentaensäure oder Docosahexaensäure hat. Vor-



zugsweise ist der Anteil an ungesättigten veresterten Fettsäuren ungefähr 30 %, besonders bevorzugt ist ein Anteil von 50 %, am meisten bevorzugt ist ein Anteil von 60 %, 70 %, 80 % oder mehr. Der Anteil an Fettsäure kann nach Überführung der Fettsäuren in die Methylester durch Umesterung gaschromatographisch bestimmt werden. Das Öl,  
5 Lipid oder Fett kann verschiedene andere gesättigte oder ungesättigte Fettsäuren, z.B. Calendulasäure, Palmitin-, Palmitolein-, Stearin-, Ölsäure etc., enthalten. Insbesondere kann je nach Ausgangspflanze der Anteil der verschiedenen Fettsäuren in dem Öl oder Fett schwanken.

- 10 Bei den im Verfahren hergestellten mehrfach ungesättigte Fettsäuren mit vorteilhaft mindestens drei, vier, fünf oder sechs, besonders vorteilhaft mit fünf oder sechs Doppelbindungen, handelt es sich wie oben beschrieben vorteilhaft um Fettsäureester beispielsweise um Sphingolipidester, Phosphoglyceridester, Lipidester, Glycolipidester, Phospholipidester, Monoacylglycerinester, Diacylglycerinester, Triacylglycerinester oder sonstige  
15 Fettsäureester, bevorzugt handelt es sich um Phospholipidester und/oder Triacylglycerinester.

- Aus den so im erfindungsgemäßen Verfahren hergestellten mehrfach ungesättigte Fettsäureestern mit vorteilhaft mindestens drei, vier, fünf oder sechs Doppelbindungen lassen sich die enthaltenen mehrfach ungesättigten Fettsäuren beispielsweise über eine  
20 Alkalibehandlung, beispielsweise mit wässriger KOH oder NaOH, oder durch saure Hydrolyse vorteilhaft in Gegenwart eines Alkohols wie Methanol oder Ethanol oder über eine enzymatische Abspaltung freisetzen und isolieren über beispielsweise Phasentrennung und anschließende Ansäuerung mit z.B. H<sub>2</sub>SO<sub>4</sub>. Die Freisetzung der Fettsäuren kann  
25 auch direkt ohne die vorhergehend beschriebene Aufarbeitung erfolgen.

- Als Substrate der im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen, die für Polypeptide mit  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase-,  $\Delta$ -5-Desaturase- und/oder  $\Delta$ -5-Elongase-Aktivität sowie ggf. Nukleinsäuresequenzen, die für Polypeptide  
30 mit  $\omega$ -3-Desaturase- und/oder  $\Delta$ -4-Desaturase-Aktivität kodieren, und/oder den weiteren verwendeten Nukleinsäuren wie den Nukleinsäuresequenzen, die für Polypeptide des Fettsäure- oder Lipidstoffwechsels ausgewählt aus der Gruppe Acyl-CoA-

Dehydrogenase(n), Acyl-ACP(= acyl carrier protein)-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Acyl-CoA:Lysophospholipid-Acyltransferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenase(n), Lipoxygenase(n), Triacylglycerol-Lipase(n), Allenoxid-Synthase(n), Hydroperoxid-Lyase(n) oder Fettsäure-Elongase(n) kodieren, eignen sich vorteilhaft C<sub>16</sub>-, C<sub>18</sub>- oder C<sub>20</sub>-Fettsäuren. Bevorzugt werden die im Verfahren als Substrate umgesetzten Fettsäuren in Form ihrer Acyl-CoA-Ester und/oder ihrer Phospholipid-Ester umgesetzt.

10

Zur Herstellung der erfindungsgemäßen langkettigen PUFAs müssen die gesättigten, einfach ungesättigten C<sub>16</sub>-Fettsäuren und/oder mehrfach ungesättigten C<sub>18</sub>-Fettsäuren zunächst je nach Substrat durch die enzymatische Aktivität einer Desaturase und/oder Elongase desaturiert und/oder elongiert oder nur desaturiert und anschließend über eine Elongase um mindestens zwei Kohlenstoffatome verlängert werden. Nach einer Elongationsrunde führt diese Enzymaktivität entweder ausgehend von C<sub>16</sub>-Fettsäuren zu C<sub>18</sub>-Fettsäuren oder ausgehend von C<sub>18</sub>-Fettsäuren zu C<sub>20</sub>-Fettsäuren, und nach zwei Elongationsrunden ausgehend von C<sub>16</sub>-Fettsäuren zu C<sub>20</sub>-Fettsäuren. Die Aktivität der im erfindungsgemäßen Verfahren verwendeten Desaturasen und Elongasen führt vorzugsweise zu C<sub>20</sub>- und/oder C<sub>22</sub>-Fettsäuren vorteilhaft mit mindestens zwei oder drei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit vier, fünf oder sechs Doppelbindungen, besonders bevorzugt zu C<sub>22</sub>-Fettsäuren mit mindestens fünf Doppelbindungen im Fettsäuremolekül. Besonders bevorzugt als Produkte des erfindungsgemäßen Verfahrens sind Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure. Die C<sub>18</sub>-Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure können durch die erfindungsgemäße enzymatische Aktivität in Form der freien Fettsäure oder in Form der Ester, wie Phospholipide, Glycolipide, Sphingolipide, Phosphoglyceride, Monoacylglycerin, Diacylglycerin oder Triacylglycerin, verlängert werden.

30

Der bevorzugte Biosyntheseort von Fettsäuren, Ölen, Lipiden oder Fette in den vorteilhaft verwendeten Pflanzen ist beispielsweise im allgemeinen der Samen oder Zellschichten des Samens, so dass eine samenspezifische Expression der im Verfahren verwen-

deten Nukleinsäuren sinnvoll ist. Es ist jedoch naheliegend, dass die Biosynthese von Fettsäuren, Ölen oder Lipiden nicht auf das Samengewebe beschränkt sein muss, sondern auch in allen übrigen Teilen der Pflanze - beispielsweise in Epidermiszellen oder in den Knollen - gewebespezifisch erfolgen kann. Vorteilhaft findet die Synthese gemäß

5 des erfinderischen Verfahrens im vegetativen (somatischen) Gewebe statt.

Durch das erfindungsgemäße Verfahren können die hergestellten mehrfach ungesättigten Fettsäuren in den im Verfahren verwendeten Pflanzen prinzipiell auf zwei Arten erhöht werden. Es kann vorteilhaft der Pool an freien mehrfach ungesättigten Fettsäuren und/oder der Anteil der über das Verfahren hergestellten veresterten mehrfach ungesättigten Fettsäuren erhöht werden. Vorteilhaft wird durch das erfindungsgemäße Verfahren

10 der Pool an veresterten mehrfach ungesättigten Fettsäuren in den transgenen Pflanzen erhöht, vorteilhaft in Form der Phosphatidylester und/oder Triacyl ester.

Die im erfindungsgemäßen Verfahren verwendeten Sequenzen werden einzeln in Ex-

15 pressionskonstrukte kloniert oder auf einem gemeinsamen rekombinanten Nukleinsäuremolekül bereitgestellt und zum Einbringen und zur Expression in Organismen verwendet. Diese Expressionskonstrukte ermöglichen eine optimale Synthese der im erfindungsgemäßen Verfahren produzierten mehrfach ungesättigten Fettsäuren.

Die im Verfahren verwendeten Nukleinsäuren können nach Einbringung in eine Pflanze oder Pflanzenzelle entweder auf einem separaten Plasmid liegen oder vorteilhaft in das Genom der Wirtszelle integriert sein. Bei Integration in das Genom kann die Integration zufallsgemäß sein oder durch derartige Rekombination erfolgen, dass das native Gen durch die eingebrachte Kopie ersetzt wird, wodurch die Produktion der gewünschten

20 Verbindung durch die Zelle moduliert wird, oder durch Verwendung eines Gens in trans, so dass das Gen mit einer funktionellen Expressionseinheit, welche mindestens eine die Expression eines Gens gewährleistende Sequenz und mindestens eine die Polyadenylierung eines funktionell transkribierten Gens gewährleistende Sequenz enthält, funktionell verbunden ist. Vorteilhaft werden die Nukleinsäuresequenzen über Multiexpressionskassetten oder Konstrukte zur multiparallelen Expression in die Pflanzen gebracht,

25 30 d.h. die Nukleinsäuresequenzen liegen in einer gemeinsamen Expressionseinheit vor.

Es kann im Nukleinsäurekonstrukt mehr als eine Nukleinsäuresequenz kodierend für ein Polypeptid mit der enzymatischen Aktivität einer  $\Delta$ -12-Desaturase,  $\Delta$ -4-Desaturase,  $\Delta$ -5-Desaturase,  $\Delta$ -6-Desaturase,  $\Delta$ -5-Elongase,  $\Delta$ -6-Elongase und/oder  $\omega$ -3-Desaturase enthalten sein. Es können auch mehrere Kopien einer Nukleinsäuresequenz kodierend für ein Polypeptid mit der enzymatischen Aktivität einer  $\Delta$ -12-Desaturase,  $\Delta$ -4-Desaturase,  $\Delta$ -5-Desaturase,  $\Delta$ -6-Desaturase,  $\Delta$ -5-Elongase,  $\Delta$ -6-Elongase und/oder  $\omega$ -3-Desaturase enthalten sein.

Zum Einbringen werden die im Verfahren verwendeten Nukleinsäuren vorteilhaft einer Amplifikation und Ligation in bekannter Weise unterworfen. Vorzugsweise geht man in Anlehnung an das Protokoll der Pfu-DNA-Polymerase oder eines Pfu/Taq-DNA-Polymerasegemisches vor. Die Primer werden in Anlehnung an die zu amplifizierende Sequenz gewählt. Zweckmäßigerweise sollten die Primer so gewählt werden, dass das Amplifikat die gesamte kodogene Sequenz vom Start- bis zum Stop-Kodon umfasst. Im Anschluss an die Amplifikation wird das Amplifikat zweckmäßigerweise analysiert. Beispielsweise kann die Analyse nach gelelektrophoretischer Auftrennung hinsichtlich Qualität und Quantität erfolgen. Im Anschluss kann das Amplifikat nach einem Standardprotokoll gereinigt werden (z.B. Qiagen). Ein Aliquot des gereinigten Amplifikats steht dann für die nachfolgende Klonierung zur Verfügung. Geeignete Klonierungsvektoren sind dem Fachmann allgemein bekannt. Hierzu gehören insbesondere Vektoren, die in mikrobiellen Systemen replizierbar sind, also vor allem Vektoren, die eine effiziente Klonierung in Hefen oder Pilze gewährleisten, und die die stabile Transformation von Pflanzen ermöglichen. Zu nennen sind insbesondere verschiedene für die T-DNA-vermittelte Transformation geeignete, binäre und co-integrierte Vektorsysteme. Derartige Vektorsysteme sind in der Regel dadurch gekennzeichnet, dass sie zumindest die für die Agrobacterium-vermittelte Transformation benötigten vir-Gene sowie die T-DNA begrenzende Sequenzen (T-DNA-Border) beinhalten. Vorzugsweise umfassen diese Vektorsysteme auch weitere cis-regulatorische Regionen wie Promotoren und Terminatoren und/oder Selektionsmarker, mit denen entsprechend transformierte Organismen identifiziert werden können. Während bei co-integrierten Vektorsystemen vir-Gene und T-DNA-Sequenzen auf demselben Vektor angeordnet sind, basieren binäre Systeme auf wenigstens zwei Vektoren, von denen einer vir-Gene, aber keine T-DNA und ein zweiter T-

DNA, jedoch kein vir-Gen trägt. Dadurch sind letztere Vektoren relativ klein, leicht zu manipulieren und sowohl in *E. coli* als auch in *Agrobacterium* zu replizieren. Zu diesen binären Vektoren gehören Vektoren der Serien pBIB-HYG, pPZP, pBecks, pGreen. Erfindungsgemäß bevorzugt verwendet werden Bin19, pBI101, pBinAR, pGPTV und pCAMBIA. Eine Übersicht über binäre Vektoren und ihre Verwendung gibt Hellens et al. (2000) Trends in Plant Science 5: 446–451. Für die Vektorpräparation können die Vektoren zunächst mit Restriktionsendonuklease(n) linearisiert und dann in geeigneter Weise enzymatisch modifiziert werden. Im Anschluss wird der Vektor gereinigt und ein Aliquot für die Klonierung eingesetzt. Bei der Klonierung wird das enzymatisch geschnittene und erforderlichenfalls gereinigte Amplifikat mit ähnlich präparierten Vektorfragmenten unter Einsatz von Ligase kloniert. Dabei kann ein bestimmtes Nukleinsäurekonstrukt bzw. Vektor- oder Plasmidkonstrukt einen oder auch mehrere kodogene Genabschnitte aufweisen. Vorzugsweise sind die kodogenen Genabschnitte in diesen Konstrukten mit regulatorischen Sequenzen funktional verknüpft. Zu den regulatorischen Sequenzen gehören insbesondere pflanzliche Sequenzen wie die oben beschriebenen Promotoren und Terminatoren. Die Konstrukte lassen sich vorteilhafterweise in Mikroorganismen, insbesondere *Escherichia coli* und *Agrobacterium tumefaciens*, unter selektiven Bedingungen stabil propagieren und ermöglichen so einen Transfer von heterologer DNA in Pflanzen.

Unter der vorteilhaften Verwendung von Klonierungsvektoren können die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen und Nukleinsäurekonstrukte in Mikroorganismen und danach in Pflanzen eingebracht werden und damit bei der Pflanzentransformation verwendet werden, wie denjenigen, die veröffentlicht sind in und dort zitiert sind: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Kapitel 6/7, S. 71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, 15-38; B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. (1991) 42: 205-225. Die im Verfahren verwendeten Nukleinsäuren, Nukleinsäurekonstrukte und/oder Vektoren lassen sich damit zur gentechnologischen Veränderung eines breiten

Spektrums an Pflanzen verwenden, so dass diese bessere und/oder effizientere Produzenten von LCPUFAs werden.

5 Durch das Einbringen eines  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase-,  $\Delta$ -5-Desaturase- und  $\Delta$ -5-Elongase-Genes in eine Pflanze allein oder in Kombination mit anderen Genen kann nicht nur der Biosynthesefluss zum Endprodukt erhöht, sondern auch die entsprechende Triacylglycerin- und/oder Phosphatidylester-Zusammensetzung erhöht oder de novo ge-  
10 und/oder neutralen Lipiden nötig sind, erhöht sein, so dass die Konzentration dieser Vorläufer, Cofaktoren oder Zwischenverbindungen innerhalb der Zellen oder innerhalb des Speicherkompartiments erhöht ist, wodurch die Fähigkeit der Zellen zur Produktion von PUFAs, wie im Folgenden beschrieben, weiter gesteigert wird. Durch Optimierung der Aktivität oder Erhöhung der Anzahl eines oder mehrerer der  $\Delta$ -6-Desaturase-,  $\Delta$ -6-  
15 Elongase-,  $\Delta$ -5-Desaturase- und/oder  $\Delta$ -5-Elongase-Gene, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Zerstören der Aktivität einer oder mehrerer Gene, die am Abbau dieser Verbindungen beteiligt sind, kann es möglich sein, die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmolekülen aus Organismen und vorteilhaft aus Pflanzen zu steigern.

20

Die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuremoleküle kodieren für Proteine oder Teile von diesen, wobei die Proteine oder das einzelne Protein oder Teile davon eine Aminosäuresequenz enthält, die ausreichend homolog zu einer Aminosäuresequenz ist, die in den Sequenzen SEQ ID NO. 65, SEQ ID NO. 2, SEQ ID NO.  
25 172 oder SEQ ID NO. 52 und ggf. SEQ ID NO. 194 oder SEQ ID NO. 78 dargestellt ist, so dass die Proteine oder Teile davon noch eine  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase-,  $\Delta$ -5-Desaturase- und/oder  $\Delta$ -5-Elongase-Aktivität sowie ggf. eine  $\Delta$ -4-Desaturase- und/oder  $\omega$ -3-Desaturase-Aktivität aufweisen. Vorzugsweise haben die Proteine oder Teile davon, die von dem Nukleinsäuremolekül/den Nukleinsäuremolekülen kodiert wird/werden, noch  
30 seine/ihre wesentliche enzymatische Aktivität und die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen oder Lipidkörperchen in Organismen vorteilhaft in Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese

Membranen teilzunehmen. Vorteilhaft sind die von den Nukleinsäuremolekülen kodierten Proteine zu mindestens etwa 60 % und bevorzugt mindestens etwa 70 %, 80 % oder 90 % und besonders bevorzugt mindestens etwa 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % oder mehr identisch zu den in SEQ ID NO. 65, SEQ ID NO. 2, SEQ ID NO. 172, SEQ ID NO. 52, SEQ ID NO. 194 oder SEQ ID NO. 78 dargestellten Aminosäuresequenzen. Im Sinne der Erfindung ist unter Homologie oder homolog, Identität oder identisch zu verstehen.

Die Homologie wurde über den gesamten Aminosäure- bzw. Nukleinsäuresequenzbereich berechnet. Für den Vergleich verschiedener Sequenzen stehen dem Fachmann eine Reihe von Programmen zur Verfügung, die auf verschiedenen Algorithmen beruhen. Dabei liefern die Algorithmen von Needleman und Wunsch oder Smith und Waterman besonders zuverlässige Ergebnisse. Für die Sequenzvergleiche wurde das Programm PileUp verwendet (J. Mol. Evolution (1987) 25: 351-360; Higgins et al. (1989) CABIOS 5: 151-153) oder die Programme Gap und BestFit (Needleman and Wunsch (1970) J. Mol. Biol. 48: 443-453 und Smith and Waterman (1981) Adv. Appl. Math. 2: 482-489), die im GCG Software-Paket (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711 (1991)) enthalten sind. Die oben in Prozent angegebenen Sequenzhomologiewerte wurden mit dem Programm GAP über den gesamten Sequenzbereich mit folgenden Einstellungen ermittelt: Gap Weight: 50, Length Weight: 3, Average Match: 10.000 und Average Mismatch: 0.000. Diese Einstellungen wurden, falls nicht anders angegeben, immer als Standardeinstellungen für Sequenzvergleiche verwendet.

Unter wesentlicher enzymatischer Aktivität der im erfindungsgemäßen Verfahren verwendeten  $\omega$ -3-Desaturase,  $\Delta$ -6-Desaturase,  $\Delta$ -6-Elongase,  $\Delta$ -5-Elongase,  $\Delta$ -4-Desaturase und/oder  $\Delta$ -5-Desaturase ist zu verstehen, dass sie im Vergleich zu den durch die Sequenz mit der SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 kodierten Proteinen/Enzymen noch eine enzymatische Aktivität von mindestens 10 %, bevorzugt von mindestens 20 %, besonders bevorzugt von mindestens 30 % und am meisten bevorzugt von mindestens 40, 50 oder 60 % aufweisen und damit am Stoffwechsel von zum Aufbau von Fettsäuren, vor-



teilhaft Fettsäureestern wie Phosphatidylestern und/oder Triacylglyceridestern, in einer Pflanze oder Pflanzenzelle notwendigen Verbindungen oder am Transport von Molekülen über Membranen teilnehmen können.

- 5 Vorteilhaft im Verfahren verwendbare Nucleinsäuren stammen aus Bakterien, Pilzen, Diatomeen, Tieren wie *Caenorhabditis* oder *Oncorhynchus* oder Pflanzen wie Algen oder Moosen wie den Gattungen *Shewanella*, *Physcomitrella*, *Thraustochytrium*, *Fusarium*, *Phytophthora*, *Ceratodon*, *Pytium irregulare*, *Mantoniella*, *Ostreococcus*, *Isochrysis*, *Aleurita*, *Muscarioides*, *Mortierella*, *Borago*, *Phaeodactylum*, *Cryptocodinium*, speziell
- 10 aus den Gattungen und Arten *Pytium irregulare*, *Oncorhynchus mykiss*, *Xenopus laevis*, *Ciona intestinalis*, *Thalassiosira pseudonona*, *Mantoniella squamata*, *Ostreococcus* sp., *Ostreococcus tauri*, *Euglena gracilis*, *Physcomitrella patens*, *Phytophthora infestans*, *Fusarium gramineum*, *Cryptocodinium cohnii*, *Ceratodon purpureus*, *Isochrysis galbana*, *Aleurita farinosa*, *Thraustochytrium* sp., *Muscarioides viallii*, *Mortierella alpina*, *Borago*
- 15 *officinalis*, *Phaeodactylum tricornutum*, *Caenorhabditis elegans* oder besonders vorteilhaft aus *Pytium irregulare*, *Thraustochytrium* sp. und/oder *Ostreococcus tauri*.

Im erfindungsgemäßen Verfahren können zusätzlich Nucleotidsequenzen verwendet werden, die für eine  $\Delta$ -12-Desaturase,  $\Delta$ -9-Elongase oder  $\Delta$ -8-Desaturase kodieren. Die

20 im Verfahren verwendeten Nucleinsäuresequenzen werden vorteilhaft in einer Expressionskassette, die die Expression der Nucleinsäuren in Pflanzen ermöglicht, eingebracht.

Die Nucleinsäuresequenzen, die für die  $\Delta$ -12-Desaturase,  $\omega$ -3-Desaturase,  $\Delta$ -9-Elongase,  $\Delta$ -6-Desaturase,  $\Delta$ -8-Desaturase,  $\Delta$ -6-Elongase,  $\Delta$ -5-Desaturase,  $\Delta$ -5-Elongase oder  $\Delta$ -4-Desaturase kodieren, werden mit einem oder mehreren Regulations-

25 signalen zur Erhöhung der Genexpression funktionell verknüpft. Diese regulatorischen Sequenzen sollen die gezielte Expression der Gene ermöglichen. Dies kann beispielsweise je nach Pflanze bedeuten, dass das Gen erst nach Induktion exprimiert und/oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird. Vorteil-

30 haft werden Sequenzen für die Expression verwendet, die eine konstitutive Expression ermöglichen, wie der CaMV35S-, CaMV36S-, CaMV35Smas-, nos-, mas-, ubi-, stpt-, lea- oder Super-Promotor. Bevorzugt erfolgt die Expression im vegetativen Gewebe wie oben

beschrieben. In einer anderen bevorzugten Ausführungsform erfolgt die Expression im Samen.

Beispielsweise handelt es sich bei diesen regulatorischen Sequenzen um Sequenzen, an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu den Regulationssequenzen, die in ihrem natürlichen Locus nicht mit den Nukleinsäuresequenzen verknüpft sind, oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht ist. Das Genkonstrukt kann außerdem vorteilhafterweise auch eine oder mehrere sogenannte "Enhancer-Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Vorteilhafte Terminatoren sind beispielweise virale Terminatoren wie der 35S-Terminator oder andere. Die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen können in einer oder mehreren Kopien in der Expressionskassette (= Genkonstrukt) enthalten sein. Vorteilhaft liegt nur jeweils eine Kopie der Gene in der Expressionskassette vor. Dieses Genkonstrukt oder die Genkonstrukte können gleichzeitig oder nacheinander in die Pflanze eingebracht werden und zusammen im Wirtsorganismus exprimiert werden. Dabei kann das Genkonstrukt oder die Genkonstrukte in einem oder mehreren Vektoren inseriert sein und frei in der Zelle vorliegen oder aber im Genom inseriert sein. Es ist vorteilhaft für die Insertion weiterer Gene in die Pflanze, wenn die zu exprimierenden Gene zusammen in einem Genkonstrukt vorliegen. Es ist aber auch möglich, jeweils ein Genkonstrukt enthaltend eine Nukleinsäuresequenz in eine Pflanze einzuführen und die so erhaltenen Pflanzen miteinander zu verkreuzen, um Nachkommen zu erhalten, die alle Genkonstrukte gemeinsam enthalten.

Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf

der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

- 5 Um eine stabile Integration der Biosynthesegene in die transgene Pflanze über mehrere Generation sicherzustellen, sollte jede der im Verfahren verwendeten Nukleinsäuren, die für die  $\Delta$ -6-Desaturase,  $\Delta$ -6-Elongase,  $\Delta$ -5-Desaturase oder  $\Delta$ -5-Elongase und ggf. für die  $\omega$ -3-Desaturase oder  $\Delta$ -4-Desaturase kodieren, unter der Kontrolle eines eigenen Promotors exprimiert werden. Dieser kann für jede der Sequenzen gleich oder unterschiedlich sein. Die Expressionskassette ist dabei vorteilhaft so aufgebaut, dass einem Promotor eine geeignete Schnittstelle zur Insertion der zu exprimierenden Nukleinsäure folgt, die vorteilhaft in einem Polylinker liegt. Hinter dem Polylinker kann ggf. ein Terminator liegen. Diese Abfolge wiederholt sich mehrfach, bevorzugt drei-, vier-, fünf- oder sechsmal, so dass bis zu sechs Gene in einem Konstrukt zusammengeführt werden und so zur Expression in die transgene Pflanze eingebracht werden können.
- 10 Die Nukleinsäuresequenzen werden zur Expression über die geeignete Schnittstelle beispielsweise im Polylinker hinter den Promotor inseriert. Vorteilhaft hat jede Nukleinsäuresequenz ihren eigenen Promotor und gegebenenfalls ihren eigenen Terminator. Es ist aber auch möglich, mehrere Nukleinsäuresequenzen hinter einem Promotor und ggf. vor einem Terminator zu inserieren. Dabei ist die Insertionsstelle bzw. die Abfolge der inserierten Nukleinsäuren in der Expressionskassette nicht von entscheidender Bedeutung, das heißt eine Nukleinsäuresequenz kann an erster oder letzter Stelle in der Kassette inseriert sein, ohne dass durch die die Position die Expression wesentlich beeinflusst wird. Es können in der Expressionskassette in einer vorteilhaften Ausführungsform unterschiedliche Promotoren wie beispielsweise der USP-, LegB4 oder DC3-Promotor und unterschiedliche Terminatoren verwendet werden. In einer weiteren vorteilhaften Ausführungsform können auch identische Promotoren wie der CaMV35S-Promotor verwendet werden.
- 25
- 30 Wie oben beschrieben sollte die Transkription der eingebrachten Gene vorteilhaft durch geeignete Terminatoren am 3'-Ende der eingebrachten Biosynthesegene (hinter dem Stopcodon) abgebrochen werden. Verwendet werden kann hier z.B. der OCS1- oder der

35SCaMV-Terminator. Wie auch für die Promotoren, sollten hier für jedes Gen unterschiedliche Terminatorsequenzen verwendet werden.

- Das Genkonstrukt kann, wie oben beschrieben, auch weitere Gene umfassen, die in die Organismen eingebracht werden sollen. Es ist möglich und vorteilhaft, in die Wirtspflanzen Regulationsgene, wie Gene für Induktoren, Repressoren oder Enzyme, welche durch ihre Enzymaktivität in die Regulation eines oder mehrerer Gene eines Biosynthesewegs eingreifen, einzubringen und darin zu exprimieren. Diese Gene können heterologen oder homologen Ursprungs sein. Weiterhin können vorteilhaft im Nukleinsäurekonstrukt bzw. Genkonstrukt weitere Biosynthesegene des Fettsäure- oder Lipidstoffwechsels enthalten sein oder aber diese Gene können auf einem weiteren oder mehreren weiteren Nukleinsäurekonstrukten liegen. Vorteilhaft werden als Biosynthesegene des Fettsäure- oder Lipidstoffwechsels ein oder mehrere Gene ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP(= acyl carrier protein)-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Acyl-CoA:Lysophospholipid-Acyltransferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenase(n), Lipoxygenase(n), Triacylglycerol-Lipase(n), Allenoxid-Synthase(n), Hydroperoxid-Lyase(n) oder Fettsäure-Elongase(n) oder Kombinationen davon verwendet. Besonders vorteilhafte Nukleinsäuresequenzen sind Biosynthesegene des Fettsäure- oder Lipidstoffwechsels ausgewählt aus der Gruppe der Acyl-CoA:Lysophospholipid-Acyltransferase,  $\Delta$ -8-Desaturase,  $\Delta$ -9-Desaturase,  $\Delta$ -12-Desaturase und/oder  $\Delta$ -9-Elongase.
- Dabei können die vorgenannten Nukleinsäuren bzw. Gene in Kombination mit anderen Elongasen und Desaturasen in Expressionskassetten, wie den vorgenannten, kloniert werden und zur Transformation von Pflanzen mit Hilfe von Agrobacterium eingesetzt werden.
- Der in dieser Beschreibung verwendete Begriff "Vektor" steht für ein Nukleinsäuremolekül, das eine andere Nukleinsäure transportieren kann, an welche es gebunden ist. Ein Vektortyp ist ein "Plasmid", eine zirkuläre doppelsträngige DNA-Schleife, in die zu-

sätzlichen DNA-Segmente ligiert werden können. Ein weiterer Vektortyp ist ein viraler Vektor, wobei zusätzliche DNA-Segmente in das virale Genom ligiert werden können. Bestimmte Vektoren können in einer Wirtszelle, in die sie eingebracht worden sind, autonom replizieren (z.B. Bakterienvektoren mit bakteriellem Replikationsursprung). Andere

5 Vektoren werden vorteilhaft beim Einbringen in die Wirtszelle in das Genom einer Wirtszelle integriert und dadurch zusammen mit dem Wirtsgenom repliziert. Zudem können bestimmte Vektoren die Expression von Genen, mit denen sie funktionsfähig verbunden sind, steuern. Diese Vektoren werden hier als "Expressionsvektoren" bezeichnet. Gewöhnlich haben Expressionsvektoren, die für DNA-Rekombinationstechniken geeignet

10 sind, die Form von Plasmiden. In der vorliegenden Beschreibung können "Plasmid" und "Vektor" austauschbar verwendet werden, da das Plasmid die am häufigsten verwendete Vektorform ist. Die Erfindung soll jedoch auch andere Expressionsvektorformen, wie virale Vektoren, die ähnliche Funktionen ausüben, umfassen. Ferner soll der Begriff Vektor auch andere Vektoren, die dem Fachmann bekannt sind, wie Phagen, Viren wie

15 SV40, CMV, TMV, Transposons, IS-Elemente, Phasmide, Phagemide, Cosmide, lineare oder zirkuläre DNA, umfassen.

Die im Verfahren vorteilhaft verwendeten rekombinanten Expressionsvektoren umfassen die im Verfahren verwendeten Nukleinsäuresequenzen oder das oben beschriebene

20 Genkonstrukt in einer Form, die sich zur Expression der verwendeten Nukleinsäuren in einer Wirtszelle eignet, was bedeutet, dass die rekombinanten Expressionsvektoren eine oder mehrere Regulationssequenzen, die auf der Basis der zur Expression zu verwendenden Wirtszellen ausgewählt ist und die mit der zu exprimierenden Nukleinsäuresequenz funktionsfähig verbunden ist, umfasst. In einem rekombinanten Expressionsvektor

25 bedeutet "funktionsfähig verbunden", dass die Nukleotidsequenz von Interesse derart an die Regulationssequenz(en) gebunden ist, dass die Expression der Nukleotidsequenz ermöglicht wird und sie aneinander gebunden sind, so dass beide Sequenzen die vorhergesagte, der Sequenz zugeschriebene Funktion erfüllen (z.B. in einem In-vitro-Transkriptions-/Translationssystem oder in einer Wirtszelle, wenn der Vektor in die

30 Wirtszelle eingebracht wird). Der Begriff "Regulationssequenz" soll Promotoren, Enhancer und andere Expressionskontrollelemente (z.B. Polyadenylierungssignale) umfassen. Diese Regulationssequenzen sind z.B. beschrieben in Goeddel: Gene Expression Tech-

nology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), oder siehe: Gruber und Crosby, in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, Hrsgb.: Glick und Thompson, Kapitel 7, 89-108, einschließlich der Literaturstellen darin. Regulationssequenzen umfassen solche, welche die konstitutive Expression einer Nukleotidsequenz in vielen Wirtszelltypen steuern, und solche, welche die direkte Expression der Nukleotidsequenz nur in bestimmten Wirtszellen unter bestimmten Bedingungen steuern. Der Fachmann weiß, dass die Gestaltung des Expressionsvektors von Faktoren, wie der Auswahl der zu transformierenden Wirtszelle, dem gewünschten Ausmaß der Expression des Proteins, usw. abhängen kann.

10

Die verwendeten rekombinanten Expressionsvektoren können zur Expression der im Verfahren verwendeten Nukleinsäuresequenzen so gestaltet sein, dass sie in prokaryotische Zwischenwirte transformiert werden können und schließlich nach Einbringung in die Pflanzen die Expression der Gene in diesen ermöglichen. Dies ist vorteilhaft, da häufig Zwischenschritte der Vektorkonstruktion der Einfachheit halber in Mikroorganismen durchgeführt werden. Beispielsweise können die  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase-,  $\Delta$ -5-Desaturase- und/oder  $\Delta$ -5-Elongase-Gene in bakteriellen Zellen, Insektenzellen (unter Verwendung von Baculovirus-Expressionsvektoren), Hefe- und anderen Pilzzellen (siehe Romanos, M.A., et al. (1992) *Yeast* 8:423-488; van den Hondel, C.A.M.J.J., et al. (1991) "Heterologous gene expression in filamentous fungi", in: *More Gene Manipulations in Fungi*, J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego; und van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: *Applied Molecular Genetics of Fungi*, Peberdy, J.F., et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge), Algen (Falciatore et al. (1999) *Marine Biotechnology*. 1: (3):239-251), Ciliaten, mit Vektoren nach einem Transformationsverfahren, wie beschrieben in WO 98/01572, sowie bevorzugt in Zellen vielzelliger Pflanzen (siehe Schmidt, R. und Willmitzer, L. (1988) "High efficiency *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants" *Plant Cell Rep.*:583-586; *Plant Molecular Biology and Biotechnology*, C Press, Boca Raton, Florida, Kapitel 6/7, S.71-119 (1993); F.F. White, B. Jenes et al., *Techniques for Gene Transfer*, in: *Transgenic Plants*, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-43; Potrykus (1991) *Annu. Rev. Plant*

30

Physiol. Plant Molec. Biol. 42: 205-225 (und darin zitierte Literaturstellen)) exprimiert werden. Geeignete Wirtszellen werden ferner erörtert in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Der rekombinante Expressionsvektor kann alternativ, zum Beispiel unter Verwendung  
5 von T7-Promotor-Regulationssequenzen und T7-Polymerase, in vitro transkribiert und translatiert werden.

Die Expression von Proteinen in Prokaryonten erfolgt meist mit Vektoren, die konstitutive oder induzierbare Promotoren enthalten, welche die Expression von Fusions- oder nicht-  
10 Fusionsproteinen steuern. Typische Fusions-Expressionsvektoren sind u.a. pGEX (Pharmacia Biotech Inc; Smith, D.B., und Johnson, K.S. (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRIT5 (Pharmacia, Piscataway, NJ), bei denen Glutathion-S-Transferase (GST), Maltose E-bindendes Protein bzw. Protein A an das rekombinante Zielprotein fusioniert wird.

15 Beispiele für geeignete induzierbare nicht-Fusions-E. coli-Expressionsvektoren sind u.a. pTrc (Amann et al. (1988) Gene 69:301-315) und pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 60-89). Die Zielgenexpression vom pTrc-Vektor beruht auf der Transkription  
20 durch Wirts-RNA-Polymerase von einem Hybrid-trp-lac-Fusionspromotor. Die Zielgenexpression aus dem pET 11d-Vektor beruht auf der Transkription von einem T7-gn10-lac-Fusions-Promotor, die von einer koexprimierten viralen RNA-Polymerase (T7 gn1) vermittelt wird. Diese virale Polymerase wird von den Wirtsstämmen BL21 (DE3) oder HMS174 (DE3) von einem residenten  $\lambda$ -Prophagen bereitgestellt, der ein T7 gn1-Gen  
25 unter der Transkriptionskontrolle des lacUV 5-Promotors birgt.

Andere in prokaryotischen Organismen geeignete Vektoren sind dem Fachmann bekannt, diese Vektoren sind beispielsweise in E. coli pLG338, pACYC184, die pBR-Reihe, wie pBR322, die pUC-Reihe, wie pUC18 oder pUC19, die M113mp-Reihe,  
30 pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1,  $\lambda$ gt11 or pBdCl, in Streptomyces pIJ101, pIJ364, pIJ702 oder pIJ361, in Bacillus pUB110, pC194 oder pBD214, in Corynebacterium pSA77 oder pAJ667.



Bei einer weiteren Ausführungsform ist der Expressionsvektor ein Hefe-Expressionsvektor. Beispiele für Vektoren zur Expression in der Hefe *S. cerevisiae* umfassen pYe-Desaturasec1 (Baldari et al. (1987) *Embo J.* 6:229-234), pMFa (Kurjan und Herskowitz  
5 (1982) *Cell* 30:933-943), pJRY88 (Schultz et al. (1987) *Gene* 54:113-123) sowie pYES2 (Invitrogen Corporation, San Diego, CA). Vektoren und Verfahren zur Konstruktion von Vektoren, die sich zur Verwendung in anderen Pilzen, wie den filamentösen Pilzen, eignen, umfassen diejenigen, die eingehend beschrieben sind in: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: *Applied Molecular Genetics of fungi*, J.F. Peberdy et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge, oder in: *More Gene Manipulations in Fungi* (J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego). Weitere geeignete Hefevektoren sind beispielsweise pAG-1, YEp6, YEp13 oder pEMBLYe23.

15 Alternativ können die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen in Insektenzellen unter Verwendung von Baculovirus-Expressionsvektoren exprimiert werden. Baculovirus-Vektoren, die zur Expression von Proteinen in gezüchteten Insektenzellen (z.B. Sf9-Zellen) verfügbar sind, umfassen die pAc-Reihe (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) und die pVL-Reihe (Lucklow und Summers (1989)  
20 *Virology* 170:31-39).

Die oben genannten Vektoren bieten nur einen kleinen Überblick über mögliche geeignete Vektoren. Weitere Plasmide sind dem Fachmann bekannt und sind zum Beispiel beschrieben in: *Cloning Vectors* (Hrsgb. Pouwels, P.H., et al., Elsevier, Amsterdam-  
25 New York-Oxford, 1985, ISBN 0 444 904018). Weitere geeignete Expressionssysteme für prokaryotische und eukaryotische Zellen siehe in den Kapiteln 16 und 17 von Sambrook, J., Fritsch, E.F., und Maniatis, T., *Molecular Cloning: A Laboratory Manual*, 2. Auflage, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

30

Auch können die im Verfahren verwendeten Gen in einzelligen Pflanzenzellen (wie Algen), siehe Falciatore et al. (1999) *Marine Biotechnology* 1 (3):239-251 und darin zitierte

Literaturangaben, und Pflanzenzellen aus höheren Pflanzen (z.B. Spermatophyten, wie Feldfrüchten) exprimiert werden. Beispiele für Pflanzen-Expressionsvektoren umfassen solche, die eingehend beschrieben sind in: Becker, D., Kemper, E., Schell, J., und Masterson, R. (1992) *Plant Mol. Biol.* 20:1195-1197; und Bevan, M.W. (1984) *Nucl. Acids Res.* 12:8711-8721; *Vectors for Gene Transfer in Higher Plants*; in: *Transgenic Plants*, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, S. 15-38.

Eine Pflanzen-Expressionskassette enthält vorzugsweise Regulationssequenzen, welche die Genexpression in Pflanzenzellen steuern können und funktionsfähig verbunden sind, so dass jede Sequenz ihre Funktion, wie Termination der Transkription, erfüllen kann, beispielsweise Polyadenylierungssignale. Bevorzugte Polyadenylierungssignale sind diejenigen, die aus *Agrobacterium tumefaciens*-T-DNA stammen, wie das als Octopinsynthese bekannte Gen 3 des Ti-Plasmids pTiACH5 (Gielen et al. (1984) *EMBO J.* 3: 835ff.) oder funktionelle Äquivalente davon, aber auch alle anderen in Pflanzen funktionell aktiven Terminatoren sind geeignet.

Da die Regulation der Pflanzengenexpression sehr oft nicht auf die Transkriptionsebene beschränkt ist, enthält eine Pflanzen-Expressionskassette vorzugsweise andere funktionsfähig verbundene Sequenzen, wie Translationsenhancer, beispielsweise die Overdrive-Sequenz, welche die 5'-untranslatierte Leader-Sequenz aus Tabakmosaikvirus, die das Protein/RNA-Verhältnis erhöht, enthält (Gallie et al. (1987) *Nucl. Acids Research* 15:8693-8711).

Die Pflanzengenexpression muss wie oben beschrieben funktionsfähig mit einem geeigneten Promotor verbunden sein, der die Genexpression steuert. Vorteilhaft nutzbare Promotoren sind konstitutive Promotoren (Benfey et al., *EMBO J.* (1989) 8: 2195-2202), wie diejenigen, die von Pflanzenviren stammen, wie 35S CAMV (Franck et al. (1980) *Cell* 21: 285-294), 19S CaMV (siehe auch US 5352605 und WO 84/02913) oder Pflanzenpromotoren, wie der in US 4,962,028 beschriebene der kleinen Untereinheit der Rubisco.

Andere bevorzugte Sequenzen für die Verwendung zur funktionsfähigen Verbindung in Pflanzengenexpressions-Kassetten sind Targeting-Sequenzen, die zur Steuerung des Genproduktes in sein entsprechendes Zellkompartiment), beispielsweise in die Vakuole, den Zellkern, alle Arten von Plastiden, wie Amyloplasten, Chloroplasten, Chromoplasten, den extrazellulären Raum, die Mitochondrien, das Endoplasmatische Retikulum, Ölkörper, Peroxisomen und andere Kompartimente von Pflanzenzellen; notwendig sind (siehe eine Übersicht in Kermodé (1996) Crit. Rev. Plant Sci. 15 (4): 285-423 und darin zitierte Literaturstellen).

10 Vektor-DNA lässt sich in prokaryotische oder eukaryotische Zellen über herkömmliche Transformations- oder Transfektionstechniken einbringen. Die Begriffe "Transformation" und "Transfektion", Konjugation und Transduktion, wie hier verwendet, sollen eine Vielzahl von im Stand der Technik bekannten Verfahren zum Einbringen fremder Nucleinsäuren (z.B. DNA) in eine Wirtszelle, einschließlich Calciumphosphat- oder Calciumchlorid-Kopräzipitation, DEAE-Dextran-vermittelte Transfektion, Lipofektion, natürliche Kompetenz, chemisch vermittelter Transfer, Elektroporation oder Teilchenbeschuss, 15 umfassen. Geeignete Verfahren zur Transformation oder Transfektion von Wirtszellen, einschließlich Pflanzenzellen, lassen sich finden in Sambrook et al. (Molecular Cloning: A Laboratory Manual., 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) und anderen Labor-Handbüchern, wie 20 Methods in Molecular Biology, 1995, Bd. 44, Agrobacterium protocols, Hrsgb: Gartland und Davey, Humana Press, Totowa, New Jersey.

Der Begriff "Nucleinsäure(molekül)", wie hier verwendet, umfasst in einer vorteilhaften 25 Ausführungsform zudem die am 3'- und am 5'-Ende des kodierenden Genbereichs gelegene untranslatierte Sequenz: mindestens 500, bevorzugt 200, besonders bevorzugt 100 Nucleotide der Sequenz stromaufwärts des 5'-Endes des kodierenden Bereichs und mindestens 100, bevorzugt 50, besonders bevorzugt 20 Nucleotide der Sequenz stromabwärts des 3'-Endes des kodierenden Genbereichs. Ein "isoliertes" Nucleinsäuremolekül 30 wird von anderen Nucleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nucleinsäure vorliegen. Eine "isolierte" Nucleinsäure hat vorzugsweise keine Sequenzen, welche die Nucleinsäure in der genomischen DNA des Organismus, aus dem

die Nukleinsäure stammt, natürlicherweise flankieren (z.B. Sequenzen, die sich an den 5'- und 3'-Enden der Nukleinsäure befinden). Bei verschiedenen Ausführungsformen kann das im Verfahren verwendete isolierte  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase- oder  $\Delta$ -5-Desaturase- sowie ggf. das  $\omega$ -3-Desaturase- oder  $\Delta$ -4-Desaturase-Molekül zum Beispiel  
5 weniger als etwa 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb oder 0,1 kb an Nukleotidsequenzen enthalten, die natürlicherweise das Nukleinsäuremolekül in der genomischen DNA der Zelle, aus der die Nukleinsäure stammt, flankieren.

Die im Verfahren verwendeten Nukleinsäuremoleküle können unter Verwendung molekularbiologischer Standardtechniken und der hier bereitgestellten Sequenzinformation  
10 isoliert werden. Auch kann mit Hilfe von Vergleichsalgorithmen beispielsweise eine homologe Sequenz oder homologe, konservierte Sequenzbereiche auf DNA- oder Aminosäureebene identifiziert werden. Diese können als Hybridisierungssonde in Standard-Hybridisierungstechniken (wie z.B. beschrieben in Sambrook et al., Molecular Cloning: A  
15 Laboratory Manual. 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) zur Isolierung weiterer im Verfahren nützlicher Nukleinsäuresequenzen verwendet werden. Überdies lassen sich die im Verfahren verwendeten Nukleinsäuremoleküle oder Teile von diesen durch Polymerasekettenreaktion  
20 isolieren, wobei Oligonukleotidprimer auf der Basis dieser Sequenz oder von Teilen davon verwendet werden (z.B. kann ein Nukleinsäuremolekül umfassend die vollständige Sequenz oder einen Teil davon durch Polymerasekettenreaktion unter Verwendung von Oligonukleotidprimern isoliert werden, die auf der Basis dieser gleichen Sequenz erstellt worden sind). Zum Beispiel lässt sich mRNA aus Zellen isolieren (z.B. durch das Guanidiniumthiocyanat-Extraktionsverfahren von Chirgwin et al. (1979) Biochemistry 18:5294-  
25 5299) und cDNA mittels Reverser Transkriptase (z.B. Moloney-MLV-Reverse-Transkriptase, erhältlich von Gibco/BRL, Bethesda, MD, oder AMV-Reverse-Transkriptase, erhältlich von Seikagaku America, Inc., St.Petersburg, FL) herstellen. Synthetische Oligonukleotidprimer zur Amplifizierung mittels Polymerasekettenreaktion lassen sich auf der Basis einer der in SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171,  
30 SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 gezeigten Sequenzen oder mit Hilfe der in SEQ ID NO. 65, SEQ ID NO. 2, SEQ ID NO. 172, SEQ ID NO. 52, SEQ ID NO. 194 oder SEQ ID NO. 78 dargestellten Aminosäuresequenzen erstellen. Eine erfin-

dungsgemäße Nucleinsäure kann unter Verwendung von cDNA oder alternativ von genomischer DNA als Matrize und geeigneten Oligonucleotidprimern nach Standard-PCR-Amplifikationstechniken amplifiziert werden. Die so amplifizierte Nucleinsäure kann in einen geeigneten Vektor kloniert werden und mittels DNA-Sequenzanalyse charakterisiert werden. Oligonucleotide können durch Standard-Syntheseverfahren, beispielsweise mit einem automatischen DNA-Synthesegerät, hergestellt werden.

Homologe der verwendeten  $\Delta$ -5-Elongase-,  $\omega$ -3-Desaturase-,  $\Delta$ -6-Desaturase-,  $\Delta$ -6-Elongase-,  $\Delta$ -4-Desaturase- oder  $\Delta$ -5-Desaturase-Nucleinsäuresequenzen mit der Sequenz SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 bedeutet beispielsweise allelische Varianten mit mindestens etwa 40, 50 oder 60 %, vorzugsweise mindestens etwa 60 oder 70 %, stärker bevorzugt mindestens etwa 70 oder 80 %, 90 % oder 95 % und noch stärker bevorzugt mindestens etwa 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % oder mehr Identität bzw. Homologie zu einer der in SEQ ID NO. 64, 66, 68 oder 70, zu einer der in SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39 oder 41, zu einer der in SEQ ID NO. 171, 173, 175, 177, 179, 181 oder 183, zu einer der in SEQ ID NO. 51, 53 oder 55, zu einer der in SEQ ID NO. 193 oder 195 oder zu einer der in oder SEQ ID NO. 77, 79, 81, 83, 85, 87, 89, 91 oder 93, insbesondere der in SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 gezeigten Nucleotidsequenzen oder ihren Homologen, Derivaten oder Analoga oder Teilen davon. Weiterhin sind isolierte Nucleinsäuremoleküle einer Nucleotidsequenz umfasst, die an eine der in SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 gezeigten Nucleotidsequenzen oder einen Teil davon, z.B. unter stringenten Bedingungen hybridisieren. Unter einem Teil gemäß der Erfindung ist dabei zu verstehen, dass mindestens 25 Basenpaare (= bp), 50 bp, 75 bp, 100 bp, 125 bp oder 150 bp, bevorzugt mindestens 175 bp, 200 bp, 225 bp, 250 bp, 275 bp oder 300 bp, besonders bevorzugt 350 bp, 400 bp, 450 bp, 500 bp oder mehr Basenpaare für die Hybridisierung verwendet werden. Es kann auch vorteilhaft die Gesamtsequenz verwendet werden. Allelische Varianten umfassen insbesondere funktionelle Varianten, die sich durch Deletion, Insertion oder Substitution von Nucleotiden aus der in SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID

NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 dargestellten Sequenz erhalten lassen, wobei aber die Enzymaktivität der davon kodierten Proteine für die Insertion im wesentlichen beibehalten wird.

- 5 Für das erfindungsgemäße Verfahren vorteilhafte Nukleinsäuremoleküle können auf der Grundlage ihrer Homologie zu den hier offenbarten  $\omega$ -3-Desaturase-,  $\Delta$ -6-Desaturase-,  $\Delta$ -5-Desaturase-,  $\Delta$ -5-Elongase-,  $\Delta$ -4-Desaturase- und/oder  $\Delta$ -6-Elongase-Nukleinsäuresequenzen unter Verwendung der Sequenzen oder eines Teils davon als Hybridisierungssonde gemäß Standard-Hybridisierungstechniken unter stringenten Hybridisierungsbedingungen isoliert werden. Dabei können beispielsweise isolierte Nukleinsäuremoleküle verwendet werden, die mindestens 15 Nukleotide lang sind und unter stringenten Bedingungen mit den Nukleinsäuremolekülen, die eine Nukleotidsequenz der SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder Seq ID No. 77 umfassen, hybridisieren. Es können auch Nukleinsäuremoleküle mit mindestens 25, 50, 100, 250 oder mehr Nukleotiden verwendet werden.

- Der Begriff "hybridisiert unter stringenten Bedingungen", wie hier verwendet, soll Hybridisierungs- und Waschbedingungen beschreiben, unter denen Nukleotidsequenzen, die mindestens 60 % homolog zueinander sind, gewöhnlich aneinander hybridisiert bleiben.
- 20 Die Bedingungen sind vorzugsweise derart, dass Sequenzen, die mindestens etwa 65 %, bevorzugt mindestens etwa 70 % und besonders bevorzugt mindestens etwa 75 % oder stärker zueinander homolog sind, gewöhnlich aneinander hybridisiert bleiben. Diese stringenten Bedingungen sind dem Fachmann bekannt und lassen sich in Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6.
- 25 finden. Ein bevorzugtes, nicht einschränkendes Beispiel für stringente Hybridisierungsbedingungen sind Hybridisierungen in 6 x Natriumchlorid/Natriumcitrat (sodium chloride/sodium citrate = SSC) bei etwa 45°C, gefolgt von einem oder mehreren Waschschritten in 0,2 x SSC, 0,1 % SDS bei 50 bis 65°C. Dem Fachmann ist bekannt, dass diese Hybridisierungsbedingungen sich je nach dem Typ der Nukleinsäure und, wenn beispielsweise organische Lösungsmittel vorliegen, hinsichtlich der Temperatur und der
- 30 Konzentration des Puffers unterscheiden. Die Temperatur liegt beispielsweise unter "Standard-Hybridisierungsbedingungen" je nach dem Typ der Nukleinsäure zwischen

42°C und 58°C in wässrigem Puffer mit einer Konzentration von 0,1 bis 5 x SSC (pH 7,2). Falls organisches Lösungsmittel, zum Beispiel 50 % Formamid, im oben genannten Puffer vorliegt, ist die Temperatur unter Standardbedingungen etwa 42°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:DNA-Hybride zum Beispiel 0,1 x SSC und 20°C bis 45°C, vorzugsweise 30°C bis 45°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:RNA-Hybride zum Beispiel 0,1 x SSC und 30°C bis 55°C, vorzugsweise 45°C bis 55°C. Die vorstehend genannten Hybridisierungstemperaturen sind beispielsweise für eine Nukleinsäure mit etwa 100 bp (= Basenpaare) Länge und einem G + C-Gehalt von 50 % in Abwesenheit von Formamid bestimmt. Der Fachmann weiß, wie die erforderlichen Hybridisierungsbedingungen anhand von Lehrbüchern, wie dem vorstehend erwähnten oder aus den folgenden Lehrbüchern Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989; Hames und Higgins (Hrsgb.) 1985, "Nucleic Acids Hybridization: A Practical Approach", IRL Press at Oxford University Press, Oxford; Brown (Hrsgb.) 1991, "Essential Molecular Biology: A Practical Approach", IRL Press at Oxford University Press, Oxford, bestimmt werden können.

Zur Bestimmung der prozentualen Homologie (= Identität) von zwei Aminosäuresequenzen (z.B. einer der Sequenzen der SEQ ID NO. 65, SEQ ID NO. 2, SEQ ID NO. 172, SEQ ID NO. 52, SEQ ID NO. 194 oder SEQ ID NO. 78) oder von zwei Nukleinsäuren (z.B. SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77) werden die Sequenzen untereinander geschrieben, um sie optimal vergleichen zu können (z.B. können Lücken in die Sequenz eines Proteins oder einer Nukleinsäure eingefügt werden, um ein optimales Alignment mit dem anderen Protein oder der anderen Nukleinsäure zu erzeugen). Die Aminosäurereste oder Nukleotide an den entsprechenden Aminosäurepositionen oder Nukleotidpositionen werden dann verglichen. Wenn eine Position in einer Sequenz durch den gleichen Aminosäurerest oder das gleiche Nukleotid wie die entsprechende Stelle in der anderen Sequenz belegt wird, dann sind die Moleküle an dieser Position homolog (d.h. Aminosäure- oder Nukleinsäure-"Homologie", wie hier verwendet, entspricht Aminosäure- oder Nukleinsäure-"Identität"). Die prozentuale Homologie zwischen den beiden Sequenzen ist eine Funktion der Anzahl an identischen Positionen, die den Sequenzen gemeinsam sind (d.h. % Homologie = Anzahl der identischen Positionen/Gesamtanzahl der Positionen



x 100). Die zur Bestimmung der Homologie verwendeten Programme bzw. Algorithmen sind oben beschrieben.

Ein isoliertes Nukleinsäuremolekül, das für eine im Verfahren verwendete  $\omega$ -3-  
5 Desaturase,  $\Delta$ -6-Desaturase,  $\Delta$ -5-Desaturase,  $\Delta$ -5-Elongase,  $\Delta$ -4-Desaturase und/oder  
 $\Delta$ -6-Elongase kodiert, die zu einer Proteinsequenz der SEQ ID NO. 65, SEQ ID NO. 2,  
SEQ ID NO. 172, SEQ ID NO. 52, SEQ ID NO. 194 oder SEQ ID NO. 78 homolog ist,  
kann durch Einbringen einer oder mehrerer Nukleotidsubstitutionen, -additionen oder  
-deletionen in eine Nukleotidsequenz der SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO.  
10 171, SEQ ID NO. 51, SEQ ID NO. 193 oder SEQ ID NO. 77 erzeugt werden, so  
dass eine oder mehrere Aminosäuresubstitutionen, -additionen oder -deletionen in das  
kodierte Protein eingebracht werden. Mutationen können in eine der Sequenzen der  
SEQ ID NO. 64, SEQ ID NO. 1, SEQ ID NO. 171, SEQ ID NO. 51, SEQ ID NO. 193 oder  
SEQ ID NO. 77 durch Standardtechniken, wie stellenspezifische Mutagenese und PCR-  
15 vermittelte Mutagenese, eingebracht werden. Vorzugsweise werden konservative Ami-  
nosäuresubstitutionen an einer oder mehreren der vorhergesagten nicht-essentiellen  
Aminosäureresten hergestellt. Bei einer "konservativen Aminosäuresubstitution" wird der  
Aminosäurerest gegen einen Aminosäurerest mit einer ähnlichen Seitenkette ausgetauscht.  
Im Fachgebiet sind Familien von Aminosäureresten mit ähnlichen Seitenketten  
20 definiert worden. Diese Familien umfassen Aminosäuren mit basischen Seitenketten  
(z.B. Lysin, Arginin, Histidin), sauren Seitenketten (z.B. Asparaginsäure, Glutaminsäure),  
ungeladenen polaren Seitenketten (z.B. Glycin, Asparagin, Glutamin, Serin, Threonin,  
Tyrosin, Cystein), unpolaren Seitenketten, (z.B. Alanin, Valin, Leucin, Isoleucin, Prolin,  
Phenylalanin, Methionin, Tryptophan), beta-verzweigten Seitenketten (z.B. Threonin,  
25 Valin, Isoleucin) und aromatischen Seitenketten (z.B. Tyrosin, Phenylalanin, Tryptophan,  
Histidin). Ein vorhergesagter nicht-essentieller Aminosäurerest in einer  $\omega$ -3-Desaturase,  
 $\Delta$ -6-Desaturase,  $\Delta$ -5-Desaturase,  $\Delta$ -5-Elongase,  $\Delta$ -4-Desaturase oder  $\Delta$ -6-Elongase  
wird somit vorzugsweise durch einen anderen Aminosäurerest aus der gleichen Seiten-  
kettenfamilie ausgetauscht. Alternativ können bei einer anderen Ausführungsform die  
30 Mutationen zufallsgemäß über die gesamte oder einen Teil der  $\omega$ -3-Desaturase,  $\Delta$ -6-  
Desaturase,  $\Delta$ -5-Desaturase,  $\Delta$ -5-Elongase,  $\Delta$ -4-Desaturase oder  $\Delta$ -6-Elongase kodie-  
renden Sequenz eingebracht werden, z.B. durch Sättigungsmutagenese, und die resul-

5 tierenden Mutanten können nach der hier beschriebenen  $\omega$ -3-Desaturase-,  $\Delta$ -6-Desaturase-,  $\Delta$ -5-Desaturase-,  $\Delta$ -5-Elongase-,  $\Delta$ -4-Desaturase- oder  $\Delta$ -6-Elongase-Aktivität durchmustert werden, um Mutanten zu identifizieren, die die  $\omega$ -3-Desaturase-,  $\Delta$ -6-Desaturase-,  $\Delta$ -5-Desaturase-,  $\Delta$ -5-Elongase-,  $\Delta$ -4-Desaturase- oder  $\Delta$ -6-Elongase-Aktivität beibehalten haben. Nach der Mutagenese kann das kodierte Protein rekombinant exprimiert werden, und die Aktivität des Proteins kann z.B. unter Verwendung der hier beschriebenen Tests bestimmt werden.

10 Diese Erfindung wird durch die nachstehenden Beispiele weiter veranschaulicht, die nicht als beschränkend aufgefasst werden sollten. Der Inhalt sämtlicher in dieser Patentanmeldung zitierten Literaturstellen, Patentanmeldungen, Patente und veröffentlichten Patentanmeldungen ist hier durch Bezugnahme aufgenommen.

15 Die folgende Tabelle zeigt die Sequenzkennzahlen, wie sie in der Prioritätsanmeldung vom 21.02.2006 mit dem deutschen Anmeldeaktenzeichen 102006008030.0 verwendet wurden sowie den entsprechenden Sequenzkennzahlen in der vorliegenden Nachanmeldung. Die durch die SEQ ID No: 1 der Prioritätsanmeldung gekennzeichnete Nukleinsäuresequenz entspricht beispielsweise der durch die SEQ ID NO: 64 der Nachanmeldung gekennzeichneten Nukleinsäuresequenz.

20

Konkordanztabelle Sequenzkennzahlen der Prioritätsanmeldung und der Sequenzkennzahlen in der Nachanmeldung:

SEQ ID NO: Prioritätsanmeldung Deutsches Anmeldeaktenzeichen 102006008030.0	SEQ ID NO: vorliegende Nachanmeldung	Organism
1	64	Ostreococcus tauri
2	65	Ostreococcus tauri
3	1	Phytium irregulare
4	2	Phytium irregulare
5	171	Traustochytrium sp.
6	172	Traustochytrium sp.
7	51	Thraustochytrium ssp.

65

8	52	Thraustochytrium ssp.
9	193	Phytophthora infestans
10	194	Phytophthora infestans
11	77	Traustochytrium sp.
12	78	Traustochytrium sp.
13	109	Ostreococcus tauri
n.a.	110	Ostreococcus tauri
14	122	Ostreococcus tauri
n.a.	123	Ostreococcus tauri
15	143	Ostreococcus tauri
16	144	Ostreococcus tauri
17	161	Cauliflower mosaic virus
18	162	Cauliflower mosaic virus
19	163	Thalassiosira pseudonana
20	164	Thalassiosira pseudonana

## Beispiele

### Beispiel 1: Allgemeine Klonierungsverfahren

5

Die Klonierungsverfahren wie z.B. Restriktionsspaltungen, Agarose-Gelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrozellulose- und Nylon-Membranen, Verknüpfen von DNA-Fragmenten, Transformation von Escherichia coli-Zellen, Anzucht von Bakterien und die Sequenzanalyse rekombinanter DNA wurden wie bei Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) beschrieben durchgeführt.

10

### Beispiel 2: Sequenzanalyse rekombinanter DNA

15

Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma ABI nach der Methode von Sanger (Sanger et al. (1977)

Proc. Natl. Acad. Sci. USA 74: 5463-5467). Fragmente resultierend aus einer Polymerase-Kettenreaktion wurden zur Vermeidung von Polymerasefehlern in zu exprimierenden Konstrukten sequenziert und überprüft.

5 Beispiel 3: Klonierung von Genen aus *Ostreococcus tauri*

In einer *Ostreococcus tauri* Sequenzdatenbank (genomische Sequenzen) konnte durch Suche nach konservierten Bereichen jeweils eine Sequenz kodierend für ein Protein mit  $\Delta$ -5-Elongaseaktivität oder  $\Delta$ -6-Elongase-Aktivität identifiziert werden. Es handelt sich  
10 dabei um die folgenden Sequenzen:

Gen-Name	SEQ ID	Aminosäuren
OtELO1.1, ( $\Delta$ -6-Elongase)	SEQ ID NO. 143	292
OtELO2.1, ( $\Delta$ -5-Elongase)	SEQ ID NO. 109	300

OtElo2.1 weist die höchste Ähnlichkeit zu einer Elongase aus *Danio rerio* auf (GenBank AAN77156; ca. 26 % Identität), während OtElo1.1 die größte Ähnlichkeit zur Elongase  
15 aus *Physcomitrella* (PSE) (ca. 36 % Identität) aufweist (Alignments wurden mit dem tBLASTn-Aalgorithmus (Altschul et al. (1990) J. Mol. Biol. 215: 403 – 410) durchgeführt).

Die Klonierung der Elongasen wurde wie folgt durchgeführt:

40 ml einer *Ostreococcus tauri* Kultur in der stationären Phase wurden abzentrifugiert und in 100  $\mu$ l Aqua bidest resuspendiert und bei  $-20^{\circ}\text{C}$  gelagert. Mit Hilfe des PCR-  
20 Verfahrens wurden die entsprechenden genomischen DNAs amplifiziert. Die entsprechenden Primerpaare wurden so ausgewählt, dass sie die Hefe-Konsensus-Sequenz für hocheffiziente Translation (Kozak (1986) Cell 44: 283-292) neben dem Startcodon trugen. Die Amplifizierung der OtElo-DNAs wurde jeweils mit 1  $\mu$ l aufgetauten Zellen, 200  $\mu$ M dNTPs, 2,5 U Taq-Polymerase und 100 pmol eines jeden Primers in einem Gesamt-

volumen von 50 µl durchgeführt. Die Bedingungen für die PCR waren wie folgt: Erste Denaturierung bei 95°C für 5 Minuten, gefolgt von 30 Zyklen bei 94°C für 30 Sekunden, 55°C für 1 Minute und 72°C für 2 Minuten sowie ein letzter Verlängerungsschritt bei 72°C für 10 Minuten.

5

#### Beispiel 4: Optimierung von Elongase-Genen aus *Ostreococcus tauri*

Elongasen aus dem Organismus *Ostreococcus tauri* wurden isoliert wie in Beispiel 3 beschrieben. Um eine Steigerung des Gehalts an C22-Fettsäuren zu erreichen, wurden die Sequenzen Seq ID No. 143 ( $\Delta 6$ -Elongase) und Seq ID No. 109 (kodierend für ein durch die Seq ID No. 110 gekennzeichnetes Protein)( $\Delta 5$ -Elongase) an die Kodonverwendung in Raps, Lein und Soja angepasst. Dazu wurde die Aminosäuresequenz der  $\Delta 6$ -Elongase und der  $\Delta 5$ -Elongase (SEQ ID NO. 144 für die  $\Delta 6$ -Elongase; SEQ ID NO. 65 für die  $\Delta 5$ -Elongase) revers translatiert, wodurch degenerierte DNA-Sequenzen erhalten wurden. Diese DNA-Sequenzen wurden mit Hilfe des Programms GeneOptimizer (Fa. Geneart, Regensburg) an die Kodonverwendung in Raps, Soja und Lein angepasst, wobei die natürliche Häufigkeit einzelner Kodons berücksichtigt wurde. Die so erhaltenen optimierten Sequenzen, die in SEQ ID NO. 64 ( $\Delta 5$ -Elongase) und SEQ ID NO. 122 (kodierend für ein durch die SEQ ID NO. 123 gekennzeichnetes Protein)( $\Delta 6$ -Elongase) angegeben sind, wurden in vitro synthetisiert.

10

15

20

#### Beispiel 5: Klonierung von Expressionsplasmiden zur heterologen Expression in Hefen

Zur Charakterisierung der Funktion der optimierten Nukleinsäuresequenzen wurden die offenen Leserahmen der jeweiligen DNAs stromabwärts des Galactose-induzierbaren GAL1-Promotors von pYES2.1/V5-His-TOPO (Invitrogen) kloniert, wodurch die Plasmide pOTE1.2 (enthaltend die  $\Delta 6$ -Elongase-Sequenz) und pOTE2.2 (enthaltend die  $\Delta 5$ -Elongase-Sequenz) erhalten wurden.

25

Übersicht zu den in den Hefe-Vektor pYES2.1/V5-His-TOPO klonierten Elongase-Sequenzen:

30

Gen-Name	SEQ ID	Aminosäuren
----------	--------	-------------

pOTE1.1, ( $\Delta$ -6-Elongase)	SEQ ID NO. 143	292
pOTE1.2, ( $\Delta$ -6-Elongase)	SEQ ID NO. 122	292, Kodon-optimiert
pOTE2.1, ( $\Delta$ -5-Elongase)	SEQ ID NO. 109	300
pOTE2.2, ( $\Delta$ -5-Elongase)	SEQ ID NO. 64	300, Kodon-optimiert

Der *Saccharomyces cerevisiae*-Stamm 334 wurde durch Elektroporation (1500 V) mit den Vektoren pOTE1.2 bzw. pOTE2.2 sowie den Vergleichskonstrukten pOTE1.1 und pOTE2.1, die die natürliche, für die  $\Delta$ 6-Elongase bzw.  $\Delta$ 5-Elongase kodierende Nukleinsäuresequenz enthalten, transformiert. Als Kontrolle wurde eine Hefe verwendet, die mit dem leeren Vektor pYES2 transformiert wurde. Die Selektion der transformierten Hefen erfolgte auf Komplet-Minimalmedium (CMdum)-Agarplatten mit 2% Glucose, aber ohne Uracil. Nach der Selektion wurden je drei Transformanten zur weiteren funktionellen Expression ausgewählt.

Für die Expression der Ot-Elongasen wurden zunächst Vorkulturen aus jeweils 5 ml CMdum-Flüssigmedium mit 2% (w/v) Raffinose aber ohne Uracil mit den ausgewählten Transformanten angeimpft und 2 Tage bei 30°C, 200 rpm inkubiert. 5 ml CMdum-Flüssigmedium (ohne Uracil) mit 2% Raffinose wurden dann mit den Vorkulturen auf eine OD<sub>600</sub> von 0,05 angeimpft. Dabei wurde der Hefekultur, die mit pOTE1.1 und pOTE1.2 transformiert worden war, jeweils 0,2mM  $\gamma$ -Linolensäure (GLA) zugegeben. Ausgehend von der Aktivität von OtELO1.1 ist eine Elongation der  $\gamma$ -Linolensäure zur Fettsäure 20:3 zu erwarten. Der Hefekultur, die mit pOTE2.1 und pOTE2.2 transformiert worden war, wurden jeweils 0,2mM Arachidonsäure bzw. Eicosapentaensäure zugegeben. Entsprechend der Aktivität von OtELO2.1 ist eine Elongation der Fettsäuren ARA bzw. EPA zu den Fettsäuren 22:4 bzw. 22:5 zu erwarten. Die Expression wurde durch die Zugabe von 2% (w/v) Galaktose induziert. Die Kulturen wurden für weitere 96 h bei 20°C inkubiert.

Beispiel 6: Expression von OtELO2.2 (wie in SEQ ID NO: 64 dargestellt) und OtELO1.2 (wie in SEQ ID NO: 122) in Hefen

5 Hefen, die wie in Beispiel 5 mit den Plasmiden pYES2, pOTE1.2, und pOTE2.1 transformiert wurden, wurden folgendermaßen analysiert:

Die Hefezellen aus den Hauptkulturen wurden durch Zentrifugation (100 x g, 5 min, 20°C) geerntet und mit 100 mM NaHCO<sub>3</sub>, pH 8,0 gewaschen, um restliches Medium und Fettsäuren zu entfernen. Aus den Hefe-Zellsedimenten wurden Fettsäuremethylester (FAMES) durch saure Methanolyse hergestellt. Hierzu wurden die Zellsedimente mit 2 ml 1 N methanolischer Schwefelsäure und 2% (v/v) Dimethoxypropan für 1 h bei 80°C inkubiert. Die Extraktion der FAMES erfolgte durch zweimalige Extraktion mit Petrolether (PE). Zur Entfernung nicht derivatisierter Fettsäuren wurden die organischen Phasen je einmal mit 2 ml 100 mM NaHCO<sub>3</sub>, pH 8,0 und mit 2 ml Aqua dest. gewaschen. Anschließend wurden die PE-Phasen mit Na<sub>2</sub>SO<sub>4</sub> getrocknet, unter Argon eingedampft und in 100 µl PE aufgenommen. Die Proben wurden auf einer DB-23-Kapillarsäule (30 m, 0,25 mm, 0,25 µm, Agilent) in einem Hewlett-Packard 6850-Gaschromatographen mit Flammenionisationsdetektor getrennt. Die Bedingungen für die GLC-Analyse waren wie folgt: Die Ofentemperatur wurde von 50°C bis 250°C mit einer Rate von 5°C/min und schließlich 10 min bei 250°C (Halten) programmiert.

Die Identifikation der Signale erfolgte durch Vergleiche der Retentionszeiten mit entsprechenden Fettsäurestandards (Sigma). Die Methodik ist beschrieben zum Beispiel in Napier and Michaelson (2001) *Lipids* 36(8):761-766; Sayanova et al. (2001) *Journal of Experimental Botany* 52(360):1581-1585, Sperling et al. (2001) *Arch. Biochem. Biophys.* 388(2):293-298 und Michaelson et al. (1998) *FEBS Letters* 439(3):215-218. Die Ergebnisse der Analysen sind in Tabelle 1 dargestellt.

30 Sowohl für pOTE1.1/pOTE1.2 als auch für pOTE2.1/2.2 konnten die entsprechenden Aktivitäten bestätigt werden. In beiden Fällen zeigte die optimierte Sequenz (pOTE1.2 bzw. pOTE2.2) Aktivität. Dabei konnte die Synthese der  $\gamma$ -Linolensäure durch pOTE1.2 gegenüber der Wildtyp-Sequenz nur geringfügig gesteigert werden. Dagegen konnte für



pOTE2.2 überraschenderweise sowohl eine Erhöhung der Aktivität als auch eine Veränderung der Spezifität beobachtet werden (Tabelle 1). Dabei hat sich die Aktivität zur Verlängerung von EPA nahezu verdoppelt, während die Verlängerung von ARA sich mehr als vervierfacht hat. Mit der Optimierung der Sequenz der  $\Delta 5$ -Elongase von *Ostreococcus tauri* konnte somit in Hefe bei gleicher Substratmenge die Ausbeute an den Vorstufen von DHA 6fach erhöht werden.

Beispiel 7: Klonierung von Expressionsplasmiden zur Samen-spezifischen Expression in Pflanzen

10

Die folgenden beschriebenen allgemeinen Bedingungen gelten für alle nachfolgenden Versuche, wenn nicht anders beschrieben.

15

Erfindungsgemäß bevorzugt verwendet werden für die folgenden Beispiele Bin19, pBI101, pBinAR, pGPTV, pCAMBIA oder pSUN. Eine Übersicht über binäre Vektoren und ihre Verwendung gibt Hellens et al, Trends in Plant Science (2000) 5: 446–451. Verwendet wurde ein pGPTV-Derivat wie in DE10205607 beschrieben. Dieser Vektor unterscheidet sich von pGPTV durch eine zusätzlich eingefügte *AscI*-Restriktionsschnittstelle.

20

Ausgangspunkt der Klonierung war der Klonierungsvektor pUC19 (Maniatis et al.). Im ersten Schritt wurde das Conlinin-Promotor-Fragment mit folgenden Primern amplifiziert:

25

Cnl1 C 5': gaattcggcgcgccgagctcctcgagcaacgggtccggcggtatagagttgggtaattcga

Cnl1 C 3': cccgggatcgatgccggcagatctccaccatttttggtggtgat

Zusammensetzung des PCR-Ansatzes (50  $\mu$ l):

5,00  $\mu$ l Template cDNA

30

5,00  $\mu$ l 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00  $\mu$ l 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

5

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

10

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym EcoRI und dann für 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Der Klonierungsvektor pUC19 wurde in gleicher Weise inkubiert. Anschließend wurden das PCR-Produkt und der 2668 bp große, geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1-C wurde durch Sequenzierung verifiziert.

15

20

Im nächsten Schritt wurde der OCS-Terminator (Genbank Accession V00088; De Greve, H. et al. (1982) J. Mol. Appl. Genet. 1 (6): 499-511) aus dem Vektor pGPVT-USP/OCS (DE 102 05 607) mit den folgenden Primern amplifiziert:

OCS\_C 5': aggcctccatggcctgctttaatgagatatgcgagacgcc

25

OCS\_C 3': cccgggccggacaatcagtaaattgaacggag

Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

30

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

5

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

10

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym StuI und dann für 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Der Vektor pUC19-Cnl1-C wurde 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1C\_OCS wurde durch Sequenzierung verifiziert.

20

Im nächsten Schritt wurde der Cnl1-B Promotor durch PCR mittels folgender Primer amplifiziert:

Cnl1-B 5': aggcctcaacgggtccggcggtatag

25

Cnl1-B 3': cccgggggtaacgctagcgggcccgatatcggatcccatttttggtggtgattggttct

Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

30

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

5

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

10

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym StuI und dann für 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Der Vektor pUC19-Cnl1-C wurde 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1C\_Cnl1B\_OCS wurde durch Sequenzierung verifiziert.

20

In einem weiteren Schritt wurde der OCS-Terminator für Cnl1B eingefügt. Dazu wurde die PCR mit folgenden Primer durchgeführt:

OCS2 5': aggcctcctgctttaatgagatatgcgagac

25 OCS2 3': cccggggcggacaatcagtaaattgaacggag

Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

30 5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

5

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

10

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym StuI und dann für 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Der Vektor

pUC19-Cnl1C\_Cnl1B\_OCS wurde für 12 h bei 25°C mit dem Restriktionsenzym SmaI inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch

15

Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1C\_Cnl1B\_OCS2 wurde durch Sequenzierung verifiziert.

20

Im nächsten Schritt wurde der Cnl1-A Promotor durch PCR mittels folgender Primer amplifiziert:

Cnl1-B 5': aggcctcaacgggtccggcggtatagag

25

Cnl1-B 3': aggccttctagactgcaggcggccgcccgcatttttggtggtgattggt

Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

30

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

75

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

5

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

10

Das PCR-Produkt wurde für 2 h bei 37°C mit dem Restriktionsenzym *StuI* inkubiert. Der Vektor pUC19-Cnl1-C wurde für 12 h bei 25°C mit dem Restriktionsenzym *SmaI* inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1C\_Cnl1B\_Cnl1A\_OCS2 wurde durch Sequenzierung verifiziert.

15

20

In einem weiteren Schritt wurde der OCS-Terminator für Cnl1A eingefügt. Dazu wurde die PCR mit folgenden Primer durchgeführt:

OCS2 5': ggcctcctgctttaatgagatatgcca

OCS2 3': aagcttggcgcgcccagctcgtcgcacggacaatcagtaaattgaacggaga

25

Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

30

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

## Reaktionsbedingungen der PCR:

Anlagerungstemperatur: 1 min 55°C

5 Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym *StuI* und  
10 dann für 2 h bei 37°C mit dem Restriktionsenzym *HindIII* inkubiert. Der Vektor  
*pUC19-Cnl1C\_Cnl1B\_Cnl1A\_OCS2* wurde für 2 h bei 37°C mit dem Restriktionsenzym  
*StuI* und für 2 h bei 37°C mit dem Restriktionsenzym *HindIII* inkubiert. Anschließend  
wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese  
aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung  
15 der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschlie-  
ßend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von  
Roche verwendet. Das entstandene Plasmid *pUC19-Cnl1C\_Cnl1B\_Cnl1A\_OCS3* wurde  
durch Sequenzierung verifiziert.

20 Das Plasmid *pUC19-Cnl1C\_Cnl1B\_Cnl1A\_OCS3* wurde im nächsten Schritt zur Klonie-  
rung der  $\Delta 6$ -,  $\Delta 5$ -Desaturase und  $\Delta 6$ -Elongase verwendet. Dazu wurde die  $\Delta 6$ -  
Desaturase aus *Phytium irregulare* (WO02/26946) mit folgenden PCR-Primern amplifi-  
ziert:

25 D6Des(Pir) 5': agatctatggtggacctcaagcctggagtg

D6Des(Pir) 3': ccatggcccgggttacatcgctgggaactcggtgat

## Zusammensetzung des PCR-Ansatzes (50 µl):

30 5,00 µl Template cDNA

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP



1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

5

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

10

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym BglIII und dann für 2 h bei 37°C mit dem Restriktionsenzym NcoI inkubiert. Der Vektor pUC19-Cnl1C\_Cnl1B\_Cnl1A\_OCS3 wurde für 2 h bei 37°C mit dem Restriktionsenzym BglIII und für 2 h bei 37°C mit dem Restriktionsenzym NcoI inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1\_d6Des(Pir) wurde durch Sequenzierung verifiziert.

15

20

Das Plasmid pUC19-Cnl1\_d6Des(Pir) wurde im nächsten Schritt zur Klonierung der Δ5-Desaturase aus *Thraustochytrium* ssp. (WO02/26946) verwendet. Dazu wurde die Δ5-Desaturase aus *Thraustochytrium* ssp. mit folgenden PCR-Primern amplifiziert:

25

D5Des(Tc) 5': gggatccatgggcaagggcagcgagggccg

D5Des(Tc) 3': ggcgccgacaccaagaagcaggactgagatc

Zusammensetzung des PCR-Ansatzes (50 µl):

30

5,00 µl Template cDNA

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

5 Reaktionsbedingungen der PCR:

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

10 Anzahl der Zyklen: 35

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym BamHI und dann für 2 h bei 37°C mit dem Restriktionsenzym EcoRV inkubiert. Der Vektor pUC19-Cnl1\_d6Des(Pir) wurde für 2 h bei 37°C mit dem Restriktionsenzym BamHI und für 2 h bei 37°C mit dem Restriktionsenzym EcoRV inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc) wurde durch Sequenzierung verifiziert.

Das Plasmid pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc) wurde im nächsten Schritt zur Klonierung der Δ6-Elongase aus *Physcomitrella patens* (WO01/59128) verwendet, wozu diese mit folgenden PCR-Primern amplifiziert wurde:

D6Elo(Pp) 5': gcgccgcatggaggtcgtggagagattctacggtg

D6Elo(Pp) 3': gcaaaaggagctaaaactgagtgatctaga

30 Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

5 0,50 µl Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

Anlagerungstemperatur: 1 min 55°C

10 Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

Das PCR-Produkt wurde zuerst für 2 h bei 37°C mit dem Restriktionsenzym NotI und  
15 dann für 2 h bei 37°C mit dem Restriktionsenzym XbaI inkubiert. Der Vektor  
pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc) wurde für 2 h bei 37°C mit dem Restriktionsenzym  
NotI und für 2 h bei 37°C mit dem Restriktionsenzym XbaI inkubiert. Anschließend wur-  
den das PCR-Produkt und der geschnittene Vektor durch Agarose-Gelelektrophorese  
aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung  
20 der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschlie-  
ßend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von  
Roche verwendet. Das entstandene Plasmid pUC19-  
Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp) wurde durch Sequenzierung verifiziert.

25 Ausgehend von pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp) wurde der binäre Vektor  
für die Pflanzentransformation hergestellt. Dazu wurde  
pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp) für 2 h bei 37°C mit dem Restriktionsen-  
zym AscI inkubiert. Der Vektor pGPTV wurde in gleicher Weise behandelt. Anschließend  
wurden das Fragment aus pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp) und der ge-  
30 schnittene pGPTV-Vektor durch Agarose-Gelelektrophorese aufgetrennt und die ent-  
sprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels  
Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und

PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pGPTV-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp) wurde durch Sequenzierung verifiziert.

- 5 Ein weiteres Konstrukt, pGPTV-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co), fand Verwendung. Dazu wurde ausgehend von pUC19-Cnl1C\_OCS mit folgenden Primern amplifiziert:

Cnl1\_OCS 5': gtcgatcaacgggtccggcggtatagagttg

- 10 Cnl1\_OCS 3': gtcgatcggacaatcagtaaattgaacggaga

Zusammensetzung des PCR-Ansatzes (50 µl):

5,00 µl Template cDNA

- 15 5,00 µl 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

5,00 µl 2mM dNTP

1,25 µl je Primer (10 pmol/µl)

0,50 µl Advantage-Polymerase (Clontech)

- 20 Reaktionsbedingungen der PCR:

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

- 25 Anzahl der Zyklen: 35

Das PCR-Produkt wurde für 2 h bei 37°C mit dem Restriktionsenzym Sall inkubiert. Der Vektor pUC19 wurde für 2 h bei 37°C mit dem Restriktionsenzym Sall inkubiert. Anschließend wurden das PCR-Produkt und der geschnittene Vektor durch Agarose-

- 30 Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid

Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1\_OCS wurde durch Sequenzierung verifiziert.

5 In einem weiteren Schritt wurde das  $\Delta 12$ -Desaturase-Gen aus *Calendula officinalis* (WO01/85968) in pUC19-Cnl1\_OCS kloniert. Dazu wurde d12Des(Co) mit folgenden Primern amplifiziert:

D12Des(Co) 5': agatctatgggtgcaggcggtcgaatgc

D12Des(Co) 3': ccatggtaaactctattacgatacc

10

Zusammensetzung des PCR-Ansatzes (50  $\mu$ l):

5,00  $\mu$ l Template cDNA

5,00  $\mu$ l 10x Puffer (Advantage-Polymerase) + 25mM MgCl<sub>2</sub>

15

5,00  $\mu$ l 2mM dNTP

1,25  $\mu$ l je Primer (10 pmol/ $\mu$ l)

0,50  $\mu$ l Advantage-Polymerase (Clontech)

Reaktionsbedingungen der PCR:

20

Anlagerungstemperatur: 1 min 55°C

Denaturierungstemperatur: 1 min 94°C

Elongationstemperatur: 2 min 72°C

Anzahl der Zyklen: 35

25

Das PCR-Produkt wurde für 2 h bei 37°C mit dem Restriktionsenzym BglII und anschließend für 2 h bei gleicher Temperatur mit NcoI inkubiert. Der Vektor pUC19-Cnl1\_OCS wurde in gleicher Weise inkubiert. Anschließend wurden das PCR-Fragment und der geschnittene Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-

30

Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1\_D12Des(Co) wurde durch Sequenzierung verifiziert.

Das Plasmid pUC19-Cnl1\_D12Des(Co), sowie das Plasmid

- 5 pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp) wurden für 2 h bei 37°C mit dem Restriktionsenzym Sall inkubiert. Anschließend wurde das Vektor-Fragment sowie der Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und Vektor-Fragment ligiert.
- 10 Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pUC19-Cnl1\_ d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co) wurde durch Sequenzierung verifiziert.

Ausgehend von pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co) wurde der

- 15 binäre Vektor für die Pflanzentransformation hergestellt. Dazu wurde pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co) für 2 h bei 37°C mit dem Restriktionsenzym Ascl inkubiert. Der Vektor pGPTV wurde in gleicher Weise behandelt. Anschließend wurden das Fragment aus
- 20 pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co) und der geschnittene pGPTV-Vektor durch Agarose-Gelelektrophorese aufgetrennt und die entsprechenden DNA-Fragmente ausgeschnitten. Die Aufreinigung der DNA erfolgte mittels Qiagen Gel Purification Kit gemäß Herstellerangaben. Anschließend wurden Vektor und PCR-Produkt ligiert. Dazu wurde das Rapid Ligation Kit von Roche verwendet. Das entstandene Plasmid pGPTV- Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co) wurde
- 25 durch Sequenzierung verifiziert.

Ein weiteres Beispiel für die Verwendung von samenspezifischen Expressionskonstrukten ist der Napin Promotor. In Wu et al. (2005) Nat. Biotech. 23:1013-1017 ist die Herstellung dieser Expressionskonstrukte in den Vektoren pGPTV oder pSUN beschrieben.

Ein weiterer für die Pflanzentransformation geeigneter Vektor ist pSUN2. Um die Zahl der im Vektor enthaltenen Expressionskassetten auf mehr als vier zu erhöhen wurde dieser Vektor in Kombination mit dem Gateway-System (Invitrogen, Karlsruhe) verwendet. Dazu wurde in den Vektor pSUN2 gemäß Herstellerangaben die Gateway-Kassette A wie folgendermaßen beschrieben, eingefügt:

Der pSUN2 Vektor (1 µg) wurde 1 h mit dem Restriktionsenzym EcoRV bei 37° inkubiert. Anschließend wurde die Gateway-Kassette A (Invitrogen, Karlsruhe) in den geschnittenen Vektor ligiert mittels des Rapid Ligation Kits von Roche, Mannheim. Das entstandene Plasmid wurde in E. coli DB3.1 Zellen (Invitrogen) transformiert. Das isolierte Plasmid pSUN-GW wurde anschließend durch Sequenzierung verifiziert.

Im zweiten Schritt wurde die Expressionskassette aus pUC19-Cnl1\_d6Des(Pir)\_d5Des(Tc)\_D6Elo(Pp)\_D12Des(Co) mittels AscI ausgeschnitten und in den in gleicherweise behandelten Vektor pSUN-GW ligiert. Das so entstandene Plasmid pSUN-4G wurde für weitere Genkonstrukte verwendet.

Dazu wurde zuerst gemäß Herstellerangaben (Invitrogen) ein pENTR-Klon modifiziert. Das Plasmid pENTR1A (Invitrogen) wurde 1 h bei 37° mit dem Restriktionsenzym EcoRI inkubiert, anschließend für 30 min mit Klenow-Enzym, sowie einem 1 µM dNTP-Mix behandelt und dann der AscI-Adapter (5'-ggcgcgcc; am 5'-Ende phosphoryliert, doppelsträngig) in den pENTR1A-Vektor ligiert. In diesen modifizierten wurde wie oben beschrieben schrittweise Gene in die Cnl-Kassette eingefügt und über AscI in den pENTR-Vektor übertragen, wodurch der pENTR-Cnl-Vektor entstand.

In einem weiteren Schritt wurde das Konstrukt pSUN-8G hergestellt. Dazu wurden 5'- und 3'-Primer für die Gene mit den SEQ ID NOs: 1, 3, 5 und 7 mit den oben beschriebenen Restriktionsschnittstellen sowie den ersten und jeweils letzten 20 Nukleotiden des offenen Leserahmens erstellt und mit den Standardbedingungen (siehe oben) amplifiziert und in den pENTR-Cnl-Vektor ligiert, der anschließend einer Rekombinationsreaktion nach Herstellerangaben mit dem Vektor pSUN-4G unterworfen wurde.



Dadurch wurde das Konstrukt pSUN-8G hergestellt, das in *Brassica juncea* bzw. *Brassica napus* transformiert wurde. Die Samen der transgenen Pflanzen wurden durch Gaschromatographie analysiert.

- 5 Ein weiteres Konstrukt, das für die Transformation von *B. juncea* und *B. napus* verwendet wurde, war das Konstrukt pSUN-9G. Dieses Konstrukt wurde entsprechend zu Wu et al. (2005) *Nat. Biotech.* 23:1013-1017 mit dem Napin Promotor dargestellt. In Modifikation zu Wu et al. 2005 wurde anstelle des Genes OmELO die kodierende Sequenz von OtELO2.2 in der beschriebenen Art eingesetzt. Das erhaltene Konstrukt pSUN-9G wurde  
10 dann in *B. juncea* bzw. *B. napus* transformiert.

#### Beispiel 8: Lipidextraktion aus Pflanzenmaterial

- Die Auswirkung der genetischen Modifikation in Pflanzen auf die Produktion einer gewünschten Verbindung (wie einer Fettsäure) kann bestimmt werden, indem die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen) ge-  
15 züchtet wird und das Medium und/oder die zellulären Komponenten auf die erhöhte Produktion des gewünschten Produktes (d.h. der Lipide oder einer Fettsäure) untersucht werden. Diese Analysetechniken sind dem Fachmann bekannt und umfassen Spektro-  
20 skopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie (siehe beispielsweise Ullman, *Encyclopedia of Industrial Chemistry*, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon A. et al. (1987) "Applications of HPLC in Biochemistry" in: *Laboratory Techniques in Biochemistry and Molecular Biology*, Bd. 17; Rehm et al. (1993) *Biotechnology*, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A. et al. (1988) *Bioseparations: downstream processing for Biotechnology*, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) *Recovery processes for biological Materials*, John Wiley and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988) *Biochemical Separations*,  
30 in: *Ullmann's Encyclopedia of Industrial Chemistry*, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) *Separation and purification techniques in biotechnology*, Noyes Publications).

Neben den oben erwähnten Verfahren werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22): 12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145 beschrieben extrahiert. Die qualitative und quantitative Lipid- oder Fettsäureanalyse ist beschrieben bei Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

Zusätzlich zur Messung des Endproduktes der Fermentation ist es auch möglich, andere Komponenten der Stoffwechselwege zu analysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamteffizienz der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (z.B. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere Ionen), Messungen der Biomassezusammensetzung und des Wachstums, Analyse der Produktion üblicher Metabolite von Biosynthesewegen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F. Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192 (ISBN: 0199635773) und darin angegebenen Literaturstellen beschrieben.

Ein Beispiel ist die Analyse von Fettsäuren (Abkürzungen: FAME, Fettsäuremethylester; GC-MS, Gas-Flüssigkeitschromatographie-Massenspektrometrie; TAG, Triacylglycerin; TLC, Dünnschichtchromatographie).

Der unzweideutige Nachweis für das Vorliegen von Fettsäureprodukten kann mittels Analyse rekombinanter Organismen nach Standard-Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: Advances on Lipid Methodology, Vierte Aufl.: Christie, Oily Press,

Dundee, 119-169; 1998, Gaschromatographie-Massenspektrometrie-Verfahren, Lipide 33:343-353).

Das zu analysierende Material kann durch Ultraschallbehandlung, Mahlen in der Glas-  
5 mühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren auf-  
gebrochen werden. Das Material muss nach dem Aufbrechen zentrifugiert werden. Das  
Sediment wird in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt  
und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit  
2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl- und Lipidverbindun-  
10 gen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester werden in  
Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapil-  
larsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 Mikrom, 0,32 mm) bei  
einem Temperaturgradienten zwischen 170°C und 240°C für 20 min und 5 min bei  
240°C unterworfen. Die Identität der erhaltenen Fettsäuremethylester muss unter Ver-  
15 wendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma),  
definiert werden.

Pflanzenmaterial wird zunächst mechanisch durch Mörsern homogenisiert, um es einer  
Extraktion zugänglicher zu machen.

20

Dann wird 10 min auf 100°C erhitzt und nach dem Abkühlen auf Eis erneut sedimentiert.  
Das Zellsediment wird mit 1 M methanolischer Schwefelsäure und 2 % Dimethoxypropan  
für 1h bei 90°C hydrolysiert und die Lipide transmethyliert. Die resultierenden Fettsäure-  
methylester (FAME) werden in Petrolether extrahiert. Die extrahierten FAME werden  
25 durch Gasflüssigkeitschromatographie mit einer Kapillarsäule (Chrompack, WCOT Fu-  
sed Silica, CP-Wax-52 CB, 25 m, 0,32 mm) und einem Temperaturgradienten von 170°C  
auf 240°C in 20 min und 5 min bei 240°C analysiert. Die Identität der Fettsäuremethyles-  
ter wird durch Vergleich mit entsprechenden FAME-Standards (Sigma) bestätigt. Die  
Identität und die Position der Doppelbindung kann durch geeignete chemische Derivati-  
30 sierung der FAME-Gemische z.B. zu 4,4-Dimethoxyoxazolin-Derivaten (Christie, 1998)  
mittels GC-MS weiter analysiert werden.

Beispiel 9: Verwendung der optimierten  $\Delta 5$ -Elongase (wie in SEQ ID NO: 64 dargestellt) aus *Ostreococcus tauri* für Konstrukte zur konstitutiven Expression

Für die Transformation von Pflanzen wurden Transformationsvektoren auf Basis von  
5 pGPTV-35S, eines Plasmids basierend auf pBIN19-35S (Bevan M. (1984) Nucl. Acids  
Res. 18:203) erzeugt. Dazu wurde zunächst in einem pUC-Vektor eine Expressionskas-  
sette bestehend aus dem Promotor-Element CaMV35S (Seq ID No. 161) sowie dem  
35S-Terminator (SEQ ID NO. 162; Franck, A. et al. (1980) Cell 21 (1): 285-294) zusam-  
10 mengesetzt. Dabei wurde der Promotor über die Restriktionsschnittstellen Sall/XbaI und  
der Terminator über die Restriktionsschnittstellen BamHI/SmaI eingesetzt. An den Ter-  
minator wurde außerdem ein Polylinker mit der XhoI Schnittstelle angehängt ('triple liga-  
tion'). Das entstandene Plasmid pUC19-35S wurde dann für die Klonierung von PUFA  
Genen eingesetzt. Es wurden parallel die offenen Leserahmen der  $\Delta 6$ -Desaturase- (SEQ  
ID NO. 1), der  $\Delta 5$ -Desaturase- (SEQ ID NO. 51) und  $\Delta 6$ -Elongase- (SEQ ID NO. 171)  
15 Sequenzen über die EcoRV-Schnittstelle in pUC19-35S Vektoren eingefügt. Die ent-  
standenen Plasmide pUC-D6, pUC-D5, pUC-E6(Tc) wurden für die Erstellung des binä-  
ren Vektors pGPTV-35S\_D6D5E6(Tc) verwendet. Dazu wurde der Vektor pGPTV mit  
dem Enzym Sall, das Plasmid pUC-D6 mit Sall/XhoI verdaut und die korrekten Fragmen-  
te ligiert. Das entstandene Plasmid pGPTV-D6 wurde anschließend mit Sall, das Plas-  
20 mid pUC-D5 mit Sall/XhoI verdaut und die korrekten Fragmente ligiert. Das entstandene  
Plasmid pGPTV-D6-D5 wurde dann ein weiteres Mal mit Sall verdaut, das Plasmid pUC-  
E6(Tc) mit Sall/XhoI und die korrekten Fragmente ligiert. Aus diesen sequentiellen Klo-  
nierungsschritten entstand der binäre Vektor pGPTV-D6D5E6(Tc), der für die Transfor-  
mation eingesetzt wurde.

25 In einer weiteren Ausführung wurde an Stelle der Sequenz d6Elo(Tc) die Sequenz von  
d6Elo(Tp) (SEQ ID NO. 163) in den Vektor pUC19-35S eingesetzt. Das entstandene  
Plasmid pUC-E6(Tp) wurde zur Herstellung des binären Vektors pGPTV-  
35S\_D6D5E6(Tp) verwendet.

30 In einer weiteren Ausführung wurde der offene Leserahmen der  $\omega 3$ Des (SEQ ID NO.  
193) in pUC19-35S kloniert. Das entstandene Plasmid pUC- $\omega 3$ Pi wurde über Sall/XhoI in

die binären Vektoren pGPTV-D6D5E6(Tc) und pGPTV-D6D5E6(Tp) übertragen. Die entstandenen Vektoren pGPTV-D6D5E6(Tc) $\omega$ 3Pi und pGPTV-D6D5E6(Tp) $\omega$ 3Pi wurden für die Pflanzentransformation eingesetzt.

- 5 In einer weiteren Ausführung wurde der offene Leserahmen der optimierten  $\Delta$ 5-Elongase aus *Ostreococcus tauri* (Seq ID No. 64), sowie der offene Leserahmen der  $\Delta$ 4-Desaturase aus *Thraustochytrium* sp. (Seq ID No. 77) in pUC19-35S kloniert. Die entstandenen Plasmide pUC-E5 und pUC-D4 wurden dann über Sall/XhoI entsprechend obiger Angaben in den Vektor pGPTV-D6D5E6(Tp) $\omega$ 3Pi übertragen. Der entstandene
- 10 Vektor pGPTV-D6D5E6(Tp) $\omega$ 3PiE5D4 wurde für die Pflanzentransformation eingesetzt.

Alle binären Vektoren wurden in *E. coli* DH5 $\alpha$ -Zellen (Invitrogen) nach Herstellerangaben transformiert. Positive Klone wurden durch PCR identifiziert und Plasmid-DNA isoliert (Qiagen Dneasy).

- 15 Beispiel 10: Transformation der konstitutiven binären Vektoren in Pflanzen

a) Erzeugung transgene *Brassica napus* bzw. *Brassica juncea* Pflanzen

Es wurde das Protokoll zur Transformation von Rapspflanzen verwendet (verändert nach Moloney et al. (1992) *Plant Cell Reports* 8:238-242).

20

Zur Erzeugung transgener Pflanzen wurde der binäre Vektor pGPTV-D6D5E6(Tp) $\omega$ 3PiE5D4 in *Agrobacterium tumefaciens* C58C1:pGV2260 transformiert (Deblaere et al. (1984) *Nucl. Acids. Res.* 13: 4777-4788). Zur Transformation von *Orychophragmus violaceus* wurde eine 1:50 Verdünnung einer Übernachtskultur einer positiv

25 transformierten Agrobakterienkolonie in Murashige-Skoog Medium (Murashige und Skoog (1962) *Physiol. Plant.* 15: 473) mit 3 % Saccharose (3MS-Medium) verwendet. Petiolen oder Hypokotyledonen frisch gekeimter steriler Pflanzen (zu je ca. 1 cm<sup>2</sup>) wurden in einer Petrischale mit einer 1:50 Agrobakterienverdünnung für 5-10 Minuten inkubiert. Es folgt eine 3-tägige Koinkubation in Dunkelheit bei 25°C auf 3MS-Medium mit

30 0,8 % Bacto-Agar. Die Kultivierung wurde anschließend mit 16 Stunden Licht / 8 Stunden Dunkelheit und in wöchentlichem Rhythmus auf MS-Medium mit 500 mg/l Claforan (Cefotaxime-Natrium), 15 mg/l Kanamycin, 20  $\mu$ M Benzylaminopurin (BAP) und 1,6 g/l Glu-

kose weitergeführt. Wachsende Sprosse wurden auf MS-Medium mit 2 % Saccharose, 250 mg/l Claforan und 0,8 % Bacto-Agar überführt. Bildeten sich nach drei Wochen keine Wurzeln, wurde dem Medium als Wachstumshormon 2-Indolbuttersäure zum Bewurzeln zugegeben.

5

Regenerierte Sprosse wurden auf 2MS-Medium mit Kanamycin und Claforan erhalten, nach Bewurzelung in Erde überführt und nach Kultivierung für zwei Wochen in einer Klimakammer oder im Gewächshaus angezogen, zur Blüte gebracht, reife Samen geerntet und mittels Lipidanalysen auf Elongase-Expression wie  $\Delta$ -6-Elongaseaktivität oder  $\Delta$ -5- oder  $\Delta$ -6-Desaturaseaktivität untersucht. Linien mit erhöhten Gehalten an mehrfach ungesättigten C20- und C22-Fettsäuren wurden so identifiziert.

10

b) Erzeugung transgener *Orychopragmus violaceus* Pflanzen

15

Es wurde das Protokoll zur Transformation von Rapspflanzen verwendet (verändert nach Moloney et al. (1992) Plant Cell Reports 8:238-242) wie unter a) beschrieben, angewendet.

Zur Erzeugung transgener Pflanzen wurde der binäre Vektor pGPTV-

20

D6D5E6(Tp) $\omega$ 3PiE5D4 in *Agrobacterium tumefaciens* C58C1:pGV2260 transformiert (Deblaere et al. (1984) Nucl. Acids. Res. 13: 4777-4788). Zur Transformation von *Orychopragmus violaceus* wurde eine 1:50 Verdünnung einer Übernachtskultur einer positiv transformierten Agrobakterienkolonie in Murashige-Skoog Medium (Murashige und Skoog (1962) Physiol. Plant. 15: 473) mit 3 % Saccharose (3MS-Medium) verwendet.

25

Petiolen oder Hypokotyledonen frisch gekeimter steriler Pflanzen (zu je ca. 1 cm<sup>2</sup>) wurden in einer Petrischale mit einer 1:50 Agrobakterienverdünnung für 5-10 Minuten inkubiert. Es folgt eine 3-tägige Koinkubation in Dunkelheit bei 25°C auf 3MS-Medium mit 0,8 % Bacto-Agar. Die Kultivierung wurde anschließend mit 16 Stunden Licht / 8 Stunden Dunkelheit und in wöchentlichem Rhythmus auf MS-Medium mit 500 mg/l Claforan (Cefotaxime-Natrium), 15 mg/l Kanamycin, 20  $\mu$ M Benzylaminopurin (BAP) und 1,6 g/l Glukose weitergeführt. Wachsende Sprosse wurden auf MS-Medium mit 2 % Saccharose, 250 mg/l Claforan und 0,8 % Bacto-Agar überführt. Bildeten sich nach drei Wochen kei-

30

ne Wurzeln, wurde dem Medium als Wachstumshormon 2-Indolbuttersäure zum Bewurzeln zugegeben.

Regenerierte Sprosse wurden auf 2MS-Medium mit Kanamycin und Claforan erhalten,  
5 nach Bewurzelung in Erde überführt und nach Kultivierung für zwei Wochen in einer Klimakammer oder im Gewächshaus angezogen, zur Blüte gebracht, reife Samen geerntet und mittels Lipidanalysen auf Elongase-Expression wie  $\Delta$ -6-Elongaseaktivität oder  $\Delta$ -5- oder  $\Delta$ -6-Desaturaseaktivität untersucht. Linien mit erhöhten Gehalten an mehrfach ungesättigten C20- und C22-Fettsäuren wurden so identifiziert.

10

#### c) Transformation von *Arabidopsis thaliana* Pflanzen

Es wurde das Protokoll von Bechthold et al. (1993) C.R. Acad. Sci. Ser. III Sci. Vie. 316: 1194-1199 angewendet.

Zur Erzeugung transgener Pflanzen wurde der erzeugte binäre Vektor pGPTV-  
15 D6D5E6(Tp) $\omega$ 3PiE5D4 in *Agrobacterium tumefaciens* C58C1:pMP90 transformiert (Deblaere et al. (1984) Nucl. Acids. Res. 13: 4777-4788) und gemäß des Protokolls von Bechthold et al. (1993) Blüten von *Arabidopsis thaliana* cv. Columbia 0 in eine Agrobakterien-Lösung mit OD<sub>600</sub>=1,0 getaucht. Der Vorgang wurde zwei Tage später nochmals wiederholt. Samen aus diesen Blüten wurden dann auf Agar-Platten mit ½ MS, 2% Saccharose und 50mg/L Kanamycin ausgelegt. Grüne Keimlinge wurden dann auf Erde  
20 transferiert.

#### Beispiel 11: Analyse von Pflanzenmaterial transgener *Orychophragmus*- bzw. *Arabidopsis*-Pflanzen

25

Die Extraktion von Blattmaterial transgener *Orychophragmus violaceus* und *Arabidopsis thaliana* Pflanzen transformiert mit pGPTV-D6D5E6(Tp) $\omega$ 3PiE5D4 sowie die gaschromatographische Analyse wurde wie in Beispiel 8 beschrieben durchgeführt. Tabelle 2 zeigt die Ergebnisse der Analysen. Die verschiedenen Fettsäuren sind in Gewichtsprozent  
30 angegeben. Für beide unterschiedliche Pflanzenarten konnte die Synthese von langkettigen, mehrfach ungesättigte Fettsäuren gezeigt werden. Überraschenderweise konnte



mit der optimierten Sequenz der  $\Delta 5$ -Elongase (wie in SEQ ID NO: 64 dargestellt) aus *Ostreococcus tauri* eine deutlich höhere Ausbeute an DHA erhalten werden wie dies zum Beispiel von Robert et al. (2005) *Functional Plant Biology* 32:473-479 für *Arabidopsis thaliana* mit 1,5% DHA berichtet wird. Für *Orychophragmus violaceus* konnte zum ersten Mal eine Synthese von langkettigen, mehrfach ungesättigten Fettsäuren erzielt werden.

#### Beispiel 11: Analyse von Samen von transgener *Brassica juncea* Linien

Die Extraktion von Samen von transgenen *Brassica juncea* Pflanzen transformiert mit pSUN-9G sowie die gaschromatographische Analyse wurde wie in Beispiel 8 beschrieben durchgeführt. Tabelle 6 zeigt die Ergebnisse der Analysen. Die verschiedenen Fettsäuren sind in Flächenprozent angegeben. Wie in Wu et al. 2005 konnte die Synthese von langkettigen mehrfach ungesättigten Fettsäuren (PUFA) gezeigt werden. Überraschenderweise resultierte die Verwendung der modifizierten Elongase-Sequenz OtE-LO2.2 wie in der durch die SEQ ID NO: 64 beschriebenen Nukleinsäuresequenz in einer dramatischen Erhöhung des Gehalts an C22 Fettsäuren. Insgesamt enthielt das Samenöl etwa 8% Gew.-% mehrfach ungesättigte C22-Fettsäuren. Im speziellen war die Fettsäure Docosahexaensäure (DHA) zu 1,9% Gew.-% im Samenöl enthalten, was einer Steigerung um Faktor 10 im Vergleich zu Wu et al. 2005 darstellte.

#### Beispiel 12: Detaillierte Analyse der Lipidklassen und Positionsanalyse von Blattmaterial von *O. violaceus*

Etwa 1 g Blattgewebe wurde für 10 Minuten bei 95°C in 4 ml of Isopropanol erhitzt, durch Polytron homogenisiert und nach der Zugabe von 1,5 ml Chloroform geschüttelt. Die Proben wurden zentrifugiert, der Überstand gesammelt und das Pellet mit Isopropanol:Chloroform 1:1 (v/v) nochmals extrahiert. Die beiden Extrakte wurden vereinigt, getrocknet und in Chloroform gelöst. Der Lipidextrakt wurde an einer Silica PrepSep-Säule (Fisher Scientific, Nepean, Canada) in neutrale Lipide, Glykolipide und Phospholipide vorfraktioniert, wobei mit Chloroform:Essigsäure 100:1 (v/v), Aceton:Essigsäure 100:1 (v/v) bzw. Methanol:Chloroform:Wasser 100:50:40 (v/v/v) eluiert wurde. Diese Fraktionen wurden auf Silica G-25 Dünnschichtchromatographieplatten (TLC; Macherey-Nagel, Dü-

ren, Germany) weiter aufgetrennt. Neutrale Lipide wurden mit Hexan:Diethylether:Essigsäure (70:30:1), Glykolipide mit Chloroform:Methanol:Ammoniak (65:25:4 v/v/v) und Phospholipide mit Chloroform:Methanol:Ammoniak:Wasser (70:30:4:1 v/v/v/v) entwickelt. Die einzelnen Lipidklassen wurden nach dem Besprühen mit Primulin unter UV-Licht identifiziert, durch Abkratzen von den Platten entfernt, und entweder für die direkte Transmethylierung verwendet oder durch ein geeignetes Lösungsmittel für die weitere Analyse extrahiert.

Entsprechend der beschriebenen Methoden konnten die verschiedenen Lipidklassen (Neutral-, Phospho- und Galaktolipide) aufgetrennt und separat analysiert werden. Desweiteren wurden die Glycolipide auf die Position der einzelnen Fettsäuren untersucht.

a) Regiospezifische Analyse der Triacylglyceride (TAG)

Drei bis fünf mg der TLC-gereinigten TAG wurden unter Stickstoff in Glasröhrchen getrocknet, in wässrigem Puffer durch kurze Ultraschallbehandlung resuspendiert (1 M Tris pH 8; 2.2% CaCl<sub>2</sub> (w/v); 0.05% Gallensalze (w/v)) und für 4 Minuten bei 40° C inkubiert. Nach der Zugabe von 0.1 ml einer Lösung von Pankreaslipase (10 mg/ml in Wasser) wurden die Proben für 3 Minuten kräftig gevortext und der Verdau durch Zugabe von 1 ml Ethanol und 1,5 ml 4 M HCl abgebrochen. Die teilweise verdauten TAG wurden zweimal mit Diethylether extrahiert, mit Wasser gewaschen, getrocknet und in einem kleinen Volumen Chloroform gelöst. Monoacylglycerole (MAG) wurden von den freien Fettsäuren und unverdauten TAGs auf einer TLC-Platte getrennt wie oben für neutrale Lipide beschrieben. Der Punkt, der den MAGs entsprach, wurde durch GC analysiert und stellte die sn-2-Position der TAG dar. Die Verteilung der Fettsäuren an den verbleibenden sn-1- and sn-3-Positionen wurde nach der folgenden Formel berechnet:  $sn-1 + sn-3 = (TAG \times 3 - MAG) / 2$ .

Diese Positionsanalyse der Triacylglyceride ergab dabei, dass EPA und DHA in den Positionen sn-2 und sn-1/3 in ähnlichen Konzentrationen vorhanden sind, während ARA insgesamt nur in geringen Mengen in den Triacylglyceriden zu finden ist und hier hauptsächlich in der Position sn-2 (Tab. 3).

## b) Stereospezifische Analyse von Phospholipiden

Aufgetrenntes und extrahiertes Phosphatidylglykol (PG), Phosphatidylethanolamin (PE) und Phosphatidylcholin (PC) wurden unter N<sub>2</sub> getrocknet und in 0.5 ml Boratpuffer (0.5M, pH 7.5, enthaltend 0.4 mM CaCl<sub>2</sub>) resuspendiert. Nach einer kurzen Ultraschallbehandlung wurden 5U der Phospholipase A2 aus dem Gift von Naja mossambica (Sigma P-7778) und 2 ml Diethylether zugegeben und die Proben wurden für 2 Stunden bei Raumtemperatur gevortext. Die Etherphase wurde getrocknet, der Verdau mit 0,3 ml 1M HCl abgebrochen, und das Reaktionsgemisch wurde mit Chloroform:Methanol (2:1 v/v) extrahiert. Die verdauten Phospholipide wurden durch TLC in Chloroform:Methanol:Ammoniak:Wasser (70:30:4:2 v/v/v/v) getrennt und Punkte, die den freigesetzten freien Fettsäuren und Lysophospholipiden entsprachen, wurden durch Kratzen entfernt und direkt transmethyliert.

15

Die Positionsanalyse der Phospholipide zeigte eine Akkumulation von EPA und DHA in der sn-2 Position von Phosphatidylcholin (PC), während DHA in Phosphatidylethanolamin (PE) in sn-1 und sn-2 Position ähnlich verteilt war. In beiden Phospholipiden war kein oder nur Spuren von ARA zu finden (Tab. 4). In Phosphatidylglycerol fanden sich geringere Konzentrationen an EPA und DHA als in den anderen untersuchten Phospholipiden, wobei auch in dieser Lipidklasse eine Akkumulation in der sn-2 Position zu beobachten war (Tab. 4, PG).

25

## c) Stereospezifische Analyse von Glykolipiden

Als weitere polare Lipidklasse wurden die Galaktolipide untersucht. Galaktolipide finden sich in den Membranen von Plastiden und bilden dort die Hauptkomponenten.

30

TLC-gereinigtes Monogalaktosyldiacylglycerol (MGDG) und Digalaktosyldiacylglycerol (DGDG) wurden unter Stickstoff getrocknet und in 0.5 ml Diethylether gelöst. Dann wurden 25 Einheiten der Lipase aus Rhizopus arrhizus (Sigma 62305), resuspendiert in 2 ml Boratpuffer (50 mM, pH 7.5 enthaltend 2 mM CaCl<sub>2</sub>), zugegeben und die Proben für 2

Stunden bei Raumtemperatur gevortext. Die Etherphase wurde getrocknet und der Verdau durch die Zugabe von 0,3 ml 1M HCl abgebrochen und die Lipide wurden mit 4 ml Chloroform:Methanol (2:1 v/v) extrahiert. Nach dem Trocknen wurden die verdauten Galaktolipide in einem kleinen Volumen Chloroform:Methanol (2:1 v/v) und zweimal auf einer vorgefertigten Silica TLC-Platte entwickelt, zuerst mit Chloroform:Methanol:Ammoniak:Wasser (70:30:4:1 v/v/v/v) bis etwa zwei Drittel der Plattenhöhe, gefolgt von der vollständigen Entwicklung in Hexan:Diethylether:Essigsäure (70:30:1). Die Punkte, die den freigesetzten freien Fettsäuren und den Lysogalaktolipiden entsprachen, wurden nach dem Besprühen mit Primulin identifiziert, abgekratzt und direkt für die GC-Analyse transmethyliert.

Auch in diesen Lipiden konnten VLCPUFA gefunden werden, wobei eine Akkumulation von EPA in der sn-2 Position zu beobachten war. DHA war nur in den Digalaktodiacylglycerolen (DGDG) zu finden und konnte nicht in den Monogalaktodiacylglycerolen (MGDG) nachgewiesen werden (Tabelle 5). Die Verteilung von VLCPUFA in Galaktolipiden, einem Kompartiment, in dem diese Fettsäuren nicht erwartet wurden, zeigt die Dynamik der Synthese und den späteren Umbau. Ernährungstechnisch sind VLCPUFA in Polarlipiden besonders wertvoll, da diese im Darm von Säugetieren besser aufgenommen werden können als die Neutrallipide.

Tabelle 1: Test der optimierten Sequenzen von pOTE1.1 und pOTE2.1 in Hefe. Entsprechend der Substratumsetzungen wurden die Konversionsraten bestimmt. Mit der optimierten Sequenz in Plasmid pOTE2.2 konnte ein deutlicher Anstieg der Aktivität erreicht werden.

Konversionsraten der <i>Ostreococcus tauri</i> Elongasen				
Genes	Substrate Product	GLA 20:3	ARA 22:4	EPA 22:5
pOTE1.1	d6Elongase(Ot)	21,1		
pOTE1.2	d6Elongase(Ot)_opt	25,6		
pOTE2.1	d5Elongase(Ot)		7,3	35,9
pOTE2.2	d5Elongase(Ot)_opt		32,7	63,1



Tabelle 2: Gaschromatographische Analyse von Blattmaterial von *Orychopragmus violaceus* und *Arabidopsis thaliana*. Die einzelnen Fettsäuren sind in Flächenprozent angegeben.

Fettsäurekomposition Blattmaterial <i>Orychopragmus violaceus</i>											
Fettsäuren	16:0	16:3	18:1	18:2	GLA	18:3	18:4	ARA	EPA	DPA	DHA
Kontrolle	20,9	8,5	3,3	16,0	0,0	47,4	0,0	0,0	0,0	0,0	0,0
Trangen	21,3	8,2	5,2	5,2	4,2	23,1	5,0	0,6	13,5	2,7	4,5

Fettsäurekomposition Blattmaterial <i>Arabidopsis thaliana</i>											
Fettsäuren	16:0	16:3	18:1	18:2	GLA	18:3	18:4	ARA	EPA	DPA	DHA
Kontrolle	12,8	10,0	3,5	14,2	0,0	54,6	0,0	0,0	0,0	0,0	0,0
Transgen	19,3	8,5	5,0	4,6	6,4	31,0	4,4	0,0	6,3	1,5	6,3



Tabelle 4: Stereospezifische Analyse der Phospholipide aus Blattmaterial von transgenen *O. violaceus* Pflanzen.

PG	16:0	16:1	18:0	18:1n-9	18:1n-7	18:2n-9	18:2n-6	18:3n-6	18:3n-3	18:4n-3	20:3n-6	20:4n-6	20:4n-3	20:5n-3	22:5n-3	22:6n-3
<b>WT</b>	<b>27,96</b>	<b>20,04</b>	<b>4,11</b>	<b>2,89</b>	<b>0,90</b>	<b>21,82</b>	<b>0,00</b>	<b>21,56</b>								
sn-1	17,26	0,53	2,61	3,82	1,91	39,01	0,00	34,44								
sn-2	38,66	39,56	5,62	1,96	0,00	4,62	0,00	8,69								
<b>Transgen</b>	<b>27,15</b>	<b>24,70</b>	<b>3,08</b>	<b>4,62</b>	<b>1,20</b>	<b>15,15</b>	<b>1,53</b>	<b>17,94</b>	<b>1,40</b>	<b>0,00</b>	<b>0,00</b>	<b>0,45</b>	<b>2,18</b>	<b>0,10</b>	<b>0,58</b>	
sn-1	21,16	3,61	4,23	7,52	2,14	27,40	0,50	31,57	0,81	0,38	1,24	0,00	0,33			
sn-2	33,15	45,79	1,94	1,71	0,27	2,90	2,57	4,30	2,00	0,51	3,13	0,27	0,83			
<b>PE</b>	<b>16,0</b>	<b>16,1</b>	<b>18,0</b>	<b>18:1n-9</b>	<b>18:1n-7</b>	<b>18:2n-9</b>	<b>18:2n-6</b>	<b>18:3n-6</b>	<b>18:3n-3</b>	<b>18:4n-3</b>	<b>20:3n-6</b>	<b>20:4n-6</b>	<b>20:4n-3</b>	<b>20:5n-3</b>	<b>22:5n-3</b>	<b>22:6n-3</b>
<b>WT</b>	<b>37,49</b>	<b>0,00</b>	<b>6,62</b>	<b>4,35</b>	<b>1,37</b>	<b>19,28</b>	<b>29,95</b>									
sn-1	54,22	0,00	7,74	3,39	3,42	12,64	13,71									
sn-2	20,77	0,00	5,51	5,31	0,00	25,93	46,18									
<b>Transgen</b>	<b>31,78</b>	<b>0,81</b>	<b>5,84</b>	<b>3,08</b>	<b>2,20</b>	<b>5,57</b>	<b>11,25</b>	<b>11,34</b>	<b>7,38</b>	<b>0,00</b>	<b>0,00</b>	<b>2,88</b>	<b>9,41</b>	<b>1,90</b>	<b>4,90</b>	
sn-1	50,17	0,33	10,86	3,22	4,94	0,35	2,63	3,27	3,59	0,56	4,42	6,18	0,38	4,19		
sn-2	13,40	1,29	0,83	2,95	0,00	1,35	8,50	19,23	19,10	12,45	1,34	12,64	3,41	5,61		
<b>PC</b>	<b>16,0</b>	<b>16:1</b>	<b>18:0</b>	<b>18:1n-9</b>	<b>18:1n-7</b>	<b>18:2n-9</b>	<b>18:2n-6</b>	<b>18:3n-6</b>	<b>18:3n-3</b>	<b>18:4n-3</b>	<b>20:3n-6</b>	<b>20:4n-6</b>	<b>20:4n-3</b>	<b>20:5n-3</b>	<b>22:5n-3</b>	<b>22:6n-3</b>
<b>WT</b>	<b>27,57</b>	<b>0,84</b>	<b>6,38</b>	<b>8,56</b>	<b>1,80</b>	<b>21,75</b>	<b>33,01</b>									
sn-1	48,05	0,44	8,65	5,05	3,41	14,52	18,04									
sn-2	7,28	1,24	4,11	12,06	0,18	28,97	47,98									
<b>Transgen</b>	<b>21,00</b>	<b>0,00</b>	<b>8,01</b>	<b>10,02</b>	<b>2,86</b>	<b>1,25</b>	<b>3,77</b>	<b>11,63</b>	<b>5,60</b>	<b>12,11</b>	<b>0,50</b>	<b>4,34</b>	<b>11,16</b>	<b>3,76</b>	<b>3,70</b>	
sn-1	45,35	0,00	14,71	5,08	5,70	0,31	3,23	3,09	4,58	2,65	0,61	4,01	8,32	0,41	1,18	
sn-2	3,36	0,00	1,30	14,96	0,02	2,20	4,31	20,18	6,62	21,56	0,38	4,66	13,99	7,12	6,22	



Tabelle 5: Stereospezifische Analyse der Galaktolipide aus Blattmaterial von transgenen *O. violaceus* Pflanzen.

MEDG	160	161	162	163	180	181n9	181n7	182n9	182n6	183n6	183n3	184n3	203n6	204n6	204n3	205n3
<b>wt</b>	<b>264</b>	<b>0,13</b>	<b>1,23</b>	<b>30,72</b>	<b>0,33</b>	<b>0,35</b>	<b>0,26</b>	<b>3,81</b>	<b>60,52</b>							
sn1	0,00	0,05	0,00	7,11	0,35	0,31	0,41	4,60	87,30							
sn2	5,34	0,21	2,55	54,34	0,31	0,39	0,12	3,01	33,74							
<b>tr</b>	<b>4,16</b>	<b>0,20</b>	<b>1,08</b>	<b>33,81</b>	<b>0,93</b>	<b>0,73</b>	<b>0,52</b>	<b>1,64</b>	<b>44,82</b>	<b>1,88</b>	<b>2,73</b>	<b>0,04</b>	<b>0,30</b>	<b>0,50</b>	<b>5,08</b>	
sn1	1,22	0,29	0,54	4,79	1,51	1,15	0,93	2,80	80,19	0,14	0,08	0,17	0,87	3,86		
sn2	7,11	0,11	1,61	62,82	0,34	0,31	0,11	0,47	9,46	3,62	5,48	0,00	0,43	0,14	6,31	

  

DEDG	160	161	162	163	180	181n9	181n7	182n9	182n6	183n6	183n3	184n3	203n6	204n6	204n3	205n3
<b>wt</b>	<b>17,67</b>	<b>0,19</b>	<b>0,38</b>	<b>2,15</b>	<b>1,61</b>	<b>0,51</b>	<b>0,94</b>	<b>5,56</b>	<b>70,71</b>							
sn1	16,84	0,25	0,50	2,52	2,21	0,55	1,75	6,07	68,74	0,00						
sn2	18,50	0,12	0,27	1,78	1,01	0,46	0,13	5,05	72,68							
<b>tr</b>	<b>18,50</b>	<b>0,00</b>	<b>0,00</b>	<b>2,62</b>	<b>2,84</b>	<b>1,36</b>	<b>1,39</b>	<b>6,28</b>	<b>54,66</b>	<b>3,55</b>	<b>0,00</b>	<b>0,00</b>	<b>0,00</b>	<b>0,00</b>	<b>2,18</b>	<b>5,36</b>
sn1	22,74	0,17	0,23	0,48	4,55	1,71	2,32	9,22	56,06	0,23	0,27	0,00	0,00	0,36	1,23	
sn2	14,27	0,00	0,00	4,77	1,12	1,00	0,46	3,33	53,26	6,88	0,00	0,00	0,00	4,01	9,49	

Tabelle 6: Gaschromatographische Bestimmung der Fettsäuren aus Samen transgener Brassica juncea Pflanzen transformiert mit dem Konstrukt pSUN-9G in Gewichtsprozent. WT beschreibt die nicht-modifizierte Wildtyp-Kontrolle.

	Lipid Profile(%)									
	16:0	18:0	18:1	18:2	$\gamma$ 18:3	$\alpha$ 18:3	18:4	20:0	20:3 (11,14,17)	20:4 (ETeA) (8,11,14,17)
BU223_PUFA184_MKP71_581A	4,4	3,0	22,5	16,9	27,0	4,9	3,2	0,6		
BU223_PUFA184_MKP71_581A	4,7	3,9	17,9	10,6	29,5	4,2	4,0	0,9		
BU223_PUFA184_MKP71_581A	4,4	3,0	18,9	13,8	30,5	4,1	3,2	0,7		
BU223_PUFA184_MKP71_581A	4,6	3,3	20,5	13,2	29,8	4,2	3,3	0,8		

  

	20:3 (11,14,17)	20:4 (ETeA) (8,11,14,17)	22:1	22:4	22:5	22:6
1,1	0,5	0,6	0,0	1,5	2,0	1,6
2,0	0,9	1,0	0,0	3,1	3,5	1,9
1,3	0,7	0,5	0,0	2,7	2,8	1,6
1,4	0,6	0,6	0,0	2,4	2,5	1,6

5

**Patentansprüche**

1. Verfahren zur Herstellung von Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure in einer transgenen Pflanze, umfassend das Bereitstellen in  
5 der Pflanze von mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität kodiert; mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert; mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität kodiert; und mindestens einer Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität und optional für eine  $\Delta 4$ -Desaturase kodiert,  
10 wobei die Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist.  
15
2. Verfahren nach Anspruch 1, wobei die Nukleinsäuresequenz zumindest an die Kodonverwendung in Raps, Soja und/oder Lein angepasst ist.
3. Verfahren nach einem der Ansprüche 1 oder 2, wobei die Nukleinsäuresequenz  
20 unter Berücksichtigung der natürlichen Häufigkeit einzelner Kodons angepasst ist.
4. Verfahren nach einem der Ansprüche 1 bis 3, wobei die veränderte Nukleinsäuresequenz der in Seq ID No. 64 angegebenen Nukleinsäuresequenz entspricht.
- 25 5. Verfahren nach einem der Ansprüche 1 bis 4, wobei die Nukleinsäuresequenzen unter der Kontrolle eines samenspezifischen Promotors exprimiert werden.
6. Verfahren nach Anspruch 5, wobei es sich bei dem Promotor um den USP-, Vicilin-, Napin-, Glp-, SBP-, Peroxiredoxin-, Legumin-, Fad3-, Conlinin- oder Oleosin-Promotor handelt.  
30

7. Verfahren nach Anspruch 6, wobei der Gehalt an mehrfachungesättigten C22-Fettsäuren im Samenöl 5 Gew.-% oder mehr des Samenölgehalts beträgt.
- 5 8. Verfahren nach einem der vorangehenden Ansprüche, wobei zusätzlich eine oder mehrere Nukleinsäuresequenzen kodierend für ein Polypeptid mit der Aktivität einer  $\omega$ 3-Desaturase und/oder einer  $\Delta$ 4-Desaturase in der Pflanze bereitgestellt werden.
9. Verfahren nach Anspruch 8, wobei der Gehalt an Docosahexaensäure im Samenöl 1  
10 Gew.% oder mehr des Samenölgehalts beträgt.
10. Verfahren nach einem der vorangehenden Ansprüche, wobei die Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure in der Pflanze vorwiegend als Ester in Phospholipiden oder Triacylglyceriden gebunden vorliegt.  
15
11. Verfahren nach einem der vorangehenden Ansprüche, wobei die Pflanze eine Ölproduzierende Pflanze ausgewählt aus der Gruppe bestehend aus *Brassica napus*, *Brassica juncea* und *Glycine max* ist.  
20
12. Verfahren nach einem der vorangehenden Ansprüche, weiter umfassend das Gewinnen der Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure in Form ihrer Öle, Lipide oder freien Fettsäuren aus der Pflanze.  
25
13. Isoliertes Nukleinsäuremolekül umfassend eine Nukleinsäuresequenz gemäß Seq ID No. 64.
14. Rekombinantes Nukleinsäuremolekül, umfassend:

## 103

- a) eine oder mehrere Kopien von mindestens einem in Pflanzenzellen aktiven Promotor,
- b) mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität kodiert,
- 5 c) mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität kodiert,
- d) mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert,
- e) mindestens eine Nukleinsäuresequenz, die für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert und die gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist, und
- 10 e) eine oder mehrere Kopien von mindestens einer Terminatorsequenz.
- 15 15. Rekombinantes Nukleinsäuremolekül nach Anspruch 14, wobei die veränderte Nukleinsäuresequenz der in Seq ID No. 64 angegebenen Nukleinsäuresequenz entspricht.
16. Rekombinantes Nukleinsäuremolekül nach einem der Ansprüche 14 oder 15, zusätzlich umfassend eine oder mehrere Nukleinsäuresequenzen kodierend für ein Polypeptid mit der Aktivität einer  $\omega 3$ -Desaturase und/oder einer  $\Delta 4$ -Desaturase.
- 20 17. Transgene Pflanze enthaltend ein rekombinantes Nukleinsäuremolekül nach einem der Ansprüche 22 bis 25 oder enthaltend die in Seq ID No. 64 angegebene Nukleinsäuresequenz.
- 25 18. Verwendung von Ölen, Lipiden oder freien Fettsäuren gemäß Anspruch 19 oder gewonnen durch ein Verfahren nach Anspruch 18 zur Herstellung von Futtermitteln, Nahrungsmitteln, Kosmetika oder Pharmazeutika.
- 30

Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren

## ZUSAMMENFASSUNG

5 Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure in transgenen Pflanzen, indem in der Pflanze bereitgestellt werden mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Desaturase-Aktivität kodiert; mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 6$ -Elongase-Aktivität kodiert; mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Desaturase-Aktivität kodiert; und mindestens eine Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, wobei die Nukleinsäuresequenz, welche für ein Polypeptid mit einer  $\Delta 5$ -Elongase-Aktivität kodiert, gegenüber der Nukleinsäuresequenz in dem Organismus, aus dem die Sequenz stammt, dadurch verändert ist, dass sie an die Kodonverwendung in einer oder mehreren Pflanzenarten angepasst ist. Für die Herstellung von Docosahexaensäure wird zusätzlich eine oder mehrere Nukleinsäuresequenzen, die für ein Polypeptid mit einer  $\Delta 4$ -Desaturase-Aktivität kodiert, in die Pflanze eingebracht.

20

PF58307.txt  
 SEQUENCE LISTING

<110> BASF Plant Science GmbH  
 <120> Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren  
 <130> PF58307  
 <140> Current patent application  
 <141> 2006 09 06  
 <160> 199  
 <170> Bi omax Patent Tool according to Patent IN 3.1 format  
 <210> 1  
 <211> 1380  
 <212> DNA  
 <213> Phytium irregulare  
 <220>  
 <221> CDS  
 <222> (1)..(1380)

<400> 1  
 at g gt g gac ct c aag cct gga gt g aag cgc ct g gt g agc t gg aag gag 48  
 Met Val Asp Leu Lys Pro Gly Val Lys Arg Leu Val Ser Trp Lys Gu  
 1 5 10 15  
 at c cgc gag cac gcg acg ccc gcg acc gcg t gg at c gt g at t cac cac 96  
 Ile Arg Glu His Ala Thr Pro Ala Thr Ala Trp Ile Val Ile His His  
 20 25 30  
 aag gt c t ac gac at c t cc aag t gg gac t cg cac ccg ggt ggc t cc gt g 144  
 Lys Val Tyr Asp Ile Ser Lys Trp Asp Ser His Pro Gly Gly Ser Val  
 35 40 45  
 at g ct c acg cag gcc ggc gag gac gcc acg gac gcc t t c gcg gt c t t c 192  
 Met Leu Thr Gl n Ala Gly Gu Asp Ala Thr Asp Ala Phe Ala Val Phe  
 50 55 60  
 cac ccg t cc t cg gcg ct c aag ct g ct c gag cag t t c t ac gt c ggc gac 240  
 His Pro Ser Ser Ala Leu Lys Leu Leu Gu Gl n Phe Tyr Val Gly Asp  
 65 70 75  
 gt g gac gaa acc t cc aag gcc gag at c gag ggg gag ccg gcg agc gac 288  
 Val Asp Glu Thr Ser Lys Ala Gu Ile Gu Gly Glu Pro Ala Ser Asp  
 85 90 95  
 gag gag cgc gcg cgc cgc gag cgc at c aac gag t t c at c gcg t cc t ac 336  
 Gu Gu Arg Ala Arg Arg Gu Arg Ile Asn Glu Phe Ile Ala Ser Tyr  
 100 105 110  
 cgc cgt ct g cgc gt c aag gt c aag ggc at g ggg ct c t ac gac gcc agc 384  
 Arg Arg Leu Arg Val Lys Val Lys Gly Met Gly Leu Tyr Asp Ala Ser  
 115 120 125  
 gcg ct c t ac t ac gcg t gg aag ct c gt g agc acg t t c ggc at c gcg gt g 432  
 Ala Leu Tyr Tyr Ala Trp Lys Leu Val Ser Thr Phe Gly Ile Ala Val  
 130 135 140  
 ct c t cg at g gcg at c t gc t t c t t c aac agt t t c gcc at g t ac at g 480  
 Leu Ser Met Ala Ile Cys Phe Phe Phe Asn Ser Phe Ala Met Tyr Met  
 145 150 155 160  
 gt c gcc ggc gt g att at g ggg ct c t t c t ac cag cag t cc gga t gg ct g 528  
 Val Ala Gly Val Ile Met Gly Leu Phe Tyr Gl n Gl n Ser Gly Trp Leu  
 165 170 175  
 gcg cac gac t t c t t g cac aac cag gt g t gc gag aac cgc acg ct c ggc 576  
 Ala His Asp Phe Leu His Asn Gl n Val Cys Glu Asn Arg Thr Leu Gly  
 180 185 190  
 aac ct t at c ggc t gc ct c gt g ggc aac gcc t gg cag ggc t t c agc at g 624  
 Asn Leu Ile Gly Cys Leu Val Gly Asn Ala Trp Gl n Gly Phe Ser Met  
 195 200 205  
 cag t gg t gg aag aac aag cac aac ct g cac cac gcg gt g ccg aac ct g 672  
 Gl n Trp Trp Lys Asn Lys His Asn Leu His His Ala Val Pro Asn Leu

Seite 1

PF58307. txt

```

210          215          220
cac agc gcc aag gac gag ggc ttc atc ggc gac ccg gac atc gac acc 720
His Ser Ala Lys Asp Gu Gly Phe Ile Gly Asp Pro Asp Ile Asp Thr
225
atg ccg ctg ctg gcg tgg tct aag gag atg gcg cgc aag gcg ttc gag 768
Met Pro Leu Leu Ala Trp Ser Lys Gu Met Ala Arg Lys Ala Phe Gu
245          250
t cg gcg cac ggc ccg ttc ttc atc cgc aac cag gcg ttc cta tac ttc 816
Ser Ala His Gly Pro Phe Phe Ile Arg Asn Gn Ala Phe Leu Tyr Phe
260          270
ccg ctg ctg ctg ct c gcg cgc ctg agc tgg ct c gcg cag t cg ttc ttc 864
Pro Leu Leu Leu Leu Ala Arg Leu Ser Trp Leu Ala Gn Ser Phe Phe
275          280
tac gtg ttc acc gag ttc t cg ttc ggc atc ttc gac aag gt c gag ttc 912
Tyr Val Phe Thr Gu Phe Ser Phe Gly Ile Phe Asp Lys Val Gu Phe
290          295          300
gac gga ccg gag aag gcg ggt ctg atc gt g cac tac atc tgg cag ct c 960
Asp Gly Pro Gu Lys Ala Gly Leu Ile Val His Tyr Ile Trp Gn Leu
305          310          315
gcg atc ccg tac ttc tgc aac atg agc ctg ttt gag ggc gt g gca tac 1008
Ala Ile Pro Tyr Phe Cys Asn Met Ser Leu Phe Gu Gly Val Ala Tyr
325          330          335
ttc ct c atg ggc cag gcg tcc tgc ggc ttg ct c ct g gcg ct g gt g ttc 1056
Phe Leu Met Gly Gn Ala Ser Cys Gly Leu Leu Leu Ala Leu Val Phe
340          345          350
agt att ggc cac aac ggc atg t cg gt g tac gag cgc gaa acc aag ccg 1104
Ser Ile Gly His Asn Gly Met Ser Val Tyr Gu Arg Gu Thr Lys Pro
355          360          365
gac ttc tgg cag ctg cag gt g acc acg acg cgc aac atc cgc gcg t cg 1152
Asp Phe Trp Gn Leu Gn Val Thr Thr Thr Arg Asn Ile Arg Ala Ser
370          375          380
gt a ttc atg gac tgg ttc acc ggt ggc ttg aac tac cag atc gac cat 1200
Val Phe Met Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gn Ile Asp His
385          390          395          400
cac ctg ttc ccg ct c gt g ccg cgc cac aac ttg cca aag gt c aac gt g 1248
His Leu Phe Pro Leu Val Pro Arg His Asn Leu Pro Lys Val Asn Val
405          410          415
ct c atc aag t cg cta tgc aag gag ttc gac atc ccg ttc cac gag acc 1296
Leu Ile Lys Ser Leu Cys Lys Gu Phe Asp Ile Pro Phe His Gu Thr
420          425          430
ggc ttc tgg gag ggc atc tac gag gt g gt g gac cac ct g gcg gac atc 1344
Gly Phe Trp Gu Gly Ile Tyr Gu Val Val Asp His Leu Ala Asp Ile
435          440          445          450
agc aag gaa ttt atc acc gag ttc cca gcg atg t aa 1380
Ser Lys Gu Phe Ile Thr Gu Phe Pro Ala Met
450          455          460

```

<210> 2  
 <211> 459  
 <212> PRT  
 <213> Phytium irregulare

```

<400> 2
Met Val Asp Leu Lys Pro Gly Val Lys Arg Leu Val Ser Trp Lys Gu
1 5 10 15
Ile Arg Gu His Ala Thr Pro Ala Thr Ala Trp Ile Val Ile His His
20 25 30
Lys Val Tyr Asp Ile Ser Lys Trp Asp Ser His Pro Gly Gly Ser Val
35 40 45
Met Leu Thr Gn Ala Gly Gu Asp Ala Thr Asp Ala Phe Ala Val Phe
50 55 60
His Pro Ser Ser Ala Leu Lys Leu Leu Gu Gn Phe Tyr Val Gly Asp
65 70 75 80
Val Asp Gu Thr Ser Lys Ala Gu Ile Gu Gly Gu Pro Ala Ser Asp
85 90 95
Gu Gu Arg Ala Arg Arg Gu Arg Ile Asn Gu Phe Ile Ala Ser Tyr
100 105 110
Arg Arg Leu Arg Val Lys Val Lys Gly Met Gly Leu Tyr Asp Ala Ser

```



PF58307.txt

115 120 125  
 Ala Leu Tyr Tyr Ala Trp Lys Leu Val Ser Thr Phe Gly Ile Ala Val  
 130 135 140  
 Leu Ser Met Ala Ile Cys Phe Phe Phe Asn Ser Ser Phe Ala Met Tyr Met  
 145 150 155  
 Val Ala Gly Val Ile Met Gly Leu Phe Tyr Gln Gln Ser Gly Trp Leu  
 165 170 175  
 Ala His Asp Phe Leu His Asn Gln Val Cys Glu Asn Arg Thr Leu Gly  
 180 185  
 Asn Leu Ile Gly Cys Leu Val Gly Asn Ala Trp Gln Gly Phe Ser Met  
 195 200 205  
 Gln Trp Trp Lys Asn Lys His Asn Leu His His Ala Val Pro Asn Leu  
 210 215  
 His Ser Ala Lys Asp Glu Gly Phe Ile Gly Asp Pro Asp Ile Asp Thr  
 225 230 235  
 Met Pro Leu Leu Ala Trp Ser Lys Glu Met Ala Arg Lys Ala Phe Glu  
 245 250 255  
 Ser Ala His Gly Pro Phe Phe Ile Arg Asn Gln Ala Phe Leu Tyr Phe  
 260 265  
 Pro Leu Leu Leu Leu Ala Arg Leu Ser Trp Leu Ala Gln Ser Phe Phe  
 275 280 285  
 Tyr Val Phe Thr Glu Phe Ser Phe Gly Ile Phe Asp Lys Val Glu Phe  
 290 295 300  
 Asp Gly Pro Glu Lys Ala Gly Leu Ile Val His Tyr Ile Trp Gln Leu  
 305 310 315 320  
 Ala Ile Pro Tyr Phe Cys Asn Met Ser Leu Phe Glu Gly Val Ala Tyr  
 325 330 335  
 Phe Leu Met Gly Gln Ala Ser Cys Gly Leu Leu Ala Leu Val Phe  
 340 345  
 Ser Ile Gly His Asn Gly Met Ser Val Tyr Glu Arg Glu Thr Lys Pro  
 355 360 365  
 Asp Phe Trp Gln Leu Gln Val Thr Thr Thr Arg Asn Ile Arg Ala Ser  
 370 375 380  
 Val Phe Met Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Asp His  
 385 390 395 400  
 His Leu Phe Pro Leu Val Pro Arg His Asn Leu Pro Lys Val Asn Val  
 405 410 415  
 Leu Ile Lys Ser Leu Cys Lys Glu Phe Asp Ile Pro Phe His Glu Thr  
 420 425 430  
 Gly Phe Trp Glu Gly Ile Tyr Glu Val Val Asp His Leu Ala Asp Ile  
 435 440 445  
 Ser Lys Glu Phe Ile Thr Glu Phe Pro Ala Met  
 450 455

<210> 3  
 <211> 1380  
 <212> DNA  
 <213> Rhizopus stolonifer

<220>  
 <221> CDS  
 <222> (1)..(1380)

<400> 3  
 atg agt aca tta gat cgt caa tct att ttc aca ata aag gaa tta gag 48  
 Met Ser Thr Leu Asp Arg Gln Ser Ile Phe Thr Ile Lys Glu Leu Glu  
 1 5 10 15  
 tcg atc agt caa cgt att cat gat gga gat gaa gag gct atg aag ttc 96  
 Ser Ile Ser Gln Arg Ile His Asp Gly Asp Glu Glu Ala Met Lys Phe  
 20 25 30  
 att att atc gac aag aag gtg tac gat gtt act gaa ttt att gaa gac 144  
 Ile Ile Ile Asp Lys Lys Val Tyr Asp Val Thr Glu Phe Ile Glu Asp  
 35 40 45  
 cac cct ggt ggt gct caa gtt ttg ttg act cat gtt ggg aaa gat gca 192  
 His Pro Gly Gly Ala Gln Val Leu Leu Thr His Val Gly Lys Asp Ala  
 50 55 60  
 tct gac gta ttt cat gcc atg cat ccc gaa tct gct tac gaa gtc ttg 240  
 Ser Asp Val Phe His Ala Met His Pro Glu Ser Ala Tyr Glu Val Leu



PF58307.txt

Phe Leu Lys Gly Thr Leu Glu Val Leu Gln Thr Leu Asp Ile Thr Ser  
 435 440 445  
 aag ct c tct ttg caa tta agc aaa aag t cg ttt t aa  
 Lys Leu Ser Leu Gln Leu Ser Lys Lys Ser Phe

1380

<210> 4  
 <211> 459  
 <212> PRT  
 <213> Rhi zopus st ol oni fer

<400> 4  
 Met Ser Thr Leu Asp Arg Gln Ser Ile Phe Thr Ile Lys Glu Leu Glu  
 1 5 10 15  
 Ser Ile Ser Gln Arg Ile His Asp Gly Asp Glu Glu Ala Met Lys Phe  
 20 25 30  
 Ile Ile Ile Asp Lys Lys Val Tyr Asp Val Thr Glu Phe Ile Glu Asp  
 35 40 45  
 His Pro Gly Gly Ala Gln Val Leu Leu Thr His Val Gly Lys Asp Ala  
 50 55 60  
 Ser Asp Val Phe His Ala Met His Pro Glu Ser Ala Tyr Glu Val Leu  
 65 70 75 80  
 Asn Asn Tyr Phe Val Gly Asp Val Gln Glu Thr Val Val Thr Glu Lys  
 85 90 95  
 Ser Ser Ser Ala Gln Phe Ala Val Glu Met Arg Gln Leu Arg Asp Gln  
 100 105 110  
 Leu Lys Lys Glu Gly Tyr Phe His Ser Ser Lys Leu Phe Tyr Ala Tyr  
 115 120 125  
 Lys Val Leu Ser Thr Leu Ala Ile Cys Ile Ala Gly Leu Ser Pro Leu  
 130 135 140  
 Tyr Ala Tyr Gly Arg Thr Ser Thr Leu Ala Val Val Ala Ser Ala Ile  
 145 150 155 160  
 Thr Val Gly Ile Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe  
 165 170 175  
 Gly His His Gln Cys Phe Glu Asp Arg Thr Trp Asn Asp Val Leu Val  
 180 185 190  
 Val Phe Leu Gly Asn Phe Cys Gln Gly Phe Ser Leu Ser Trp Trp Lys  
 195 200 205  
 Asn Lys His Asn Thr His His Ala Ser Thr Asn Val His Gly Gln Asp  
 210 215 220  
 Pro Asp Ile Asp Thr Ala Pro Val Leu Leu Trp Asp Glu Tyr Ala Ser  
 225 230 235 240  
 Ala Ala Tyr Tyr Ala Ser Leu Asp Gln Glu Pro Thr Met Val Ser Arg  
 245 250 255  
 Phe Leu Ala Glu Gln Val Leu Pro His Gln Thr Arg Tyr Phe Phe Phe  
 260 265 270  
 Ile Leu Ala Phe Ala Arg Leu Ser Trp Ala Leu Gln Ser Leu Ser Tyr  
 275 280 285  
 Ser Phe Lys Lys Glu Ser Ile Asn Lys Ser Arg Gln Leu Asn Leu Phe  
 290 295 300  
 Glu Arg Val Cys Ile Val Gly His Trp Ala Leu Ser Ala Phe Cys Ile  
 305 310 315 320  
 Tyr Ser Trp Cys Ser Asn Val Tyr His Met Val Leu Phe Phe Leu Val  
 325 330 335  
 Ser Gln Ala Thr Thr Gly Tyr Thr Leu Ala Leu Val Phe Ala Leu Asn  
 340 345 350  
 His Asn Gly Met Pro Val Ile Thr Glu Glu Lys Ala Glu Ser Met Glu  
 355 360 365  
 Phe Phe Glu Ile Gln Val Ile Thr Gly Arg Asp Val Thr Leu Ser Pro  
 370 375 380  
 Leu Gly Asp Trp Phe Met Gly Gly Leu Asn Tyr Gln Ile Glu His His  
 385 390 395 400  
 Val Phe Pro Asn Met Pro Arg His Asn Leu Pro Thr Val Lys Pro Met  
 405 410 415  
 Val Lys Ser Leu Cys Gln Lys Tyr Asp Ile Asn Tyr His Asp Thr Gly  
 420 425 430  
 Phe Leu Lys Gly Thr Leu Glu Val Leu Gln Thr Leu Asp Ile Thr Ser  
 435 440 445

Lys Leu Ser Leu G n Leu Ser Lys Lys Ser Phe  
 450 455

<210> 5  
 <211> 1380  
 <212> DNA  
 <213> Rhizopus stolonifer

<220>  
 <221> CDS  
 <222> (1)..(1380)

<400> 5  
 at g agt aca tta gat cgt caa tct att ttc aca ata aag gaa tta gag 48  
 Met Ser Thr Leu Asp Arg G n Ser Ile Phe Thr Ile Lys G u Leu G u  
 1 5 10 15  
 t cg at c agt caa cgt att cat gat gga gat gaa gag gct at g aag ttc 96  
 Ser Ile Ser G n Arg Ile His Asp G y Asp G u G u Ala Met Lys Phe  
 20 25 30  
 att att at c gac aag aag gt g tac gat gtt act gaa ttt att gaa gac 144  
 Ile Ile Ile Asp Lys Lys Val Tyr Asp Val Thr G u Phe Ile G u Asp  
 35 40 45  
 cac cct ggt ggt gct caa gtt ttg ttg act cat gtt ggg aaa gat gca 192  
 His Pro Gly Gly Ala G n Val Leu Leu Thr His Val Gly Lys Asp Ala  
 50 55 60  
 t ct gac gta ttt cat gcc at g cat ccc gaa t ct gct t ac gaa gt c tt g 240  
 Ser Asp Val Phe His Ala Met His Pro G u Ser Ala Tyr G u Val Leu  
 65 70 75  
 aac aat t ac ttc gta ggc gat gtt caa gaa aca gta gt c act gaa aaa 288  
 Asn Asn Tyr Phe Val Gly Asp Val G n G u Thr Val Val Thr G u Lys  
 85 90 95  
 t ct t ct t ct gcc caa ttc gct gta gaa at g cgt cag ct t cgg gat caa 336  
 Ser Ser Ser Ala G n Phe Ala Val G u Met Arg G n Leu Arg Asp G n  
 100 105 110  
 ttg aaa aag gaa ggt tat ttc cac tcc agc aaa ct c ttt tat gct t ac 384  
 Leu Lys Lys G u Gly Tyr Phe His Ser Ser Lys Leu Phe Tyr Ala Tyr  
 115 120 125  
 aaa gtt ct t t ct act ttg gct at c tgc at a gct ggt ct c t ct ct t tt g 432  
 Lys Val Leu Ser Thr Leu Ala Ile Cys Ile Ala Gly Leu Ser Leu Leu  
 130 135 140  
 tat gct tat ggt cgt act tcc act ttg gct gt c gt c gca t ct gct att 480  
 Tyr Ala Tyr Gly Arg Thr Ser Thr Leu Ala Val Val Ala Ser Ala Ile  
 145 150 155 160  
 act gtt ggt at c ttt tgg caa cag tgt ggt tgg ct g gct cac gat ttc 528  
 Thr Val Gly Ile Phe Trp G n G n Cys Gly Trp Leu Ala His Asp Phe  
 165 170 175  
 gga cat cat caa tgc ttc gaa gac cgc act tgg aac gat gtt ct c gt c 576  
 Gly His His G n Cys Phe G u Asp Arg Thr Trp Asn Asp Val Leu Val  
 180 185 190  
 gta ttt ct c gga aac ttt tgt caa gga ttt tca tta tca tgg tgg aag 624  
 Val Phe Leu Gly Asn Phe Cys G n Gly Phe Ser Leu Ser Trp Trp Lys  
 195 200 205  
 aac aag cac aac aca cat cac gca agt aca aat gtt cat gga cag gat 672  
 Asn Lys His Asn Thr His His Ala Ser Thr Asn Val His Gly G n Asp  
 210 215 220  
 ccc gac att gac aca gcc cct gtt ttg ttg tgg gat gaa tat gct t ct 720  
 Pro Asp Ile Asp Thr Ala Pro Val Leu Leu Trp Asp G u Tyr Ala Ser  
 225 230 235 240  
 gca gct t ac tat gct t ct ct c gat caa gaa ccc at g gt a t ct cga 768  
 Ala Ala Tyr Tyr Ala Ser Leu Asp G n G u Pro Thr Met Val Ser Arg  
 245 250 255  
 ttc ct t gct gaa caa gt c ttg ccc cac caa acc cgt tat ttt ttc ttt 816  
 Phe Leu Ala G u G n Val Leu Pro His G n Thr Arg Tyr Phe Phe Phe  
 260 265 270  
 at c ct c gct ttt gct cgt ttg t ct tgg gct ct t caa t ct ttg agc t ac 864  
 Ile Leu Ala Phe Ala Arg Leu Ser Trp Ala Leu G n Ser Leu Ser Tyr  
 275 280 285  
 t ct ttc aag aag gag tca att aac aag tcc cgt caa ct t aac ct t ttt 912

PF58307.txt

Ser Phe Lys Lys Glu Ser Ile Asn Lys Ser Arg Gln Leu Asn Leu Phe  
 290 295 300  
 gaa cgt gtt tgt att gtt ggt cac tgg gct ct c ttt gct ttt tgt at c 960  
 Gu Arg Val Cys lle Val Gy His Trp Ala Leu Phe Ala Phe Cys lle 320  
 305  
 tac tct tgg tgt agc aac gtc tat cac atg gtt ttg ttt ttc ttg gtt 1008  
 Tyr Ser Trp Cys Ser Asn Val Tyr His Met Val Leu Phe Phe Leu Val 335  
 325  
 agt cag gct acc aca ggc tac act tta gct ctt gtc ttt gct ttg aac 1056  
 Ser Gn Ala Thr Thr Gy Tyr Thr Leu Ala Leu Val Phe Ala Leu Asn 350  
 340  
 cac aat ggt atg cct gtt at c act gaa gaa aag gct gaa t c g at g gaa 1104  
 His Asn Gy Met Pro Val lle Thr Gu Gu Lys Ala Gu Ser Met Gu 365  
 355  
 ttt ttt gag att caa gtt at c act ggt cgt gat gtt act ctt tcc ccc 1152  
 Phe Phe Gu lle Gn Val lle Thr Gy Arg Asp Val Thr Leu Ser Pro 375  
 370  
 ttg ggt gat tgg ttc atg ggt ggt ttg aac tat caa at c gaa cat cat 1200  
 Leu Gy Asp Trp Phe Met Gy Gy Leu Asn Tyr Gn lle Gu His His 390  
 385 400  
 gtt ttc ccc aac atg cct cgt cac aac ttg cct aca gtt aag ccc atg 1248  
 Val Phe Pro Asn Met Pro Arg His Asn Leu Pro Thr Val Lys Pro Met 415  
 405  
 gtg aaa t c g ctt tgt cag aaa tat gac att aac tat cac gat act gga 1296  
 Val Lys Ser Leu Cys Gn Lys Tyr Asp lle Asn Tyr His Asp Thr Gy 425  
 420 430  
 ttc tta aag gga aca tta gaa gta tta caa aca ctg gat at t aca tcc 1344  
 Phe Leu Lys Gy Thr Leu Gu Val Leu Gn Thr Leu Asp lle Thr Ser 445  
 435  
 aag ct c tct ttg caa tta agc aaa aag t c g ttt taa 1380  
 Lys Leu Ser Leu Gn Leu Ser Lys Lys Ser Phe 455  
 450

<210> 6  
 <211> 459  
 <212> PRT  
 <213> Rhi zopus st ol onif er

<400> 6  
 Met Ser Thr Leu Asp Arg Gn Ser Ile Phe Thr lle Lys Gu Leu Gu  
 1 5 10 15  
 Ser Ile Ser Gn Arg lle His Asp Gy Asp Gu Gu Ala Met Lys Phe  
 20 25 30  
 lle lle lle Asp Lys Lys Val Tyr Asp Val Thr Gu Phe lle Gu Asp  
 35 40 45  
 His Pro Gy Gy Ala Gn Val Leu Leu Thr His Val Gy Lys Asp Ala  
 50 55 60  
 Ser Asp Val Phe His Ala Met His Pro Gu Ser Ala Tyr Gu Val Leu  
 65 70 75 80  
 Asn Asn Tyr Phe Val Gy Asp Val Gn Gu Thr Val Val Thr Gu Lys  
 85 90 95  
 Ser Ser Ser Ala Gn Phe Ala Val Gu Met Arg Gn Leu Arg Asp Gn  
 100 105 110  
 Leu Lys Lys Gu Gy Tyr Phe His Ser Ser Lys Leu Phe Tyr Ala Tyr  
 115 120 125  
 Lys Val Leu Ser Thr Leu Ala lle Cys lle Ala Gy Leu Ser Leu Leu  
 130 135 140  
 Tyr Ala Tyr Gy Arg Thr Ser Thr Leu Ala Val Val Ala Ser Ala lle  
 145 150 155 160  
 Thr Val Gy lle Phe Trp Gn Gn Cys Gy Trp Leu Ala His Asp Phe  
 165 170 175  
 Gy His His Gn Cys Phe Gu Asp Arg Thr Trp Asn Asp Val Leu Val  
 180 185 190  
 Val Phe Leu Gy Asn Phe Cys Gn Gy Phe Ser Leu Ser Trp Trp Lys  
 195 200 205  
 Asn Lys His Asn Thr His His Ala Ser Thr Asn Val His Gy Gn Asp  
 210 215 220  
 Pro Asp lle Asp Thr Ala Pro Val Leu Leu Trp Asp Gu Tyr Ala Ser

PF58307. txt

225 Ala Ala Tyr Tyr Ala Ser Leu Asp Gln Glu Pro Thr Met Val Ser Arg 240  
 235  
 Phe Leu Ala Glu Val Leu Pro His Glu Thr Arg Tyr Phe Phe Phe 255  
 245  
 Ile Leu Ala Phe Ala Arg Leu Ser Trp Ala Leu Gln Ser Leu Ser Tyr 270  
 260  
 Ser Phe Lys Lys Glu Ser Ile Asn Lys Ser Arg Gln Leu Asn Leu Phe 285  
 275  
 Glu Arg Val Cys Ile Val Asn Val Tyr His Trp Ala Leu Phe Ala Phe Cys Ile 300  
 290  
 Tyr Ser Trp Cys Ser Asn Val Tyr His Met Val Leu Phe Phe Leu Val 315  
 305  
 Ser Gln Ala Thr Thr Gly Tyr Thr Leu Ala Leu Val Phe Ala Phe Leu Asn 330  
 310  
 His Asn Gly Met Pro Val Ile Thr Glu Glu Lys Ala Glu Ser Met Glu 335  
 325  
 Phe Phe Glu Ile Gln Val Ile Thr Gly Arg Asp Val Thr Leu Ser Pro 350  
 335  
 Leu Gly Asp Trp Phe Met Gly Gly Leu Asn Tyr Gln Ile Glu His His 365  
 340  
 Val Phe Pro Asn Met Pro Arg His Asn Leu Pro Thr Val Lys Pro Met 380  
 355  
 Val Lys Ser Leu Cys Gln Lys Tyr Asp Ile Asn Tyr His Asp Thr Gly 395  
 370  
 Phe Leu Lys Gly Thr Leu Glu Val Leu Gln Thr Leu Asp Ile Thr Ser 400  
 385  
 Lys Leu Ser Leu Gln Leu Ser Lys Lys Ser Phe 415  
 420  
 435  
 440  
 450

<210> 7

<211> 1401

<212> DNA

<213> Cunnin ghamel l a echi nul at a

<220>

<221> CDS

<222> ( 1) . . ( 1401)

<400> 7

at g t ca ggg caa act cga gtt ttt aag cga tcc gaa gtg tcc gat t cg 48  
 Met Ser Gly Gln Thr Arg Val Phe Lys Arg Ser Glu Val Ser Asp Ser  
 1 5 10 15  
 ct c aag gcg t ac caa gcg ggc gac aag aac gcc gac aag ttc ct g att 96  
 Leu Lys Ala Tyr Gln Ala Gly Asp Lys Asn Ala Asp Lys Phe Leu Ile  
 20 25 30  
 gt c gac aac aag gt g t ac gac at c acg gat ttc att gct gat cat cct 144  
 Val Asp Asn Lys Val Tyr Asp Ile Thr Asp Phe Ile Ala Asp His Pro  
 35 40 45  
 ggc ggt gct cag gt c at c t ct acg cac att ggc aag gat gcc t ct gat 192  
 Gly Gly Ala Gln Val Ile Ser Thr His Ile Gly Lys Asp Ala Ser Asp  
 50 55 60  
 gt g ttt cat gcg at g cat ccc gag t ct gcg t ac gaa tt g ct t gcg aat 240  
 Val Phe His Ala Met His Pro Gu Ser Ala Tyr Glu Leu Leu Ala Asn  
 65 70 75 80  
 t gc t ac gt t ggc gat ct t gct gcg gac cat gct ggt gt g cag ggc gaa 288  
 Cys Tyr Val Gly Asp Leu Ala Ala Asp His Ala Gly Val Gln Gly Gu  
 85 90 95  
 tt g gt g aat ggt gt a cac aaa aag tcc aag gca ttt gcg gac gag at g 336  
 Leu Val Asn Gly Val His Lys Lys Ser Lys Ala Phe Ala Asp Gu Met  
 100 105 110  
 cga t cg tt g cgc gag cgt ct t gag acg gag ggt gcc ttt aat ggc agt 384  
 Arg Ser Leu Arg Glu Arg Leu Gu Thr Glu Gly Ala Phe Asn Gly Ser  
 115 120 125  
 gt t ccg ttt t ac att t ac aag gt g gt g t cg acg tt g gcc att ggt gcg 432  
 Val Pro Phe Tyr Ile Tyr Lys Val Val Ser Thr Leu Ala Ile Gly Ala  
 130 135 140  
 acg ggt ct g gcc at g ct g t ac t at ggc ggc cat t cg acg t cg gt c gt g 480





PF58307. txt

<400> 8

Met Ser Gly Gln Thr Arg Val Phe Lys Arg Ser Glu Val Ser Asp Ser  
 1 5 10 15  
 Leu Lys Ala Tyr Gln Ala Gly Asp Lys Asn Ala Asp Lys Phe Leu Ile  
 20 25 30  
 Val Asp Asn Lys Val Tyr Asp Ile Thr Asp Phe Ile Ala Asp His Pro  
 35 40 45  
 Gly Gly Ala Gln Val Ile Ser Thr His Ile Gly Lys Asp Ala Ser Asp  
 50 55 60  
 Val Phe His Ala Met His Pro Glu Ser Ala Tyr Glu Leu Leu Ala Asn  
 65 70 75 80  
 Cys Tyr Val Gly Asp Leu Ala Ala Asp His Ala Gly Val Gln Gly Glu  
 85 90 95  
 Leu Val Asn Gly Val His Lys Lys Ser Lys Ala Phe Ala Asp Asn Gly Ser  
 100 105 110  
 Arg Ser Leu Arg Glu Arg Leu Glu Thr Glu Gly Ala Phe Asn Gly Ser  
 115 120 125  
 Val Pro Phe Tyr Ile Tyr Lys Val Val Ser Thr Leu Ala Ile Gly Ala  
 130 135 140  
 Thr Gly Leu Ala Met Leu Tyr Tyr Gly Gly His Ser Thr Ser Val Val  
 145 150 155 160  
 Leu Ala Ala Ala Val Val Gly Leu Phe Trp Gln Gln Cys Gly Trp  
 165 170 175  
 Leu Ala His Asp Phe Gly His His Gln Ala Phe Ala Asp His Thr Val  
 180 185 190  
 Asn Asp Val Met Ile Ala Phe Leu Gly Gly Phe Cys Gln Gly Phe Ser  
 195 200 205  
 Leu Ser Trp Trp Lys Asn Lys His Asn Thr His His Ala Ser Thr Asn  
 210 215 220  
 Val His Gly His Asp Pro Asp Ile Asp Thr Ala Pro Val Leu Leu Trp  
 225 230 235 240  
 Asp Glu Phe Ala Thr Ala Asn Phe Tyr Gly Asn Leu Glu Gly Gln Lys  
 245 250 255  
 Asp Ser Ala Phe Ser Arg Phe Ile Ala Glu His Val Leu Pro Tyr Gln  
 260 265 270  
 Thr Arg Tyr Tyr Phe Phe Val Leu Gly Phe Ala Arg Leu Ser Trp Ala  
 275 280 285  
 Ile Gln Ser Leu Gln Tyr Ser Phe Thr Val Gly Thr Leu Asn Lys Ser  
 290 295 300  
 Lys Thr Leu Asn Leu Phe Glu Arg Thr Met Leu Val Ser His Trp Ile  
 305 310 315 320  
 Leu Phe Thr Thr Trp Thr Leu Leu Phe Ile Asn Ser Trp Thr Asn Met  
 325 330 335  
 Val Met Phe Phe Val Val Ser Gln Ala Thr Thr Gly Tyr Ala Leu Ala  
 340 345 350  
 Leu Val Phe Ala Met Asn His Ser Gly Met Pro Val Leu Thr Gln Glu  
 355 360 365  
 Gln Ala Gln Lys Met Glu Phe Tyr Glu Ile Gln Val Val Thr Ala Arg  
 370 375 380  
 Asp Val Thr Leu Gly Ala Leu Gly Asp Trp Phe Cys Gly Gly Leu Asn  
 385 390 395 400  
 Tyr Gln Ile Glu His His Val Phe Pro Asp Met Pro Arg His Tyr Leu  
 405 410 415  
 Pro Lys Val Lys Pro Gln Val Lys Ala Leu Cys Lys Lys His Asn Ile  
 420 425 430  
 Leu Tyr His Asp Thr Ser Ala Leu Arg Gly Thr Leu Glu Val Leu Gln  
 435 440 445  
 Thr Leu Asp Val Val Gln Lys Leu Cys Ala Ser Ser Leu Lys Arg Cys  
 450 455 460  
 Phe Arg  
 465

<210> 9

<211> 1398

<212> DNA

<213> G lossomast i x chrysopl ast a

<220>



<221> CDS

<222> (1)..(1398)

<400> 9

at g	ggt	aaa	gga	ggc	cgg	gac	gcc	ggc	gct	gt c	ggc	ggc	gaa	gcg	gag	48
Met	Gly	Lys	Gly	Gly	Arg	Asp	Ala	Gly	Ala	Val	Gly	Gly	Glu	Ala	Glu	
1			5					10					15			
aag	acg	ct g	ccc	aag	ttt	acg	ttg	gag	gag	at c	cag	aag	cat	cgc	acg	96
Lys	Thr	Leu	Pro	Lys	Phe	Thr	Leu	Glu	Glu	Ile	Gln	Lys	His	Arg	Thr	
			20					25					30			
ccg	cag	gat	gcg	tgg	at g	gt g	cac	cac	aac	aag	gt g	t ac	gat	gt g	t cc	144
Pro	Gln	Asp	Ala	Trp	Met	Val	His	His	Asn	Lys	Val	Tyr	Asp	Val	Ser	
			35				40					45				
aac	t at	at g	gac	cac	ccc	ggc	ggt	ct g	gt g	at c	ttt	agc	cac	gcg	ggc	192
Asn	Tyr	Met	Asp	His	Pro	Gly	Gly	Leu	Val	Ile	Phe	Ser	His	Ala	Gly	
	50					55					60					
gat	gac	at g	acc	gac	gt c	ttc	gcg	gcg	ttc	cac	ccg	ccc	t cg	gcg	ttc	240
Asp	Asp	Met	Thr	Asp	Val	Phe	Ala	Ala	Phe	His	Pro	Pro	Ser	Ala	Phe	
	65				70				75					80		
aac	ttc	at g	gac	aag	ttc	ttg	att	ggc	gt g	gt g	gac	agc	aag	ggc	t cg	288
Asn	Phe	Met	Asp	Lys	Phe	Leu	Ile	Gly	Val	Val	Asp	Ser	Lys	Gly	Ser	
				85				90					95			
t cg	ccg	cag	ct c	cag	aag	gac	gcg	t cc	cag	gcc	agc	ttc	gag	aag	gca	336
Ser	Pro	Gln	Leu	Gln	Lys	Asp	Ala	Ser	Gln	Ala	Ser	Phe	Glu	Lys	Ala	
			100					105					110			
t ac	cgc	aat	ct g	cga	gt c	cag	ct c	aag	aag	gcc	ggc	at g	ttc	aag	gcc	384
Tyr	Arg	Asn	Leu	Arg	Val	Gln	Leu	Lys	Lys	Ala	Gly	Met	Phe	Lys	Ala	
			115				120					125				
t cg	agc	ct c	ttc	t ac	acc	t ac	aag	gt g	ct c	t cg	acc	ct g	gcg	ct g	t gc	432
Ser	Ser	Leu	Phe	Tyr	Thr	Tyr	Lys	Val	Leu	Ser	Thr	Leu	Ala	Leu	Cys	
	130					135				140						
ct c	gt g	t cc	t gg	ggc	ct c	gt g	ct g	ggg	t cg	gac	cac	ttc	ggt	gt g	cac	480
Leu	Val	Ser	Trp	Gly	Leu	Val	Leu	Gly	Ser	Asp	His	Phe	Gly	Val	His	
	145			150				155		160					160	
ct c	gt g	ggc	gct	ct c	ttc	ct c	gcg	ct c	ttc	t gg	cag	cag	t gc	ggc	t gg	528
Leu	Val	Gly	Ala	Leu	Phe	Leu	Ala	Leu	Phe	Trp	Gln	Gln	Cys	Gly	Trp	
			165					170					175			
ttg	gcc	cac	gac	ttc	ct g	cac	cac	cag	gt c	ttc	cag	aac	cgg	gct	cac	576
Leu	Ala	His	Asp	Phe	Leu	His	His	Gln	Val	Phe	Gln	Asn	Arg	Ala	His	
			180					185					190			
ggc	gac	ct c	gcc	ggt	at c	at g	at c	ggc	aac	gt g	t gg	cag	ggc	ttc	t cc	624
Gly	Asp	Leu	Ala	Gly	Ile	Met	Ile	Gly	Asn	Val	Trp	Gln	Gly	Phe	Ser	
			195				200					205				
gt g	gcc	t gg	t gg	aag	aac	aag	cac	aac	acc	cac	cac	t cg	gt g	ccg	aac	672
Val	Ala	Trp	Trp	Lys	Asn	Lys	His	Asn	Thr	His	His	Ser	Val	Pro	Asn	
			210			215						220				
ct c	t ac	gag	t cg	cag	ccc	gac	gcc	gcc	gac	ggc	gat	ccg	gac	at c	gac	720
Leu	Tyr	Glu	Ser	Gln	Pro	Asp	Ala	Ala	Asp	Gly	Asp	Pro	Asp	Ile	Asp	
	225			230					235					240		
acc	at g	ccc	ct c	ct c	gcc	t gg	t cg	ct g	cgc	at g	gcc	aag	aac	gcg	gac	768
Thr	Met	Pro	Leu	Leu	Ala	Trp	Ser	Leu	Arg	Met	Ala	Lys	Asn	Ala	Asp	
				245				250						255		
aac	gcg	ct a	t cc	cgg	t gg	ttc	gt g	t cc	cac	cag	gcc	ttc	t gc	t ac	ttc	816
Asn	Ala	Leu	Ser	Arg	Trp	Phe	Val	Ser	His	Gln	Ala	Phe	Cys	Tyr	Phe	
			260					265					270			
ccc	at c	ct t	ggc	ct c	gcg	cgg	ct c	t cg	t gg	ct c	gag	ggg	t cc	ttc	t cc	864
Pro	Ile	Leu	Gly	Leu	Ala	Arg	Leu	Ser	Trp	Leu	Glu	Gly	Ser	Phe	Ser	
			275				280					285				
ttc	gt g	ttc	t cc	aac	ccg	ttg	gca	t gg	aag	acg	aag	aac	ct t	gat	gt a	912
Phe	Val	Phe	Ser	Asn	Pro	Leu	Ala	Trp	Lys	Thr	Lys	Asn	Leu	Asp	Val	
	290					295					300					
gcc	aag	cag	ct c	gt g	acc	aac	ccc	ct g	ttg	gag	cag	gcc	ggc	ct c	ct g	960
Ala	Lys	Gln	Leu	Val	Thr	Asn	Pro	Leu	Leu	Glu	Gln	Ala	Gly	Leu	Leu	
	305				310					315					320	
gt c	cac	t ac	gcc	t gg	gt c	ttc	gcc	ct c	t gc	gcg	t gc	acg	ggc	t cc	ct g	1008
Val	His	Tyr	Ala	Trp	Val	Phe	Ala	Leu	Cys	Ala	Cys	Thr	Gly	Ser	Leu	
				325					330					335		
gt g	cgc	gcc	ct g	gcg	ttc	ttc	ttc	gt g	gcc	acc	t gc	acc	agc	ggc	ct t	1056

PF58307.txt

Val	Arg	Ala	Leu	Ala	Phe	Phe	Phe	Val	Ala	Thr	Cys	Thr	Ser	Gly	Leu	
ctc	ctc	gcc	atc	gtc	ttc	ggc	ctc	ggc	cac	aac	ggc	atg	gcg	ctc	tac	1104
Leu	Leu	Ala	Ile	Val	Phe	Gly	Leu	Gly	His	Asn	Gly	Met	Ala	Leu	Tyr	
		340					360					365				
gag	gcc	aat	gcc	cgg	ccg	gac	ttc	tgg	aag	ctc	cag	gtg	acc	acc	acg	1152
Glu	Ala	Asn	Ala	Arg	Pro	Asp	Phe	Trp	Lys	Leu	Gln	Val	Thr	Thr	Thr	
		370				375					380					
cgg	aac	atc	acc	ggc	agc	ccc	ttt	gtg	cac	tgg	ttc	tgc	ggc	ggc	ctc	1200
Arg	Asn	Ile	Thr	Gly	Ser	Pro	Phe	Val	His	Trp	Phe	Cys	Gly	Gly	Leu	
385					390					395					400	
cag	ttc	cag	gtg	gag	cac	cac	ctc	ttc	ccc	tcg	ctg	ccc	cgg	cac	aac	1248
Gln	Phe	Gln	Val	Glu	His	His	Leu	Phe	Pro	Ser	Leu	Pro	Arg	His	Asn	
				405					410					415		
ctc	ccc	cgg	gcc	cac	gag	atc	gtg	acg	gcc	ttc	tgc	aag	gag	cag	ggc	1296
Leu	Pro	Arg	Ala	His	Glu	Ile	Val	Thr	Ala	Phe	Cys	Lys	Glu	Gln	Gly	
			420				425						430			
gtc	aag	tac	cac	gag	gcc	gac	ctc	ctt	acc	ggg	acc	aag	gag	att	ctt	1344
Val	Lys	Tyr	His	Glu	Ala	Asp	Leu	Leu	Thr	Gly	Thr	Lys	Glu	Ile	Leu	
		435					440					445				
tgc	tgc	ttg	tgc	gag	gta	acg	acg	gag	ttc	ctc	gac	gag	ttc	cca	gcc	1392
Ser	Cys	Leu	Ser	Glu	Val	Thr	Thr	Glu	Phe	Leu	Asp	Glu	Phe	Pro	Ala	
		450				455					460					
atg	taa															1398
Met																
465																

<210> 10  
 <211> 465  
 <212> PRT  
 <213> Gossomastix chrysoptasta

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

PF58307. txt

275 280 285  
 Phe Val Phe Ser Asn Pro Leu Ala Trp Lys Thr Lys Asn Leu Asp Val  
 290 295 300  
 Ala Lys Gln Leu Val Thr Asn Pro Leu Leu Glu Gln Ala Gly Leu Leu  
 305 310 315  
 Val His Tyr Ala Trp Val Phe Ala Leu Cys Ala Cys Thr Gly Ser Leu  
 325 330 335  
 Val Arg Ala Leu Ala Phe Phe Phe Val Ala Thr Cys Thr Ser Gly Leu  
 340 345 350  
 Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met Ala Leu Tyr  
 355 360 365  
 Glu Ala Asn Ala Arg Pro Asp Phe Trp Lys Leu Gln Val Thr Thr Thr  
 370 375 380  
 Arg Asn Ile Thr Gly Ser Pro Phe Val His Trp Phe Cys Gly Gly Leu  
 385 390 395 400  
 Gln Phe Gln Val Glu His His Leu Phe Pro Ser Leu Pro Arg His Asn  
 405 410 415  
 Leu Pro Arg Ala His Glu Ile Val Thr Ala Phe Cys Lys Glu Gln Gly  
 420 425 430  
 Val Lys Tyr His Glu Ala Asp Leu Leu Thr Gly Thr Lys Glu Ile Leu  
 435 440 445  
 Ser Cys Leu Ser Glu Val Thr Thr Glu Phe Leu Asp Glu Phe Pro Ala  
 450 455 460  
 Met  
 465

<210> 11  
 <211> 1455  
 <212> DNA  
 <213> Thal assi osi ra pseudonana

<220>  
 <221> CDS  
 <222> (1)..(1455)

<400> 11  
 at g gga aaa gga gga gac gca gcc gca gct acc aag cgt agt gga gca 48  
 Met Gly Lys Gly Gly Asp Ala Ala Ala Ala Thr Lys Arg Ser Gly Ala  
 1 5 10 15  
 ttg aaa ttg gcg gag aag ccg cag aag tac act tgg cag gag gtg aag 96  
 Leu Lys Leu Ala Glu Lys Pro Gln Lys Tyr Thr Trp Gln Glu Val Lys  
 20 25 30  
 aag cac atc acc ccc gac gat gcc tgg gta gtc cac caa aac aaa gtc 144  
 Lys His Ile Thr Pro Asp Asp Ala Trp Val Val His Gln Asn Lys Val  
 35 40 45  
 tac gac gtc tcc aac tgg tac gac cac ccc ggt gga gcc gtg gtg ttc 192  
 Tyr Asp Val Ser Asn Trp Tyr Asp His Pro Gly Gly Ala Val Val Phe  
 50 55 60  
 acc cac gcc gga gac gac atg acg gac atc ttc gcc gcc ttc cac gcc 240  
 Thr His Ala Gly Asp Asp Met Thr Asp Ile Phe Ala Ala Phe His Ala  
 65 70 75 80  
 caa ggc tct cag gcc atg atg aag aag ttt tac att gga gat ttg att 288  
 Gln Gly Ser Gln Ala Met Met Lys Lys Phe Tyr Ile Gly Asp Leu Ile  
 85 90 95  
 ccg gag agt gtg gag cat aag gat caa aga cag ttg gat ttc gag aag 336  
 Pro Glu Ser Val Glu His Lys Asp Gln Arg Gln Leu Asp Phe Glu Lys  
 100 105 110  
 gga tat cgt gat tta cgg gcc aag ctt gtc atg atg ggg atg ttc aag 384  
 Gly Tyr Arg Asp Leu Arg Ala Lys Leu Val Met Met Gly Met Phe Lys  
 115 120 125  
 tcg agt aag atg tat tat gca tac aag tgc tgc ttc aat atg tgc atg 432  
 Ser Ser Lys Met Tyr Tyr Ala Tyr Lys Cys Ser Phe Asn Met Cys Met  
 130 135 140  
 tgg ttg gtg gcg gtg gcc atg gtg tac tac tgc gac agt ttg gca atg 480  
 Trp Leu Val Ala Val Ala Met Val Tyr Tyr Ser Asp Ser Leu Ala Met  
 145 150 155 160  
 cac att gga tgc gct ctg ttg ttg gga ttg ttc tgg cag cag tgt gga 528  
 His Ile Gly Ser Ala Leu Leu Leu Gly Leu Phe Trp Gln Gln Cys Gly

PF58307. txt

				165					170					175					
t gg	ct t	g cg	c ac	g ac	t t t	ct t	c ac	c ac	caa	gt c	t t t	a ag	caa	c ga	a ag				576
Tr p	Leu	Al a	Hi s	Asp	Phe	Leu	Hi s	Hi s	G n	Val	Phe	Lys	G n	Ar g	Lys				
			180						185					190					
t ac	g ga	gat	ct c	gt t	ggc	at c	t t t	t gg	g ga	gat	ct c	at g	c ag	g gg	t t c				624
Tyr	G y	Asp	Leu	Val	G y	I le	Phe	Tr p	G y	Asp	Leu	Met	G n	G y	Phe				
		195					200					205							
t cg	at g	c ag	t gg	t gg	a ag	a ac	a ag	c ac	aat	ggc	c ac	cat	g ct	gt t	ccc				672
Ser	Met	G n	Tr p	Tr p	Lys	Asn	Lys	Hi s	Asn	G y	Hi s	Al a	Al a	Val	Pro				
	210					215					220								
a ac	t t g	c ac	a ac	t ct	t cc	t t g	g ac	ag t	c ag	gat	gg t	gat	ccc	gat	at t				720
Asn	Leu	Hi s	Asn	Ser	Ser	Leu	Asp	Ser	G n	Asp	G y	Asp	Pro	Asp	I le				240
225				230					235						240				
gat	acc	at g	cca	ct c	ct t	g ct	t gg	ag t	ct c	a ag	c ag	g ct	c ag	ag t	t t c				768
Asp	Thr	Met	Pro	Leu	Leu	Al a	Tr p	Ser	Leu	Lys	G n	Al a	G n	Ser	Phe				
				245					250					255					
aga	g ag	at c	aat	a ag	g ga	a ag	g ac	ag t	acc	t t c	gt c	a ag	t ac	g ct	at c				816
Arg	G u	I le	Asn	Lys	G y	Lys	Asp	Ser	Thr	Phe	Val	Lys	Tyr	Al a	I le				
			260					265					270						
aaa	t t c	c ag	g ca	t t c	aca	t ac	t t c	ccc	at c	ct c	ct c	t t g	g ct	c gc	at c				864
Lys	Phe	G n	Al a	Phe	Thr	Tyr	Phe	Pro	I le	Leu	Leu	Leu	Al a	Arg	I le				
	275						280					285							
t ct	t gg	t t g	aat	g aa	t cc	t t c	aaa	act	g ca	t t c	g ga	ct c	g ga	g ct	g cc				912
Ser	Tr p	Leu	Asn	G u	Ser	Phe	Lys	Thr	Al a	Phe	G y	Leu	G y	Al a	Al a				
	290					295					300								
t cg	g ag	aat	g cc	a ag	t t g	g ag	t t g	g ag	a ag	c gt	g ga	ct t	c ag	t ac	cca				960
Ser	G u	Asn	Al a	Lys	Leu	G u	Leu	G u	Lys	Arg	G y	Leu	G n	Tyr	Pro				320
305					310					315									
ct t	t t g	g ag	a ag	ct t	g ga	at c	acc	ct t	c ac	t ac	act	t gg	at g	t t c	gt c				1008
Leu	Leu	G u	Lys	Leu	G y	I le	Thr	Leu	Hi s	Tyr	Thr	Tr p	Met	Phe	Val				
				325					330					335					
ct c	t ct	t cc	g ga	t t t	g ga	agg	t gg	t ct	ct t	cca	t at	t cc	at c	at g	t at				1056
Leu	Ser	Ser	G y	Phe	G y	Arg	Tr p	Ser	Leu	Pro	Tyr	Ser	I le	Met	Tyr				
			340					345					350						
t t c	t t c	act	g cc	aca	t gc	t cc	t cg	g ga	ct t	t t c	ct c	g ca	t t g	gt c	t t t				1104
Phe	Phe	Thr	Al a	Thr	Cys	Ser	Ser	G y	Leu	Phe	Leu	Al a	Leu	Val	Phe				
		355					360					365							
g ga	t t g	g ga	c ac	a ac	gg t	at g	t ca	gt g	t ac	gat	g cc	acc	acc	c ga	c ct				1152
G y	Leu	G y	Hi s	Asn	G y	Met	Ser	Val	Tyr	Asp	Al a	Thr	Thr	Arg	Pro				
	370				375						380								
g ac	t t c	t gg	caa	ct c	caa	gt c	acc	act	aca	c gt	a ac	at c	at t	gg t	g ga				1200
Asp	Phe	Tr p	G n	Leu	G n	Val	Thr	Thr	Thr	Arg	Asn	I le	I le	G y	G y				400
385					390					395					400				
c ac	ggc	at t	ccc	caa	t t c	t t t	gt g	gat	t gg	t t c	t gc	gg t	g ga	t t g	caa				1248
Hi s	G y	I le	Pro	G n	Phe	Phe	Val	Asp	Tr p	Phe	Cys	G y	G y	Leu	G n				415
				405					410					415					
t ac	caa	gt g	gat	c ac	c ac	ct c	t t c	ccc	at g	at g	c ct	aga	a ac	aat	at c				1296
Tyr	G n	Val	Asp	Hi s	Hi s	Leu	Phe	Pro	Met	Met	Pro	Arg	Asn	Asn	I le				
			420					425				430							
g cg	a ag	t gc	c ac	a ag	ct t	gt g	g ag	t ca	t t c	t gt	a ag	g ag	t gg	gg t	gt g				1344
Al a	Lys	Cys	Hi s	Lys	Leu	Val	G u	Ser	Phe	Cys	Lys	G u	Tr p	G y	Val				
		435				440						445							
a ag	t ac	cat	g ag	g ct	gat	at g	t gg	gat	gg t	acc	gt g	g aa	gt g	t t g	caa				1392
Lys	Tyr	Hi s	G u	Al a	Asp	Met	Tr p	Asp	G y	Thr	Val	G u	Val	Leu	G n				
	450					455					460								
cat	ct c	t cc	a ag	gt g	t cg	gat	gat	t t c	ct t	gt g	g ag	at g	gt g	a ag	gat				1440
Hi s	Leu	Ser	Lys	Val	Ser	Asp	Asp	Phe	Leu	Val	G u	Met	Val	Lys	Asp				480
465					470				475										
t t c	c ct	g cc	at g	t aa															1455
Phe	Pro	Al a	Met																

<210> 12  
 <211> 484  
 <212> PRT  
 <213> Thal assi osi r a pseudonana

<400> 12

PF58307.txt

Met Gly Lys Gly Gly Asp Ala Ala Ala Ala Thr Lys Arg Ser Gly Ala  
1 5 10 15  
Leu Lys Leu Ala Gu Lys Pro Gn Lys Tyr Thr Trp Gn Gu Val Lys  
20 25 30  
Lys His Ile Thr Pro Asp Asp Ala Trp Val Val His Gn Asn Lys Val  
35 40 45  
Tyr Asp Val Ser Asn Trp Tyr Asp His Pro Gly Gly Ala Val Val Phe  
50 55 60  
Thr His Ala Gly Asp Asp Met Thr Asp Ile Phe Ala Ala Phe His Ala  
65 70 75 80  
Gn Gly Ser Gn Ala Met Met Lys Lys Phe Tyr Ile Gly Asp Leu Ile  
85 90 95  
Pro Gu Ser Val Gu His Lys Asp Gn Arg Gn Leu Asp Phe Gu Lys  
100 105 110  
Gly Tyr Arg Asp Leu Arg Ala Lys Leu Val Met Met Gly Met Phe Lys  
115 120 125  
Ser Ser Lys Met Tyr Tyr Ala Tyr Lys Cys Ser Phe Asn Met Cys Met  
130 135 140  
Trp Leu Val Ala Val Ala Met Val Tyr Tyr Ser Asp Ser Leu Ala Met  
145 150 155 160  
His Ile Gly Ser Ala Leu Leu Leu Gly Leu Phe Trp Gn Gn Cys Gly  
165 170 175  
Trp Leu Ala His Asp Phe Leu His His Gn Val Phe Lys Gn Arg Lys  
180 185 190  
Tyr Gly Asp Leu Val Gly Ile Phe Trp Gly Asp Leu Met Gn Gly Phe  
195 200 205  
Ser Met Gn Trp Trp Lys Asn Lys His Asn Gly His Ala Val Pro  
210 215 220  
Asn Leu His Asn Ser Ser Leu Asp Ser Gn Asp Gly Asp Pro Asp Ile  
225 230 235 240  
Asp Thr Met Pro Leu Leu Ala Trp Ser Leu Lys Gn Ala Gn Ser Phe  
245 250 255  
Arg Gu Ile Asn Lys Gly Lys Asp Ser Thr Phe Val Lys Tyr Ala Ile  
260 265 270  
Lys Phe Gn Ala Phe Thr Tyr Phe Pro Ile Leu Leu Leu Ala Arg Ile  
275 280 285  
Ser Trp Leu Asn Gu Ser Phe Lys Thr Ala Phe Gly Leu Gly Ala Ala  
290 295 300  
Ser Gu Asn Ala Lys Leu Gu Leu Gu Lys Arg Gly Leu Gn Tyr Pro  
305 310 315 320  
Leu Leu Gu Lys Leu Gly Ile Thr Leu His Tyr Thr Trp Met Phe Val  
325 330 335  
Leu Ser Ser Gly Phe Gly Arg Trp Ser Leu Pro Tyr Ser Ile Met Tyr  
340 345 350  
Phe Phe Thr Ala Thr Cys Ser Ser Gly Leu Phe Leu Ala Leu Val Phe  
355 360 365  
Gly Leu Gly His Asn Gly Met Ser Val Tyr Asp Ala Thr Thr Arg Pro  
370 375 380  
Asp Phe Trp Gn Leu Gn Val Thr Thr Thr Arg Asn Ile Ile Gly Gly  
385 390 395 400  
His Gly Ile Pro Gn Phe Val Asp Trp Phe Cys Gly Gly Leu Gn  
405 410 415  
Tyr Gn Val Asp His His Leu Phe Pro Met Met Pro Arg Asn Asn Ile  
420 425 430  
Ala Lys Cys His Lys Leu Val Gu Ser Phe Cys Lys Gu Trp Gly Val  
435 440 445  
Lys Tyr His Gu Ala Asp Met Trp Asp Gly Thr Val Gu Val Leu Gn  
450 455 460  
His Leu Ser Lys Val Ser Asp Asp Phe Leu Val Gu Met Val Lys Asp  
465 470 475 480  
Phe Pro Ala Met

<210> 13  
<211> 1446  
<212> DNA  
<213> Marchantia polynorpha

<220>  
 <221> CDS  
 <222> (1).. (1446)

<400> 13  
 at g gcc t cg t cc acc acc acc gcc gt g aag caa t ct t cg ggt ggg ct g 48  
 Met Ala Ser Ser Thr Thr Thr Ala Val Lys Gl n Ser Ser Gly Gly Leu  
 1 5 10 15  
 t gg t cg aaa t gg ggc acc ggc agc aac t t g agc t t c gt g t cg t cg cgc aag 96  
 Tr p Ser Lys Tr p Gly Thr Gly Ser Asn Leu Ser Phe Val Ser Arg Lys  
 20 25 30  
 gag cag cag cag cag cag cag cag agc t ct ccc gag gcg t cg act ccc 144  
 Gu Gl n Gl n Gl n Gl n Gl n Gl n Gl n Ser Ser Pro Gu Ala Ser Thr Pro  
 35 40 45  
 gcg gcg cag cag gag aaa t cc at c agt aga gaa t cc at c ccc gag ggc 192  
 Ala Ala Gl n Gl n Gu Lys Ser Ile Ser Arg Gl u Ser Ile Pro Gu Gly  
 50 55 60  
 t t c t t g acc gt g gag gag gt g t cg aag cac gac aat ccg agc gac t gc 240  
 Phe Leu Thr Val Gu Gu Val Ser Lys His Asp Asn Pro Ser Asp Cys  
 65 70 75 80  
 t gg at c gt c at c aac gac aag gt g t ac gac gt g agc gca t t c ggg aag 288  
 Tr p Ile Val Ile Asn Asp Lys Val Tyr Asp Val Ser Ala Phe Gly Lys  
 85 90 95  
 acg cat ccg ggc ggc cct gt g at c t t c acg cag gcc ggc cgc gac gcc 336  
 Thr His Pro Gly Gly Pro Val Ile Phe Thr Gl n Ala Gly Arg Asp Ala  
 100 105 110  
 acg gat t ct t t c aag gt t t t c cac t cc gcc aag gcg t gg cag t t t ct c 384  
 Thr Asp Ser Phe Lys Val Phe His Ser Ala Lys Ala Tr p Gl n Phe Leu  
 115 120 125  
 cag gac ct g t ac at c gga gat ct g t ac aat gcc gag cca gt g t cg gag 432  
 Gl n Asp Leu Tyr Ile Gly Asp Leu Tyr Asn Ala Gl u Pro Val Ser Gu  
 130 135 140  
 ct g gt g aag gat t ac cga gac ct g agg acg gcg t t c at g cgt t ct cag 480  
 Leu Val Lys Asp Tyr Arg Asp Leu Arg Thr Ala Phe Met Arg Ser Gl n  
 145 150 155 160  
 ct a t t c aag agc agt aaa at g t ac t ac gt g acc aag t gc gt c aca aat 528  
 Leu Phe Lys Ser Ser Lys Met Tyr Tyr Val Thr Lys Cys Val Thr Asn  
 165 170 175  
 t t t gca att ct t gcc gcc agt ct c gca gt c at c gcg t gg agc cag acg 576  
 Phe Ala Ile Leu Ala Ala Ser Leu Ala Val Ile Ala Tr p Ser Gl n Thr  
 180 185 190  
 t at ct g gcg gtt t t g t gc t cc agt t t c ct g t t g gct ct c t t c t gg cag 624  
 Tyr Leu Ala Val Leu Cys Ser Ser Phe Leu Leu Ala Leu Phe Tr p Gl n  
 195 200 205  
 caa t gt gga t gg t t a t cg cac gat t t t ct c cac cac cag gt g acc gag 672  
 Gl n Cys Gly Tr p Leu Ser His Asp Phe Leu His His Gl n Val Thr Gu  
 210 215 220  
 aac cga t cg ct c aac acg t ac t t c ggc ggc ct g t t c t gg ggt aac t t c 720  
 Asn Arg Ser Leu Asn Thr Tyr Phe Gly Gly Leu Phe Tr p Gly Asn Phe  
 225 230 235 240  
 gcc cag ggc t ac agc gt g gga t gg t gg aag acc aag cac aat gt g cac 768  
 Ala Gl n Gly Tyr Ser Val Gly Tr p Tr p Lys Thr Lys His Asn Val His  
 245 250 255  
 cac gcg gcc acg aac gaa t gc gac gac aag t at cag ccc at c gat ccc 816  
 His Ala Ala Thr Asn Gl u Cys Asp Asp Lys Tyr Gl n Pro Ile Asp Pro  
 260 265 270  
 gac at c gac acc gt g ccc ct g ct c gcc t gg agc aag gaa at c t t g gcc 864  
 Asp Ile Asp Thr Val Pro Leu Leu Ala Tr p Ser Lys Gl u Ile Leu Ala  
 275 280 285  
 acc gt c gac gac caa t t c t t c cga t cg at c at c agc gt g cag cac ct t 912  
 Thr Val Asp Asp Gl n Phe Phe Arg Ser Ile Ile Ser Val Gl n His Leu  
 290 295 300  
 ct g t t c t t c ccg ct c ct c t t c t t g gca aga t t c agc t gg ct g cat t cg 960  
 Leu Phe Phe Pro Leu Leu Phe Leu Ala Arg Phe Ser Tr p Leu His Ser  
 305 310 315 320  
 agt t gg gcc cac gcc agc aac t t c gag at g cct cgg t ac at g aga t gg 1008  
 Ser Tr p Ala His Ala Ser Asn Phe Gu Met Pro Arg Tyr Met Arg Tr p  
 325 330 335

PF58307. txt

gcg gag aag gcc t cg ct c ct c ggg cac t ac ggc gcc t ca at c ggc gcc 1056  
 Al a Gu Lys Al a Ser Leu Leu Gly Hi s Tyr Gly Al a Ser Ile Gly Al a  
 340 345 350

gcc ttc tac att ttg ccc at c ccc cag gcc at c tgc tgg ct c ttc ttg 1104  
 Al a Phe Tyr Ile Leu Pro Ile Pro G n Al a Ile Cys Trp Leu Phe Leu  
 355 360 365

t cg caa ctg ttt tgc ggc gct ct g ct c agc att gt c ttc gt g at c agc 1152  
 Ser G n Leu Phe Cys Gly Al a Leu Leu Ser Ile Val Phe Val Ile Ser  
 370 375 380

cac aat ggc at g gat gt g t ac aac gac ccc cgg gac ttc gt g acg gcc 1200  
 Hi s Asn Gly Met Asp Val Tyr Asn Asp Pro Arg Asp Phe Val Thr Al a  
 385 390 395 400

caa gt c acc t cg acc aga aac at c gaa ggc aac ttc ttc aac gac tgg 1248  
 G n Val Thr Ser Thr Arg Asn Ile Gu Gly Asn Phe Phe Asn Asp Trp  
 405 410 415

ttc acc gga ggc ct g aac agg cag at t gag cac cat ct g ttt ccg tct 1296  
 Phe Thr Gly Gly Leu Asn Arg G n Ile Gu Hi s Hi s Leu Phe Pro Ser  
 420 425 430

ct t ccg agg cac aac ct c gcc aag gt c gcg cca cac gt c aag gcg ct c 1344  
 Leu Pro Arg Hi s Asn Leu Al a Lys Val Al a Pro Hi s Val Lys Al a Leu  
 435 440 445

tgc gcc aag cac ggt ttg cat t ac gaa gaa ttg agt ct g ggc acg gga 1392  
 Cys Al a Lys Hi s Gly Leu Hi s Tyr Gu Gu Leu Ser Leu Gly Thr Gly  
 450 455 460

gt c tgt cgt gt c ttc aat cgg ct a gt a gag gt a gca t ac gct gcg aaa 1440  
 Val Cys Arg Val Phe Asn Arg Leu Val Gu Val Al a Tyr Al a Al a Lys  
 465 470 475 480

gt a tag 1446  
 Val

<210> 14  
 <211> 481  
 <212> PRT  
 <213> Marchantia polymorpha

<400> 14  
 Met Al a Ser Ser Thr Thr Thr Al a Val Lys G n Ser Ser Gly Gly Leu  
 1 5 10 15  
 Trp Ser Lys Trp Gly Thr Gly Ser Asn Leu Ser Phe Val Ser Arg Lys  
 20 25 30  
 Gu G n G n G n G n G n G n G n Ser Ser Pro Gu Al a Ser Thr Pro  
 35 40 45  
 Al a Al a G n G n Gu Lys Ser Ile Ser Arg Gu Ser Ile Pro Gu Gly  
 50 55 60  
 Phe Leu Thr Val Gu Gu Val Ser Lys Hi s Asp Asn Pro Ser Asp Cys  
 65 70 75 80  
 Trp Ile Val Ile Asn Asp Lys Val Tyr Asp Val Ser Al a Phe Gly Lys  
 85 90 95  
 Thr Hi s Pro Gly Gly Pro Val Ile Phe Thr G n Al a Gly Arg Asp Al a  
 100 105 110  
 Thr Asp Ser Phe Lys Val Phe Hi s Ser Al a Lys Al a Trp G n Phe Leu  
 115 120 125  
 G n Asp Leu Tyr Ile Gly Asp Leu Tyr Asn Al a Gu Pro Val Ser Gu  
 130 135 140  
 Leu Val Lys Asp Tyr Arg Asp Leu Arg Thr Al a Phe Met Arg Ser G n  
 145 150 155 160  
 Leu Phe Lys Ser Ser Lys Met Tyr Tyr Val Thr Lys Cys Val Thr Asn  
 165 170 175  
 Phe Al a Ile Leu Al a Al a Ser Leu Al a Val Ile Al a Trp Ser G n Thr  
 180 185 190  
 Tyr Leu Al a Val Leu Cys Ser Ser Phe Leu Leu Al a Leu Phe Trp G n  
 195 200 205  
 G n Cys Gly Trp Leu Ser Hi s Asp Phe Leu Hi s Hi s G n Val Thr Gu  
 210 215 220  
 Asn Arg Ser Leu Asn Thr Tyr Phe Gly Gly Leu Phe Trp Gly Asn Phe  
 225 230 235 240  
 Al a G n Gly Tyr Ser Val Gly Trp Trp Lys Thr Lys Hi s Asn Val Hi s



PF58307. txt

245  
 H i s A l a A l a T h r A s n G u C y s A s p A s p L y s T y r G l n P r o I l e A s p P r o  
 250  
 A s p I l e A s p T h r V a l P r o L e u L e u A l a T r p S e r L y s G l u I l e L e u A l a  
 260  
 T h r V a l A s p A s p G l n P h e P h e A r g S e r I l e I l e S e r V a l G l n H i s L e u  
 275  
 L e u P h e P h e P r o L e u L e u P h e L e u A l a A r g P h e S e r T r p L e u H i s S e r  
 285  
 S e r T r p A l a H i s A l a S e r A s n P h e G l u M e t P r o A r g T y r M e t A r g T r p  
 290  
 A l a G l u L y s A l a S e r L e u L e u G l y H i s T y r G l y A l a S e r I l e G l y A l a  
 300  
 A l a P h e T y r I l e L e u P r o I l e P r o G l n A l a I l e C y s T r p L e u P h e L e u  
 315  
 S e r G l n L e u P h e C y s G l y A l a L e u L e u S e r I l e V a l P h e V a l I l e S e r  
 330  
 H i s A s n G l y M e t A s p V a l T y r A s n A s p P r o A r g A s p P h e V a l T h r A l a  
 345  
 G l n V a l T h r S e r T h r A r g A s n I l e G l u G l y A s n P h e P h e A s n A s p T r p  
 360  
 P h e T h r G l y G l y L e u A s n A r g G l n I l e G l u H i s H i s L e u P h e P r o S e r  
 375  
 L e u P r o A r g H i s A s n L e u A l a L y s V a l A l a P r o H i s V a l L y s A l a L e u  
 380  
 C y s A l a L y s H i s G l y L e u H i s T y r G l u G l u L e u S e r L e u G l y T h r G l y  
 395  
 V a l C y s A r g V a l P h e A s n A r g L e u V a l G l u V a l A l a T y r A l a A l a L y s  
 410  
 V a l  
 420  
 435  
 445  
 455  
 465  
 V a l  
 470  
 475  
 480

<210> 15  
 <211> 1377  
 <212> DNA  
 <213> Rhizopus oryzae

<220>  
 <221> CDS  
 <222> (1)..(1377)

<400> 15  
 a t g a g t a c a t c a g a t c g t c a a t c a g t t t t c a c a t t a a a g a g t t g g a a 48  
 M e t S e r T h r S e r A s p A r g G l n S e r V a l P h e T h r L e u L y s G l u L e u G l u  
 1  
 c t g a t t a a t c a a a a a c a t c g a g a t g g a g a t a a a a g t g c g a t g a a g t t c 96  
 L e u I l e A s n G l n L y s H i s A r g A s p G l y A s p L y s S e r A l a M e t L y s P h e  
 20  
 a t t a t c a t t g a t c g t a a g g t g t a c g a t g t g a c t g a a t t c t t g g a a g a t 144  
 I l e I l e I l e A s p A r g L y s V a l T y r A s p V a l T h r G l u P h e L e u G l u A s p  
 35  
 c a t c c t g g t g g t g c a c a a g t a t t a c t g a c a c a c g t t g g a a a a g a t g c a 192  
 H i s P r o G l y G l y A l a G l n V a l L e u L e u T h r H i s V a l G l y L y s A s p A l a  
 50  
 t c t g a t g t a t t t c a t g c c a t g c a t c c c g a g t c a g c a t a t g a a a t c t t g 240  
 S e r A s p V a l P h e H i s A l a M e t H i s P r o G l u S e r A l a T y r G l u I l e L e u  
 65  
 a a c a a t t a t t t t g t a g g a g a t g t a a a a g a t g c a c a t g t a a a g a g a c c 288  
 A s n A s n T y r P h e V a l G l y A s p V a l L y s A s p A l a H i s V a l L y s G l u T h r  
 85  
 c c t t c t g c t c a a t t t g c t t c a g a a a t g c g t c a a c t t c g g g a t c a a t t g 336  
 P r o S e r A l a G l n P h e A l a S e r G l u M e t A r g G l n L e u A r g A s p G l n L e u  
 100  
 a a a a a a g a a g g t t a t t t t c a t t c t a g t a a a g c t t a t t a t t a c a a g 384  
 L y s L y s G l u G l y T y r P h e H i s S e r A r g L y s A l a T y r T y r T y r V a l T y r L y s  
 115  
 g t c c t c t c t a c t c t t g c t c t t t g t g c a g c t g g t c t t a c t c t t t t g t a t 432  
 V a l L e u S e r T h r L e u A l a L e u C y s A l a A l a G l y L e u T h r L e u L e u T y r



PF58307. txt

130	gct	t at	gga	cat	aca	t ct	act	t t a	gct	g t t	g t t	gca	t ct	gct	a t t	a t t	480
145	Al a	Tyr	G l y	H i s	Thr	Ser	Thr	Leu	Al a	Val	Val	Al a	Ser	Al a	I l l e	I l l e	
150	g t t	g g t	a t c	t t t	t g g	caa	caa	t g t	g g t	t g g	t t g	gct	cat	gat	t t t	g g a	528
165	Val	G l y	I l l e	Phe	Tr p	G n	G n	Cys	G l y	Tr p	Leu	Al a	H i s	Asp	Phe	G l y	
180	cat	cat	caa	t gc	t t t	gaa	gat	cgc	t ct	t gg	aat	gat	g t t	ct t	g t t	g t t	576
190	H i s	H i s	G n	Cys	Phe	G u	Asp	Arg	Ser	Tr p	Asn	Asp	Val	Leu	Val	Val	
205	t t c	ct t	gga	aac	t t t	t gt	caa	ggt	t t t	t ca	t t g	t ca	t gg	t gg	aag	aat	624
210	Phe	Leu	G l y	Asn	Phe	Cys	G n	G l y	Phe	Ser	Leu	Ser	Tr p	Tr p	Lys	Asn	
225	aaa	cat	aac	act	cac	cac	gct	agt	acg	aat	g t t	cat	gga	cat	gat	ccc	672
230	Lys	H i s	Asn	Thr	H i s	H i s	Al a	Ser	Thr	Asn	Val	H i s	G l y	H i s	Asp	Pro	
240	gat	a t t	gac	act	gct	cct	gt c	t t a	t t a	t gg	gat	gaa	t at	gca	t ct	gca	720
255	Asp	I l l e	Asp	Thr	Al a	Pro	Val	Leu	Leu	Tr p	Asp	G u	Tyr	Al a	Ser	Al a	
270	gct	t at	t at	gcc	t ca	ct g	gat	gaa	gaa	ccc	aca	at g	at t	t ct	cga	t t c	768
285	Al a	Tyr	Tyr	Al a	Ser	Leu	Asp	G u	G u	Pro	Thr	Met	I l l e	Ser	Arg	Phe	
300	ct t	gct	gaa	agt	gt c	t t g	cct	cat	caa	act	cgt	t at	t ac	t t c	t t t	gt t	816
315	Leu	Al a	G u	Ser	Val	Leu	Pro	H i s	G n	Thr	Arg	Tyr	Tyr	Phe	Phe	Val	
330	ct c	ggc	t t t	gct	cgt	t t a	t ca	t gg	gcc	at c	caa	t cc	t t a	ct c	t ac	t ca	864
345	Leu	G l y	Phe	Al a	Arg	Leu	Ser	Tr p	Al a	I l l e	G n	Ser	Leu	Leu	Tyr	Ser	
360	t t c	aaa	caa	ggt	gct	at t	aac	aag	t ct	cat	caa	ct c	aac	ct c	t t t	gaa	912
375	Phe	Lys	G n	G l y	Al a	I l l e	Asn	Lys	Ser	H i s	G n	Leu	Asn	Leu	Phe	G u	
390	cgc	t t t	tgt	ct t	gt t	agt	cac	t gg	act	t t a	t t c	act	t ac	tgt	aca	ct t	960
405	Arg	Phe	Cys	Leu	Val	Ser	H i s	Tr p	Thr	Leu	Phe	Thr	Tyr	Cys	Thr	Leu	
420	gcc	t gg	tgt	agc	aac	gt c	t at	cac	at g	at t	t t g	t t c	t t t	t t g	gt t	agt	1008
435	Al a	Tr p	Cys	Ser	Asn	Val	Tyr	H i s	Met	I l l e	Leu	Phe	Phe	Leu	Val	Ser	
450	caa	gca	act	acc	ggt	t ac	aca	t t g	gcc	ct t	gt a	t t t	gct	t t g	aat	cac	1056
465	G n	Al a	Thr	Thr	G l y	Tyr	Thr	Leu	Al a	Leu	Val	Phe	Al a	Leu	Asn	H i s	
480	aat	ggt	at g	cct	gt g	at t	act	gaa	gaa	aag	gcc	gag	t ca	at g	gaa	t t t	1104
495	Asn	G l y	Met	Pro	Val	I l l e	Thr	G u	G u	Lys	Al a	G u	Ser	Met	G u	Phe	
510	t t c	gaa	at t	caa	gt g	at t	aca	ggt	cgt	gat	gt a	aca	ct c	t ct	cct	t t a	1152
525	Phe	G u	I l l e	G n	Val	I l l e	Thr	G l y	Arg	Asp	Val	Thr	Leu	Ser	Pro	Leu	
540	ggt	gat	t gg	t t c	at g	ggt	gga	t t g	aac	t at	caa	at c	gag	cat	cat	gt t	1200
555	G l y	Asp	Tr p	Phe	Met	G l y	G l y	Leu	Asn	Tyr	G n	I l l e	G u	H i s	H i s	Val	
570	t t c	cct	aat	at g	cct	cgt	cat	aat	t t a	cct	aaa	gt a	aag	cca	at g	gt c	1248
585	Phe	Pro	Asn	Met	Pro	Arg	H i s	Asn	Leu	Pro	Lys	Val	Lys	Pro	Met	Val	
600	aag	t ca	ct t	tgt	aag	aaa	t at	gat	at t	aat	t at	cac	gat	act	gga	t t c	1296
615	Lys	Ser	Leu	Cys	Lys	Lys	Tyr	Asp	I l l e	Asn	Tyr	H i s	Asp	Thr	G l y	Phe	
630	t t g	aaa	gga	aca	t t g	gaa	gt a	ct a	aag	aca	t t a	gat	at t	act	t cc	aag	1344
645	Leu	Lys	G l y	Thr	Leu	G u	Val	Leu	Lys	Thr	Leu	Asp	I l l e	Thr	Ser	Lys	
660	t t g	t ct	t t g	caa	t t g	agc	aaa	aag	t ca	t t t	t aa						1377
675	Leu	Ser	Leu	G n	Leu	Ser	Lys	Lys	Ser	Phe							

<210> 16  
 <211> 458  
 <212> PRT  
 <213> Rhi zopus oryzae

<400> 16

PF58307. txt

Met Ser Thr Ser Asp Arg Gln Ser Val Phe Thr Leu Lys Glu Leu Glu  
1 5 10 15  
Leu Ile Asn Gln Lys His Arg Asp Gly Asp Lys Ser Ala Met Lys Phe  
20 25 30  
Ile Ile Ile Asp Arg Lys Val Tyr Thr Glu Phe Leu Glu Asp  
35 40 45  
His Pro Gly Gly Ala Gln Val Leu Leu Thr His Val Gly Lys Asp Ala  
50 55 60  
Ser Asp Val Phe His Ala Met His Pro Glu Ser Ala Tyr Glu Ile Leu  
65 70 75 80  
Asn Asn Tyr Phe Val Gly Asp Val Lys Asp Ala His Val Lys Glu Thr  
85 90 95  
Pro Ser Ala Gln Phe Ala Ser Glu Met Arg Gln Leu Arg Asp Gln Leu  
100 105 110  
Lys Lys Glu Gly Tyr Phe His Ser Ser Lys Ala Tyr Tyr Val Tyr Lys  
115 120 125  
Val Leu Ser Thr Leu Ala Leu Cys Ala Ala Gly Leu Thr Leu Leu Tyr  
130 135 140  
Ala Tyr Gly His Thr Ser Thr Leu Ala Val Val Ala Ser Ala Ile Ile  
145 150 155 160  
Val Gly Ile Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Gly  
165 170 175  
His His Gln Cys Phe Glu Asp Arg Ser Trp Asn Asp Val Leu Val Val  
180 185 190  
Phe Leu Gly Asn Phe Cys Gln Gly Phe Ser Leu Ser Trp Trp Lys Asn  
195 200 205  
Lys His Asn Thr His His Ala Ser Thr Asn Val His Gly His Asp Pro  
210 215 220  
Asp Ile Asp Thr Ala Pro Val Leu Leu Trp Asp Glu Tyr Ala Ser Ala  
225 230 235 240  
Ala Tyr Tyr Ala Ser Leu Asp Glu Glu Pro Thr Met Ile Ser Arg Phe  
245 250 255  
Leu Ala Glu Ser Val Leu Pro His Gln Thr Arg Tyr Tyr Phe Phe Val  
260 265 270  
Leu Gly Phe Ala Arg Leu Ser Trp Ala Ile Gln Ser Leu Leu Tyr Ser  
275 280 285  
Phe Lys Gln Gly Ala Ile Asn Lys Ser His Gln Leu Asn Leu Phe Glu  
290 295 300  
Arg Phe Cys Leu Val Ser His Trp Thr Leu Phe Thr Tyr Cys Thr Leu  
305 310 315 320  
Ala Trp Cys Ser Asn Val Tyr His Met Ile Leu Phe Phe Leu Val Ser  
325 330 335  
Gln Ala Thr Thr Gly Tyr Thr Leu Ala Leu Val Phe Ala Leu Asn His  
340 345 350  
Asn Gly Met Pro Val Ile Thr Glu Glu Lys Ala Glu Ser Met Glu Phe  
355 360 365  
Phe Glu Ile Gln Val Ile Thr Gly Arg Asp Val Thr Leu Ser Pro Leu  
370 375 380  
Gly Asp Trp Phe Met Gly Gly Leu Asn Tyr Gln Ile Glu His His Val  
385 390 395 400  
Phe Pro Asn Met Pro Arg His Asn Leu Pro Lys Val Lys Pro Met Val  
405 410 415  
Lys Ser Leu Cys Lys Lys Tyr Asp Ile Asn Tyr His Asp Thr Gly Phe  
420 425 430  
Leu Lys Gly Thr Leu Glu Val Leu Lys Thr Leu Asp Ile Thr Ser Lys  
435 440 445  
Leu Ser Leu Gln Leu Ser Lys Lys Ser Phe  
450 455

<210> 17  
<211> 1404  
<212> DNA  
<213> Mucor rouxi i

<220>  
<221> CDS  
<222> (1).. (1404)

PF58307. txt

<400> 17

at g	agc	agc	gac	gt a	gga	gca	aca	gt a	ccc	cat	ttt	t at	act	cgt	gct	48
Met	Ser	Ser	Asp	Val	G y	Al a	Thr	Val	Pro	Hi s	Phe	Tyr	Thr	Arg	Al a	
1			5					10						15		
gaa	ttg	gct	gat	at c	cat	caa	gat	gt t	ct g	gac	aag	aaa	cct	gaa	gct	96
G u	Leu	Al a	Asp	l l e	Hi s	G n	Asp	Val	Leu	Asp	Lys	Lys	Pro	G u	Al a	
			20					25					30			
aga	aaa	ct a	att	gt c	gt c	gag	aat	aag	gt a	t ac	gat	at c	aca	gat	tt c	144
Arg	Lys	Leu	l l e	Val	Val	G u	Asn	Lys	Val	Tyr	Asp	l l e	Thr	Asp	Phe	
		35					40					45				
gt a	ttt	gat	cac	cct	ggc	ggt	gag	cga	gt a	ct t	ttg	act	caa	gag	ggc	192
Val	Phe	Asp	Hi s	Pro	G y	G y	G u	Arg	Val	Leu	Leu	Thr	G n	G u	G y	
	50				55						60					
aga	gac	gct	aca	gat	gt c	ttt	cat	gaa	at g	cat	ccc	t ct	gct	t ac		240
Arg	Asp	Al a	Thr	Asp	Val	Phe	Hi s	G u	Met	Hi s	Pro	Pro	Ser	Al a	Tyr	
65				70				75						80		
gaa	ttg	ct g	gcg	aat	t gc	t at	gt t	ggc	gat	t gt	gag	ccc	aag	ct g	cct	288
G u	Leu	Leu	Al a	Asn	Cys	Tyr	Val	G y	Asp	Cys	G u	Pro	Lys	Leu	Pro	
				85				90						95		
at c	gat	t ca	acc	gat	aag	aag	gca	ttg	aat	t ct	gct	gct	ttt	gct	caa	336
l l e	Asp	Ser	Thr	Asp	Lys	Lys	Al a	Leu	Asn	Ser	Al a	Al a	Phe	Al a	G n	
			100					105					110			
gaa	att	cgt	gat	ct c	cga	gat	aaa	tt a	gaa	aaa	caa	ggc	t at	tt c	gac	384
G u	l l e	Arg	Asp	Leu	Arg	Asp	Lys	Leu	G u	Lys	G n	G y	Tyr	Phe	Asp	
		115					120					125				
gct	agt	act	ggt	ttc	t ac	at c	t at	aaa	gt t	t cc	act	acc	ct a	ct t	gt c	432
Al a	Ser	Thr	G y	Phe	Tyr	l l e	Tyr	Lys	Val	Ser	Thr	Thr	Leu	Leu	Val	
		130				135					140					
t gt	att	gt a	ggt	tt a	gct	at c	ct c	aaa	gct	t gg	ggt	aga	gag	t ct	act	480
Cys	l l e	Val	G y	Leu	Al a	l l e	Leu	Lys	Al a	Trp	G y	Arg	G u	Ser	Thr	
145				150						155				160		
ttg	gct	gt g	ttt	att	gct	t ct	tt a	gt t	gg t	ct t	ttt	t gg	cag	cag		528
Leu	Al a	Val	Phe	l l e	Al a	Al a	Ser	Leu	Val	G y	Leu	Phe	Trp	G n	G n	
			165					170					175			
t gc	ggt	t gg	ct t	gcc	cat	gat	t at	gct	cac	t at	caa	gt c	at c	aaa	gac	576
Cys	G y	Trp	Leu	Al a	Hi s	Asp	Tyr	Al a	Hi s	Tyr	G n	Val	l l e	Lys	Asp	
			180					185					190			
ccc	aat	gt c	aat	aat	ct c	ttt	tt a	gt t	act	ttt	ggc	aac	ct g	gt t	caa	624
Pro	Asn	Val	Asn	Asn	Leu	Phe	Leu	Val	Thr	Phe	G y	Asn	Leu	Val	G n	
		195					200				205					
ggt	ttc	t ct	ct t	t ca	t gg	t gg	aaa	aac	aag	cac	aat	acc	cac	cac	gct	672
G y	Phe	Ser	Leu	Ser	Trp	Trp	Lys	Asn	Lys	Hi s	Asn	Thr	Hi s	Hi s	Al a	
		210				215					220					
agc	acc	aat	gt c	t ct	ggc	gaa	gat	ccc	gat	at c	gat	act	gcc	ccc	at t	720
Ser	Thr	Asn	Val	Ser	G y	G u	Asp	Pro	Asp	l l e	Asp	Thr	Al a	Pro	l l e	
225					230					235				240		
ttg	ct a	t gg	gac	gag	ttt	gca	gt c	gct	aat	tt c	t at	gga	t cc	tt g	aag	768
Leu	Leu	Trp	Asp	G u	Phe	Al a	Val	Al a	Asn	Phe	Tyr	G y	Ser	Leu	Lys	
			245							250			255			
gac	aat	gcc	agt	gga	ttt	gac	aga	ttc	att	gct	gaa	cac	att	ttg	ccc	816
Asp	Asn	Al a	Ser	G y	Phe	Asp	Arg	Phe	l l e	Al a	G u	Hi s	l l e	Leu	Pro	
			260					265					270			
t at	caa	act	cgt	t ac	t ac	ttc	ttc	att	ct t	ggt	ttc	gct	cgt	acc	t ct	864
Tyr	G n	Thr	Arg	Tyr	Tyr	Phe	Phe	l l e	Leu	G y	Phe	Al a	Arg	Thr	Ser	
			275					280					285			
t gg	gct	at c	caa	t ct	att	at c	t at	t cc	ttc	aag	aac	gaa	aca	ttg	aat	912
Trp	Al a	l l e	G n	Ser	l l e	l l e	Tyr	Ser	Phe	Lys	Asn	G u	Thr	Leu	Asn	
		290				295					300					
aaa	t ca	aag	ct c	ttg	t cc	t gg	t gt	gag	cgc	at c	ttt	ttg	att	gt t	cat	960
Lys	Ser	Lys	Leu	Leu	Ser	Trp	Cys	G u	Arg	l l e	Phe	Leu	l l e	Val	Hi s	
305					310					315					320	
t gg	gt c	ttt	ttc	act	t ac	t gc	act	at t	gcc	t gg	at c	agc	t ct	at c	aga	1008
Trp	Val	Phe	Phe	Thr	Tyr	Cys	Thr	l l e	Al a	Trp	l l e	Ser	Ser	l l e	Arg	
				325						330				335		
aac	at c	gcc	at g	ttc	ttt	gt t	gt t	agc	caa	at c	act	act	ggt	t ac	ttg	1056
Asn	l l e	Al a	Met	Phe	Phe	Val	Val	Ser	G n	l l e	Thr	Thr	G y	Tyr	Leu	
			340					345					350			
ct c	gcc	at c	gt c	ttt	gct	at g	aac	cat	aat	ggc	at g	cct	gt t	t ac	agc	1104

Seite 21

PF58307.txt

Leu Ala Ile Val Phe Ala Met Asn His Asn Gly Met Pro Val Tyr Ser  
 355 360 365  
 ccc gaa gaa gca aac cat acc gag ttt tat gaa ttg cag tgt atc act 1152  
 Pro Gu Gu Ala Asn His Thr Gu Phe Tyr Gu Leu Gn Oys Ile Thr  
 370 375 380  
 ggt cgc gat gtc aac tgc act gta ttt ggc gat tgg ct c at g ggt gga 1200  
 Gly Arg Asp Val Asn Cys Thr Val Phe Gly Asp Trp Leu Met Gly Gly 400  
 385  
 ttg aat tat caa att gag cac cat ct t ttc ccc gaa at g cct cga cat 1248  
 Leu Asn Tyr Gn Ile Gu His His Leu Phe Pro Gu Met Pro Arg His  
 405 410 415  
 cat tta tcc aag gtg aaa tcc at g gt c aaa ccc at t gct caa aag tat 1296  
 His Leu Ser Lys Val Lys Ser Met Val Lys Pro Ile Ala Gn Lys Tyr 430  
 420 425 430  
 aat atc cct tac cat gat acc aca gt c at t ggt ggt acc at t gaa gt c 1344  
 Asn Ile Pro Tyr His Asp Thr Thr Val Ile Gly Gly Thr Ile Gu Val 445  
 435 440 445  
 ttg caa acc ttg gat ttt gtt caa aag at t t cg cag aaa ttc agc aaa 1392  
 Leu Gn Thr Leu Asp Phe Val Gn Lys Ile Ser Gn Lys Phe Ser Lys 450 460  
 455  
 aaa at g ct c taa 1404  
 Lys Met Leu  
 465

<210> 18  
 <211> 467  
 <212> PRT  
 <213> Mucor rouxi i

<400> 18  
 Met Ser Ser Asp Val Gly Ala Thr Val Pro His Phe Tyr Thr Arg Ala  
 1 5 10 15  
 Gu Leu Ala Asp Ile His Gn Asp Val Leu Asp Lys Lys Pro Gu Ala  
 20 25 30  
 Arg Lys Leu Ile Val Val Gu Asn Lys Val Tyr Asp Ile Thr Asp Phe  
 35 40 45  
 Val Phe Asp His Pro Gly Gly Gu Arg Val Leu Leu Thr Gn Gu Gly  
 50 55 60  
 Arg Asp Ala Thr Asp Val Phe His Gu Met His Pro Pro Ser Ala Tyr  
 65 70 75 80  
 Gu Leu Leu Ala Asn Cys Tyr Val Gly Asp Cys Gu Pro Lys Leu Pro  
 85 90 95  
 Ile Asp Ser Thr Asp Lys Lys Ala Leu Asn Ser Ala Ala Phe Ala Gn  
 100 105 110  
 Gu Ile Arg Asp Leu Arg Asp Lys Leu Gu Lys Gn Gly Tyr Phe Asp  
 115 120 125  
 Ala Ser Thr Gly Phe Tyr Ile Tyr Lys Val Ser Thr Thr Leu Leu Val  
 130 135 140  
 Cys Ile Val Gly Leu Ala Ile Leu Lys Ala Trp Gly Arg Gu Ser Thr  
 145 150 155 160  
 Leu Ala Val Phe Ile Ala Ala Ser Leu Val Gly Leu Phe Trp Gn Gn  
 165 170 175  
 Cys Gly Trp Leu Ala His Asp Tyr Ala His Tyr Gn Val Ile Lys Asp  
 180 185 190  
 Pro Asn Val Asn Asn Leu Phe Leu Val Thr Phe Gly Asn Thr Leu Val Gn  
 195 200 205  
 Gly Phe Ser Leu Ser Trp Trp Lys Asn Lys His Asn Thr His His Ala  
 210 215 220  
 Ser Thr Asn Val Ser Gly Gu Asp Pro Asp Ile Asp Thr Ala Pro Ile  
 225 230 235 240  
 Leu Leu Trp Asp Gu Phe Ala Val Ala Asn Phe Tyr Gly Ser Leu Lys  
 245 250 255  
 Asp Asn Ala Ser Gly Phe Asp Arg Phe Ile Ala Gu His Ile Leu Pro  
 260 265 270  
 Tyr Gn Thr Arg Tyr Tyr Phe Phe Ile Leu Gly Phe Ala Arg Thr Ser  
 275 280 285  
 Trp Ala Ile Gn Ser Ile Ile Tyr Ser Phe Lys Asn Gu Thr Leu Asn  
 290 295 300

PF58307.txt

Lys Ser Lys Leu Leu Ser Trp Cys G u Arg Ile Phe Leu Ile Val His  
 305 310 315 320  
 Trp Val Phe Phe Thr Tyr Cys Thr Ile Ala Trp Ile Ser Ser Ile Arg  
 325 330 335  
 Asn Ile Ala Met Phe Phe Val Val Ser Ile Thr Thr Gly Tyr Leu  
 340 345 350  
 Leu Ala Ile Val Phe Ala Met Asn His Asn Gly Met Pro Val Tyr Ser  
 355 360 365  
 Pro G u G u Ala Asn His Thr G u Phe Tyr G u Leu G n Cys Ile Thr  
 370 375 380  
 G y Arg Asp Val Asn Cys Thr Val Phe G y Asp Trp Leu Met G y G y  
 385 390 395 400  
 Leu Asn Tyr G n Ile G u His His Leu Phe Pro G u Met Pro Arg His  
 405 410 415  
 His Leu Ser Lys Val Lys Ser Met Val Lys Pro Ile Ala G n Lys Tyr  
 420 425 430  
 Asn Ile Pro Tyr His Asp Thr Thr Val Ile Gly Gly Thr Ile G u Val  
 435 440 445  
 Leu G n Thr Leu Asp Phe Val G n Lys Ile Ser G n Lys Phe Ser Lys  
 450 455 460  
 Lys Met Leu  
 465

<210> 19  
 <211> 1374  
 <212> DNA  
 <213> Mbrt ierell a alpi na

<220>  
 <221> CDS  
 <222> (1).. (1374)

<400> 19  
 at g gct gct gct ccc agt gt g agg acg ttt act cgg gcc gag att ttg 48  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala G u Ile Leu  
 1 5 10 15  
 aat gcc gag gct ct g aat gag ggc aag aag gat gcc gag gca ccc ttc 96  
 Asn Ala G u Ala Leu Asn G u G y Lys Lys Asp Ala G u Ala Pro Phe  
 20 25 30  
 ct g at g at c at c gac aac aag gt g t ac gat gt c cgc gag ttc gt c cct 144  
 Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg G u Phe Val Pro  
 35 40 45  
 gat cat ccc ggt gga agt gt g at t ct c acg cac gt t ggc aag gac ggc 192  
 Asp His Pro G y G y Ser Val Ile Leu Thr His Val G y Lys Asp G y  
 50 55 60  
 act gac gt c ttt gac act ttt cac ccc gag gct gct tgg gag act ctt 240  
 Thr Asp Val Phe Asp Thr Phe His Pro G u Ala Ala Trp G u Thr Leu  
 65 70 75 80  
 gcc aac ttt t ac gtt ggt gat att gac gag agc gac cgc gat at c aag 288  
 Ala Asn Phe Tyr Val G y Asp Ile Asp G u Ser Asp Arg Asp Ile Lys  
 85 90 95  
 aat gat gac ttt gcg gcc gag gt c cgc aag ct g cgt acc ttg ttc cag 336  
 Asn Asp Asp Phe Ala Ala G u Val Arg Lys Leu Arg Thr Leu Phe G n  
 100 105 110  
 t ct ctt ggt t ac t ac gat t ct t cc aag gca t ac t ac gcc ttc aag gt c 384  
 Ser Leu G y Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val  
 115 120 125  
 t cg ttc aac ct c tgc at c tgg ggt ttg t cg acg gt c att gt g gcc aag 432  
 Ser Phe Asn Leu Cys Ile Trp G y Leu Ser Thr Val Ile Val Ala Lys  
 130 135 140  
 t gg ggc cag acc t cg acc ct c gcc aac gt g ct c t cg gct gcg ctt ttg 480  
 Trp G y G n Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu  
 145 150 155 160  
 ggt ct g ttc tgg cag ct g gga tgg ttg gct cac gac ttt ttg cat 528  
 G y Leu Phe Trp G n G n Cys G y Trp Leu Ala His Asp Phe Leu His  
 165 170 175  
 cac cag gt c ttc cag gac cgt ttc tgg ggt gat ctt ttc ggc gcc ttc 576  
 His G n Val Phe G n Asp Arg Phe Trp G y Asp Leu Phe G y Ala Phe

PF58307. txt

			180					185					190				
t t g	g g a	g g t	g t c	t g c	c a g	g g c	t t c	t c g	t c c	t c g	t g g	t g g	a a g	g a c	a a g		624
Leu	Gly	Gly	Val	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	Lys		
		195					200					205					
c a c	a a c	a c t	c a c	c a c	g c c	g c c	c c c	a a c	g t c	c a c	g g c	g a a	g a t	c c c	g a c		672
His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp		
	210					215					220						
a t t	g a c	a c c	c a c	c c t	c t g	t t g	a c c	t g g	a g t	g a g	c a t	g c g	t t g	g a g	a t g		720
Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met		
	225				230					235				240			
t t c	t c g	g a t	g t c	c c a	g a t	g a g	g a g	c t g	a c c	c g c	a t g	t g g	t c g	c g t	t t c		768
Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe		
				245				250						255			
a t g	g t c	c t g	a a c	c a g	a c c	t g g	t t t	t a c	t t c	c c c	a t t	c t c	t c g	t t t	g c c		816
Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala		
			260				265						270				
c g t	c t c	t c c	t g g	t g c	c t c	c a g	t c c	a t t	c t c	t t t	g t g	c t g	c c t	a a c	g g t		864
Arg	Leu	Ser	Trp	Oys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	Gly		
		275					280						285				
c a g	g c c	c a c	a a g	c c c	t c g	g g c	g c g	c g t	g t g	c c c	a t c	t c g	t t g	g t c	g a g		912
Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu		
	290					295				300							
c a g	c t g	t c g	c t t	g c g	a t g	c a c	t g g	a c c	t g g	t a c	c t c	g c c	a c c	a t g	t t c		960
Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe		
	305				310					315					320		
c t g	t t c	a t c	a a g	g a t	c c c	g t c	a a c	a t g	c t g	g t g	t a c	t t t	t t g	g t g	t c g		1008
Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Leu	Val	Tyr	Phe	Leu	Val	Ser		
				325				330						335			
c a g	c c g	g t g	t g c	g g a	a a c	t t g	t t g	g c g	a t c	g t g	t t c	t c g	c t c	a a c	c a c		1056
Gln	Pro	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His		
			340					345					350				
a a c	g g t	a t g	c c t	g t g	a t c	t c g	a a g	g a g	g c g	g t c	g a t	a t g	g a t	t t c			1104
Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	Met	Asp	Phe		
		355					360				365						
t t c	a c g	a a g	c a g	a t c	a t c	a c g	g g t	c g t	g a t	g t c	c a c	c c g	g g t	c t a	t t t		1152
Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	His	Pro	Gly	Leu	Phe		
	370					375					380						
g c c	a a c	t g g	t t c	a c g	g g t	g g a	t t g	a a c	t a t	c a g	a t c	g a g	c a c	c a c	t t g		1200
Ala	Asn	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	His	Leu		
	385				390				395						400		
t t c	c c t	t c g	a t g	c c t	c g c	c a c	a a c	t t t	t c a	a a g	a t c	c a g	c c t	g c t	g t c		1248
Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	Gln	Pro	Ala	Val		
				405				410						415			
g a g	a c c	c t g	t g c	a a a	a a g	t a c	a a t	g t c	c g a	t a c	c a c	a c c	a c c	g g t	a t g		1296
Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Asn	Val	Arg	Tyr	His	Thr	Thr	Gly	Met		
			420					425					430				
a t c	g a g	g g a	a c t	g c a	g a g	g t c	t t t	a g c	c g t	c t g	a a c	g a g	g t c	t c c	a a g		1344
Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	Val	Ser	Lys		
		435					440					445					
g c t	g c c	t c c	a a g	a t g	g g c	a a g	g c a	c a g	t a a								1374
Ala	Ala	Ser	Lys	Met	Gly	Lys	Ala	Gln									
	450				455												

<210> 20  
 <211> 457  
 <212> PRT  
 <213> Mbrtierella alpina

<400> 20  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu Ile Leu  
 1 5 10 15  
 Asn Ala Glu Ala Leu Asn Glu Gly Lys Lys Asp Ala Glu Ala Pro Phe  
 20 25 30  
 Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe Val Pro  
 35 40 45  
 Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly  
 50 55 60  
 Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu Thr Leu  
 Seite 24

PF58307.txt

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

<210> 21  
<211> 1377  
<212> DNA  
<213> Mbrt ierel la al pi na

<220>  
<221> CDS  
<222> (1)..(1377)

<400> 21  
atg acc acc agc gac cca tct gtc aga gcg ttc aca cgc tca gaa gtg 48  
Met Thr Thr Ser Asp Pro Ser Val Arg Ala Phe Thr Arg Ser Gu Val  
1 5 10 15  
ttg cac gcc gat gcc ttg aac gag ggc aaa aag aac gcc gag gca ccg 96  
Leu His Ala Asp Ala Leu Asn Gu Gy Lys Lys Asn Ala Gu Ala Pro  
20 25 30  
ttt ctg atg atc atc gac aac aag gtc tac gat gtg cgc gag ttt atc 144  
Phe Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Gu Phe Ile  
45 50 55 60 65 70 75 80  
Sei te 25



PF58307.txt

35			40			45										
ccc	gac	cat	cct	ggt	ggg	agc	gtc	att	t t g	acc	cac	gt a	ggc	aag	gac	192
Pro	Asp	His	Pro	Gly	Gly	Ser	Val	Ile	Leu	Thr	His	Val	Gly	Lys	Asp	
	50					55					60					
ggc	acc	gac	gtt	ttc	gag	acc	ttc	cat	cct	gag	gct	gct	tgg	gag	acg	240
Gly	Thr	Asp	Val	Phe	Glu	Thr	Phe	His	Pro	Glu	Ala	Ala	Trp	Glu	Thr	
65					70					75					80	
ctc	gcc	aat	ttt	tat	gtc	ggt	gac	att	gt a	gaa	tcc	gat	cgc	gcc	atc	288
Leu	Ala	Asn	Phe	Tyr	Val	Gly	Asp	Ile	Val	Glu	Ser	Asp	Arg	Ala	Ile	
				85					90					95		
gag	aac	gac	gag	ttt	gca	gct	gag	gtt	cgt	aag	ctg	cgg	aca	ttg	ttt	336
Glu	Asn	Asp	Glu	Phe	Ala	Ala	Glu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	
			100					105					110			
tat	tct	t t g	ggc	tac	tac	gac	tca	tcc	aag	gtt	tac	tac	gcc	ttc	aag	384
Tyr	Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Val	Tyr	Tyr	Ala	Phe	Lys	
		115					120					125				
gtc	tcg	ttc	aac	ctc	tgc	atc	tgg	ggc	ctg	tct	gca	ttc	att	gtt	gcc	432
Val	Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Ala	Phe	Ile	Val	Ala	
	130					135					140					
aaa	tgg	ggc	cag	acc	tcg	acc	ctc	gca	aac	gtg	ata	tca	gcc	tca	ctc	480
Lys	Trp	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Val	Ile	Ser	Ala	Ser	Leu	
145					150				155						160	
ctg	ggt	gtc	ttt	tgg	caa	cag	tgc	ggt	tgg	ctc	gcc	cat	gat	ttc	ttg	528
Leu	Gly	Val	Phe	Trp	Gln	Gln	Cys	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	
			165					170						175		
cac	cat	cag	gtc	ttt	cac	gat	cga	ttc	tgg	ggc	gat	ctg	ttc	ggt	gca	576
His	His	Gln	Val	Phe	His	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	Ala	
			180					185					190			
ttt	ctc	ggc	gga	gtc	tgt	caa	ggt	ttc	tcc	tcg	tcc	tgg	tgg	aag	gac	624
Phe	Leu	Gly	Gly	Val	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	
	195						200					205				
aaa	cac	aac	acc	cac	cac	gcg	gcg	ccc	aat	gtc	cat	gga	gag	gat	ccc	672
Lys	His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	
	210					215					220					
gat	atc	gac	aca	cat	ccg	ctt	t t g	acg	tgg	agt	gag	cat	gcg	ctc	gag	720
Asp	Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	
225					230					235				240		
atg	ttt	tcg	gat	gtg	ccc	gat	gag	gag	ctt	acc	caa	atg	tgg	tcc	cgg	768
Met	Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Gln	Met	Trp	Ser	Arg	
			245					250					255			
ttt	atg	gtt	ctg	aac	cag	gcc	tgg	ttt	tac	ttt	ccc	att	ctg	tca	ttt	816
Phe	Met	Val	Leu	Asn	Gln	Ala	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	
			260					265					270			
gcc	cgc	ctg	tcc	tgg	tgc	atc	cag	tcg	att	ctt	ttt	gtg	cta	ccg	aac	864
Ala	Arg	Leu	Ser	Trp	Cys	Ile	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	
		275					280					285				
gga	cag	gca	cac	aaa	cct	gcg	ggg	gct	cgg	gtt	ccc	atc	tcg	ctg	gtg	912
Gly	Gln	Ala	His	Lys	Pro	Ala	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	
290				295							300					
gag	caa	t t g	tcg	t t g	gcg	atg	cac	tgg	acc	tgg	tac	ctg	gca	acc	atg	960
Glu	Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	
305				310						315					320	
ttc	ctg	ttc	atc	aag	gat	ccc	gtc	aac	atg	atg	gtg	tat	ttc	t t g	gtc	1008
Phe	Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Met	Val	Tyr	Phe	Leu	Val	
				325				330						335		
tcg	caa	gct	gtc	tgc	ggc	aac	ctg	tta	gcg	att	gtg	ttc	tcg	ctg	aac	1056
Ser	Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	
		340						345					350			
cat	aac	ggt	atg	cct	gtg	atc	tcg	cag	gag	gaa	gcg	gtc	gag	atg	gat	1104
His	Asn	Gly	Met	Pro	Val	Ile	Ser	Gln	Glu	Glu	Ala	Val	Glu	Met	Asp	
		355					360					365				
ttc	ttc	aca	aag	cag	atc	att	acg	ggt	cgt	gat	gtc	tac	ccg	ggt	tgg	1152
Phe	Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	Tyr	Pro	Gly	Trp	
		370					375					380				
ttt	gca	gac	tgg	ttc	acg	ggt	gga	t t g	aac	tat	cag	att	gaa	cac	cat	1200
Phe	Ala	Asp	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	His	
385					390					395					400	
ctg	ttc	ccg	tcg	atg	cct	cga	cac	cat	ttc	tca	aag	atc	cag	ccc	gcg	1248



PF58307. txt

Leu	Phe	Pro	Ser	Met 405	Pro	Arg	His	His	Phe 410	Ser	Lys	Ile	Gln	Pro 415	Ala		
gtt	gaa	tcg	ctg	tgc	aag	aag	tac	ggg	gtc	cga	tac	cat	acg	acg	ggg		1296
Val	Gu	Ser	Leu 420	Cys	Lys	Lys	Tyr	Gly 425	Val	Arg	Tyr	His	Thr 430	Thr	Gly		
atg	att	gct	ggc	acc	gca	gag	gtc	ttt	tcg	cga	ctg	aac	gag	gtg	tcc		1344
Met	Ile	Ala 435	Gly	Thr	Ala	Gu	Val 440	Phe	Ser	Arg	Leu	Asn 445	Gu	Val	Ser		
cag	gct	gca	agc	aag	ctc	ggc	aag	tct	gct	tga							1377
Gln	Ala 450	Ala	Ser	Lys	Leu	Gly 455	Lys	Ser	Ala								

<210> 22  
 <211> 458  
 <212> PRT  
 <213> Mbrtierella alpina

<400> 22

Met 1	Thr	Thr	Ser	Asp 5	Pro	Ser	Val	Arg	Ala 10	Phe	Thr	Arg	Ser	Gu 15	Val		
Leu	His	Ala	Asp 20	Ala	Leu	Asn	Gu	Gly 25	Lys	Lys	Asn	Ala	Gu 30	Ala	Pro		
Phe	Leu	Met 35	Ile	Ile	Asp	Asn	Lys 40	Val	Tyr	Asp	Val	Arg 45	Gu	Phe	Ile		
Pro	Asp 50	His	Pro	Gly	Gly	Ser 55	Val	Ile	Leu	Thr	His 60	Val	Gly	Lys	Asp		
Gly 65	Thr	Asp	Val	Phe 70	Gu	Thr	Phe	His	Pro	Gu 75	Ala	Ala	Trp	Gu	Thr 80		
Leu	Ala	Asn	Phe 85	Tyr	Val	Gly	Asp	Ile	Val	Gu 90	Ser	Asp	Arg	Ala	Ile 95		
Gu	Asn	Asp	Gu 100	Phe	Ala	Ala	Gu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe		
Tyr	Ser	Leu 115	Gly	Tyr	Tyr	Asp	Ser 120	Ser	Lys	Val	Tyr	Tyr 125	Ala	Phe	Lys		
Val	Ser 130	Phe	Asn	Leu	Cys	Ile	Trp 135	Gly	Leu	Ser	Ala	Phe 140	Ile	Val	Ala		
Lys 145	Trp	Gly	Gln	Thr	Ser 150	Thr	Leu	Ala	Asn	Val 155	Ile	Ser	Ala	Ser	Leu 160		
Leu	Gly	Val	Phe 165	Trp	Gln	Gln	Cys	Gly	Trp 170	Leu	Ala	His	Asp	Phe 175	Leu		
His	His	Gln	Val 180	Phe	His	Asp	Arg	Phe 185	Trp	Gly	Asp	Leu	Phe 190	Gly	Ala		
Phe	Leu	Gly 195	Gly	Val	Cys	Gln	Gly 200	Phe	Ser	Ser	Ser	Trp 205	Trp	Lys	Asp		
Lys	His 210	Asn	Thr	His	His	Ala 215	Ala	Pro	Asn	Val	His 220	Gly	Gu	Asp	Pro		
Asp 225	Ile	Asp	Thr	His	Pro 230	Leu	Leu	Thr	Trp	Ser 235	Gu	His	Ala	Leu	Gu 240		
Met	Phe	Ser	Asp 245	Val	Pro	Asp	Gu	Gu	Leu	Thr 250	Gln	Met	Trp	Ser 255	Arg		
Phe	Met	Val	Leu 260	Asn	Gln	Ala	Trp	Phe 265	Tyr	Phe	Pro	Ile	Leu	Ser 270	Phe		
Ala	Arg	Leu 275	Ser	Trp	Cys	Ile	Gln 280	Ser	Ile	Leu	Phe	Val 285	Leu	Pro	Asn		
Gly	Gln	Ala 290	His	Lys	Pro	Ala 295	Gly	Ala	Arg	Val	Pro 300	Ile	Ser	Leu	Val		
Gu 305	Gln	Leu	Ser	Leu	Ala 310	Met	His	Trp	Thr	Trp 315	Tyr	Leu	Ala	Thr	Met 320		
Phe	Leu	Phe	Ile	Lys 325	Asp	Pro	Val	Asn	Met 330	Met	Val	Tyr	Phe	Leu 335	Val		
Ser	Gln	Ala	Val 340	Cys	Gly	Asn	Leu	Leu	Ala 345	Ile	Val	Phe	Ser 350	Leu	Asn		
His	Asn	Gly 355	Met	Pro	Val	Ile	Ser 360	Gln	Gu	Gu	Ala	Val 365	Gu	Met	Asp		
Phe 370	Phe	Thr	Lys	Gln	Ile	Ile	Thr 375	Gly	Arg	Asp	Val 380	Tyr	Pro	Gly	Trp		
Phe 385	Ala	Asp	Trp	Phe	Thr 390	Gly	Gly	Leu	Asn	Tyr 395	Gln	Ile	Gu	His	His 400		

PF58307. txt

Leu Phe Pro Ser Met Pro Arg His His Phe Ser Lys Ile Gl n Pro Ala  
 405 410 415  
 Val Gl u Ser Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly  
 420 425  
 Met Ile Ala Ala Ser Lys Leu Gly Lys Ser Ala  
 435 440 445  
 Gl n Ala Ala Ser Lys Leu Gly Lys Ser Ala  
 450 455

<210> 23  
 <211> 1434  
 <212> DNA  
 <213> Phaeodactylum tricornutum

<220>  
 <221> CDS  
 <222> (1)..(1434)

<400> 23  
 at g ggc aaa gga ggg gac gct cgg gcc t c g aag ggc t c a acg gcg gct 48  
 Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala  
 1 5 10 15  
 cgc aag at c agt tgg cag gaa gt c aag acc cac gcg t ct ccg gag gac 96  
 Arg Lys Ile Ser Trp Gl n Gl u Val Lys Thr His Ala Ser Pro Gl u Asp  
 20 25 30  
 gcc t gg at c att cac t cc aat aag gt c t ac gac gt g t cc aac t gg cac 144  
 Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His  
 35 40 45  
 gaa cat ccc gga ggc gcc gt c att t t c acg cac gcc ggt gac gac at g 192  
 Gl u His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met  
 50 55 60  
 acg gac att t t c gct gcc t t t cac gca ccc gga t c g cag t c g ct c at g 240  
 Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gl n Ser Leu Met  
 65 70 75 80  
 aag aag t t c t ac att ggc gaa t t g ct c ccg gaa acc acc ggc aag gag 288  
 Lys Lys Phe Tyr Ile Gly Gl u Leu Leu Pro Gl u Thr Thr Gly Lys Gl u  
 85 90 95  
 ccg cag caa at c gcc t t t gaa aag ggc t ac cgc gat ct g cgc t cc aaa 336  
 Pro Gl n Gl n Ile Ala Phe Gl u Lys Gly Tyr Arg Asp Leu Arg Ser Lys  
 100 105 110  
 ct c at c at g at g ggc at g t t c aag t cc aac aag t gg t t c t ac gt c t ac 384  
 Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr  
 115 120 125  
 aag t gc ct c agc aac at g gcc att t gg gcc gcc gcc t gt gct ct c gt c 432  
 Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val  
 130 135 140  
 t t t t ac t c g gac cgc t t c t gg gt a cac ct g gcc agc gcc gt c at g ct g 480  
 Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu  
 145 150 155 160  
 gga aca t t c t t t cag cag t c g gga t gg t t g gca cac gac t t t ct g cac 528  
 Gly Thr Phe Phe Gl n Gl n Ser Gly Trp Leu Ala His Asp Phe Leu His  
 165 170 175  
 cac cag gt c t t c acc aag cgc aag cac ggg gat ct c gga gga ct c t t t 576  
 His Gl n Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe  
 180 185 190  
 t gg ggg aac ct c at g cag ggt t ac t cc gt a cag t gg t gg aaa aac aag 624  
 Trp Gly Asn Leu Met Gl n Gly Tyr Ser Val Gl n Trp Trp Lys Asn Lys  
 195 200 205  
 cac aac gga cac cac gcc gt c ccc aac ct c cac t gc t cc t cc gca gt c 672  
 His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val  
 210 215 220  
 gcg caa gat ggg gac ccg gac at c gat acc at g ccc ct t ct c gcc t gg 720  
 Ala Gl n Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp  
 225 230 235 240  
 t cc gt c cag caa gcc cag t ct t ac cgg gaa ct c caa gcc gac gga aag 768  
 Ser Val Gl n Gl n Ala Gl n Ser Tyr Arg Gl u Leu Gl n Ala Asp Gly Lys  
 245 250 255  
 gat t c g ggt t t g gt c aag t t c at g at c cgt aac caa t cc t ac t t t t ac 816

PF58307. txt

Asp	Ser	Gly	Leu	Val	Lys	Phe	Met	Ile	Arg	Asn	Gln	Ser	Tyr	Phe	Tyr		
260			260					265					270				
ttt	ccc	atc	ttg	ttg	ctc	gcc	cgc	ctg	tcg	tgg	ttg	aac	gag	tcc	ttc		864
Phe	Pro	Ile	Leu	Leu	Leu	Ala	Arg	Leu	Ser	Trp	Leu	Asn	Glu	Ser	Phe		
		275					280					285					
aag	tgc	gcc	ttt	ggg	ctt	gga	gct	gcg	tcg	gag	aac	gct	gct	ctc	gaa		912
Lys	Cys	Ala	Phe	Gly	Leu	Gly	Ala	Ala	Ser	Glu	Asn	Ala	Ala	Leu	Glu		
	290			295						300							
ctc	aag	gcc	aag	ggt	ctt	cag	tac	ccc	ctt	ttg	gaa	aag	gct	ggc	atc		960
Leu	Lys	Ala	Lys	Gly	Leu	Gln	Tyr	Pro	Leu	Leu	Glu	Lys	Ala	Gly	Ile		
305				310						315					320		
ctg	ctg	cac	tac	gct	tgg	atg	ctt	aca	gtt	tcg	tcc	ggc	ttt	gga	cgc		1008
Leu	Leu	His	Tyr	Ala	Trp	Met	Leu	Thr	Val	Ser	Ser	Gly	Phe	Gly	Arg		
				325					330					335			
ttc	tcg	ttc	gcg	tac	acc	gca	ttt	tac	ttt	cta	acc	gcg	acc	gcg	tcc		1056
Phe	Ser	Phe	Ala	Tyr	Thr	Ala	Phe	Tyr	Phe	Leu	Thr	Ala	Thr	Ala	Ser		
			340					345					350				
tgt	gga	ttc	ttg	ctc	gcc	att	gtc	ttt	ggc	ctc	ggc	cac	aac	ggc	atg		1104
Cys	Gly	Phe	Leu	Leu	Ala	Ile	Val	Phe	Gly	Leu	Gly	His	Asn	Gly	Met		
	355						360					365					
gcc	acc	tac	aat	gcc	gac	gcc	cgt	ccg	gac	ttc	tgg	aag	ctc	caa	gtc		1152
Ala	Thr	Tyr	Asn	Ala	Asp	Ala	Arg	Pro	Asp	Phe	Trp	Lys	Leu	Gln	Val		
	370					375					380						
acc	acg	act	cgc	aac	gtc	acg	ggc	gga	cac	ggg	ttc	ccc	caa	gcc	ttt		1200
Thr	Thr	Thr	Arg	Asn	Val	Thr	Gly	Gly	His	Gly	Phe	Pro	Gln	Ala	Phe		
	385				390					395				400			
gtc	gac	tgg	ttc	tgt	ggt	ggc	ctc	cag	tac	caa	gtc	gac	cac	cac	tta		1248
Val	Asp	Trp	Phe	Cys	Gly	Gly	Leu	Gln	Tyr	Gln	Val	Asp	His	His	Leu		
				405				410						415			
ttc	ccc	agc	ctg	ccc	cga	cac	aat	ctg	gcc	aag	aca	cac	gca	ctg	gtc		1296
Phe	Pro	Ser	Leu	Pro	Arg	His	Asn	Leu	Ala	Lys	Thr	His	Ala	Leu	Val		
			420					425					430				
gaa	tgc	ttc	tgc	aag	gag	tgg	ggt	gtc	cag	tac	cac	gaa	gcc	gac	ctt		1344
Glu	Ser	Phe	Cys	Lys	Glu	Trp	Gly	Val	Gln	Tyr	His	Glu	Ala	Asp	Leu		
	435					440						445					
gtg	gac	ggg	acc	atg	gaa	gtc	ttg	cac	cat	ttg	ggc	agc	gtg	gcc	ggc		1392
Val	Asp	Gly	Thr	Met	Glu	Val	Leu	His	His	Leu	Gly	Ser	Val	Ala	Gly		
	450				455					460							
gaa	ttc	gtc	gtg	gat	ttt	gta	cgc	gat	gga	ccc	gcc	atg	t aa				1434
Glu	Phe	Val	Val	Asp	Phe	Val	Arg	Asp	Gly	Pro	Ala	Met					
	465				470					475							

<210> 24  
 <211> 477  
 <212> PRT  
 <213> Phaeodactylum tricorutum

<400> 24  
 Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala  
 1 5 10 15  
 Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp  
 20 25 30  
 Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His  
 35 40 45  
 Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met  
 50 55 60  
 Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met  
 65 70 75 80  
 Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu  
 85 90 95  
 Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys  
 100 105 110  
 Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr  
 115 120 125  
 Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val  
 130 135 140  
 Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu  
 145 150 155 160

PF58307. txt

Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His  
 165 170 175  
 His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe  
 180 185  
 Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys  
 195 200  
 His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val  
 210 215  
 Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp  
 225 230 235  
 Ser Val Gln Gln Ala Gln Ser Tyr Arg Gu Leu Gln Ala Asp Gly Lys  
 245 250 255  
 Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr  
 260 265 270  
 Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Gu Ser Phe  
 275 280 285  
 Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Gu Asn Ala Ala Leu Gu  
 290 295 300  
 Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Gu Lys Ala Gly Ile  
 305 310 315 320  
 Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg  
 325 330 335  
 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser  
 340 345 350  
 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met  
 355 360 365  
 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val  
 370 375 380  
 Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe  
 385 390 395 400  
 Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu  
 405 410 415  
 Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val  
 420 425 430 435  
 Gu Ser Phe Cys Lys Gu Trp Gly Val Gln Tyr His Gu Ala Asp Leu  
 440 445  
 Val Asp Gly Thr Met Gu Val Leu His His Leu Gly Ser Val Ala Gly  
 450 455 460  
 Gu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met  
 465 470 475

<210> 25  
 <211> 1374  
 <212> DNA  
 <213> Mbrtierella isabellina

<220>  
 <221> CDS  
 <222> (1).. (1374)

<400> 25  
 atg gct gct gct ccc agt gtg agg acg ttt act cgg gcc gag att ctg 48  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Gu Ile Leu  
 1 5 10 15  
 aat gcc gag gcc ctg aat gag ggc aag aag gat gcc gag gca cct ttc 96  
 Asn Ala Gu Ala Leu Asn Gu Gly Lys Lys Asp Ala Gu Ala Pro Phe  
 20 25 30  
 ttg atg atc atc gac aac aag gtg tat gat gtc cgc gag ttt gtc cct 144  
 Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Gu Phe Val Pro  
 35 40 45  
 gat cat ccc ggt gga agt gtt att ctg acg cac gtt ggc aag gac ggc 192  
 Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly  
 50 55 60  
 act gac gtc ttt gac act ttc cac ccc gag gct gct tgg gag aca ctg 240  
 Thr Asp Val Phe Asp Thr Phe His Pro Gu Ala Ala Trp Gu Thr Leu  
 65 70 75 80  
 gcc aac ttt tac gtt ggt gat att gat gag agc gac cgt gcc atc aag 288  
 Ala Asn Phe Tyr Val Gly Asp Ile Asp Gu Ser Asp Arg Ala Ile Lys

PF58307. txt

			85	90										95			
aac	gat	gac	ttt	gcg	gcc	gag	gtc	cgc	aag	ctg	cgt	act	ttg	ttc	cag	336	
Asn	Asp	Asp	Phe	Ala	Ala	Glu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	Gln		
			100											110			
tct	ctg	ggt	tac	tac	gat	tcc	tcc	aag	gca	tac	tac	gcc	ttc	aag	gtc	384	
Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Ala	Tyr	Tyr	Ala	Phe	Lys	Val		
			115											125			
tca	ttc	aac	ctc	tgc	atc	tgg	ggc	ctg	tcg	acg	ttc	att	gtt	gcc	aag	432	
Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Thr	Phe	Ile	Val	Ala	Lys		
			130											140			
tgg	ggc	cag	acc	tgc	acc	ctc	gcc	aac	gtg	ctc	tgc	gct	gcg	ctc	ttg	480	
Trp	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Val	Leu	Ser	Ala	Ala	Leu	Leu		
			145											155			
ggt	ctc	ttc	tgg	cag	cag	tgt	gga	tgg	ttg	gcg	cac	gac	ttt	ttg	cac	528	
Gly	Leu	Phe	Trp	Gln	Gln	Cys	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His		
			165											170			
cac	cag	gtc	ttc	cag	gac	cgt	ttc	tgg	ggt	gat	ctt	ttc	ggc	gcc	ttc	576	
His	Gln	Val	Phe	Gln	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	Ala	Phe		
			180											190			
ttg	gga	ggt	gtc	tgc	cag	ggc	ttc	tgc	tcc	tca	tgg	tgg	aag	gac	aag	624	
Leu	Gly	Gly	Val	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	Lys		
			195											205			
cac	aac	act	cac	cac	gcc	gcc	ccc	aac	gtc	cac	ggc	gag	gat	ccc	gac	672	
His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp		
			210											220			
att	gac	act	cac	cct	ctg	ttg	acg	tgg	agt	gag	cat	gct	ttg	gag	atg	720	
Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met		
			225											235			
ttc	tgc	gac	gtc	cct	gac	gag	gag	ctg	acc	cgc	atg	tgg	tgc	cgc	ttc	768	
Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe		
			245											255			
atg	gtc	ctt	aac	cag	acc	tgg	ttc	tac	ttt	ccc	att	ctc	tgc	ttt	gcc	816	
Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala		
			260											270			
cgt	ctc	tcc	tgg	tgc	ctc	cag	tcc	atc	ctc	ttt	gtt	ctg	cct	aac	ggt	864	
Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	Gly		
			275											285			
cag	gcc	cac	aag	ccc	tct	gga	gcc	cgt	gtg	ccc	att	tcc	ttg	gtc	gag	912	
Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu		
			290											300			
cag	ctg	tct	ctt	gcc	gtg	cac	tgg	acc	tgg	tac	ctc	gcc	acc	atg	ttc	960	
Gln	Leu	Ser	Leu	Ala	Val	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe		
			305											315			
ttg	ttc	att	aag	gac	ccc	gtc	aac	atg	atg	gtg	tac	ttt	ttg	gtg	tct	1008	
Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Met	Val	Tyr	Phe	Leu	Val	Ser		
			325											330			
cag	gct	gtt	tgc	ggt	aac	ctg	ttg	gcg	att	gtg	ttc	tgc	ctc	aac	cac	1056	
Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His		
			340											350			
aac	ggt	atg	cct	gtg	atc	tcc	aag	gag	gaa	gcc	gtc	gac	atg	gat	ttc	1104	
Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	Met	Asp	Phe		
			355											365			
ttc	acc	aag	cag	atc	atc	acg	ggt	cgt	gat	gtt	cac	cct	ggt	ctg	ttc	1152	
Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	His	Pro	Gly	Leu	Phe		
			370											380			
gcc	aac	tgg	ttc	acg	ggt	gga	ttg	aac	tac	cag	att	gag	cac	cac	ttg	1200	
Ala	Asn	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	His	Leu		
			385											395			
ttc	cct	tgc	atg	cct	cgc	cac	aac	ttt	tca	aag	atc	cag	cct	gct	gtc	1248	
Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	Gln	Pro	Ala	Val		
			405											415			
gag	acc	ttg	tgc	aaa	aag	tat	ggt	gtc	cga	tac	cac	acc	act	ggc	atg	1296	
Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Gly	Val	Arg	Tyr	His	Thr	Thr	Gly	Met		
			420											430			
atc	gag	gga	act	gca	gag	gtc	ttt	agc	cgt	ttg	aac	gag	gtc	tcc	aag	1344	
Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	Val	Ser	Lys		
			435											445			
gcc	gcc	tcc	aag	atg	ggt	aag	gcg	cag	t aa							1374	

Al a Al a Ser Lys Met Gly Lys Al a G n  
 450 455

<210> 26  
 <211> 457  
 <212> PRT  
 <213> Mbrtierella isabellina

<400> 26  
 Met Al a Al a Al a Pro Ser Val Arg Thr Phe Thr Arg Al a Gl u Ile Leu  
 1 5 10 15  
 Asn Al a Gl u Al a Leu Asn Gl u Gly Lys Lys Asp Al a Gl u Al a Pro Phe  
 20 25 30  
 Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Gl u Phe Val Pro  
 35 40  
 Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly  
 50 55 60  
 Thr Asp Val Phe Asp Thr Phe His Pro Gl u Al a Trp Gl u Thr Leu  
 65 70 75 80  
 Al a Asn Phe Tyr Val Gly Asp Ile Asp Gl u Ser Asp Arg Al a Ile Lys  
 85 90 95  
 Asn Asp Asp Phe Al a Al a Gl u Val Arg Lys Leu Arg Thr Leu Phe G n  
 100 105 110  
 Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Al a Tyr Tyr Al a Phe Lys Val  
 115 120 125  
 Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Phe Ile Val Al a Lys  
 130 135 140  
 Trp Gly G n Thr Ser Thr Leu Al a Asn Val Leu Ser Al a Al a Leu Leu  
 145 150 155 160  
 Gly Leu Phe Trp G n G n Cys Gly Trp Leu Al a His Asp Phe Leu His  
 165 170 175  
 His G n Val Phe G n Asp Arg Phe Trp Gly Asp Leu Phe Gly Al a Phe  
 180 185 190  
 Leu Gly Gly Val Cys G n Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys  
 195 200 205  
 His Asn Thr His His Al a Al a Pro Asn Val His Gly Gl u Asp Pro Asp  
 210 215 220  
 Ile Asp Thr His Pro Leu Leu Thr Trp Ser Gl u His Al a Leu Gl u Met  
 225 230 235 240  
 Phe Ser Asp Val Pro Asp Gl u Gl u Leu Thr Arg Met Trp Ser Arg Phe  
 245 250 255  
 Met Val Leu Asn G n Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Al a  
 260 265 270  
 Arg Leu Ser Trp Cys Leu G n Ser Ile Leu Phe Val Leu Pro Asn Gly  
 275 280 285  
 G n Al a His Lys Pro Ser Gly Al a Arg Val Pro Ile Ser Leu Val Gl u  
 290 295 300  
 G n Leu Ser Leu Al a Val His Trp Thr Trp Tyr Leu Al a Thr Met Phe  
 305 310 315 320  
 Leu Phe Ile Lys Asp Pro Val Asn Met Met Val Tyr Phe Leu Val Ser  
 325 330 335  
 G n Al a Val Cys Gly Asn Leu Leu Al a Ile Val Phe Ser Leu Asn His  
 340 345 350  
 Asn Gly Met Pro Val Ile Ser Lys Gl u Gl u Al a Val Asp Met Asp Phe  
 355 360 365  
 Phe Thr Lys G n Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe  
 370 375 380  
 Al a Asn Trp Phe Thr Gly Gly Leu Asn Tyr G n Ile Gl u His His Leu  
 385 390 395 400  
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile G n Pro Al a Val  
 405 410 415  
 Gl u Thr Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly Met  
 420 425 430  
 Ile Gl u Gly Thr Al a Gl u Val Phe Ser Arg Leu Asn Gl u Val Ser Lys  
 435 440 445  
 Al a Al a Ser Lys Met Gly Lys Al a G n  
 450 455

<210> 27  
 <211> 1374  
 <212> DNA  
 <213> Mbr t i e r e l l a a l p i n a

<220>  
 <221> CDS  
 <222> (1)..(1374)

<400> 27  
 at g gct gct gct ccc agt gt g agg acg t t t act cgg gcc gag at t ct g 48  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Gu l l e Leu  
 1 5 10 15  
 aat gcc gag gcc ct g aat gag ggc aag aag gat gcc gag gca cct t t c 96  
 Asn Ala Gu Ala Leu Asn Gu Gy Lys Lys Asp Ala Gu Ala Pro Phe  
 20 25 30  
 t t g at g at c at c gac aac aag gt g t at gat gt c cgc gag t t t gt c cct 144  
 Leu Met l l e l l e Asp Asn Lys Val Tyr Asp Val Arg Gu Phe Val Pro  
 35 40 45  
 gat cat ccc ggt gga agt gt t at t ct c acg cac gt t ggc aag gac ggc 192  
 Asp His Pro Gy Gy Ser Val l l e Leu Thr His Val Gy Lys Asp Gy  
 50 55 60  
 act gac gt c t t t gac act t t c cac ccc gag gct gct t gg gag aca ct t 240  
 Thr Asp Val Phe Asp Thr Phe His Pro Gu Ala Ala Trp Gu Thr Leu  
 65 70 75 80  
 gcc aac t t t t ac gt t ggt gat at t gat gag agc gac cgt gcc at c aag 288  
 Ala Asn Phe Tyr Val Gy Asp l l e Asp Gu Ser Asp Arg Ala l l e Lys  
 85 90 95  
 aac gat gac t t t gcg gcc gag gt c cgc aag ct g cgt act t t g t t c cag 336  
 Asn Asp Asp Phe Ala Ala Gu Val Arg Lys Leu Arg Thr Leu Phe Gn  
 100 105  
 t ct ct g ggt t ac t ac gat t cc t cc aag gca t ac t ac gcc t t c aag gt c 384  
 Ser Leu Gy Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val  
 115 120 125  
 t ca t t c aac ct c t gc at c t gg ggc ct g t cg acg t t c at t gt t gcc aag 432  
 Ser Phe Asn Leu Cys l l e Tr p Gy Leu Ser Thr Phe l l e Val Ala Lys  
 130 135 140  
 t gg ggc cag acc t cg acc ct c gcc aac gag ct c t cg gct gcg ct c t t g 480  
 Tr p Gy Gn Thr Ser Thr Leu Ala Asn Gu Leu Ser Ala Ala Leu Leu  
 145 150 155 160  
 ggt ct c t t c t gg cag cag aga gga t gg t t g gcg cac gac t t t t t g cac 528  
 Gy Leu Phe Tr p Gn Gn Arg Gy Tr p Leu Ala His Asp Phe Leu His  
 165 170 175  
 cac cag gt c t t c cag gac cgt t t c t gg gga gat ct t t t c ggc gcc t t c 576  
 His Gn Val Phe Gn Asp Arg Phe Tr p Gy Asp Leu Phe Gy Ala Phe  
 180 185 190  
 t t g gga gga gac t gc cag gcc t t c t cg t cc t ca t gg t gg aag gac aag 624  
 Leu Gy Gy Asp Cys Gn Gy Phe Ser Ser Ser Tr p Tr p Lys Asp Lys  
 195 200 205  
 cac aac act cac cac gcc gcc ccc aac gt c cac ggc gag gat ccc gac 672  
 His Asn Thr His His Ala Ala Pro Asn Val His Gy Gu Asp Pro Asp  
 210 215 220  
 at t gac act cac cct ct g t t g acg t gg agt gag cat gct t t g gag at g 720  
 l l e Asp Thr His Pro Leu Leu Thr Tr p Ser Gu His Ala Leu Gu Met  
 225 230 235 240  
 t t c t cg gac gt c cct gac gag gag ct g acc cgc at g t gg t cg cgc t t c 768  
 Phe Ser Asp Val Pro Asp Gu Gu Leu Thr Arg Met Tr p Ser Arg Phe  
 245 250 255  
 at g gt c ct t aac cag acc t gg t t c t ac t t t ccc at t ct c t cg t t t gcc 816  
 Met Val Leu Asn Gn Thr Tr p Phe Tyr Phe Pro l l e Leu Ser Phe Ala  
 260 265 270  
 cgt ct c t cc t gg t gc ct c cag t cc at c ct c t t t gt t ct g cct aac ggt 864  
 Arg Leu Ser Tr p Cys Leu Gn Ser l l e Leu Phe Val Leu Pro Asn Gy  
 275 280 285  
 cag gcc cac aag ccc t ct gga gcc cgt gt g ccc at t t cc t t g gt c gag 912  
 Gn Ala His Lys Pro Ser Gy Ala Arg Val Pro l l e Ser Leu Val Gu  
 290 295 300  
 cag ct g t ct ct t gcc at g cac t gg acc t gg t ac ct c gcc acc at g t t c 960



PF58307. txt

G n	Leu	Ser	Leu	Al a	Met	Hi s	Tr p	Thr	Tr p	Tyr	Leu	Al a	Thr	Met	Phe		
305					310					315					320		
t t g	t t c	a t t	a a g	g a c	c c c	g t c	a a c	a t g	a t g	g t g	t a c	t t t	t t g	g t g	t c t		1008
Leu	Phe	I l e	Lys	Asp	Pr o	Val	Asn	Met	Met	Val	Tyr	Phe	Leu	Val	Ser		
				325						330				335			
c a g	g c t	g t t	t g c	g g t	a a c	c t g	t t g	g c g	a t t	g t g	t t c	t c g	c t c	a a c	c a c		1056
G n	Al a	Val	Cys	G y	Asn	Leu	Leu	Al a	I l e	Val	Phe	Ser	Leu	Asn	Hi s		
			340					345					350				
a a c	g g t	a t g	c c t	g t g	a t c	t c c	a a g	g a g	g a a	g c c	g t c	g a c	a t g	g a t	t t c		1104
Asn	G y	Met	Pr o	Val	I l e	Ser	Lys	G u	G u	Al a	Val	Asp	Met	Asp	Phe		
		355					360					365					
t t c	a c c	a a g	c a g	a t c	a t c	a c g	g g t	c g t	g a t	g t t	c a c	c c t	g g t	c t g	t t c		1152
Phe	Thr	Lys	G n	I l e	I l e	Thr	G y	Arg	Asp	Val	Hi s	Pr o	G y	Leu	Phe		
	370					375					380						
g c c	a a c	t g g	t t c	a c g	g g t	g g a	t t g	a a c	t a c	c a g	a t t	g a g	c a c	c a c	t t g		1200
Al a	Asn	Tr p	Phe	Thr	G y	G y	Leu	Asn	Tyr	G n	I l e	G u	Hi s	Hi s	Leu		
385				390						395					400		
t t c	c c t	t c g	a t g	c c t	c g c	c a c	a a c	t t t	t c a	a a g	a t c	c a g	c c t	g c t	g t c		1248
Phe	Pr o	Ser	Met	Pr o	Arg	Hi s	Asn	Phe	Ser	Lys	I l e	G n	Pr o	G t	Al a	Val	
				405				410						415			
g a g	a c c	t t g	t g c	a a a	a a g	t a t	g g t	g t c	c g a	t a c	c a c	a c c	a c t	g g c	a t g		1296
G u	Thr	Leu	Cys	Lys	Lys	Tyr	G y	Val	Arg	Tyr	Hi s	Thr	Thr	G y	Met		
			420				425						430				
a t c	g a g	g g a	a c t	g c a	g a g	g t c	t t t	a g c	c g t	t t g	a a c	g a g	g t c	t c c	a a g		1344
I l e	G u	G y	Thr	Al a	G u	Val	Phe	Ser	Arg	Leu	Asn	G u	Val	Ser	Lys		
		435				440						445					
g c c	g c c	t c c	a a g	a t g	g g t	a a g	g c g	c a g	t a a								1374
Al a	Al a	Ser	Lys	Met	G y	Lys	Al a	G n									
	450				455												

<210> 28  
 <211> 457  
 <212> PRT  
 <213> Mbr t i e r e l l a a l p i n a

<400> 28

Met	Al a	Al a	Al a	Pr o	Ser	Val	Arg	Thr	Phe	Thr	Arg	Al a	G u	I l e	Leu		
1				5					10					15			
Asn	Al a	G u	Al a	Leu	Asn	G u	G y	Lys	Lys	Asp	Al a	G u	Al a	Pr o	Phe		
			20					25					30				
Leu	Met	I l e	Asp	Asn	Lys	Val	Tyr	Asp	Val	Arg	G u	Phe	Val	Pr o			
		35				40						45					
Asp	Hi s	Pr o	G y	G y	Ser	Val	I l e	Leu	Thr	Hi s	Val	G y	Lys	Asp	G y		
	50				55						60						
Thr	Asp	Val	Phe	Asp	Thr	Phe	Hi s	Pr o	G u	Al a	Tr p	G u	Thr	Leu			
65				70						75				80			
Al a	Asn	Phe	Tyr	Val	G y	Asp	I l e	Asp	G u	Ser	Asp	Arg	Al a	I l e	Lys		
				85					90					95			
Asn	Asp	Asp	Phe	Al a	Al a	G u	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	G n		
			100					105						110			
Ser	Leu	G y	Tyr	Tyr	Asp	Ser	Ser	Lys	Al a	Tyr	Tyr	Al a	Phe	Lys	Val		
		115				120						125					
Ser	Phe	Asn	Leu	Cys	I l e	Tr p	G y	Leu	Ser	Thr	Phe	I l e	Val	Al a	Lys		
	130					135					140						
Tr p	G y	G n	Thr	Ser	Thr	Leu	Al a	Asn	G u	Leu	Ser	Al a	Al a	Leu	Leu		
145					150					155				160			
G y	Leu	Phe	Tr p	G n	G n	Arg	G y	Tr p	Leu	Al a	Hi s	Asp	Phe	Leu	Hi s		
				165					170					175			
Hi s	G n	Val	Phe	G n	Asp	Arg	Phe	Tr p	G y	Asp	Leu	Phe	G y	Al a	Phe		
			180					185					190				
Leu	G y	G y	Asp	Cys	G n	G y	Phe	Ser	Ser	Ser	Tr p	Tr p	Lys	Asp	Lys		
		195				200					205						
Hi s	Asn	Thr	Hi s	Hi s	Al a	Al a	Pr o	Asn	Val	Hi s	G y	G u	Asp	Pr o	Asp		
	210					215						220					
I l e	Asp	Thr	Hi s	Pr o	Leu	Leu	Thr	Tr p	Ser	G u	Hi s	Al a	Leu	G u	Met		
225					230					235				240			
Phe	Ser	Asp	Val	Pr o	Asp	G u	G u	Leu	Thr	Arg	Met	Tr p	Ser	Arg	Phe		
				245					250					255			



PF58307. txt

Met Val Leu Asn Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala  
 260 275 280 285 290 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450  
 Arg Leu Ser Trp Cys Leu Gln Ser Ile Leu Phe Val Leu Pro Asn Gly  
 Gln Ala His Lys Pro Ser Gly Ala Arg Val Pro Ile Ser Leu Val Gu  
 Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met Phe  
 Leu Phe Ile Lys Asp Pro Val Asn Met Met Val Tyr Phe Leu Val Ser  
 Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn His  
 Asn Gly Met Pro Val Ile Ser Lys Gu Gu Ala Val Asp Met Asp Phe  
 Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe  
 Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Gu His His Leu  
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val  
 Gu Thr Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly Met  
 Ile Gu Gly Thr Ala Gu Val Phe Ser Arg Leu Asn Gu Val Ser Lys  
 Ala Ala Ser Lys Met Gly Lys Ala Gln

<210> 29  
 <211> 1380  
 <212> DNA  
 <213> Pythium irregulare

<220>  
 <221> CDS  
 <222> (1)..(1380)

<400> 29  
 atg gtg gac ctg aag cct gga gtg aag cgc ctg gtg agc tgg aag gag 48  
 Met Val Asp Leu Lys Pro Gly Val Lys Arg Leu Val Ser Trp Lys Gu  
 1 5 10 15  
 atc cgc gag cac gcg acg ccc gcg acc gcg tgg atc gtg att cac cac 96  
 Ile Arg Glu His Ala Thr Pro Ala Thr Ala Trp Ile Val Ile His His  
 20 25 30  
 aag gtc tac gac atc tcc aag tgg gac tgc cac ccg ggt ggc tcc gtg 144  
 Lys Val Tyr Asp Ile Ser Lys Trp Asp Ser His Pro Gly Gly Ser Val  
 35 40 45  
 atg ctg acg cag gcc ggc gag gac gcc acg gac gcc ttc gcg gtc ttc 192  
 Met Leu Thr Gln Ala Gly Glu Asp Ala Thr Asp Ala Phe Ala Val Phe  
 50 55 60  
 cac ccg tcc tgc gcg ctg aag ctg ctg gag cag ttc tac gtc ggc gac 240  
 His Pro Ser Ser Ala Leu Lys Leu Leu Glu Gln Phe Tyr Val Gly Asp  
 65 70 75 80  
 gtg gac gaa acc tcc aag gcc gag atc gag ggg gag ccg gcg agc gac 288  
 Val Asp Glu Thr Ser Lys Ala Glu Ile Glu Gly Glu Pro Ala Ser Asp  
 85 90 95  
 gag gag cgc gcg cgc cgc gag cgc atc aac gag ttc atc gcg tcc tac 336  
 Gu Glu Arg Ala Arg Arg Glu Arg Ile Asn Gu Phe Ile Ala Ser Tyr  
 100 105 110  
 cgt cgt ctg cgc gtc aag gtc aag ggc atg ggg ctg tac gac gcc agc 384  
 Arg Arg Leu Arg Val Lys Val Lys Gly Met Gly Leu Tyr Asp Ala Ser  
 115 120 125  
 gcg ctg tac tac gcg tgg aag ctg gtg agc acg ttc ggc atc gcg gtg 432  
 Ala Leu Tyr Tyr Ala Trp Lys Leu Val Ser Thr Phe Gly Ile Ala Val  
 130 135 140  
 ctg tgc atg gcg atc tgc ttc ttc ttc aac agt ttc gcc atg tac atg 480  
 Leu Ser Met Ala Ile Cys Phe Phe Phe Asn Ser Phe Ala Met Tyr Met  
 145 150 155 160  
 gtc gcc ggc gtg att atg ggg ctg ttc tac cag cag tcc gga tgg ctg 528

Seite 35

PF58307.txt

Val Ala Gly Val Ile Met Gly Leu Phe Tyr Gln Gln Ser Gly Trp Leu  
 165 170 175  
 gcg cac gac ttc ttg cac aac cag gtg tgc gag aac cgc acg ctg ggc 576  
 Ala His Asp Phe Leu His Asn Gln Val Cys Glu Asn Arg Thr Leu Gly  
 180 185 190  
 aac ctt atc ggc tgc ctg gtg ggc aac gcc tgg cag ggc ttc agc gtg 624  
 Asn Leu Ile Gly Cys Leu Val Gly Asn Ala Trp Gln Gly Phe Ser Val  
 195 200 205  
 cag tgg tgg aag aac aag cac aac ctg cac cac gcg gtg ccg aac ctg 672  
 Gln Trp Trp Lys Asn Lys His Asn Leu His His Ala Val Pro Asn Leu  
 210 215 220  
 cac agc gcc aag gac gag ggc ttc atc ggc gac ccg gac atc gac acc 720  
 His Ser Ala Lys Asp Glu Gly Phe Ile Gly Asp Pro Asp Ile Asp Thr  
 225 230 235 240  
 atg ccg ctg ctg gcg tgg tct aag gag atg gcg cgc aag gcg ttc gag 768  
 Met Pro Leu Leu Ala Trp Ser Lys Glu Met Ala Arg Lys Ala Phe Glu  
 245 250 255  
 tgc gcg cac ggc ccg ttc ttc atc cgc aac cag gcg ttc cta tac ttc 816  
 Ser Ala His Gly Pro Phe Phe Ile Arg Asn Gln Ala Phe Leu Tyr Phe  
 260 265 270  
 ccg ctg ctg ctg ctg gcg cgc ctg agc tgg ctg gcg cag tgc ttc ttc 864  
 Pro Leu Leu Leu Leu Ala Arg Leu Ser Trp Leu Ala Gln Ser Phe Phe  
 275 280 285  
 tac gtg ttc acc gag ttc tgc ttc ggc atc ttc gac aag gtg gag ttc 912  
 Tyr Val Phe Thr Glu Phe Ser Phe Gly Ile Phe Asp Lys Val Glu Phe  
 290 295 300  
 gac gga ccg gag aag gcg ggt ctg atc gtg cac tac atc tgg cag ctg 960  
 Asp Gly Pro Glu Lys Ala Gly Leu Ile Val His Tyr Ile Trp Gln Leu  
 305 310 315 320  
 gcg atc ccg tac ttc tgc aac atg agc ctg ttt gag ggc gtg gca tac 1008  
 Ala Ile Pro Tyr Phe Cys Asn Met Ser Leu Phe Glu Gly Val Ala Tyr  
 325 330 335  
 ttc ctg atg ggc cag gcg tcc tgc ggc ttg ctg ctg gcg ctg gtg ttc 1056  
 Phe Leu Met Gly Gln Ala Ser Cys Gly Leu Leu Leu Ala Leu Val Phe  
 340 345 350  
 agt att ggc cac aac ggc atg tgc gtg tac gag cgc gaa acc aag ccg 1104  
 Ser Ile Gly His Asn Gly Met Ser Val Tyr Glu Arg Gln Thr Lys Pro  
 355 360 365  
 gac ttc tgg cag ctg cag gtg acc acg acg cgc aac atc cgc gcg tgc 1152  
 Asp Phe Trp Gln Leu Gln Val Thr Thr Thr Arg Asn Ile Arg Ala Ser  
 370 375 380  
 gta ttc atg gac tgg ttc acc ggt ggc ttg aac tac cag atc gac cat 1200  
 Val Phe Met Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Asp His  
 385 390 395 400  
 cac ctg ttc ccg ctg gtg ccg cgc cac aac ttg cca aag gtg aac gtg 1248  
 His Leu Phe Pro Leu Val Pro Arg His Asn Leu Pro Lys Val Asn Val  
 405 410 415  
 ctg atc aag tgc cta tgc aag gag ttc gac atc ccg ttc cac gag acc 1296  
 Leu Ile Lys Ser Leu Cys Lys Glu Phe Asp Ile Pro Phe His Glu Thr  
 420 425 430  
 ggc ttc tgg gag ggc atc tac gag gtg gac cac ctg gcg gac atc 1344  
 Gly Phe Trp Glu Gly Ile Tyr Glu Val Val Asp His Leu Ala Asp Ile  
 435 440 445  
 agc aag gaa ttc atc acc gag ttc cca gcg atg taa 1380  
 Ser Lys Glu Phe Ile Thr Glu Phe Pro Ala Met

<210> 30  
 <211> 459  
 <212> PRT  
 <213> Pythium irregularis

<400> 30  
 Met Val Asp Leu Lys Pro Gly Val Lys Arg Leu Val Ser Trp Lys Glu  
 1 5 10 15  
 Ile Arg Glu His Ala Thr Pro Ala Thr Ala Trp Ile Val Ile His His  
 20 25 30  
 Lys Val Tyr Asp Ile Ser Lys Trp Asp Ser His Pro Gly Gly Ser Val  
 Sei te 36

PF58307. txt

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

<210> 31  
 <211> 1374  
 <212> DNA  
 <213> Mbr t i e r e l l a a l p i n a

<220>  
 <221> CDS  
 <222> ( 1 ) . . ( 1374)

<400> 31  
 at g gct gct gct ccc agt gtg agg acg ttt act cgg gcc gag att ctg 48  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Gu Ile Leu  
 1 5 10 15  
 aat gcc gag gcc ctg aat gag ggc aag aag gat gcc gag gca cct ttc 96  
 Sei t e 37

PF58307. txt

Asn	Ala	Glu	Ala	Leu	Asn	Glu	Gly	Lys	Lys	Asp	Ala	Glu	Ala	Pro	Phe		
ttg	atg	atc	atc	gac	aac	aag	gtg	tat	gat	gtc	cgc	gag	ttt	gtc	cct		144
Leu	Met	Ile	Ile	Asp	Asn	Lys	Val	Tyr	Asp	Val	Arg	Glu	Phe	Val	Pro		
		35					40					45					
gat	cat	ccc	ggt	gga	agt	gtt	att	ctc	acg	cac	gtt	ggc	aag	gac	ggc		192
Asp	His	Pro	Gly	Gly	Ser	Val	Ile	Leu	Thr	His	Val	Gly	Lys	Asp	Gly		
	50				55					60							
act	gac	gtc	ttt	gac	act	ttc	cac	ccc	gag	gct	tgg	gag	aca	ctt			240
Thr	Asp	Val	Phe	Asp	Thr	Phe	His	Pro	Glu	Ala	Ala	Trp	Glu	Thr	Leu		
	65				70					75				80			
gcc	aac	ttt	tac	gtt	gga	gat	att	gat	gag	agc	gac	cgt	gcc	atc	aag		288
Ala	Asn	Phe	Tyr	Val	Gly	Asp	Ile	Asp	Glu	Ser	Asp	Arg	Ala	Ile	Lys		
				85					90					95			
aac	gat	gac	ttt	gcg	gcc	gag	gtc	cgc	aag	ctg	cgt	act	ttg	ttc	cag		336
Asn	Asp	Asp	Phe	Ala	Ala	Glu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	Gln		
			100					105					110				
tct	ctg	ggt	tac	tac	gat	tcc	tcc	aag	gca	tac	tac	gcc	ttc	aag	gtc		384
Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Ala	Tyr	Tyr	Ala	Phe	Lys	Val		
		115				120						125					
tca	ttc	aac	ctc	tgc	atc	tgg	ggc	ctg	tcg	acg	ttc	att	gtt	gcc	aag		432
Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Thr	Phe	Ile	Val	Ala	Lys		
		130				135					140						
agg	ggc	cag	acc	tcg	acc	ctc	gcc	aac	gag	ctc	tcg	gct	gcg	ctc	ttg		480
Arg	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Glu	Leu	Ser	Ala	Ala	Leu	Leu		
	145			150					155					160			
ggt	ctc	ttc	tgg	cag	cag	aga	gga	tgg	ttg	gcg	cac	gac	ttt	ttg	cac		528
Gly	Leu	Phe	Trp	Gln	Gln	Arg	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His		
				165				170						175			
cac	cag	gtc	ttc	cag	gac	cgt	ttc	tgg	gga	gat	ctt	ttc	ggc	gcc	ttc		576
His	Gln	Val	Phe	Gln	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	Ala	Phe		
			180					185					190				
ttg	gga	gga	gac	tgc	cag	ggc	ttc	tcg	tcc	tca	tgg	tgg	aag	gac	aag		624
Leu	Gly	Gly	Asp	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	Lys		
		195				200					205						
cac	aac	act	cac	cac	gcc	gcc	ccc	aac	gtc	cac	ggc	gag	gat	ccc	gac		672
His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp		
		210				215					220						
att	gac	act	cac	cct	ctg	ttg	acg	tgg	agt	gag	cat	gct	ttg	gag	atg		720
Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met		
		225			230				235					240			
ttc	tcg	gac	gtc	cct	gac	gag	gag	ctg	acc	cgc	atg	tgg	tcg	cgc	ttc		768
Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe		
				245					250					255			
atg	gtc	ctt	aac	cag	acc	tgg	ttc	tac	ttt	ccc	att	ctc	tcg	ttt	gcc		816
Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala		
			260				265						270				
cgt	ctc	tcc	tgg	tgc	ctc	cag	tcc	atc	ctc	ttt	gtt	ctg	cct	aac	ggt		864
Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	Gly		
		275				280							285				
cag	gcc	cac	aag	ccc	tct	gga	ggc	cgt	gtg	ccc	att	tcc	ttg	gtc	gag		912
Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu		
		290				295					300						
cag	ctg	tct	ctt	gcc	atg	cac	tgg	acc	tgg	tac	ctc	gcc	acc	atg	ttc		960
Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe		
		305			310				315						320		
ttg	ttc	att	aag	gac	ccc	gtc	aac	atg	atg	gtg	tac	ttt	ttg	gtg	tct		1008
Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Met	Val	Tyr	Phe	Leu	Val	Ser		
				325				330						335			
cag	gct	gtt	tgc	ggt	aac	ctg	ttg	gcg	att	gtg	ttc	tcg	ctc	aac	cac		1056
Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His		
			340					345					350				
aac	ggt	atg	cct	gtg	atc	tcc	aag	gag	gaa	gcc	gtc	gac	atg	gat	ttc		1104
Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	Met	Asp	Phe		
		355					360					365					
ttc	acc	aag	cag	atc	atc	acg	ggt	cgt	gat	gtt	cac	cct	ggt	ctg	ttc		1152
Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	His	Pro	Gly	Leu	Phe		
		370				375					380						

PF58307.txt

gcc	aac	tgg	ttc	acg	ggt	gga	ttg	aac	tac	cag	att	gag	cac	cac	ttg	1200
Ala	Asn	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	His	Leu	
385					390					395					400	
ttc	cct	tcg	atg	cct	cgc	cac	aac	ttt	tca	aag	atc	cag	cct	gct	gtc	1248
Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	Gln	Pro	Ala	Val	
				405				410						415		
gag	acc	ttg	tgc	aaa	aag	tat	ggt	gtc	cga	tac	cac	acc	act	ggc	atg	1296
Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Gly	Val	Arg	Tyr	His	Thr	Thr	Gly	Met	
			420					425					430			
atc	gag	gga	act	gca	gag	gtc	ttt	agc	cgt	ttg	aac	gag	gtc	tcc	aag	1344
Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	Val	Ser	Lys	
		435				440						445				
gcc	gcc	tcc	aag	atg	ggt	aag	gcg	cag	taa							1374
Ala	Ala	Ser	Lys	Met	Gly	Lys	Ala	Gln								
	450				455											

<210> 32  
 <211> 457  
 <212> PRT  
 <213> Mbrt ierell a alpi na

<400> 32

Met	Ala	Ala	Ala	Pro	Ser	Val	Arg	Thr	Phe	Thr	Arg	Ala	Glu	Ile	Leu
1				5					10					15	
Asn	Ala	Glu	Ala	Leu	Asn	Glu	Gly	Lys	Lys	Asp	Ala	Glu	Ala	Pro	Phe
			20					25					30		
Leu	Met	Ile	Ile	Asp	Asn	Lys	Val	Tyr	Asp	Val	Arg	Glu	Phe	Val	Pro
		35					40					45			
Asp	His	Pro	Gly	Gly	Ser	Val	Ile	Leu	Thr	His	Val	Gly	Lys	Asp	Gly
	50				55						60				
Thr	Asp	Val	Phe	Asp	Thr	Phe	His	Pro	Glu	Ala	Ala	Trp	Glu	Thr	Leu
65					70					75					80
Ala	Asn	Phe	Tyr	Val	Gly	Asp	Ile	Asp	Glu	Ser	Asp	Arg	Ala	Ile	Lys
				85					90					95	
Asn	Asp	Asp	Phe	Ala	Ala	Glu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	Gln
			100					105					110		
Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Ala	Tyr	Tyr	Ala	Phe	Lys	Val
		115					120					125			
Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Thr	Phe	Ile	Val	Ala	Lys
	130					135					140				
Arg	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Glu	Leu	Ser	Ala	Ala	Leu	Leu
145					150					155					160
Gly	Leu	Phe	Trp	Gln	Gln	Arg	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His
				165					170					175	
His	Gln	Val	Phe	Gln	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	Ala	Phe
			180					185					190		
Leu	Gly	Gly	Asp	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	Lys
		195					200					205			
His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp
	210					215					220				
Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met
225					230					235					240
Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe
				245					250					255	
Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala
			260					265					270		
Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	Gly
		275					280					285			
Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu
		290				295						300			
Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe
305					310					315					320
Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Met	Val	Tyr	Phe	Leu	Val	Ser
				325					330					335	
Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His
			340					345					350		
Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	Met	Asp	Phe
		355					360						365		

PF58307.txt

Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe  
 370 375 380  
 Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Gu His His Leu  
 385 390 395 400  
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val  
 405 410 415  
 Gu Thr Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly Met  
 420 425 430  
 Ile Gu Gly Thr Ala Gu Val Phe Ser Arg Leu Asn Gu Val Ser Lys  
 435 440 445  
 Ala Ala Ser Lys Met Gly Lys Ala Gln  
 450 455

<210> 33  
 <211> 1374  
 <212> DNA  
 <213> Mbrtierella isabellina

<220>  
 <221> CDS  
 <222> (1)..(1374)

<400> 33  
 atg gct gct gct ccc agt gtg agg acg ttt act cgg gcc gag att ctg 48  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Gu Ile Leu  
 1 5 10 15  
 aat gcc gag gcc ctg aat gag ggc aag aag gat gcc gag gca cct ttc 96  
 Asn Ala Gu Ala Leu Asn Gu Gly Lys Lys Asp Ala Gu Ala Pro Phe  
 20 25 30  
 ttg atg atc atc gac aac aag gtg tat gat gtc cgc gag ttt gtc cct 144  
 Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Gu Phe Val Pro  
 35 40 45  
 gat cat ccc ggt gga agt gtt att ctg acg cac gtt ggc aag gac ggc 192  
 Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly  
 50 55 60  
 act gac gtc ttt gac act ttc cac ccc gag gct gct tgg gag aca ctt 240  
 Thr Asp Val Phe Asp Thr Phe His Pro Gu Ala Ala Trp Gu Thr Leu  
 65 70 75 80  
 gcc aac ttt tat gtt ggt gat att gat gag agc gac cgt gcc atc aaa 288  
 Ala Asn Phe Tyr Val Gly Asp Ile Asp Gu Ser Asp Arg Ala Ile Lys  
 85 90 95  
 aac gat gac ttt gcg gcc gag gtc cgc aag ctg cgt act ttg ttc cag 336  
 Asn Asp Asp Phe Ala Ala Gu Val Arg Lys Leu Arg Thr Leu Phe Gln  
 100 105 110  
 tct ctg ggt tac tac gat tcc tcc aag gca tac tac gcc ttc aag gtc 384  
 Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val  
 115 120 125  
 tca ttc aac ctg tgc atc tgg ggc ctg tcg acg ttc att gtt gcc aag 432  
 Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Phe Ile Val Ala Lys  
 130 135 140  
 tgg ggc cag acc tcg acc ctg gcc aac gta ctg tcg gct gcg ctg ttg 480  
 Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu  
 145 150 155 160  
 ggt ctg ttc tgg cag cag tgt gga tgg ttg gcg cac gac ttt ttg cac 528  
 Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu His  
 165 170 175  
 cac cag gtc ttc cag gac cgt ttc tgg ggt gat ctt ttc ggc gcc ttc 576  
 His Gln Val Phe Gln Asp Arg Phe Trp Gly Asp Leu Phe Gly Ala Phe  
 180 185 190  
 ttg gga ggt gtc tgc cag gcc ttc tcg tcc tca tgg tgg aag gac aag 624  
 Leu Gly Gly Val Cys Gln Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys  
 195 200 205  
 cac aac act cac cac gcc gcc ccc aac gtc cac ggc gag gat ccc gac 672  
 His Asn Thr His His Ala Ala Pro Asn Val His Gly Gu Asp Pro Asp  
 210 215 220  
 att gac act cac cct ctg ttg acg tgg agt gag cat gct ttg gag atg 720  
 Ile Asp Thr His Pro Leu Leu Thr Trp Ser Gu His Ala Leu Gu Met  
 225 230 235 240

PF58307. txt

t t c	t c g	g a c	g t c	c c t	g a c	g a g	g a g	c t g	a c c	c g c	a t g	t g g	t c g	c g c	t t c	768
P h e	S e r	A s p	V a l	P r o	A s p	G u	G u	L e u	T h r	A r g	M e t	T r p	S e r	A r g	P h e	
				245					250					255		
a t g	g t c	c t t	a a c	c a g	a c c	t g g	t t c	t a c	t t t	c c c	a t t	c t c	t c g	t t t	g c c	816
M e t	V a l	L e u	A s n	G n	T h r	T r p	P h e	T y r	P h e	P r o	I l e	L e u	S e r	P h e	A l a	
			260					265					270			
c g t	c t c	t c c	t g g	t g c	c t c	c a g	t c c	a t c	c t c	c t t	g t t	c t g	c c t	a a c	g g t	864
A r g	L e u	S e r	T r p	C y s	L e u	G n	S e r	I l e	L e u	L e u	V a l	L e u	P r o	A s n	G y	
		275					280						285			
c a g	g c c	c a c	a a g	c c c	t c t	g g a	g c c	c g t	g t g	t c c	a t t	t c c	t t g	g t c	g a g	912
G n	A l a	H i s	L y s	P r o	S e r	G y	A l a	A r g	V a l	S e r	I l e	S e r	L e u	V a l	G u	
	290					295				300						
c a g	c t g	t c t	c t t	g c c	a t g	c a c	t g g	a c c	t g g	t a c	c t c	g c c	a c c	a t g	t t c	960
G n	L e u	S e r	L e u	A l a	M e t	H i s	T r p	T h r	T r p	T y r	L e u	A l a	T h r	M e t	P h e	
	305			310						315					320	
t t g	t t c	a t t	a a g	g a c	c c c	g t c	a a c	a t g	a t g	g t g	t a c	t t t	t t g	g t g	t c t	1008
L e u	P h e	I l e	L y s	A s p	P r o	V a l	A s n	M e t	M e t	V a l	T y r	P h e	L e u	V a l	S e r	
				325					330					335		
c a g	g c t	g t t	t g c	g g t	a a c	c t g	t t g	g c g	a t t	g t g	t t c	t c g	c t c	a a c	c a c	1056
G n	A l a	V a l	C y s	G y	A s n	L e u	L e u	A l a	I l e	V a l	P h e	S e r	L e u	A s n	H i s	
		340						345					350			
a a c	g g t	a t g	c c t	g t g	a t c	t c c	a a g	g a g	g a a	g c c	g t c	g a c	a t g	g a t	t t c	1104
A s n	G y	M e t	P r o	V a l	I l e	S e r	L y s	G u	G u	A l a	V a l	A s p	M e t	A s p	P h e	
		355					360					365				
t t c	a c c	a a g	c a g	a t c	a t c	a c g	g g t	c g t	g a t	g t t	c a c	c c t	g g t	c t g	t t c	1152
P h e	T h r	L y s	G n	I l e	I l e	T h r	G y	A r g	A s p	V a l	H i s	P r o	G y	L e u	P h e	
	370					375					380					
g c c	a a t	t g g	t t c	a c g	g g t	g g a	t t g	a a c	t a c	c a g	a t t	g a g	c a c	c a c	t t g	1200
A l a	A s n	T r p	P h e	T h r	G y	G y	L e u	A s n	T y r	G n	I l e	G u	H i s	H i s	L e u	
	385			390						395					400	
t t c	c c t	t c g	a t g	c c t	c g c	c a c	a a c	t t t	t c a	a a g	a t c	c a g	c c t	g c t	g t c	1248
P h e	P r o	S e r	M e t	P r o	A r g	H i s	A s n	P h e	S e r	L y s	I l e	G n	P r o	A l a	V a l	
				405					410					415		
g a g	a c c	t t g	t g c	a a a	a a g	t a t	g g t	g t c	c g a	t a c	c a c	a c c	a c t	g g c	a t g	1296
G u	T h r	L e u	C y s	L y s	L y s	T y r	G y	V a l	A r g	T y r	H i s	T h r	T h r	G y	M e t	
			420					425					430			
a t c	g a g	g g a	a c t	g c a	g a g	g t c	t t t	a g c	c g t	t t g	a a c	g a g	g t c	t c c	a a g	1344
I l e	G u	G y	T h r	A l a	G u	V a l	P h e	S e r	A r g	L e u	A s n	G u	V a l	S e r	L y s	
		435					440					445				
g c c	g c c	t c c	a a g	a t g	g g t	a a g	g c g	c a g	t a a							1374
A l a	A l a	S e r	L y s	M e t	G y	L y s	A l a	G n								
	450					455										

<210> 34  
 <211> 457  
 <212> PRT  
 <213> Mbrtierella isabellina

<400> 34  
 M e t A l a A l a A l a P r o S e r V a l A r g T h r P h e T h r A r g A l a G u I l e L e u  
 1 5 10 15  
 A s n A l a G u A l a L e u A s n G u G y L y s L y s A s p A l a G u A l a P r o P h e  
 20 25 30  
 L e u M e t I l e I l e A s p A s n L y s V a l T y r A s p V a l A r g G u P h e V a l P r o  
 35 40 45  
 A s p H i s P r o G y G y S e r V a l I l e L e u T h r H i s V a l G y L y s A s p G y  
 50 55 60  
 T h r A s p V a l P h e A s p T h r P h e H i s P r o G u A l a A l a T r p G u T h r L e u  
 65 70 75 80  
 A l a A s n P h e T y r V a l G y A s p I l e A s p G u S e r A s p A r g A l a I l e L y s  
 85 90 95  
 A s n A s p A s p P h e A l a A l a G u V a l A r g L y s L e u A r g T h r L e u P h e G n  
 100 105 110  
 S e r L e u G y T y r T y r A s p S e r S e r L y s A l a T y r T y r A l a P h e L y s V a l  
 115 120 125  
 S e r P h e A s n L e u C y s I l e T r p G y L e u S e r T h r P h e I l e V a l A l a L y s  
 130 135 140  
 T r p G y G n T h r S e r T h r L e u A l a A s n V a l L e u S e r A l a A l a L e u L e u  
 S e i t e 41



PF58307.txt

145 150 155 160  
 Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu His  
 His Gln Val Phe 165 Gln Asp Arg Phe Trp 170 Gly Asp Leu Phe Gly 175 Ala Phe  
 Leu Gly Gly Val 180 Cys Gln Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys  
 His Asn Thr His His Ala Ala Pro Asn Val His Gly 205 Asp Pro Asp  
 210  
 Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu Glu Met  
 225 230 235  
 Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Arg Met Trp Ser Arg Phe  
 245 250 255  
 Met Val Leu Asn 260 Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala  
 Arg Leu Ser Trp Cys Leu Gln Ser Ile Leu Leu Val Leu Pro Asn Gly  
 275 280  
 Gln Ala His Lys Pro Ser Gly Ala Arg Val Ser Ile Ser Leu Val Glu  
 290 295  
 Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met Phe  
 305 310 315 320  
 Leu Phe Ile Lys Asp Pro Val Asn Met Met Val Tyr Phe Leu Val Ser  
 325 330 335  
 Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn His  
 340 345 350  
 Asn Gly Met Pro Val Ile Ser Lys Glu Glu Ala Val Asp Met Asp Phe  
 355 360 365  
 Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe  
 370 375 380  
 Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu  
 385 390 395 400  
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val  
 405 410 415  
 Glu Thr Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly Met  
 420 425 430 435  
 Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser Lys  
 440 445 450  
 Ala Ala Ser Lys Met Gly Lys Ala Gln  
 450 455

<210> 35  
 <211> 1563  
 <212> DNA  
 <213> Ceratodon purpureus

<220>  
 <221> CDS  
 <222> (1)..(1563)

<400> 35  
 atg gtg tcc cag ggc ggc ggt ct c t cg cag ggt t cc att gaa gaa aac 48  
 Met Val Ser Gln Gly Gly Gly Leu Ser Gln Gly Ser Ile Glu Glu Asn  
 1 5 10 15  
 att gac gtt gag cac ttg gca acg atg ccc ct c gt c agt gac t t c ct a 96  
 Ile Asp Val Glu His Leu Ala Thr Met Pro Leu Val Ser Asp Phe Leu  
 20 25 30  
 aat gtc ctg gga acg act ttg ggc cag tgg agt ct t t cc act aca t t c 144  
 Asn Val Leu Gly Thr Thr Leu Gly Gln Trp Ser Leu Ser Thr Thr Phe  
 35 40 45  
 gct t t c aag agg ct c acg act aag aaa cac agt t cg gac at c t cg gt g 192  
 Ala Phe Lys Arg Leu Thr Thr Lys Lys His Ser Ser Asp Ile Ser Val  
 50 55 60  
 gag gca caa aaa gaa t cg gt t gcg cgg ggg cca gt t gag aat at t t ct 240  
 Glu Ala Gln Lys Glu Ser Val Ala Arg Gly Pro Val Glu Asn Ile Ser  
 65 70 75  
 caa t cg gtt gcg cag ccc at c agg cgg agg t gg gt g cag gat aaa aag 288  
 Gln Ser Val Ala Gln Pro Ile Arg Arg Arg Trp Val Gln Asp Lys Lys  
 85 90 95



PF58307. txt

ccg	gtt	act	tac	agc	ctg	aag	gat	gta	gct	tcg	cac	gat	atg	ccc	cag	336
Pro	Val	Thr	Tyr	Ser	Leu	Lys	Asp	Val	Ala	Ser	His	Asp	Met	Pro	Gln	
			100					105					110			
gac	tgc	tgg	att	ata	atc	aaa	gag	aag	gtg	tat	gat	gtg	agc	acc	ttc	384
Asp	Cys	Trp	Ile	Ile	Ile	Lys	Glu	Lys	Val	Tyr	Asp	Val	Ser	Thr	Phe	
		115					120					125				
gct	gag	cag	cac	cct	gga	ggc	acg	gtt	atc	aac	acc	tac	ttc	gga	cga	432
Ala	Glu	Gln	His	Pro	Gly	Gly	Thr	Val	Ile	Asn	Thr	Tyr	Phe	Gly	Arg	
		130					135					140				
gac	gcc	aca	gat	gtt	ttc	tct	act	ttc	cac	gca	tcc	acc	tca	tgg	aag	480
Asp	Ala	Thr	Asp	Val	Phe	Ser	Thr	Phe	His	Ala	Ser	Thr	Ser	Trp	Lys	
145					150				155					160		
att	ctt	cag	aat	ttc	tac	atc	ggg	aac	ctt	gtt	agg	gag	gag	ccg	act	528
Ile	Leu	Gln	Asn	Phe	Tyr	Ile	Gly	Asn	Leu	Val	Arg	Glu	Glu	Pro	Thr	
				165				170						175		
ttg	gag	ctg	ctg	aag	gag	tac	aga	gag	ttg	aga	gcc	ctt	ttc	ttg	aga	576
Leu	Glu	Leu	Leu	Lys	Glu	Tyr	Arg	Glu	Leu	Arg	Ala	Leu	Phe	Leu	Arg	
				180				185					190			
gaa	cag	ctt	ttc	aag	agt	tcc	aaa	tcc	tac	tac	ctt	ttc	aag	act	ctc	624
Glu	Gln	Leu	Phe	Lys	Ser	Ser	Lys	Ser	Tyr	Tyr	Leu	Phe	Lys	Thr	Leu	
		195					200					205				
ata	aat	gtt	tcc	att	gtt	gcc	aca	agc	att	gcg	ata	atc	agt	ctg	tac	672
Ile	Asn	Val	Ser	Ile	Val	Ala	Thr	Ser	Ile	Ala	Ile	Ile	Ser	Leu	Tyr	
		210				215				220						
aag	tct	tac	cgg	gcg	gtt	ctg	tta	tca	gcc	agt	ttg	atg	ggc	ttg	ttt	720
Lys	Ser	Tyr	Arg	Ala	Val	Leu	Leu	Ser	Ala	Ser	Leu	Met	Gly	Leu	Phe	
225				230					235					240		
att	caa	cag	tgc	gga	tgg	ttg	tct	cac	gat	ttt	cta	cac	cat	cag	gta	768
Ile	Gln	Gln	Cys	Gly	Trp	Leu	Ser	His	Asp	Phe	Leu	His	His	Gln	Val	
				245					250					255		
ttt	gag	aca	cgc	tgg	ctc	aat	gac	gtt	gtt	ggc	tat	gtg	gtc	ggc	aac	816
Phe	Glu	Thr	Arg	Trp	Leu	Asn	Asp	Val	Val	Gly	Tyr	Val	Val	Gly	Asn	
				260				265					270			
gtt	gtt	ctg	gga	ttc	agt	gtc	tcg	tgg	tgg	aag	acc	aag	cac	aac	ctg	864
Val	Val	Leu	Gly	Phe	Ser	Val	Ser	Trp	Trp	Lys	Thr	Lys	His	Asn	Leu	
		275					280					285				
cat	cat	gct	gct	ccg	aat	gaa	tgc	gac	caa	aag	tac	aca	ccg	att	gat	912
His	His	Ala	Ala	Pro	Asn	Glu	Cys	Asp	Gln	Lys	Tyr	Thr	Pro	Ile	Asp	
		290				295					300					
gag	gat	att	gat	act	ctc	ccc	atc	att	gct	tgg	agt	aaa	gat	ctc	ttg	960
Glu	Asp	Ile	Asp	Thr	Leu	Pro	Ile	Ile	Ala	Trp	Ser	Lys	Asp	Leu	Leu	
305					310					315				320		
gcc	act	gtt	gag	agc	aag	acc	atg	ttg	cga	gtt	ctt	cag	tac	cag	cac	1008
Ala	Thr	Val	Glu	Ser	Lys	Thr	Met	Leu	Arg	Val	Leu	Gln	Tyr	Gln	His	
				325					330					335		
cta	ttc	ttt	ttg	gtt	ctt	ttg	acg	ttt	gcc	gcg	agt	tgg	cta	ttt	ttt	1056
Leu	Phe	Phe	Leu	Val	Leu	Leu	Thr	Phe	Ala	Arg	Ala	Ser	Trp	Leu	Phe	
				340				345					350			
tgg	agc	gcg	gcc	ttc	act	ctc	agg	ccc	gag	ttg	acc	ctt	ggc	gag	aag	1104
Trp	Ser	Ala	Ala	Phe	Thr	Leu	Arg	Pro	Glu	Leu	Thr	Leu	Gly	Glu	Lys	
		355					360					365				
ctt	ttg	gag	agg	gga	acg	atg	gct	ttg	cac	tac	att	tgg	ttt	aat	agt	1152
Leu	Leu	Glu	Arg	Gly	Thr	Met	Ala	Leu	His	Tyr	Ile	Trp	Phe	Asn	Ser	
		370				375						380				
gtt	gcg	ttt	tat	ctg	ctc	ccc	gga	tgg	aaa	cca	gtt	gta	tgg	atg	gtg	1200
Val	Ala	Phe	Tyr	Leu	Leu	Pro	Gly	Trp	Lys	Pro	Val	Val	Trp	Met	Val	
385					390					395				400		
gtc	agc	gag	ctc	atg	tct	ggt	ttc	ctg	ctg	gga	tac	gta	ttt	gta	ctc	1248
Val	Ser	Glu	Leu	Met	Ser	Gly	Phe	Leu	Leu	Gly	Tyr	Val	Phe	Val	Leu	
				405				410					415			
agt	cac	aat	gga	atg	gag	gtg	tac	aat	acg	tca	aag	gac	ttc	gtg	aat	1296
Ser	His	Asn	Gly	Met	Glu	Val	Tyr	Asn	Thr	Ser	Lys	Asp	Phe	Val	Asn	
			420					425					430			
gcc	cag	att	gca	tcg	act	cgc	gac	atc	aaa	gca	ggg	gtg	ttt	aat	gat	1344
Ala	Gln	Ile	Ala	Ser	Thr	Arg	Asp	Ile	Lys	Ala	Gly	Val	Phe	Asn	Asp	
		435					440					445				
tgg	ttc	acc	gga	ggt	ctc	aac	aga	cag	att	gag	cat	cat	cta	ttt	cca	1392
Trp	Phe	Thr	Gly	Gly	Leu	Asn	Arg	Gln	Ile	Glu	His	His	Leu	Phe	Pro	

PF58307. txt

450	acg	atg	ccc	agg	cac	aac	ctt	aat	aaa	att	tct	cct	cac	gtg	gag	act	1440
	Thr	Met	Pro	Arg	His	Asn	Leu	Asn	Lys	Ile	Ser	Pro	His	Val	Glu	Thr	
465	ttg	tgc	aag	aag	cat	gga	ctg	gtc	tac	gaa	gac	gtg	agc	atg	gct	tcg	1488
	Leu	Cys	Lys	Lys	His	Gly	Leu	Val	Tyr	Glu	Asp	Val	Ser	Met	Ala	Ser	
					485					490					495		
	ggc	act	tac	cgg	gtt	ttg	aaa	aca	ctt	aag	gac	gtt	gcc	gat	gct	gct	1536
	Gly	Thr	Tyr	Arg	Val	Leu	Lys	Thr	Leu	Lys	Asp	Val	Ala	Asp	Ala	Ala	
				500					505					510			
	tca	cac	cag	cag	ctt	gct	gcg	agt	tga								1563
	Ser	His	Gln	Gln	Leu	Ala	Ala	Ser									
			515					520									

<210> 36  
 <211> 520  
 <212> PRT  
 <213> Ceratodon purpureus

<400> 36  
 Met Val Ser Gln Gly Gly Gly Leu Ser Gln Gly Ser Ile Glu Glu Asn  
 1  
 Ile Asp Val Glu His Leu Ala Thr Met Pro Leu Val Ser Asp Phe Leu  
 20  
 Asn Val Leu Gly Thr Thr Leu Gly Gln Trp Ser Leu Ser Thr Thr Phe  
 35  
 Ala Phe Lys Arg Leu Thr Thr Lys Lys His Ser Ser Asp Ile Ser Val  
 50  
 Glu Ala Gln Lys Glu Ser Val Ala Arg Gly Pro Val Glu Asn Ile Ser  
 65  
 Gln Ser Val Ala Gln Pro Ile Arg Arg Arg Trp Val Gln Asp Lys Lys  
 80  
 Pro Val Thr Tyr Ser Leu Lys Asp Val Ala Ser His Asp Met Pro Gln  
 100  
 Asp Cys Trp Ile Ile Ile Lys Glu Lys Val Tyr Asp Val Ser Thr Phe  
 115  
 Ala Glu Gln His Pro Gly Gly Thr Val Ile Asn Thr Tyr Phe Gly Arg  
 130  
 Asp Ala Thr Asp Val Phe Ser Thr Phe His Ala Ser Thr Ser Trp Lys  
 145  
 Ile Leu Gln Asn Phe Tyr Ile Gly Asn Leu Val Arg Glu Glu Pro Thr  
 160  
 Leu Glu Leu Leu Lys Glu Tyr Arg Glu Leu Arg Ala Leu Phe Leu Arg  
 175  
 Gu Gln Leu Phe Lys Ser Ser Lys Ser Tyr Tyr Leu Phe Lys Thr Leu  
 190  
 Ile Asn Val Ser Ile Val Ala Thr Ser Ile Ala Ile Ile Ser Leu Tyr  
 205  
 Lys Ser Tyr Arg Ala Val Leu Leu Ser Ala Ser Leu Met Gly Leu Phe  
 220  
 Ile Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His His Gln Val  
 235  
 Phe Glu Thr Arg Trp Leu Asn Asp Val Val Gly Tyr Val Val Gly Asn  
 250  
 Val Val Leu Gly Phe Ser Val Ser Trp Trp Lys Thr Lys His Asn Leu  
 265  
 His His Ala Ala Pro Asn Glu Cys Asp Gln Lys Tyr Thr Pro Ile Asp  
 280  
 Gu Asp Ile Asp Thr Leu Pro Ile Ile Ala Trp Ser Lys Asp Leu Leu  
 295  
 Ala Thr Val Glu Ser Lys Thr Met Leu Arg Val Leu Gln Tyr Gln His  
 310  
 Leu Phe Phe Leu Val Leu Leu Thr Phe Ala Arg Ala Ser Trp Leu Phe  
 325  
 Trp Ser Ala Ala Phe Thr Leu Arg Pro Gu Leu Thr Leu Gly Glu Lys  
 340  
 Leu Leu Glu Arg Gly Thr Met Ala Leu His Tyr Ile Trp Phe Asn Ser  
 355  
 370  
 375

PF58307. txt

Val Ala Phe Tyr Leu Leu Pro Gly Trp Lys Pro Val Val Trp Met Val  
 385 390 395 405 410 415 420 425 430 435 440 445 450 460 465 470 475 480 485 490 495 500 510 515  
 Val Ser Glu Leu Met Ser Gly Phe Leu Leu Gly Tyr Val Phe Val Leu  
 Ser His Asn Gly Met Glu Val Tyr Asn Thr Ser Lys Asp Phe Val Asn  
 Ala Gln Ile Ala Ser Thr Arg Asp Ile Lys Ala Gly Val Phe Asn Asp  
 Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro  
 Thr Met Pro Arg His Asn Leu Asn Lys Ile Ser Pro His Val Glu Thr  
 Leu Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Met Ala Ser  
 Gly Thr Tyr Arg Val Leu Lys Thr Leu Lys Asp Val Ala Asp Ala Ala  
 Ser His Gln Gln Leu Ala Ala Ser 520

<210> 37  
 <211> 1452  
 <212> DNA  
 <213> Ceratodon purpureus

<220>  
 <221> CDS  
 <222> (1)..(1452)

<400> 37  
 atg gcc ctg gtt acc gac ttt ctg aac ttt ctg ggc acg aca tgg agc 48  
 Met Ala Leu Val Thr Asp Phe Leu Asn Phe Leu Gly Thr Thr Trp Ser  
 1 5 10 15  
 aag tac agc gtg tac acc cat agc tat gct gga aac tat ggg cct act 96  
 Lys Tyr Ser Val Tyr Thr His Ser Tyr Ala Gly Asn Tyr Gly Pro Thr  
 20 25 30  
 ttg aag cac gcc aaa aag gtt tct gct caa ggt aaa act gcg gga cag 144  
 Leu Lys His Ala Lys Lys Val Ser Ala Gln Gly Lys Thr Ala Gly Gln  
 35 40 45  
 aca ctg aga cag aga tcg gtg cag gac aaa aag cca ggc act tac tct 192  
 Thr Leu Arg Gln Arg Ser Val Gln Asp Lys Lys Pro Gly Thr Tyr Ser  
 50 55 60  
 ctg gcc gat gtt gct tct cac gac agg cct gga gac tgc tgg atg atc 240  
 Leu Ala Asp Val Ala Ser His Asp Arg Pro Gly Asp Cys Trp Met Ile  
 65 70 75 80  
 gtc aaa gag aag gtg tat gat atg agc cgt ttt gcg gac gac cac cct 288  
 Val Lys Glu Lys Val Tyr Asp Ile Ser Arg Phe Ala Asp Asp His Pro  
 85 90 95  
 gga ggg acg gta att agc acc tac ttt ggg cgg gat ggc aca gac gtt 336  
 Gly Gly Thr Val Ile Ser Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val  
 100 105  
 ttc gca aca ttc cat cca cct gcc gca tgg aag caa ctg aat gac tac 384  
 Phe Ala Thr Phe His Pro Pro Ala Ala Trp Lys Gln Leu Asn Asp Tyr  
 115 120 125  
 tac att gga gac ctt gct agg gaa gag ccc ctt gat gaa ttg ctt aaa 432  
 Tyr Ile Gly Asp Leu Ala Arg Glu Glu Pro Leu Asp Glu Leu Leu Lys  
 130 135 140  
 gac tac aga gat atg aga gcc gag ttt gtt aga gaa ggg ctt ttc aag 480  
 Asp Tyr Arg Asp Met Arg Ala Glu Phe Val Arg Glu Gly Leu Phe Lys  
 145 150 155 160  
 agt tcc aag gcc tgg ttc ctg ctt cag act ctg att aat gca gct ctg 528  
 Ser Ser Lys Ala Trp Phe Leu Leu Gln Thr Leu Ile Asn Ala Ala Leu  
 165 170 175  
 ttt gct gcg agc att gcg act atc tgt tac gac aag agt tac tgg gct 576  
 Phe Ala Ala Ser Ile Ala Thr Ile Cys Tyr Asp Lys Ser Tyr Trp Ala  
 180 185 190  
 att gtg ctg tca gcc agt ttg atg ggt ctg ttc gtc caa cag tgt gga 624  
 Ile Val Leu Ser Ala Ser Leu Met Gly Leu Phe Val Gln Gln Cys Gly  
 195 200 205

PF58307.txt

t gg	ct t	gcc	cat	gat	t t c	ct t	cat	caa	cag	gt c	t t t	gag	aac	cg t	acc	672
Tr p	Leu	Al a	Hi s	Asp	Phe	Leu	Hi s	G n	G n	Val	Phe	G u	Asn	Arg	Thr	
	210					215					220					
gcg	aac	t cc	t t c	t t t	ggc	t at	t t g	t t c	ggc	aat	t gc	gt g	ct t	ggc	t t t	720
Al a	Asn	Ser	Phe	Phe	G y	Tyr	Leu	Phe	G y	Asn	Cys	Val	Leu	G y	Phe	
225					230				235						240	
agt	gt a	t ca	t gg	t gg	agg	acg	aag	cac	aac	at t	cat	cat	act	gct	ccg	768
Ser	Val	Ser	Tr p	Tr p	Arg	Thr	Lys	Hi s	Asn	I le	Hi s	Hi s	Thr	Al a	Pro	
				245					250					255		
aat	gag	t gc	gac	gaa	cag	t ac	aca	cct	ct a	gac	gaa	gac	at t	gat	act	816
Asn	G u	Cys	Asp	G u	G n	Tyr	Thr	Pro	Leu	Asp	G u	Asp	I le	Asp	Thr	
			260					265					270			
ct c	ccc	at c	at t	gcc	t gg	agc	aag	gaa	at t	t t g	gcc	acc	gt t	gag	agc	864
Leu	Pro	I le	I le	Al a	Tr p	Ser	Lys	G u	I le	Leu	Al a	Thr	Val	G u	Ser	
		275					280					285				
aag	aga	at t	t t g	cga	gt g	ct t	caa	t at	cag	cac	t ac	at g	at t	ct g	cct	912
Lys	Arg	I le	Leu	Arg	Val	Leu	G n	Tyr	G n	Hi s	Tyr	Met	I le	Leu	Pro	
	290					295					300					
ct a	t t g	t t c	at g	gcc	cgg	t ac	agt	t gg	act	t t t	gga	agt	t t g	ct c	t t c	960
Leu	Leu	Phe	Met	Al a	Arg	Tyr	Ser	Tr p	Thr	Phe	G y	Ser	Leu	Leu	Phe	
305					310					315					320	
aca	t t c	aat	cct	gat	t t g	agc	acg	acc	aag	gga	t t g	at a	gag	aag	gga	1008
Thr	Phe	Asn	Pro	Asp	Leu	Ser	Thr	Thr	Lys	G y	Leu	I le	G u	Lys	G y	
				325					330					335		
aca	gt t	gct	t t t	cac	t ac	gcc	t gg	t t c	agt	t gg	gct	gcg	t t c	cat	at t	1056
Thr	Val	Al a	Phe	Hi s	Tyr	Al a	Tr p	Phe	Ser	Tr p	Al a	Al a	Phe	Hi s	I le	
			340					345					350			
t t g	ccg	ggt	gt c	gct	aag	cct	ct t	cg c	t gg	at g	gt a	gca	act	gag	ct t	1104
Leu	Pro	G y	Val	Al a	Lys	Pro	Leu	Al a	Tr p	Met	Val	Al a	Thr	G u	Leu	
		355				360						365				
gt g	gcc	ggt	t t g	t t g	t t g	gga	t t c	gt g	t t t	acg	t t g	agt	cac	aat	gga	1152
Val	Al a	G y	Leu	Leu	Leu	G y	Phe	Val	Phe	Thr	Leu	Ser	Hi s	Asn	G y	
	370					375					380					
aag	gag	gt t	t ac	aat	gaa	t cg	aag	gac	t t c	gt g	aga	gcc	cag	gt t	at t	1200
Lys	G u	Val	Tyr	Asn	G u	Ser	Lys	Asp	Phe	Val	Arg	Al a	G n	Val	I le	
385					390					395				400		
acc	acc	cg t	aac	acc	aag	cga	ggc	t gg	t t c	aac	gat	t gg	t t c	act	ggg	1248
Thr	Thr	Arg	Asn	Thr	Lys	Arg	G y	Tr p	Phe	Asn	Asp	Tr p	Phe	Thr	G y	
				405					410					415		
gga	ct c	gac	acc	cag	at t	gag	cat	cac	ct g	t t t	cca	aca	at g	ccc	agg	1296
G y	Leu	Asp	Thr	G n	I le	G u	Hi s	Hi s	Leu	Phe	Pro	Thr	Met	Pro	Arg	
			420					425					430			
cac	aac	t ac	ccc	aag	at c	gca	cct	cag	gt c	gag	gct	ct t	t gc	aag	aag	1344
Hi s	Asn	Tyr	Pro	Lys	I le	Al a	Pro	G n	Val	G u	Al a	Leu	Cys	Lys	Lys	
		435				440					445					
cac	ggc	ct c	gag	t ac	gat	aat	gt c	t cc	gt c	gt t	ggt	gcc	t ct	gt c	gcg	1392
Hi s	G y	Leu	G u	Tyr	Asp	Asn	Val	Ser	Val	Val	G y	Al a	Ser	Val	Al a	
	450				455					460						
gt t	gt g	aag	gcg	ct c	aag	gaa	at t	gct	gat	gaa	gcg	t ca	at t	cg g	ct t	1440
Val	Val	Lys	Al a	Leu	Lys	G u	I le	Al a	Asp	G u	Al a	Ser	I le	Arg	Leu	
465					470					475					480	
cac	gct	cac	t aa													1452
Hi s	Al a	Hi s														

<210> 38  
 <211> 483  
 <212> PRT  
 <213> Cer at odon pur pur eus

<400> 38  
 Met Al a Leu Val Thr Asp Phe Leu Asn Phe Leu G y Thr Thr Tr p Ser  
 1 5 10  
 Lys Tyr Ser Val Tyr Thr Hi s Ser Tyr Al a G y Asn Tyr G y Pro Thr  
 20 25 30  
 Leu Lys Hi s Al a Lys Lys Val Ser Al a G n G y Lys Thr Al a G y G n  
 35 40 45  
 Thr Leu Arg G n Arg Ser Val G n Asp Lys Lys Pro G y Thr Tyr Ser  
 Sei te 46

PF58307.txt

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

<210> 39  
 <211> 1374  
 <212> DNA  
 <213> Mbr tierella alpina

<220>  
 <221> CDS  
 <222> (1)..(1374)

<400> 39  
 atg gct gct gct ccc agt gtg agg acg ttt act cgg gcc gag gtt ttg  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu Val Leu  
 Sei te 47

PF58307. txt

1	aat	gcc	gag	gct	ct g	aat	gag	ggc	aag	aag	gat	gcc	gag	gca	ccc	t t c	96
	Asn	Al a	Gl u	Al a	Leu	Asn	Gl u	Gl y	Lys	Lys	Asp	Al a	Gl u	Al a	Pro	Phe	
				20					25					30			
	t t g	at g	at c	at c	gac	aac	aag	gt g	t ac	gat	gt c	cgc	gag	t t c	gt c	cct	144
	Leu	Met	l l e	l l e	Asp	Asn	Lys	Val	Tyr	Asp	Val	Arg	Gl u	Phe	Val	Pro	
				35					40					45			
	gat	cat	ccc	ggt	gga	agt	gt g	at t	ct c	acg	cac	gt t	ggc	aag	gac	ggc	192
	Asp	Hi s	Pro	Gl y	Gl y	Ser	Val	l l e	Leu	Thr	Hi s	Val	Gl y	Lys	Asp	Gl y	
		50						55				60					
	act	gac	gt c	t t t	gac	act	t t t	cac	ccc	gag	gct	gct	t gg	gag	act	ct t	240
	Thr	Asp	Val	Phe	Asp	Thr	Phe	Hi s	Pro	Gl u	Al a	Al a	Tr p	Gl u	Thr	Leu	
	65					70				75					80		
	gcc	aac	t t t	t ac	gt t	ggt	gat	at t	gac	gag	agc	gac	cgc	gat	at c	aag	288
	Al a	Asn	Phe	Tyr	Val	Gl y	Asp	l l e	Asp	Gl u	Ser	Asp	Arg	Asp	l l e	Lys	
					85					90					95		
	aat	gat	gac	t t t	gcg	gcc	gag	gt c	cgc	aag	ct g	cgt	acc	t t g	t t c	cag	336
	Asn	Asp	Asp	Phe	Al a	Al a	Gl u	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	Gl n	
				100					105					110			
	t ct	ct t	ggt	t ac	t ac	gat	t ct	t cc	aag	gca	t ac	t ac	gcc	t t c	aag	gt c	384
	Ser	Leu	Gl y	Tyr	Tyr	Asp	Ser	Ser	Lys	Al a	Tyr	Tyr	Al a	Phe	Lys	Val	
				115				120					125				
	t cg	t t c	aac	ct c	t gc	at c	t gg	ggt	t t g	t cg	acg	gt c	at t	gt g	gcc	aag	432
	Ser	Phe	Asn	Leu	Cys	l l e	Tr p	Gl y	Leu	Ser	Thr	Val	l l e	Val	Al a	Lys	
		130					135					140					
	t gg	ggc	cag	acc	t cg	acc	ct c	gcc	aac	gt g	ct c	t cg	gct	gcg	ct t	t t g	480
	Tr p	Gl y	Gl n	Thr	Ser	Thr	Leu	Al a	Asn	Val	Leu	Ser	Al a	Al a	Leu	Leu	
	145					150				155					160		
	ggt	ct g	t t c	t gg	cag	cag	t gc	gga	t gg	t t g	gct	cac	gac	t t t	t t g	cat	528
	Gl y	Leu	Phe	Tr p	Gl n	Gl n	Cys	Gl y	Tr p	Leu	Al a	Hi s	Asp	Phe	Leu	Hi s	
					165				170						175		
	cac	cag	gt c	t t c	cag	gac	cgt	t t c	t gg	ggt	gat	ct t	t t c	ggc	gcc	t t c	576
	Hi s	Gl n	Val	Phe	Gl n	Asp	Arg	Phe	Tr p	Gl y	Asp	Leu	Phe	Gl y	Al a	Phe	
				180					185					190			
	t t g	gga	ggt	gt c	t gc	cag	ggc	t t c	t cg	t cc	t cg	t gg	t gg	aag	gac	aag	624
	Leu	Gl y	Gl y	Val	Cys	Gl n	Gl y	Phe	Ser	Ser	Ser	Tr p	Tr p	Lys	Asp	Lys	
				195				200						205			
	cac	aac	act	cac	cac	gcc	gcc	ccc	aac	gt c	cac	ggc	gag	gat	ccc	gac	672
	Hi s	Asn	Thr	Hi s	Hi s	Al a	Al a	Pro	Asn	Val	Hi s	Gl y	Gl u	Asp	Pro	Asp	
						215						220					
	at t	gac	acc	cac	cct	ct g	t t g	acc	t gg	agt	gag	cat	gcg	t t g	gag	at g	720
	l l e	Asp	Thr	Hi s	Pro	Leu	Leu	Thr	Tr p	Ser	Gl u	Hi s	Al a	Leu	Gl u	Met	
	225					230					235					240	
	t t c	t cg	gat	gt c	cca	gat	gag	gag	ct g	acc	cgc	at g	t gg	t cg	cgt	t t c	768
	Phe	Ser	Asp	Val	Pro	Asp	Gl u	Gl u	Leu	Thr	Arg	Met	Tr p	Ser	Arg	Phe	
					245					250					255		
	at g	gt c	ct g	aac	cag	acc	t gg	t t t	t ac	t t c	ccc	at t	ct c	t cg	t t t	gcc	816
	Met	Val	Leu	Asn	Gl n	Thr	Tr p	Phe	Tyr	Phe	Pro	l l e	Leu	Ser	Phe	Al a	
				260					265					270			
	cgt	ct c	t cc	t gg	t gc	ct c	cag	t cc	at t	ct c	t t t	gt g	ct g	cct	aac	ggt	864
	Arg	Leu	Ser	Tr p	Cys	Leu	Gl n	Ser	l l e	Leu	Phe	Val	Leu	Pro	Asn	Gl y	
				275				280						285			
	cag	gcc	cac	aag	ccc	t cg	ggc	gcg	cgt	gt g	ccc	at c	t cg	t t g	gt c	gag	912
	Gl n	Al a	Hi s	Lys	Pro	Ser	Gl y	Al a	Arg	Val	Pro	l l e	Ser	Leu	Val	Gl u	
							295					300					
	cag	ct g	t cg	ct t	gcg	at g	cac	t gg	acc	t gg	t ac	ct c	gcc	acc	at g	t t c	960
	Gl n	Leu	Ser	Leu	Al a	Met	Hi s	Tr p	Thr	Tr p	Tyr	Leu	Al a	Thr	Met	Phe	
	305					310					315					320	
	ct g	t t c	at c	aag	gat	ccc	gt c	aac	at g	ct g	gt g	t ac	t t t	t t g	gt g	t cg	1008
	Leu	Phe	l l e	Lys	Asp	Pro	Val	Asn	Met	Leu	Val	Tyr	Phe	Leu	Val	Ser	
					325					330					335		
	cag	gcg	gt g	t gc	gga	aac	t t g	t t g	gcg	at c	gt g	t t c	t cg	ct c	aac	cac	1056
	Gl n	Al a	Val	Cys	Gl y	Asn	Leu	Leu	Al a	l l e	Val	Phe	Ser	Leu	Asn	Hi s	
				340					345					350			
	aac	ggt	at g	cct	gt g	at c	t cg	aag	gag	gag	gcg	gt c	gat	at g	gat	t t c	1104
	Asn	Gl y	Met	Pro	Val	l l e	Ser	Lys	Gl u	Gl u	Al a	Val	Asp	Met	Asp	Phe	
				355				360					365				
	t t c	acg	aag	cag	at c	at c	acg	ggt	cgt	gat	gt c	cac	ccg	ggt	ct a	t t t	1152

PF58307.txt

Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	His	Pro	Gly	Leu	Phe	
	370					375					380					
gcc	aac	tgg	ttc	acg	ggg	gga	ttg	aac	tat	cag	atc	gag	cac	cac	ttg	1200
Ala	Asn	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	His	Leu	
	385				390					395					400	
ttc	cct	tcg	atg	cct	cgc	cac	aac	ttt	tca	aag	atc	cag	cct	gct	gtc	1248
Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	Gln	Pro	Ala	Val	
				405				410						415		
gag	acc	ctg	tgc	aaa	aag	tac	aat	gtc	cga	tac	cac	acc	acc	ggg	atg	1296
Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Asn	Val	Arg	Tyr	His	Thr	Thr	Gly	Met	
			420					425					430			
atc	gag	gga	act	gca	gag	gtc	ttt	agc	cgt	ctg	aac	gag	gtc	tcc	aag	1344
Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	Val	Ser	Lys	
		435				440						445				
gct	gcc	tcc	aag	atg	ggg	aag	gcg	cag	t aa							1374
Ala	Ala	Ser	Lys	Met	Gly	Lys	Ala	Gln								
	450				455											

<210> 40  
 <211> 457  
 <212> PRT  
 <213> Mbrtierella alpina

<400> 40

Met	Ala	Ala	Ala	Pro	Ser	Val	Arg	Thr	Phe	Thr	Arg	Ala	Glu	Val	Leu	
1				5					10					15		
Asn	Ala	Glu	Ala	Leu	Asn	Glu	Gly	Lys	Lys	Asp	Ala	Glu	Ala	Pro	Phe	
			20					25					30			
Leu	Met	Ile	Ile	Asp	Asn	Lys	Val	Tyr	Asp	Val	Arg	Glu	Phe	Val	Pro	
		35				40						45				
Asp	His	Pro	Gly	Gly	Ser	Val	Ile	Leu	Thr	His	Val	Gly	Lys	Asp	Gly	
	50					55					60					
Thr	Asp	Val	Phe	Asp	Thr	Phe	His	Pro	Glu	Ala	Ala	Trp	Glu	Thr	Leu	
65					70					75					80	
Ala	Asn	Phe	Tyr	Val	Gly	Asp	Ile	Asp	Glu	Ser	Asp	Arg	Asp	Ile	Lys	
				85					90					95		
Asn	Asp	Asp	Phe	Ala	Ala	Glu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	Gln	
			100					105					110			
Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Ala	Tyr	Tyr	Ala	Phe	Lys	Val	
		115				120						125				
Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Thr	Val	Ile	Val	Ala	Lys	
	130					135					140					
Trp	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Val	Leu	Ser	Ala	Ala	Leu	Leu	
145					150					155					160	
Gly	Leu	Phe	Trp	Gln	Cys	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His		
				165				170					175			
His	Gln	Val	Phe	Gln	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	Ala	Phe	
			180					185					190			
Leu	Gly	Gly	Val	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	Lys	
		195					200					205				
His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp	
	210				215						220					
Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met	
225					230					235					240	
Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe	
				245					250					255		
Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala	
			260				265						270			
Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	Gly	
		275					280						285			
Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu	
	290					295						300				
Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe	
305					310					315					320	
Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Leu	Val	Tyr	Phe	Leu	Val	Ser	
				325					330					335		
Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His	
			340					345						350		



PF58307.txt

Asn Gly Met Pro Val Ile Ser Lys Gu Gu Ala Val Asp Met Asp Phe  
 355 360 365  
 Phe Thr Lys G n Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe  
 370 375 380  
 Ala Asn Trp Phe Thr Gly Leu Asn Tyr G n Ile Gu His His Leu  
 385 390 395 400  
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile G n Pro Ala Val  
 405 410 415  
 Gu Thr Leu Cys Lys Lys Tyr Asn Val Arg Tyr His Thr Thr Gly Met  
 420 425 430  
 Ile Gu Gly Thr Ala Gu Val Phe Ser Arg Leu Asn Gu Val Ser Lys  
 435 440 445  
 Ala Ala Ser Lys Met Gly Lys Ala G n  
 450 455

<210> 41  
 <211> 1374  
 <212> DNA  
 <213> Mbr t i e r e l l a a l p i n a

<220>  
 <221> CDS  
 <222> (1)..(1374)

<400> 41  
 atg gct gct gct ccc agt gtg agg acg ttt act cgg gcc gag att ttg 48  
 Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Gu Ile Leu  
 1 5 10 15  
 aat gcc gag gcc ctg aat gag ggc aag aag gat gcc gag gca ccc ttt 96  
 Asn Ala Gu Ala Leu Asn Gu Gy Lys Lys Asp Ala Gu Ala Pro Phe  
 20 25 30  
 ctg atg atc att gac aac aag gtg tac gat gtc cgc gag ttt gtc cct 144  
 Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Gu Phe Val Pro  
 35 40 45  
 gat cat ccc ggt gga agt gtg att ct c acg cac gtt ggc aag gac ggc 192  
 Asp His Pro Gy Gy Ser Val Ile Leu Thr His Val Gy Lys Asp Gy  
 50 55 60  
 act gac gtc ttt gac act ttc cac ccc gag gct gct tgg gag act ctt 240  
 Thr Asp Val Phe Asp Thr Phe His Pro Gu Ala Ala Trp Gu Thr Leu  
 65 70 75 80  
 gcc aac ttt tac gtt ggt gat att gat gag agc gat cgt gcc atc aag 288  
 Ala Asn Phe Tyr Val Gy Asp Ile Asp Gu Ser Asp Arg Ala Ile Lys  
 85 90 95  
 aat gat gac ttt gcg gcc gag gtt cgc aag ctg cgc acc ttg ttc cag 336  
 Asn Asp Asp Phe Ala Ala Gu Val Arg Lys Leu Arg Thr Leu Phe G n  
 100 105 110  
 tcc ctt ggc tac tac gac tgc tcc aag gca tac tat gcc ttc aag gtc 384  
 Ser Leu Gy Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val  
 115 120 125  
 tgc ttc aac ct c tgc at c tgg ggc ttg tgc act ttc att gtt gcc aag 432  
 Ser Phe Asn Leu Cys Ile Trp Gy Leu Ser Thr Phe Ile Val Ala Lys  
 130 135 140  
 tgg ggc cag acc tgc acc ct c gcc aac gtg ct c tgc gct gcg ct c ttg 480  
 Trp Gy G n Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu  
 145 150 155 160  
 ggt ct c ttc tgg cag cag tgc gga tgg ttg gcg cac gac ttt ttg cac 528  
 Gy Leu Phe Trp G n G n Cys Gy Trp Leu Ala His Asp Phe Leu His  
 165 170 175  
 cac cag gtc ttc cag gac cgt ttc tgg ggt gat ct t ttc ggc gcc ttc 576  
 His G n Val Phe G n Asp Arg Phe Trp Gy Asp Leu Phe Gy Ala Phe  
 180 185 190  
 ttg gga ggt gtc tgc cag ggt ttc tgc tcc tcc tgg tgg aag gac aag 624  
 Leu Gy Gy Val Cys G n Gy Phe Ser Ser Ser Trp Trp Lys Asp Lys  
 195 200 205  
 cac aac act cac cac gct gct ccc aac gt c cac ggc gag gat ccc gac 672  
 His Asn Thr His His Ala Ala Pro Asn Val His Gy Gu Asp Pro Asp  
 210 215 220  
 att gac act cac cct ct g ttg acc tgg agt gag cat gct ct g gag at g 720



PF58307.txt

I l e	A s p	T h r	H i s	P r o	L e u	L e u	T h r	T r p	S e r	G l u	H i s	A l a	L e u	G l u	M e t	
225					230					235					240	
t t c	t c g	g a t	g t t	c c t	g a c	g a g	g a g	c t g	a c c	c g t	a t g	t g g	t c g	c g c	t t c	768
P h e	S e r	A s p	V a l	P r o	A s p	G l u	G l u	L e u	T h r	A r g	M e t	T r p	S e r	A r g	P h e	
				245						250				255		
a t g	g t c	c t c	a a c	c a g	a c c	t g g	t t c	t a c	t t c	c c c	a t t	c t c	t c g	t t t	g c c	816
M e t	V a l	L e u	A s n	G n	T h r	T r p	P h e	T y r	P h e	P r o	I l e	L e u	S e r	P h e	A l a	
			260					265					270			
c g t	c t g	t c c	t g g	t g c	c t c	c a g	t c c	a t t	a t g	t t t	g t t	c t g	c c c	a a c	g g t	864
A r g	L e u	S e r	T r p	C y s	L e u	G n	S e r	I l e	M e t	P h e	V a l	L e u	P r o	A s n	G l y	
			275				280					285				
c a g	g c c	c a c	a a g	c c c	t c t	g g a	g c g	c g t	g t g	c c c	a t t	t c g	t t g	g t c	g a g	912
G n	A l a	H i s	L y s	P r o	S e r	G l y	A l a	A r g	V a l	P r o	I l e	S e r	L e u	V a l	G l u	
	290					295				300						
c a g	c t g	t c t	c t g	g c t	a t g	c a c	t g g	a c c	t g g	t a c	c t c	g c c	a c c	a t g	t t c	960
G n	L e u	S e r	L e u	A l a	M e t	H i s	T r p	T h r	T r p	T y r	L e u	A l a	T h r	M e t	P h e	
	305			310						315					320	
c t g	t t c	a t t	a a g	g a t	c c c	g t c	a a c	a t g	a t t	g t g	t a c	t t t	t t g	g t g	t c g	1008
L e u	P h e	I l e	L y s	A s p	P r o	V a l	A s n	M e t	I l e	V a l	T y r	P h e	L e u	V a l	S e r	
				325					330					335		
c a g	g c t	g t t	t g c	g g c	a a c	t t g	t t g	g c g	a t t	g t g	t t c	t c g	c t c	a a c	c a c	1056
G n	A l a	V a l	C y s	G l y	A s n	L e u	L e u	A l a	I l e	V a l	P h e	S e r	L e u	A s n	H i s	
			340					345					350			
a a c	g g c	a t g	c c t	g t g	a t c	t c c	a a g	g a g	g a a	g c g	g t c	g a t	a t g	g a c	t t c	1104
A s n	G l y	M e t	P r o	V a l	I l e	S e r	L y s	G l u	G l u	A l a	V a l	A s p	M e t	A s p	P h e	
	355					360						365				
t t c	a c c	a a g	c a g	a t c	a t c	a c g	g g t	c g t	g a t	g t t	c a c	c c t	g g t	c t g	t t t	1152
P h e	T h r	L y s	G n	I l e	I l e	T h r	G l y	A r g	A s p	V a l	H i s	P r o	G l y	L e u	P h e	
	370					375					380					
g c c	a a c	t g g	t t c	a c g	g g t	g g a	t t g	a a c	t a c	c a g	a t t	g a g	c a c	c a c	t t g	1200
A l a	A s n	T r p	P h e	T h r	G l y	G l y	L e u	A s n	T y r	G n	I l e	G l u	H i s	H i s	L e u	
	385				390					395					400	
t t c	c c t	t c g	a t g	c c c	c g c	c a c	a a c	t t t	t c a	a a g	a t c	c a g	c c t	g c t	g t c	1248
P h e	P r o	S e r	M e t	P r o	A r g	H i s	A s n	P h e	S e r	L y s	I l e	G n	P r o	A l a	V a l	
				405				410						415		
g a g	a c t	t t g	t g c	a a a	a a g	t a c	g g t	g t c	c g a	t a c	c a t	a c c	a c t	g g t	a t g	1296
G l u	T h r	L e u	C y s	L y s	L y s	T y r	G l y	V a l	A r g	T y r	H i s	T h r	T h r	G l y	M e t	
			420				425						430			
a t c	g a g	g g a	a c t	g c a	g a g	g t c	t t t	a g c	c g t	t t g	a a c	g a g	g t c	t c c	a a g	1344
I l e	G l u	G l y	T h r	A l a	G l u	V a l	P h e	S e r	A r g	L e u	A s n	G l u	V a l	S e r	L y s	
		435				440						445				
g c g	g c c	t c c	a a g	a t g	g g c	a a g	g c a	c a g	t a a							1374
A l a	A l a	S e r	L y s	M e t	G l y	L y s	A l a	G n								
	450				455											

<210> 42  
 <211> 457  
 <212> PRT  
 <213> Mbr t i e r e l l a a l p i n a

<400> 42  
 M e t A l a A l a A l a P r o S e r V a l A r g T h r P h e T h r A r g A l a G l u I l e L e u  
 1 5 10 15  
 A s n A l a G l u A l a L e u A s n G l u G l y L y s L y s A s p A l a G l u A l a P r o P h e  
 20 30  
 L e u M e t I l e I l e A s p A s n L y s V a l T y r A s p V a l A r g G l u P h e V a l P r o  
 35 40 45  
 A s p H i s P r o G l y G l y S e r V a l I l e L e u T h r H i s V a l G l y L y s A s p G l y  
 50 55 60  
 T h r A s p V a l P h e A s p T h r P h e H i s P r o G l u A l a A l a T r p G l u T h r L e u  
 65 70 75 80  
 A l a A s n P h e T y r V a l G l y A s p I l e A s p G l u S e r A s p A r g A l a I l e L y s  
 85 90 95  
 A s n A s p A s p P h e A l a A l a G l u V a l A r g L y s L e u A r g T h r L e u P h e G n  
 100 105 110  
 S e r L e u G l y T y r T y r A s p S e r S e r L y s A l a T y r T y r A l a P h e L y s V a l  
 115 120 125  
 S e r P h e A s n L e u C y s I l e T r p G l y L e u S e r T h r P h e I l e V a l A l a L y s  
 S e i t e 51

PF58307.txt

Trp	130	Gly	Gln	Thr	Ser	Thr	135	Leu	Ala	Asn	Val	Leu	140	Ser	Ala	Ala	Leu	Leu
145	Gly	Leu	Phe	Trp	Gln	150	Cys	Gly	Trp	Leu	155	Ala	His	Asp	Phe	Leu	160	His
His	Gln	Val	Phe	Gln	165	Asp	Arg	Phe	Trp	Gly	170	Asp	Leu	Phe	Gly	175	Ala	Phe
Leu	Gly	Gly	180	Val	Cys	Gln	Gly	Phe	Ser	Ser	185	Ser	Trp	Trp	Lys	190	Asp	Lys
His	Asn	Thr	195	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	200	Glu	Asp	Pro	Asp	Asp
210	Ile	Asp	Thr	His	Pro	Leu	215	Leu	Thr	Trp	Ser	Glu	220	His	Ala	Leu	Glu	Met
225	Phe	Ser	Asp	Val	Pro	230	Asp	Glu	Glu	Leu	Thr	235	Arg	Met	Trp	Ser	Arg	Phe
Met	Val	Leu	Asn	Gln	Thr	245	Trp	Phe	Tyr	Phe	Pro	Ile	250	Leu	Ser	Phe	Ala	Ala
Arg	Leu	Ser	260	Trp	Cys	Leu	Gln	Ser	Ile	Met	Phe	Val	265	Leu	Pro	Asn	Gly	Gly
Gln	Ala	His	Lys	Pro	Ser	Gly	275	Ala	Arg	Val	Pro	Ile	280	Ser	Leu	Val	Glu	Glu
Gln	Leu	Ser	Leu	Ala	Met	His	285	Trp	Thr	Trp	Tyr	Leu	290	Ala	Thr	Met	Phe	Phe
305	Leu	Phe	Ile	Lys	Asp	Pro	310	Val	Asn	Met	Ile	Val	315	Tyr	Phe	Leu	Val	Ser
Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	320	Leu	Asn	His	His	His
Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	325	Met	Asp	Phe	Phe	Phe
Phe	Thr	Lys	Gln	Ile	Ile	Thr	330	Gly	Arg	Asp	Val	His	335	Pro	Gly	Leu	Phe	Phe
Ala	Asn	Trp	Phe	Thr	Gly	Gly	340	Leu	Asn	Tyr	Gln	Ile	345	Glu	His	His	Leu	Leu
385	Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	350	Gln	Pro	Ala	Val	Val
Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Gly	Val	Arg	Tyr	His	Thr	355	Thr	Thr	Gly	Met	Met
Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	360	Val	Ser	Lys	Lys	Lys
Ala	Ala	Ser	Lys	Met	Gly	Lys	Ala	Gln					365					
450						455							440					

<210> 43  
 <211> 444  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (5)..(10)  
 <223> xaa in position 5 to 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is any amino acid  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(20)  
 <223> xaa in position 19 to 20 is any amino acid

<220>  
<221> Variant  
<222> (22)..(23)  
<223> xaa in position 22 to 23 is any amino acid  
<220>  
<221> Variant  
<222> (25)..(25)  
<223> xaa in position 25 is any amino acid  
<220>  
<221> Variant  
<222> (27)..(27)  
<223> xaa in position 27 is any amino acid  
<220>  
<221> Variant  
<222> (30)..(30)  
<223> xaa in position 30 is any amino acid  
<220>  
<221> Variant  
<222> (32)..(34)  
<223> xaa in position 32 to 34 is any amino acid  
<220>  
<221> Variant  
<222> (37)..(38)  
<223> xaa in position 37 to 38 is any amino acid  
<220>  
<221> Variant  
<222> (40)..(42)  
<223> xaa in position 40 to 42 is any amino acid  
<220>  
<221> Variant  
<222> (44)..(46)  
<223> xaa in position 44 to 46 is any amino acid  
<220>  
<221> Variant  
<222> (48)..(48)  
<223> xaa in position 48 is any amino acid  
<220>  
<221> Variant  
<222> (51)..(59)  
<223> xaa in position 51 to 59 is any amino acid  
<220>  
<221> Variant  
<222> (60)..(77)  
<223> xaa in position 60 to 77 is any or no amino acid  
<220>  
<221> Variant  
<222> (79)..(82)  
<223> xaa in position 79 to 82 is any amino acid  
<220>  
<221> Variant  
<222> (84)..(84)  
<223> xaa in position 84 is any amino acid  
<220>  
<221> Variant  
<222> (87)..(92)  
<223> xaa in position 87 to 92 is any amino acid  
<220>  
<221> Variant  
<222> (94)..(97)  
<223> xaa in position 94 to 97 is any amino acid  
<220>  
<221> Variant  
<222> (99)..(101)  
<223> xaa in position 99 to 101 is any amino acid  
<220>  
<221> Variant  
<222> (103)..(104)  
<223> xaa in position 103 to 104 is any amino acid

<220>  
<221> Vari ant  
<222> ( 106).. ( 132)  
<223> xaa in position 106 to 132 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 133).. ( 134)  
<223> xaa in position 133 to 134 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 136).. ( 138)  
<223> xaa in position 136 to 138 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 140).. ( 140)  
<223> xaa in position 140 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 142).. ( 142)  
<223> xaa in position 142 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 145).. ( 145)  
<223> xaa in position 145 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 153).. ( 153)  
<223> xaa in position 153 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 159).. ( 160)  
<223> xaa in position 159 to 160 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 162).. ( 171)  
<223> xaa in position 162 to 171 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 172).. ( 172)  
<223> xaa in position 172 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 174).. ( 176)  
<223> xaa in position 174 to 176 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 181).. ( 182)  
<223> xaa in position 181 to 182 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 186).. ( 186)  
<223> xaa in position 186 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 190).. ( 190)  
<223> xaa in position 190 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 194).. ( 195)  
<223> xaa in position 194 to 195 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 197).. ( 200)  
<223> xaa in position 197 to 200 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 201).. ( 207)  
<223> xaa in position 201 to 207 is any or no ami no aci d

<220>  
<221> Variant  
<222> (214)..(214)  
<223> xaa in position 214 is any amino acid  
<220>  
<221> Variant  
<222> (216)..(216)  
<223> xaa in position 216 is any amino acid  
<220>  
<221> Variant  
<222> (218)..(218)  
<223> xaa in position 218 is any amino acid  
<220>  
<221> Variant  
<222> (220)..(238)  
<223> xaa in position 220 to 238 is any amino acid  
<220>  
<221> Variant  
<222> (239)..(250)  
<223> xaa in position 239 to 250 is any or no amino acid  
<220>  
<221> Variant  
<222> (252)..(255)  
<223> xaa in position 252 to 255 is any amino acid  
<220>  
<221> Variant  
<222> (257)..(258)  
<223> xaa in position 257 to 258 is any amino acid  
<220>  
<221> Variant  
<222> (260)..(261)  
<223> xaa in position 260 to 261 is any amino acid  
<220>  
<221> Variant  
<222> (264)..(264)  
<223> xaa in position 264 is any amino acid  
<220>  
<221> Variant  
<222> (267)..(269)  
<223> xaa in position 267 to 269 is any amino acid  
<220>  
<221> Variant  
<222> (271)..(284)  
<223> xaa in position 271 to 284 is any amino acid  
<220>  
<221> Variant  
<222> (285)..(296)  
<223> xaa in position 285 to 296 is any or no amino acid  
<220>  
<221> Variant  
<222> (298)..(298)  
<223> xaa in position 298 is any amino acid  
<220>  
<221> Variant  
<222> (300)..(305)  
<223> xaa in position 300 to 305 is any amino acid  
<220>  
<221> Variant  
<222> (307)..(326)  
<223> xaa in position 307 to 326 is any amino acid  
<220>  
<221> Variant  
<222> (327)..(328)  
<223> xaa in position 327 to 328 is any or no amino acid  
<220>  
<221> Variant  
<222> (330)..(336)  
<223> xaa in position 330 to 336 is any amino acid

<220>  
<221> Vari ant  
<222> ( 338) .. ( 339)  
<223> xaa in position 338 to 339 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 342) .. ( 342)  
<223> xaa in position 342 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 345) .. ( 347)  
<223> xaa in position 345 to 347 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 352) .. ( 352)  
<223> xaa in position 352 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 354) .. ( 359)  
<223> xaa in position 354 to 359 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 360) .. ( 363)  
<223> xaa in position 360 to 363 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 365) .. ( 367)  
<223> xaa in position 365 to 367 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 369) .. ( 370)  
<223> xaa in position 369 to 370 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 372) .. ( 372)  
<223> xaa in position 372 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 374) .. ( 379)  
<223> xaa in position 374 to 379 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 380) .. ( 384)  
<223> xaa in position 380 to 384 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 386) .. ( 387)  
<223> xaa in position 386 to 387 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 390) .. ( 390)  
<223> xaa in position 390 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 404) .. ( 405)  
<223> xaa in position 404 to 405 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 410) .. ( 411)  
<223> xaa in position 410 to 411 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 413) .. ( 416)  
<223> xaa in position 413 to 416 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 418) .. ( 419)  
<223> xaa in position 418 to 419 is any ami no aci d

<220>  
 <221> Variant  
 <222> (423)..(427)  
 <223> xaa in position 423 to 427 is any amino acid  
 <220>  
 <221> Variant  
 <222> (430)..(435)  
 <223> xaa in position 430 to 435 is any amino acid  
 <220>  
 <221> Variant  
 <222> (438)..(438)  
 <223> xaa in position 438 is any amino acid  
 <220>  
 <221> Variant  
 <222> (441)..(443)  
 <223> xaa in position 441 to 443 is any amino acid

<400> 43  
 Lys Val Tyr Asp xaa xaa xaa xaa xaa xaa xaa His Pro Gly Gly xaa  
 1 5 10 15  
 xaa Val xaa xaa Thr xaa xaa Gly xaa Asp xaa Thr Asp xaa Phe xaa  
 20 25 30  
 xaa xaa His Pro xaa xaa Ala xaa xaa xaa Leu xaa xaa xaa Tyr xaa  
 35 40 45  
 Gly Asp xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 50 55 60  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Phe xaa xaa  
 65 70 75 80  
 xaa xaa Arg xaa Leu Arg xaa xaa xaa xaa xaa xaa Gly xaa xaa xaa  
 85 90 95  
 xaa Ser xaa xaa xaa Tyr xaa xaa Lys xaa xaa xaa xaa xaa xaa xaa  
 100 105 110  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 115 120 125  
 xaa xaa xaa xaa xaa xaa Ala xaa xaa xaa Gly xaa Phe xaa Gn Gn  
 130 135 140  
 xaa Gly Trp Leu Ala His Asp Phe xaa His His Gn Val Phe xaa xaa  
 145 150 155 160  
 Arg xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa xaa xaa  
 165 170 175  
 Gn Gly Phe Ser xaa xaa Trp Trp Lys xaa Lys His Asn xaa His His  
 180 185 190  
 Ala xaa xaa Asn xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Asp  
 195 200 205  
 Pro Asp Ile Asp Thr xaa Pro xaa Leu xaa Trp xaa xaa xaa xaa xaa  
 210 215 220  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 225 230 235 240  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa xaa xaa xaa Phe  
 245 250 255  
 xaa xaa Leu xaa xaa Ala Arg xaa Ser Trp xaa xaa xaa Ser xaa xaa  
 260 265 270  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 275 280 285  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 290 295 300  
 xaa His xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 305 310 315 320  
 xaa xaa xaa xaa xaa xaa xaa xaa Phe xaa xaa xaa xaa xaa xaa xaa  
 325 330 335  
 Gly xaa xaa Leu Ala xaa Val Phe xaa xaa xaa His Asn Gly Met xaa  
 340 345 350  
 Val xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Phe xaa xaa xaa Gn  
 355 360 365  
 xaa xaa Thr xaa Arg xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 370 375 380  
 Phe xaa xaa Trp Phe xaa Gly Gly Leu Asn Tyr Gn Ile Gu His His  
 385 390 395 400





<223> xaa in position 26 is Glu or Lys  
 <220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is Ala, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (30)..(30)  
 <223> xaa in position 30 is any or no amino acid  
  
 <400> 44  
 Leu xaa xaa Gln Ile xaa His His xaa Phe Pro xaa xaa Pro Arg His  
   1          5          10          15  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Leu Cys xaa Lys  
           20          25          30

<210> 45  
 <211> 27  
 <212> PRT  
 <213> Artificial sequence  
  
 <220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Asp, Gly or Asn  
 <220>  
 <221> Variant  
 <222> (3)..(4)  
 <223> xaa in position 3 to 4 is any amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is any amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is Ala, Gln or Ser  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Asp or Asn  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Ala, Ser or Val  
 <220>  
 <221> Variant  
 <222> (23)..(23)  
 <223> xaa in position 23 is Pro or Thr  
 <220>  
 <221> Variant  
 <222> (25)..(25)  
 <223> xaa in position 25 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (26)..(26)  
 <223> xaa in position 26 is any amino acid  
 <220>  
 <221> Variant

&lt;222&gt; (27)..(27)

&lt;223&gt; xaa in position 27 is Glu, Gly, Asn or Ser

&lt;400&gt; 45

Gly xaa xaa xaa Gln Gly Phe Ser xaa xaa Trp Trp Lys xaa Lys His  
 1 5 10 15  
 Asn xaa His His xaa xaa xaa Asn xaa xaa xaa  
 20 25

&lt;210&gt; 46

&lt;211&gt; 31

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (2)..(3)

&lt;223&gt; xaa in position 2 to 3 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (5)..(5)

&lt;223&gt; xaa in position 5 is Glu, Lys or Arg

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (7)..(7)

&lt;223&gt; xaa in position 7 is Ala or Gly

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (8)..(8)

&lt;223&gt; xaa in position 8 is Ser or Thr

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (10)..(10)

&lt;223&gt; xaa in position 10 is Ala, Ser or Val

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (12)..(13)

&lt;223&gt; xaa in position 12 to 13 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (14)..(14)

&lt;223&gt; xaa in position 14 is Phe or Met

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (16)..(16)

&lt;223&gt; xaa in position 16 is Pro or Ser

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (17)..(18)

&lt;223&gt; xaa in position 17 to 18 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (20)..(22)

&lt;223&gt; xaa in position 20 to 22 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (24)..(25)

&lt;223&gt; xaa in position 24 to 25 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (26)..(26)

&lt;223&gt; xaa in position 26 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (28)..(28)

&lt;223&gt; xaa in position 28 is any amino acid

&lt;220&gt;

<221> Variant  
 <222> (29)..(29)  
 <223> xaa in position 29 is any or no amino acid

<400> 46  
 Thr xaa xaa Gly xaa Asp xaa xaa Asp xaa Phe xaa xaa xaa His xaa  
 1 5 10 15  
 xaa xaa Ala xaa xaa xaa Leu xaa xaa xaa Tyr xaa xaa Gly Asp  
 20 25 30

<210> 47  
 <211> 26  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Phe, Leu or Val  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is any amino acid  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is Ala, Gly or Ser  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Gln or Thr  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Ala, Cys or Pro  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Ser, Thr or Val  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Cys, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (10)..(11)  
 <223> xaa in position 10 to 11 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (17)..(18)  
 <223> xaa in position 17 to 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(20)  
 <223> xaa in position 19 to 20 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (22)..(23)  
 <223> xaa in position 22 to 23 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (26)..(26)  
 <223> xaa in position 26 is Ala, Pro or Ser

&lt;400&gt; 47

Phe xaa xaa xaa xaa xaa xaa xaa Gly xaa xaa Leu Ala xaa Val Phe  
 1 5 10 15  
 xaa xaa xaa Asn xaa xaa Gly Met xaa  
 20 25

&lt;210&gt; 48

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (6)..(6)

&lt;223&gt; xaa in position 6 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (8)..(8)

&lt;223&gt; xaa in position 8 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (11)..(11)

&lt;223&gt; xaa in position 11 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (13)..(13)

&lt;223&gt; xaa in position 13 is Asp or Ser

&lt;400&gt; 48

Pro Asp Ile Asp Thr xaa Pro xaa Leu Leu xaa Trp xaa  
 1 5 10

&lt;210&gt; 49

&lt;211&gt; 17

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (2)..(3)

&lt;223&gt; xaa in position 2 to 3 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (4)..(4)

&lt;223&gt; xaa in position 4 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (7)..(8)

&lt;223&gt; xaa in position 7 to 8 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (9)..(9)

&lt;223&gt; xaa in position 9 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (11)..(11)

&lt;223&gt; xaa in position 11 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (12)..(12)

&lt;223&gt; xaa in position 12 is Phe or Leu

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (15)..(15)

&lt;223&gt; xaa in position 15 is Ile or Leu

Seite 62

<400> 49  
 G n xaa xaa xaa Tyr Phe xaa xaa xaa Leu xaa xaa Al a Arg xaa Ser  
 1 5 10 15  
 Trp

<210> 50  
 <211> 20  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Asp, Glu or Thr  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is Ala or Thr  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is Asp, Gly, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (5)..(7)  
 <223> xaa in position 5 to 7 is any amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(10)  
 <223> xaa in position 9 to 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Phe, Leu or Val  
 <220>  
 <221> Variant  
 <222> (14)..(15)  
 <223> xaa in position 14 to 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Ala, Asp, Gly, Asn or Ser  
 <220>  
 <221> Variant  
 <222> (18)..(19)  
 <223> xaa in position 18 to 19 is any amino acid  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Ala, Gln or Ser

<400> 50  
 H s xaa xaa xaa xaa xaa xaa Gly xaa xaa Glu Val xaa xaa xaa Leu  
 1 5 10 15  
 xaa xaa xaa xaa  
 20

<210> 51  
 <211> 1320  
 <212> DNA  
 <213> *Thraustochytrium* ssp.

<220>  
 <221> CDS  
 <222> (1)..(1320)

```

<400> 51
at g ggc aag ggc agc gag ggc cgc agc gcg gcg cgc gag at g acg gcc      48
Met Gly Lys Gly Ser Gu Gly Arg Ser Ala Ala Arg Gu Met Thr Ala
1 5
gag gcg aac ggc gac aag cgg aaa acg at t ct g at c gag ggc gt c ct g      96
Gu Ala Asn Gly Asp Lys Arg Lys Thr lle Leu lle Gu Gly Val Leu
20
t ac gac gcg acg aac t t t aag cac cgc ggc ggt t cg at c at c aac t t c      144
Tyr Asp Ala Thr Asn Phe Lys His Pro Gly Gly Ser lle lle Asn Phe
35
t t g acc gag ggc gag gcc ggc gt g gac gcg acg cag gcg t ac cgc gag      192
Leu Thr Gu Gly Gu Ala Gly Val Asp Ala Thr Gn Ala Tyr Arg Gu
50
t t t cat cag cgg t cc ggc aag gcc gac aag t ac ct c aag t cg ct g ccg      240
Phe His Gn Arg Ser Gly Lys Ala Asp Lys Tyr Leu Lys Ser Leu Pro
65
aag ct g gat gcg t cc aag gt g gag t cg cgg t t c t cg gcc aaa gag cag      288
Lys Leu Asp Ala Ser Lys Val Gu Ser Arg Phe Ser Ala Lys Gu Gn
85
gcg cgg cgc gac gcc at g acg cgc gac t ac gcg gcc t t t cgc gag gag      336
Ala Arg Arg Asp Ala Met Thr Arg Asp Tyr Ala Ala Phe Arg Gu Gu
100
ct c gt c gcc gag ggg t ac t t t gac ccg t cg at c ccg cac at g at t t ac      384
Leu Val Ala Gu Gly Tyr Phe Asp Pro Ser lle Pro His Met lle Tyr
115
cgc gt c gt g gag at c gt g gcg ct c t t c gcg ct c t cg t t c t gg ct c at g      432
Arg Val Val Gu lle Val Ala Leu Phe Ala Leu Ser Phe Trp Leu Met
130
t cc aag gcc t cg ccc acc t cg ct c gt g ct g ggc gt g gt g at g aac ggc      480
Ser Lys Ala Ser Pro Thr Ser Leu Val Leu Gly Val Val Met Asn Gly
145
at t gcg cag ggc cgc t gc ggc t gg gt c at g cac gag at g ggc cac ggg      528
lle Ala Gn Gly Arg Cys Gly Trp Val Met His Gu Met Gly His Gly
165
t cg t t c acg ggc gt c at c t gg ct c gac gac cgg at g t gc gag t t c t t c      576
Ser Phe Thr Gly Val lle Trp Leu Asp Asp Arg Met Cys Gu Phe Phe
180
t ac ggc gt c ggc t gc ggc at g agc ggg cac t ac t gg aag aac cag cac      624
Tyr Gly Val Gly Cys Gly Met Ser Gly His Tyr Trp Lys Asn Gn His
195
agc aag cac cac gcc gcg ccc aac cgc ct c gag cac gat gt c ct c      672
Ser Lys His His Ala Ala Pro Asn Arg Leu Gu His Asp Val Asp Leu
210
aac acg ct g ccc ct g gt c gcc t t t aac gag cgc gt c gt g cgc aag gt c      720
Asn Thr Leu Pro Leu Val Ala Phe Asn Gu Arg Val Val Arg Lys Val
225
aag ccg gga t cg ct g ct g gcg ct c t gg ct g cgc gt g cag gcg t ac ct c      768
Lys Pro Gly Ser Leu Leu Ala Leu Trp Leu Arg Val Gn Ala Tyr Leu
245
t t t gcg ccc gt c t cg t gc ct g ct c at c ggc t t t gg acg ct c t ac      816
Phe Ala Pro Val Ser Cys Leu Leu lle Gly Leu Gly Trp Thr Leu Tyr
260
ct g cac ccg cgc t ac at g ct g cgc acc aag cgg cac at g gag t t c gt c      864
Leu His Pro Arg Tyr Met Leu Arg Thr Lys Arg His at g Gu Phe Val
275
t gg at c t t c gcg cgc t ac at t ggc t gg t t c t cg ct c at g ggc gct ct c      912
Trp lle Phe Ala Arg Tyr lle Gly Trp Phe Ser Leu Met Gly Ala Leu
290
ggc t ac t cg ccg ggc acc t cg gt c ggg at g t ac ct g t gc t cg t t c ggc      960
Gly Tyr Ser Pro Gly Thr Ser Val Gly Met Tyr Leu Cys Ser Phe Gly
305
ct c ggc t gc at t t ac at t t t c ct g cag t t c gcc gt c agc cac acg cac      1008
Leu Gly Cys lle Tyr lle Phe Leu Gn Phe Ala Val Ser His Thr His
325
ct g ccg gt g acc aac ccg gag gac cag ct g cac t gg ct c gag t ac gcg      1056
Leu Pro Val Thr Asn Pro Gu Asp Gn Leu His Trp Leu Gu Tyr Ala
340
345
350

```

PF58307. txt

gcc gac cac acg gt g aac att agc acc aag t cc t gg ct c gt c acg t gg 1104  
Al a Asp Hi s Thr Val Asn lle Ser Thr Lys Ser Trp Leu Val Thr Trp  
355 360 365  
t gg at g t cg aac ct g aac t t t cag at c gag cac cac ct c t t c ccc acg 1152  
Tr p Met Ser Asn Leu Asn Phe G n lle G u Hi s Hi s Leu Phe Pro Thr  
370 375 380  
gcg ccg cag t t c cgc t t c aag gaa at c agt cct cgc gt c gag gcc ct c 1200  
Al a Pro G n Phe Arg Phe Lys G u lle Ser Pro Arg Val G u Al a Leu  
385 390 395 400  
t t c aag cgc cac aac ct c ccg t ac t ac gac ct g ccc t ac acg agc gcg 1248  
Phe Lys Arg Hi s Asn Leu Pro Tyr Tyr Asp Leu Pro Tyr Thr Ser Al a  
405 410 415  
gt c t cg acc acc t t t gcc aat ct t t at t cc gt c ggc cac t cg gt c ggc 1296  
Val Ser Thr Thr Phe Al a Asn Leu Tyr Ser Val G y Hi s 430  
420 425  
gcc gac acc aag aag cag gac t ga 1320  
Al a Asp Thr Lys Lys G n Asp 440  
435 440

<210> 52  
<211> 439  
<212> PRT  
<213> Thr aust ochyt r i um ssp.

<400> 52  
Met G y Lys G y Ser G u G y Arg Ser Al a Al a Arg G u Met Thr Al a  
1 5 10 15  
G u Al a Asn G y Asp Lys Arg Lys Thr lle Leu lle G u G y Val Leu  
20 25 30  
Tyr Asp Al a Thr Asn Phe Lys Hi s Pro G y G y Ser lle lle Asn Phe  
35 40 45  
Leu Thr G u G y G u Al a G y Val Asp Al a Thr G n Al a Tyr Arg G u  
50 55 60  
Phe Hi s G n Arg Ser G y Lys Al a Asp Lys Tyr Leu Lys Ser Leu Pro  
65 70 75 80  
Lys Leu Asp Al a Ser Lys Val G u Ser Arg Phe Ser Al a Lys G u G n  
85 90 95  
Al a Arg Arg Asp Al a Met Thr Arg Asp Tyr Al a Al a Phe Arg G u G u  
100 105 110  
Leu Val Al a G u G y Tyr Phe Asp Pro Ser lle Pro Hi s Met lle Tyr  
115 120 125  
Arg Val Val G u lle Val Al a Leu Phe Al a Leu Ser Phe Trp Leu Met  
130 135 140  
Ser Lys Al a Ser Pro Thr Ser Leu Val Leu G y Val Val Met Asn G y  
145 150 155 160  
lle Al a G n G y Arg Cys G y Trp Val Met Hi s G u Met G y Hi s G y  
165 170 175 180  
Ser Phe Thr G y Val lle Trp Leu Asp Asp Arg Met Cys G u Phe Phe  
180 185 190  
Tyr G y Val G y Cys G y Met Ser G y Hi s Tyr Trp Lys Asn G n Hi s  
195 200 205  
Ser Lys Hi s Hi s Al a Al a Pro Asn Arg Leu G u Hi s Asp Val Asp Leu  
210 215 220  
Asn Thr Leu Pro Leu Val Al a Phe Asn G u Arg Val Val Arg Lys Val  
225 230 235 240  
Lys Pro G y Ser Leu Leu Al a Leu Trp Leu Arg Val G n Al a Tyr Leu  
245 250 255  
Phe Al a Pro Val Ser Cys Leu Leu lle G y Leu G y Trp Thr Leu Tyr  
260 265 270  
Leu Hi s Pro Arg Tyr Met Leu Arg Thr Lys Arg Hi s Met G u Phe Val  
275 280 285  
Trp lle Phe Al a Arg Tyr lle G y Trp Phe Ser Leu Met G y Al a Leu  
290 295 300  
G y Tyr Ser Pro G y Thr Ser Val G y Met Tyr Leu Cys Ser Phe G y  
305 310 315 320  
Leu G y Cys lle Tyr lle Phe Leu G n Phe Al a Val Ser Hi s Thr Hi s  
325 330 335  
Leu Pro Val Thr Asn Pro G u Asp G n Leu Hi s Trp Leu G u Tyr Al a  
Sei te 65

PF58307. txt

Al a Asp Hi s Thr Val Asn Ile Ser Thr Lys Ser Trp Leu Val Thr Trp  
 340 345 350  
 355 Thr Lys Ser Trp Leu Val Thr Trp  
 Tr p Met Ser Asn Leu Asn Phe G n Ile G u Hi s Hi s 365 365  
 370 375 380  
 Al a Pro G n Phe Arg Phe Lys G u Ile Ser Pro Arg Val G u Al a Leu  
 385 390 395 400  
 Phe Lys Arg Hi s Asn Leu Pro Tyr Tyr Asp Leu Pro Tyr Thr Ser Al a  
 405 410 415  
 Val Ser Thr Thr Phe Al a Asn Leu Tyr Ser Val G y Hi s Ser Val G y  
 420 425 430  
 Al a Asp Thr Lys Lys G n Asp  
 435

<210> 53  
 <211> 1371  
 <212> DNA  
 <213> *Ostreococcus tauri*

<220>  
 <221> CDS  
 <222> (1).. (1371)

<400> 53  
 at g tgc gt g gag acg gaa aat aac gat ggg at c ccc acg gt g gag at c 48  
 Met Cys Val G u Thr G u Asn Asn Asp G y Ile Pro Thr Val G u Ile  
 1 5 10  
 gcg ttc gac ggt gag cgc gag cgg gcg gag gca aac gt g aag ct g tcc 96  
 Al a Phe Asp G y G u Arg G u Arg Al a G u Al a Asn Val Lys Leu Ser  
 20 25 30  
 gcg gag aag at g gag ccg gcg gcg ct g gcg aag acg ttc gcg agg cgg 144  
 Al a G u Lys Met G u Pro Al a Al a Leu Al a Lys Thr Phe Al a Arg Arg  
 35 40 45  
 t ac gt c gt g at c gag ggg gt g gag t ac gat gt g acg gat ttt aag cac 192  
 Tyr Val Val Ile G u G y Val G u Tyr Asp Val Thr Asp Phe Lys Hi s  
 50 55 60  
 ccg gga gga acg gtt att ttc tat gcg ttg tca aac acc ggg gcg gac 240  
 Pro G y G y Thr Val Ile Phe Tyr Al a Leu Ser Asn Thr G y Al a Asp  
 65 70 75 80  
 gcg acg gaa gcg ttc aag gag ttt cat cat cgg tcg aga aag gcg agg 288  
 Al a Thr G u Al a Phe Lys G u Phe Hi s Hi s Arg Ser Arg Lys Al a Arg  
 85 90 95  
 aaa gcc ttg gcg gcg ct c ccg tct cga ccg gcc aag acg gcc aag gt g 336  
 Lys Al a Leu Al a Al a Leu Pro Ser Arg Pro Al a Lys Thr Al a Lys Val  
 100 105 110  
 gac gac gcg gag at g ct c caa gat ttc gcc aag tgg cgg aaa gaa ttg 384  
 Asp Asp Al a G u Met Leu G n Asp Phe Al a Lys Trp Arg Lys G u Leu  
 115 120 125  
 gag aga gat gga ttc ttc aag ccc tct ccg gcg cac gt g gcg tat cgc 432  
 G u Arg Asp G y Phe Phe Lys Pro Ser Pro Al a Hi s Val Al a Tyr Arg  
 130 135 140  
 ttc gcc gag ct c gcg gcg at g tac gct ct c ggg acg tac ct g at g tac 480  
 Phe Al a G u Leu Al a Al a Met Tyr Al a Leu G y Thr Tyr Leu Met Tyr  
 145 150 155 160  
 gct cga tac gt c gt c tcc tcc gt g ct c gt g tac gct tgc ttt ttc ggc 528  
 Al a Arg Tyr Val Val Ser Ser Val Leu Val Tyr Al a Cys Phe Phe G y  
 165 170 175  
 gcc cga tgc ggt tgg gt g cag cac gag ggc gga cac agc tcc ct g acg 576  
 Al a Arg Cys G y Trp Val G n Hi s G u G y G y Hi s Ser Ser Leu Thr  
 180 185 190  
 ggc aac att tgg tgg gac aag cgc at c cag gcc ttc aca gcc ggg ttc 624  
 G y Asn Ile Trp Trp Asp Lys Arg Ile G n Al a Phe Thr Al a G y Phe  
 195 200 205  
 ggt ct c gcc ggt agc ggc gac at g tgg aac tcc at g cac aag cat 672  
 G y Leu Al a G y Ser G y Asp Met Trp Asn Ser Met Hi s Asn Lys Hi s  
 210 215 220  
 cac gcg acg cct caa aag gtt cgt cac gac at g gat ct g gac acc acc 720  
 Hi s Al a Thr Pro G n Lys Val Arg Hi s Asp Met Asp Leu Asp Thr Thr  
 225 230 235



PF58307. txt

225	ccc	gcg	gt g	gcg	t t c	t t c	aac	acc	gcg	gt g	gaa	gac	aat	cgt	ccc	cgt	768
	Pro	Al a	Val	Al a	Phe	Phe	Asn	Thr	Al a	Val	Gl u	Asp	Asn	Arg	Pro	Arg	
					245					250					255		
	ggc	t t t	agc	aag	t ac	t gg	t t g	cgc	ct t	cag	gcg	t gg	acc	t t c	at c	ccc	816
	Gl y	Phe	Ser	Lys	Tyr	Trp	Leu	Arg	Leu	Gl n	Al a	Trp	Thr	Phe	I l e	Pro	
				260					265					270			
	gt g	acg	t cc	ggc	t t g	gt g	ct c	ct t	t t c	t gg	at g	t t t	t t c	ct c	cac	ccc	864
	Val	Thr	Ser	Gl y	Leu	Val	Leu	Leu	Phe	Trp	Met	Phe	Leu	Hi s	Pro		
				275					280				285				
	t cc	aag	gct	t t g	aag	ggc	ggc	aag	t ac	gaa	gag	t t g	gt g	t gg	at g	ct c	912
	Ser	Lys	Al a	Leu	Lys	Gl y	Gl y	Lys	Tyr	Gl u	Gl u	Leu	Val	Trp	Met	Leu	
				290			295					300					
	gcc	gcg	cac	gt c	at c	cgc	acg	t gg	acg	at c	aag	gcg	gt g	acc	gga	t t c	960
	Al a	Al a	Hi s	Val	I l e	Arg	Thr	Trp	Thr	I l e	Lys	Al a	Val	Thr	Gl y	Phe	
	305					310					315					320	
	acc	gcg	at g	cag	t cc	t ac	ggc	t t a	t t t	t t g	gcg	acg	agc	t gg	gt g	agc	1008
	Thr	Al a	Met	Gl n	Ser	Tyr	Gl y	Leu	Phe	Leu	Al a	Thr	Ser	Trp	Val	Ser	
					325					330					335		
	ggc	t gc	t at	ct g	t t t	gca	cac	t t c	t cc	acg	t cg	cac	acg	cac	ct g	gat	1056
	Gl y	Cys	Tyr	Leu	Phe	Al a	Hi s	Phe	Ser	Thr	Ser	Hi s	Thr	Hi s	Leu	Asp	
				340					345					350			
	gt g	gt g	ccc	gcg	gac	gag	cat	ct c	t cc	t gg	gt t	cga	t ac	gcc	gt c	gat	1104
	Val	Val	Pro	Al a	Asp	Gl u	Hi s	Leu	Ser	Trp	Val	Arg	Tyr	Al a	Val	Asp	
			355					360					365				
	cac	acg	at c	gac	at c	gat	ccg	agt	caa	ggc	t gg	gt g	aac	t gg	t t g	at g	1152
	Hi s	Thr	I l e	Asp	I l e	Asp	Pro	Ser	Gl n	Gl y	Trp	Val	Asn	Trp	Leu	Met	
							375					380					
	ggc	t ac	ct c	aac	t gc	caa	gt c	at c	cac	cac	ct c	t t t	ccg	agc	at g	ccg	1200
	Gl y	Tyr	Leu	Asn	Cys	Gl n	Val	I l e	Hi s	Hi s	Leu	Phe	Pro	Ser	Met	Pro	
					390						395				400		
	cag	t t c	cgc	cag	ccc	gag	gt a	t ct	cgc	cgc	t t c	gt c	gcc	t t t	gcg	aaa	1248
	Gl n	Phe	Arg	Gl n	Pro	Gl u	Val	Ser	Arg	Arg	Phe	Val	Al a	Phe	Al a	Lys	
					405					410					415		
	aag	t gg	aac	ct c	aac	t ac	aag	gt c	at g	acc	t ac	gcc	ggc	ggc	t gg	aag	1296
	Lys	Trp	Asn	Leu	Asn	Tyr	Lys	Val	Met	Thr	Tyr	Al a	Gl y	Al a	Trp	Lys	
				420					425					430			
	gca	acg	ct c	gga	aac	ct c	gac	aac	gt g	ggc	aag	cac	t ac	t ac	gt g	cac	1344
	Al a	Thr	Leu	Gl y	Asn	Leu	Asp	Asn	Val	Gl y	Lys	Hi s	Tyr	Tyr	Val	Hi s	
				435				440					445				
	ggc	caa	cac	t cc	gga	aag	acg	gcg	t aa								1371
	Gl y	Gl n	Hi s	Ser	Gl y	Lys	Thr	Al a									
				450			455										

<210> 54  
 <211> 456  
 <212> PRT  
 <213> *Ostreococcus tauri*

<400> 54  
 Met Cys Val Gl u Thr Gl u Asn Asn Asp Gl y I l e Pro Thr Val Gl u I l e  
 1 5 10 15  
 Al a Phe Asp Gl y Gl u Arg Gl u Arg Al a Gl u Al a Asn Val Lys Leu Ser  
 20 25 30  
 Al a Gl u Lys Met Gl u Pro Al a Al a Leu Al a Lys Thr Phe Al a Arg Arg  
 35 40 45  
 Tyr Val Val I l e Gl u Gl y Val Gl u Tyr Asp Val Thr Asp Phe Lys Hi s  
 50 55 60  
 Pro Gl y Gl y Thr Val I l e Phe Tyr Al a Leu Ser Asn Thr Gl y Al a Asp  
 65 70 75 80  
 Al a Thr Gl u Al a Phe Lys Gl u Phe Hi s Hi s Arg Ser Arg Lys Al a Arg  
 85 90 95  
 Lys Al a Leu Al a Al a Leu Pro Ser Arg Pro Al a Lys Thr Al a Lys Val  
 100 105 110  
 Asp Asp Al a Gl u Met Leu Gl n Asp Phe Al a Lys Trp Arg Lys Gl u Leu  
 115 120 125  
 Gl u Arg Asp Gl y Phe Phe Lys Pro Ser Pro Al a Hi s Val Al a Tyr Arg  
 130 135 140

PF58307. txt

Phe Ala Glu Leu Ala Ala Met Tyr Ala Leu Gly Thr Tyr Leu Met Tyr  
 145 Ala Arg Tyr Val Val Ser Ser Val Leu Val Tyr Ala Cys Phe Phe Gly  
 Ala Arg Cys Gly Trp Val Gn His Gu Gly Gly His Ser Ser Leu Thr  
 Gly Asn Ile Trp Trp Asp Lys Arg Ile Gn Ala Phe Thr Ala Gly Phe  
 Gly Leu Ala Gly Ser Gly Asp Met Trp Asn Ser Met His Asn Lys His  
 His Ala Thr Pro Gn Lys Val Arg His Asp Met Asp Leu Asp Thr Thr  
 225 Pro Ala Val Ala Phe Phe Asn Thr Ala Val Gu Asp Asn Arg Pro Arg  
 Gly Phe Ser Lys Tyr Trp Leu Arg Leu Gn Ala Trp Thr Phe Ile Pro  
 Val Thr Ser Gly Leu Val Leu Leu Phe Trp Met Phe Phe Leu His Pro  
 Ser Lys Ala Leu Lys Gly Gly Lys Tyr Gu Gu Leu Val Trp Met Leu  
 Ala Ala His Val Ile Arg Thr Trp Thr Ile Lys Ala Val Thr Gly Phe  
 305 Thr Ala Met Gn Ser Tyr Gly Leu Phe Leu Ala Thr Ser Trp Val Ser  
 Gly Cys Tyr Leu Phe Ala His Phe Ser Thr Ser His Thr His Leu Asp  
 Val Val Pro Ala Asp Gu His Leu Ser Trp Val Arg Tyr Ala Val Asp  
 His Thr Ile Asp Ile Asp Pro Ser Gn Gly Trp Val Asn Trp Leu Met  
 Gly Tyr Leu Asn Cys Gn Val Ile His His Leu Phe Pro Ser Met Pro  
 385 Gn Phe Arg Gn Pro Gu Val Ser Arg Arg Phe Val Ala Phe Ala Lys  
 Lys Trp Asn Leu Asn Tyr Lys Val Met Thr Tyr Ala Gly Ala Trp Lys  
 Ala Thr Leu Gly Asn Leu Asp Asn Val Gly Lys His Tyr Val His  
 Gly Gn His Ser Gly Lys Thr Ala  
 450 455

<210> 55  
 <211> 1254  
 <212> DNA  
 <213> Leishmani a major

<220>  
 <221> CDS  
 <222> (1).. (1254)

<400> 55  
 atg gcc ctg gac aat gtc cgt ccg cac caa ccg aac gag gtg ctg atc 48  
 Met Ala Leu Asp Asn Val Arg Pro His Gn Pro Asn Gu Val Leu Ile  
 1 gat ggc gtc ctt tac gac tgc acc gat ttc cgg cat ccg ggt ggc agc 96  
 Asp Gly Val Leu Tyr Asp Cys Thr Asp Phe Arg His Pro Gly Gly Ser  
 20 att ctg aaa tac tac ctg ggc agc ggc gac gcc acc gag acg tac caa 144  
 Ile Leu Lys Tyr Tyr Leu Gly Ser Gly Asp Ala Thr Thr Tyr Gn  
 35 cag ttc cac ttg aag ctg ccc agg gcg gac aag tat ctg aag cgg ctg 192  
 Gn Phe His Leu Lys Leu Pro Arg Ala Asp Lys Tyr Leu Lys Arg Leu  
 50 ccg aat cgc ccg gcg ccg cca cag cac agc gtc aac gtg gat gag cag 240  
 Pro Asn Arg Pro Ala Pro Pro Gn His Ser Val Asn Val Asp Gu Gn  
 65 aag cga ttg gag aag ctg tgc cgg gac ttc aag gcg ctg cag gat gcg 288  
 Lys Arg Leu Gu Lys Leu Ser Arg Asp Phe Lys Ala Leu Gn Asp Ala  
 Sei te 68

PF58307. txt

				85					90					95					
tgc	gt a	gag	gag	ggc	ct t	t t t	aac	gcc	agc	t gg	ccg	cac	at c	gt c	t ac				336
Cys	Val	Gl u	Gl u	Gly	Leu	Phe	Asn	Al a	Ser	Tr p	Pro	His	Ile	Val	Tyr				
			100					105					110						
cg	t t t	t ct	gag	ct g	at c	ct g	at g	gcc	at c	gg	ct t	t ac	at g	ct c					384
Arg	Phe	Ser	Gl u	Leu	Ile	Leu	Met	His	Al a	Ile	Gly	Leu	Tyr	Met	Leu				
		115					120				125								
t t c	cgt	ct t	ccg	at c	ct g	t gg	ccc	gt c	gcg	ct g	gt g	at c	ct t	gga	gt g				432
Phe	Arg	Leu	Pro	Ile	Leu	Tr p	Pro	Val	Al a	Leu	Val	Ile	Leu	Gly	Val				
	130					135					140								
gcg	gag	ggg	cga	tgt	ggc	t gg	t gg	at g	cac	gag	gcc	gg	cac	t ac	agc				480
Al a	Gl u	Gl y	Arg	Cys	Gly	Tr p	Tr p	Met	His	Gl u	Al a	Gly	His	Tyr	Ser				160
	145			150					155										
gt c	aca	ggc	att	ccg	t gg	t t g	gac	at t	aaa	at a	cag	gag	gt a	ct c	t ac				528
Val	Thr	Gly	Ile	Pro	Tr p	Leu	Asp	Ile	Lys	Ile	Gln	Gl u	Val	Leu	Tyr				
			165						170					175					
gga	ct t	ggc	gat	gga	at g	agc	gcg	t cg	t gg	t gg	cgg	t cg	cag	cat	aac				576
Gly	Leu	Gly	Asp	Gly	Met	Ser	Al a	Ser	Tr p	Tr p	Arg	Ser	Gln	His	Asn				
			180					185					190						
aag	cat	cac	gct	act	ccg	cag	aag	cac	cgg	cac	gac	gt g	gac	ct t	gag				624
Lys	His	His	Al a	Thr	Pro	Gln	Lys	His	Arg	His	Asp	Val	Asp	Leu	Gl u				
		195					200					205							
acg	ct g	cct	ct c	gt c	gcc	t t c	aac	aag	at c	at c	gcg	cgc	cgc	ggc	aag				672
Thr	Leu	Pro	Leu	Val	Al a	Phe	Asn	Lys	Ile	Ile	Al a	Arg	Arg	Gly	Lys				
	210					215					220								
agg	aac	gcg	agc	att	cg	cg	t gg	at c	t cc	ct g	cag	at g	t t c	ct c	t t c				720
Arg	Asn	Al a	Ser	Ile	Arg	Arg	Tr p	Ile	Ser	Leu	Gln	Met	Phe	Leu	Phe				240
	225				230					235									
ggc	ccc	gt c	acc	tgc	t cc	ct t	gt c	gcc	ct c	t ac	t gg	cag	ct c	t t c	ct a				768
Gly	Pro	Val	Thr	Cys	Ser	Leu	Val	Al a	Leu	Tyr	Tr p	Gln	Leu	Phe	Leu				
			245					250					255						
cac	gt c	cgc	cac	gcc	at g	cg	act	cag	cg	t ac	aca	gag	ggc	t ct	gcc				816
His	Val	Arg	His	Al a	Met	Arg	Thr	Gln	Arg	Tyr	Thr	Gl u	Gly	Ser	Al a				
			260					265					270						
at c	ct g	tgc	cgc	t gg	at c	gt g	gt c	ggc	gt t	at c	tgt	cac	cag	ct g	cag				864
Ile	Leu	Cys	Arg	Tr p	Ile	Val	Val	Gly	Val	Ile	Cys	His	Gln	Leu	Gln				
		275					280					285							
gt c	t cg	t t c	t gg	caa	ggc	ct c	ggc	ggc	gt t	ct c	t t c	t cc	cag	gcc	t t c				912
Val	Ser	Phe	Tr p	Gln	Gly	Leu	Gly	Gly	Val	Leu	Phe	Ser	Gln	Al a	Phe				
	290				295						300								
agc	gcc	gcc	t ac	at c	t t c	at a	aac	t t c	gcc	ct c	aac	cac	t ct	cac	ct g				960
Ser	Al a	Al a	Tyr	Ile	Phe	Ile	Asn	Phe	Al a	Leu	Asn	His	Ser	His	Leu				320
	305				310				315										
ccg	at g	ct t	ccg	gaa	gac	gaa	cac	gcg	cac	t t c	gt c	gag	t ac	gcg	gcc				1008
Pro	Met	Leu	Pro	Gl u	Asp	Gl u	His	Al a	His	Phe	Val	Gl u	Tyr	Al a	Al a				
				325					330					335					
at c	t ac	acc	at g	aac	gt g	aca	ccg	t cg	t gg	t t c	gt g	acg	t gg	t t c	at g				1056
Ile	Tyr	Thr	Met	Asn	Val	Thr	Pro	Ser	Tr p	Phe	Val	Thr	Tr p	Phe	Met				
		340						345					350						
ggc	t ac	ct t	aac	t ac	cag	gt g	gaa	cac	ct c	t t c	cct	acc	at g	cca					1104
Gly	Tyr	Leu	Asn	Tyr	Gln	Val	Gl u	His	His	Leu	Phe	Pro	Thr	Met	Pro				
		355				360					365								
cag	t t c	cgc	t t c	gt c	caa	ct g	gcg	ccg	cga	gt g	cgg	aaa	ct t	t t t	gag				1152
Gln	Phe	Arg	Phe	Val	Gln	Leu	Al a	Pro	Arg	Val	Arg	Lys	Leu	Phe	Gl u				
	370					375					380								
gaa	aac	ggc	ct c	aag	t ac	gat	t cg	cgt	ccg	t ac	at g	gag	t cg	ct c	cag				1200
Gl u	Asn	Gl y	Leu	Lys	Tyr	Asp	Ser	Arg	Pro	Tyr	Met	Gl u	Ser	Leu	Gln				400
	385			390					395										
aaa	acc	t t c	aag	aac	ct c	ggc	gac	gt g	gcc	gag	t t c	at c	gt t	gct	ggg				1248
Lys	Thr	Phe	Lys	Asn	Leu	Gly	Asp	Val	Al a	Gl u	Phe	Ile	Val	Al a	Gly				
				405					410					415					
aac	t aa																		1254
Asn																			

<210> 56  
 <211> 417  
 <212> PRT

&lt;213&gt; Lei shmani a maj or

&lt;400&gt; 56

Met Ala Leu Asp Asn Val Arg Pro His Gln Pro Asn Glu Val Leu Ile  
 1 5 10 15  
 Asp Gly Val Leu Tyr Asp Cys Thr Asp Phe Arg His Pro Gly Gly Ser  
 20 25 30  
 Ile Leu Lys Tyr Tyr Leu Gly Ser Gly Asp Ala Thr Glu Thr Tyr Gln  
 35 40 45  
 Gln Phe His Leu Lys Leu Pro Arg Ala Asp Lys Tyr Leu Lys Arg Leu  
 50 55 60  
 Pro Asn Arg Pro Ala Pro Pro Gln His Ser Val Asn Val Asp Glu Gln  
 65 70 75 80  
 Lys Arg Leu Glu Lys Leu Ser Arg Asp Phe Lys Ala Leu Gln Asp Ala  
 85 90 95  
 Cys Val Glu Glu Gly Leu Phe Asn Ala Ser Trp Pro His Ile Val Tyr  
 100 105 110  
 Arg Phe Ser Glu Leu Ile Leu Met His Ala Ile Gly Leu Tyr Met Leu  
 115 120 125  
 Phe Arg Leu Pro Ile Leu Trp Pro Val Ala Leu Val Ile Leu Gly Val  
 130 135 140  
 Ala Glu Gly Arg Cys Gly Trp Trp Met His Glu Ala Gly His Tyr Ser  
 145 150 155 160  
 Val Thr Gly Ile Pro Trp Leu Asp Ile Lys Ile Gln Glu Val Leu Tyr  
 165 170 175  
 Gly Leu Gly Asp Gly Met Ser Ala Ser Trp Trp Arg Ser Gln His Asn  
 180 185 190  
 Lys His His Ala Thr Pro Gln Lys His Arg His Asp Val Asp Leu Glu  
 195 200 205  
 Thr Leu Pro Leu Val Ala Phe Asn Lys Ile Ile Ala Arg Arg Gly Lys  
 210 215 220  
 Arg Asn Ala Ser Ile Arg Arg Trp Ile Ser Leu Gln Met Phe Leu Phe  
 225 230 235 240  
 Gly Pro Val Thr Cys Ser Leu Val Ala Leu Tyr Trp Gln Leu Phe Leu  
 245 250 255  
 His Val Arg His Ala Met Arg Thr Gln Arg Tyr Thr Glu Gly Ser Ala  
 260 265 270  
 Ile Leu Cys Arg Trp Ile Val Val Gly Val Ile Cys His Gln Leu Gln  
 275 280 285  
 Val Ser Phe Trp Gln Gly Leu Gly Gly Val Leu Phe Ser Gln Ala Phe  
 290 295 300  
 Ser Ala Ala Tyr Ile Phe Ile Asn Phe Ala Leu Asn His Ser His Leu  
 305 310 315 320  
 Pro Met Leu Pro Glu Asp Glu His Ala His Phe Val Glu Tyr Ala Ala  
 325 330 335  
 Ile Tyr Thr Met Asn Val Thr Pro Ser Trp Phe Val Thr Trp Phe Met  
 340 345 350  
 Gly Tyr Leu Asn Tyr Gln Val Glu His His Leu Phe Pro Thr Met Pro  
 355 360 365  
 Gln Phe Arg Phe Val Gln Leu Ala Pro Arg Val Arg Lys Leu Phe Glu  
 370 375 380  
 Glu Asn Gly Leu Lys Tyr Asp Ser Arg Pro Tyr Met Glu Ser Leu Gln  
 385 390 395 400  
 Lys Thr Phe Lys Asn Leu Gly Asp Val Ala Glu Phe Ile Val Ala Gly  
 405 410 415  
 Asn

&lt;210&gt; 57

&lt;211&gt; 401

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (2)..(2)

&lt;223&gt; xaa in position 2 is any amino acid

<220>  
<221> Variant  
<222> (5)..(5)  
<223> xaa in position 5 is any amino acid  
<220>  
<221> Variant  
<222> (8)..(8)  
<223> xaa in position 8 is any amino acid  
<220>  
<221> Variant  
<222> (10)..(10)  
<223> xaa in position 10 is any amino acid  
<220>  
<221> Variant  
<222> (12)..(12)  
<223> xaa in position 12 is any amino acid  
<220>  
<221> Variant  
<222> (17)..(26)  
<223> xaa in position 17 to 26 is any amino acid  
<220>  
<221> Variant  
<222> (27)..(29)  
<223> xaa in position 27 to 29 is any or no amino acid  
<220>  
<221> Variant  
<222> (33)..(37)  
<223> xaa in position 33 to 37 is any amino acid  
<220>  
<221> Variant  
<222> (40)..(44)  
<223> xaa in position 40 to 44 is any amino acid  
<220>  
<221> Variant  
<222> (46)..(46)  
<223> xaa in position 46 is any amino acid  
<220>  
<221> Variant  
<222> (48)..(48)  
<223> xaa in position 48 is any amino acid  
<220>  
<221> Variant  
<222> (50)..(51)  
<223> xaa in position 50 to 51 is any amino acid  
<220>  
<221> Variant  
<222> (54)..(56)  
<223> xaa in position 54 to 56 is any amino acid  
<220>  
<221> Variant  
<222> (58)..(69)  
<223> xaa in position 58 to 69 is any amino acid  
<220>  
<221> Variant  
<222> (70)..(77)  
<223> xaa in position 70 to 77 is any or no amino acid  
<220>  
<221> Variant  
<222> (79)..(89)  
<223> xaa in position 79 to 89 is any amino acid  
<220>  
<221> Variant  
<222> (91)..(91)  
<223> xaa in position 91 is any amino acid  
<220>  
<221> Variant  
<222> (93)..(94)  
<223> xaa in position 93 to 94 is any amino acid

<220>  
<221> Variant  
<222> ( 96).. ( 97)  
<223> xaa in position 96 to 97 is any amino acid  
<220>  
<221> Variant  
<222> ( 99).. ( 100)  
<223> xaa in position 99 to 100 is any amino acid  
<220>  
<221> Variant  
<222> ( 103).. ( 104)  
<223> xaa in position 103 to 104 is any amino acid  
<220>  
<221> Variant  
<222> ( 106).. ( 110)  
<223> xaa in position 106 to 110 is any amino acid  
<220>  
<221> Variant  
<222> ( 112).. ( 127)  
<223> xaa in position 112 to 127 is any amino acid  
<220>  
<221> Variant  
<222> ( 128).. ( 129)  
<223> xaa in position 128 to 129 is any or no amino acid  
<220>  
<221> Variant  
<222> ( 131).. ( 137)  
<223> xaa in position 131 to 137 is any amino acid  
<220>  
<221> Variant  
<222> ( 142).. ( 143)  
<223> xaa in position 142 to 143 is any amino acid  
<220>  
<221> Variant  
<222> ( 146).. ( 146)  
<223> xaa in position 146 is any amino acid  
<220>  
<221> Variant  
<222> ( 149).. ( 149)  
<223> xaa in position 149 is any amino acid  
<220>  
<221> Variant  
<222> ( 151).. ( 151)  
<223> xaa in position 151 is any amino acid  
<220>  
<221> Variant  
<222> ( 154).. ( 155)  
<223> xaa in position 154 to 155 is any amino acid  
<220>  
<221> Variant  
<222> ( 157).. ( 157)  
<223> xaa in position 157 is any amino acid  
<220>  
<221> Variant  
<222> ( 159).. ( 166)  
<223> xaa in position 159 to 166 is any amino acid  
<220>  
<221> Variant  
<222> ( 168).. ( 168)  
<223> xaa in position 168 is any amino acid  
<220>  
<221> Variant  
<222> ( 170).. ( 172)  
<223> xaa in position 170 to 172 is any amino acid  
<220>  
<221> Variant  
<222> ( 174).. ( 176)  
<223> xaa in position 174 to 176 is any amino acid

<220>  
<221> Vari ant  
<222> ( 178) .. ( 180)  
<223> xaa in position 178 to 180 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 182) .. ( 182)  
<223> xaa in position 182 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 187) .. ( 187)  
<223> xaa in position 187 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 189) .. ( 192)  
<223> xaa in position 189 to 192 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 195) .. ( 195)  
<223> xaa in position 195 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 198) .. ( 198)  
<223> xaa in position 198 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 200) .. ( 200)  
<223> xaa in position 200 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 202) .. ( 202)  
<223> xaa in position 202 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 206) .. ( 221)  
<223> xaa in position 206 to 221 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 223) .. ( 225)  
<223> xaa in position 223 to 225 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 227) .. ( 229)  
<223> xaa in position 227 to 229 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 231) .. ( 231)  
<223> xaa in position 231 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 234) .. ( 236)  
<223> xaa in position 234 to 236 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 238) .. ( 239)  
<223> xaa in position 238 to 239 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 241) .. ( 241)  
<223> xaa in position 241 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 243) .. ( 245)  
<223> xaa in position 243 to 245 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 248) .. ( 258)  
<223> xaa in position 248 to 258 is any ami no aci d  
Seite 73

<220>  
<221> Vari ant  
<222> ( 260).. ( 285)  
<223> xaa in position 260 to 285 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 286).. ( 286)  
<223> xaa in position 286 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 288).. ( 289)  
<223> xaa in position 288 to 289 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 291).. ( 298)  
<223> xaa in position 291 to 298 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 300).. ( 300)  
<223> xaa in position 300 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 302).. ( 303)  
<223> xaa in position 302 to 303 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 305).. ( 307)  
<223> xaa in position 305 to 307 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 309).. ( 309)  
<223> xaa in position 309 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 312).. ( 324)  
<223> xaa in position 312 to 324 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 327).. ( 329)  
<223> xaa in position 327 to 329 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 331).. ( 338)  
<223> xaa in position 331 to 338 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 339).. ( 339)  
<223> xaa in position 339 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 341).. ( 341)  
<223> xaa in position 341 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 343).. ( 343)  
<223> xaa in position 343 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 345).. ( 346)  
<223> xaa in position 345 to 346 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 349).. ( 349)  
<223> xaa in position 349 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 351).. ( 352)  
<223> xaa in position 351 to 352 is any ami no aci d  
Seite 74



<220>  
 <221> Variant  
 <222> (358)..(359)  
 <223> xaa in position 358 to 359 is any amino acid  
 <220>  
 <221> Variant  
 <222> (364)..(369)  
 <223> xaa in position 364 to 369 is any amino acid  
 <220>  
 <221> Variant  
 <222> (371)..(379)  
 <223> xaa in position 371 to 379 is any amino acid  
 <220>  
 <221> Variant  
 <222> (381)..(381)  
 <223> xaa in position 381 is any amino acid  
 <220>  
 <221> Variant  
 <222> (383)..(386)  
 <223> xaa in position 383 to 386 is any amino acid  
 <220>  
 <221> Variant  
 <222> (388)..(393)  
 <223> xaa in position 388 to 393 is any amino acid  
 <220>  
 <221> Variant  
 <222> (395)..(396)  
 <223> xaa in position 395 to 396 is any amino acid  
 <220>  
 <221> Variant  
 <222> (399)..(400)  
 <223> xaa in position 399 to 400 is any amino acid

<400> 57  
 Ile xaa Gly Val xaa Tyr Asp xaa Thr xaa Phe xaa His Pro Gly Gly  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Asp Ala Thr  
 20 25 30  
 xaa xaa xaa xaa xaa Phe His xaa xaa xaa xaa xaa Ala xaa Lys xaa  
 35 40 45  
 Leu xaa xaa Leu Pro xaa xaa xaa Ala xaa xaa xaa xaa xaa xaa xaa  
 50 55 60  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Asp xaa xaa  
 65 70 75 80  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa Phe xaa xaa Ser xaa  
 85 90 95  
 xaa His xaa xaa Tyr Arg xaa xaa Gu xaa xaa xaa xaa xaa Ala xaa  
 100 105 110  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 115 120 125  
 xaa Val xaa xaa xaa xaa xaa xaa xaa Arg Cys Gly Trp xaa xaa His  
 130 135 140  
 Gu xaa Gly His xaa Ser xaa Thr Gly xaa xaa Trp xaa Asp xaa xaa  
 145 150 155 160  
 xaa xaa xaa xaa xaa xaa Gly xaa Gly xaa xaa Ser xaa xaa xaa  
 165 170 175  
 Trp xaa xaa xaa His xaa Lys His His Ala xaa Pro xaa xaa xaa xaa  
 180 185 190  
 His Asp xaa Asp Leu xaa Thr xaa Pro xaa Val Ala Phe xaa xaa xaa  
 195 200 205  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Trp xaa xaa  
 210 215 220  
 xaa Gn xaa xaa xaa Phe xaa Pro Val xaa xaa xaa Leu xaa xaa Leu  
 225 230 235 240  
 xaa Trp xaa xaa xaa Leu His xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 245 250 255  
 xaa xaa Gu xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 260 265 270

PF58307. txt

```

xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa
                275                280                285
xaa Leu xaa xaa xaa xaa xaa xaa xaa xaa xaa Tyr xaa Phe xaa xaa Phe
   290                295                300
xaa xaa xaa His xaa His xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa
305                310                315                320
xaa xaa xaa xaa Tyr Ala xaa xaa xaa Thr xaa xaa xaa xaa xaa xaa
                325                330                335
xaa xaa xaa Val xaa Trp xaa Met xaa xaa Leu Asn xaa G n xaa xaa
                340                345                350
His His Leu Phe Pro xaa xaa Pro G n Phe Arg xaa xaa xaa xaa xaa
   355                360                365
xaa Arg xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Leu xaa Tyr xaa xaa
   370                375                380
xaa xaa Tyr xaa xaa xaa xaa xaa xaa xaa Thr xaa xaa Asn Leu xaa xaa
385                390                395                400
Val

```

- <210> 58
- <211> 59
- <212> PRT
- <213> Artificial sequence
  
- <220>
- <221> Variant
- <222> (3)..(4)
- <223> xaa in position 3 to 4 is any amino acid
- <220>
- <221> Variant
- <222> (7)..(7)
- <223> xaa in position 7 is any amino acid
- <220>
- <221> Variant
- <222> (10)..(10)
- <223> xaa in position 10 is any amino acid
- <220>
- <221> Variant
- <222> (12)..(12)
- <223> xaa in position 12 is Phe, Leu or Val
- <220>
- <221> Variant
- <222> (15)..(15)
- <223> xaa in position 15 is any or no amino acid
- <220>
- <221> Variant
- <222> (17)..(17)
- <223> xaa in position 17 is any or no amino acid
- <220>
- <221> Variant
- <222> (19)..(19)
- <223> xaa in position 19 is any amino acid
- <220>
- <221> Variant
- <222> (21)..(21)
- <223> xaa in position 21 is any amino acid
- <220>
- <221> Variant
- <222> (22)..(22)
- <223> xaa in position 22 is Lys or Arg
- <220>
- <221> Variant
- <222> (23)..(23)
- <223> xaa in position 23 is Ile or Met
- <220>
- <221> Variant
- <222> (24)..(24)

<223> xaa in position 24 is any amino acid  
<220>  
<221> Variant  
<222> (25)..(25)  
<223> xaa in position 25 is Ala or Glu  
<220>  
<221> Variant  
<222> (26)..(26)  
<223> xaa in position 26 is Phe or Val  
<220>  
<221> Variant  
<222> (27)..(28)  
<223> xaa in position 27 to 28 is any amino acid  
<220>  
<221> Variant  
<222> (30)..(30)  
<223> xaa in position 30 is Phe, Leu or Val  
<220>  
<221> Variant  
<222> (32)..(32)  
<223> xaa in position 32 is any amino acid  
<220>  
<221> Variant  
<222> (33)..(33)  
<223> xaa in position 33 is Ala or Gly  
<220>  
<221> Variant  
<222> (34)..(34)  
<223> xaa in position 34 is any amino acid  
<220>  
<221> Variant  
<222> (36)..(36)  
<223> xaa in position 36 is Ala or Gly  
<220>  
<221> Variant  
<222> (37)..(38)  
<223> xaa in position 37 to 38 is any amino acid  
<220>  
<221> Variant  
<222> (40)..(40)  
<223> xaa in position 40 is any amino acid  
<220>  
<221> Variant  
<222> (41)..(41)  
<223> xaa in position 41 is Asn or Ser  
<220>  
<221> Variant  
<222> (42)..(42)  
<223> xaa in position 42 is any amino acid  
<220>  
<221> Variant  
<222> (44)..(44)  
<223> xaa in position 44 is Asn or Ser  
<220>  
<221> Variant  
<222> (49)..(49)  
<223> xaa in position 49 is Ala or Thr  
<220>  
<221> Variant  
<222> (51)..(51)  
<223> xaa in position 51 is Asn or Gln  
<220>  
<221> Variant  
<222> (52)..(52)  
<223> xaa in position 52 is Lys or Arg  
<220>  
<221> Variant  
<222> (53)..(53)

<223> xaa in position 53 is any amino acid  
 <220>  
 <221> Variant  
 <222> (54)..(54)  
 <223> xaa in position 54 is Gu or Arg  
 <220>  
 <221> Variant  
 <222> (57)..(57)  
 <223> xaa in position 57 is Met or Val  
  
 <400> 58  
 Gly Trp xaa xaa His Gu xaa Gly His xaa Ser xaa Thr Gly xaa Ile  
   1          5          10          15  
 xaa Trp xaa Asp xaa xaa xaa xaa xaa xaa Gly xaa Gly xaa  
           20          25          30  
 xaa xaa Ser xaa xaa xaa Trp xaa xaa xaa His xaa Lys His His Ala  
           35          40          45  
 xaa Pro xaa xaa xaa xaa His Asp xaa Asp Leu  
   50                  55

<210> 59  
 <211> 50  
 <212> PRT  
 <213> Artificial sequence  
  
 <220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Asn or Thr  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is any amino acid  
 <220>  
 <221> Variant  
 <222> (7)..(8)  
 <223> xaa in position 7 to 8 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(12)  
 <223> xaa in position 11 to 12 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Ser or Thr  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is any amino acid  
 <220>  
 <221> Variant

<222> (27)..(28)  
<223> xaa in position 27 to 28 is any amino acid  
<220>  
<221> Variant  
<222> (29)..(29)  
<223> xaa in position 29 is Glu or Gln  
<220>  
<221> Variant  
<222> (30)..(30)  
<223> xaa in position 30 is Ile, Leu or Val  
<220>  
<221> Variant  
<222> (31)..(31)  
<223> xaa in position 31 is Ala or Ser  
<220>  
<221> Variant  
<222> (32)..(32)  
<223> xaa in position 32 is any amino acid  
<220>  
<221> Variant  
<222> (34)..(34)  
<223> xaa in position 34 is Phe or Val  
<220>  
<221> Variant  
<222> (35)..(36)  
<223> xaa in position 35 to 36 is any amino acid  
<220>  
<221> Variant  
<222> (37)..(37)  
<223> xaa in position 37 is Phe or Leu  
<220>  
<221> Variant  
<222> (38)..(38)  
<223> xaa in position 38 is any amino acid  
<220>  
<221> Variant  
<222> (39)..(39)  
<223> xaa in position 39 is Glu or Lys  
<220>  
<221> Variant  
<222> (40)..(40)  
<223> xaa in position 40 is Glu, Lys or Arg  
<220>  
<221> Variant  
<222> (41)..(41)  
<223> xaa in position 41 is any amino acid  
<220>  
<221> Variant  
<222> (42)..(42)  
<223> xaa in position 42 is Gly or Asn  
<220>  
<221> Variant  
<222> (44)..(44)  
<223> xaa in position 44 is any amino acid  
<220>  
<221> Variant  
<222> (46)..(46)  
<223> xaa in position 46 is any amino acid  
<220>  
<221> Variant  
<222> (47)..(47)  
<223> xaa in position 47 is Asp, Ser or Val  
<220>  
<221> Variant  
<222> (48)..(48)  
<223> xaa in position 48 is any amino acid  
<220>  
<221> Variant

&lt;222&gt; (49)..(49)

&lt;223&gt; xaa in position 49 is Pro or Thr

&lt;400&gt; 59

Val xaa Trp xaa Met xaa xaa xaa Asn xaa xaa xaa Gln xaa xaa His  
 1 5 10 15  
 His Leu Phe Pro xaa xaa Pro Gln Phe Arg xaa xaa xaa xaa xaa xaa  
 20 25 30  
 Arg xaa xaa xaa xaa xaa xaa xaa xaa Leu xaa Tyr xaa xaa xaa  
 35 40 45  
 xaa Tyr  
 50

&lt;210&gt; 60

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (2)..(2)

&lt;223&gt; xaa in position 2 is Asp or Glu

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (5)..(5)

&lt;223&gt; xaa in position 5 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (8)..(8)

&lt;223&gt; xaa in position 8 is Ala, Cys or Val

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (10)..(10)

&lt;223&gt; xaa in position 10 is Asp or Asn

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (12)..(12)

&lt;223&gt; xaa in position 12 is Lys or Arg

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (17)..(17)

&lt;223&gt; xaa in position 17 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (18)..(18)

&lt;223&gt; xaa in position 18 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (20)..(20)

&lt;223&gt; xaa in position 20 is Phe, Ile or Leu

&lt;400&gt; 60

Ile xaa Gly Val xaa Tyr Asp xaa Thr xaa Phe xaa His Pro Gly Gly  
 1 5 10 15  
 xaa xaa Ile xaa  
 20

&lt;210&gt; 61

&lt;211&gt; 61

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (2)..(2)

&lt;223&gt; xaa in position 2 is any amino acid

<220>  
<221> Variant  
<222> (4)..(4)  
<223> xaa in position 4 is Ala or Leu  
<220>  
<221> Variant  
<222> (8)..(10)  
<223> xaa in position 8 to 10 is any amino acid  
<220>  
<221> Variant  
<222> (11)..(11)  
<223> xaa in position 11 is Ala, Ile or Val  
<220>  
<221> Variant  
<222> (12)..(12)  
<223> xaa in position 12 is Ala or Val  
<220>  
<221> Variant  
<222> (13)..(13)  
<223> xaa in position 13 is Glu or Arg  
<220>  
<221> Variant  
<222> (14)..(14)  
<223> xaa in position 14 is Asp, Lys or Arg  
<220>  
<221> Variant  
<222> (15)..(15)  
<223> xaa in position 15 is Gly, Asn or Val  
<220>  
<221> Variant  
<222> (16)..(16)  
<223> xaa in position 16 is Lys or Arg  
<220>  
<221> Variant  
<222> (17)..(18)  
<223> xaa in position 17 to 18 is any amino acid  
<220>  
<221> Variant  
<222> (19)..(19)  
<223> xaa in position 19 is Ala, Gly or Ser  
<220>  
<221> Variant  
<222> (20)..(23)  
<223> xaa in position 20 to 23 is any amino acid  
<220>  
<221> Variant  
<222> (25)..(26)  
<223> xaa in position 25 to 26 is any or no amino acid  
<220>  
<221> Variant  
<222> (28)..(29)  
<223> xaa in position 28 to 29 is any or no amino acid  
<220>  
<221> Variant  
<222> (31)..(31)  
<223> xaa in position 31 is any amino acid  
<220>  
<221> Variant  
<222> (32)..(32)  
<223> xaa in position 32 is Phe, Trp or Tyr  
<220>  
<221> Variant  
<222> (33)..(33)  
<223> xaa in position 33 is any amino acid  
<220>  
<221> Variant  
<222> (35)..(35)  
<223> xaa in position 35 is Ala, Gly or Ile

<220>  
<221> Variant  
<222> (38)..(38)  
<223> xaa in position 38 is Ser or Thr  
<220>  
<221> Variant  
<222> (39)..(39)  
<223> xaa in position 39 is Cys or Ser  
<220>  
<221> Variant  
<222> (40)..(40)  
<223> xaa in position 40 is any amino acid  
<220>  
<221> Variant  
<222> (42)..(42)  
<223> xaa in position 42 is Ile or Val  
<220>  
<221> Variant  
<222> (43)..(43)  
<223> xaa in position 43 is Ala, Gly or Leu  
<220>  
<221> Variant  
<222> (45)..(45)  
<223> xaa in position 45 is any amino acid  
<220>  
<221> Variant  
<222> (47)..(47)  
<223> xaa in position 47 is any amino acid  
<220>  
<221> Variant  
<222> (48)..(48)  
<223> xaa in position 48 is Phe or Leu  
<220>  
<221> Variant  
<222> (49)..(49)  
<223> xaa in position 49 is Phe or Tyr  
<220>  
<221> Variant  
<222> (52)..(52)  
<223> xaa in position 52 is Pro or Val  
<220>  
<221> Variant  
<222> (53)..(53)  
<223> xaa in position 53 is Arg or Ser  
<220>  
<221> Variant  
<222> (54)..(55)  
<223> xaa in position 54 to 55 is any amino acid  
<220>  
<221> Variant  
<222> (56)..(56)  
<223> xaa in position 56 is Leu or Met  
<220>  
<221> Variant  
<222> (57)..(57)  
<223> xaa in position 57 is Lys or Arg  
<220>  
<221> Variant  
<222> (58)..(58)  
<223> xaa in position 58 is Gly or Thr  
<220>  
<221> Variant  
<222> (59)..(59)  
<223> xaa in position 59 is any amino acid  
<220>  
<221> Variant  
<222> (60)..(60)  
<223> xaa in position 60 is Lys or Arg



<220>  
 <221> Variant  
 <222> (61)..(61)  
 <223> xaa in position 61 is His or Tyr  
  
 <400> 61  
 Thr xaa Pro xaa Val Ala Phe xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa Trp xaa xaa Leu xaa xaa Gln xaa xaa  
 20 25 30  
 xaa Phe xaa Pro Val xaa xaa xaa Leu xaa xaa Leu xaa Trp xaa xaa  
 35 40 45  
 xaa Leu His xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 50 55 60

<210> 62  
 <211> 34  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(5)  
 <223> xaa in position 2 to 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (8)..(9)  
 <223> xaa in position 8 to 9 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is Asp or Pro  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is Met or Val  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Asn or Pro  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (16)..(17)  
 <223> xaa in position 16 to 17 is Asp or Glu  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is Ala or Leu  
 <220>  
 <221> Variant  
 <222> (20)..(20)

```

<223> xaa in position 20 is any amino acid
<220>
<221> Variant
<222> (21)..(21)
<223> xaa in position 21 is Phe or Trp
<220>
<221> Variant
<222> (22)..(22)
<223> xaa in position 22 is Leu or Val
<220>
<221> Variant
<222> (23)..(23)
<223> xaa in position 23 is Glu or Arg
<220>
<221> Variant
<222> (26)..(26)
<223> xaa in position 26 is Ala or Val
<220>
<221> Variant
<222> (27)..(27)
<223> xaa in position 27 is any amino acid
<220>
<221> Variant
<222> (28)..(28)
<223> xaa in position 28 is His or Tyr
<220>
<221> Variant
<222> (30)..(30)
<223> xaa in position 30 is Ile, Met or Val
<220>
<221> Variant
<222> (31)..(31)
<223> xaa in position 31 is Asp or Asn
<220>
<221> Variant
<222> (32)..(32)
<223> xaa in position 32 is Ile or Val
<220>
<221> Variant
<222> (33)..(33)
<223> xaa in position 33 is Asp, Ser or Thr
<220>
<221> Variant
<222> (34)..(34)
<223> xaa in position 34 is Pro or Thr

<400> 62
Ala xaa xaa xaa xaa xaa His xaa xaa Leu xaa xaa xaa xaa xaa xaa
  1           5           10           15
xaa xaa xaa xaa xaa xaa xaa Tyr Ala xaa xaa xaa Thr xaa xaa xaa
           20           25           30
xaa xaa

```

```

<210> 63
<211> 28
<212> PRT
<213> Artificial sequence

```

```

<220>
<221> Variant
<222> (4)..(4)
<223> xaa in position 4 is Gu or Gn
<220>
<221> Variant
<222> (5)..(5)
<223> xaa in position 5 is Ala or Thr

```

<220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Phe or Tyr  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is any amino acid  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Gu or Gn  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is any amino acid  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is Lys or Arg  
 <220>  
 <221> Variant  
 <222> (13)..(14)  
 <223> xaa in position 13 to 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Lys or Arg  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Asp or Arg  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is any amino acid  
 <220>  
 <221> Variant  
 <222> (21)..(22)  
 <223> xaa in position 21 to 22 is any amino acid  
 <220>  
 <221> Variant  
 <222> (25)..(26)  
 <223> xaa in position 25 to 26 is any amino acid  
 <220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is Asp or Pro

<400> 63  
 Asp Ala Thr xaa xaa xaa xaa xaa Phe His xaa xaa xaa xaa xaa Ala  
 1 5 10 15  
 xaa Lys xaa Leu xaa xaa Leu Pro xaa xaa xaa Ala  
 20 25

<210> 64  
 <211> 903  
 <212> DNA  
 <213> *Ostreococcus tauri*

<220>  
 <221> CDS  
 <222> (1)..(903)

<400> 64  
 atg tct gct tct gga gct ttg ttg cct gct att gct ttc gct gct tac  
 Met Ser Ala Ser Gly Ala Leu Leu Pro Ala Ile Ala Phe Ala Ala Tyr  
 1 5 10 15

PF58307.txt

gct tac gct acc tac gct tat gct ttc gag tgg tct cat gct aac gga 96  
Ala Tyr Ala Thr Tyr Ala Tyr Ala Phe Gu Trp Ser His Ala Asn Gly  
20 25 30  
atc gat aac gtg gat gct aga gag tgg att gga gct ttg tct ttg aga 144  
Ile Asp Asn Val Asp Ala Arg Gu Trp Ile Gly Ala Leu Ser Leu Arg  
35 40 45  
ctc cct gca att gct acc acc atg tac ctc ttg ttc tgc ctt gtg gga 192  
Leu Pro Ala Ile Ala Thr Thr Met Tyr Leu Leu Phe Cys Leu Val Gly  
50 55  
cct aga ttg atg gct aag agg gag gct ttt gat cct aag gga ttc atg 240  
Pro Arg Leu Met Ala Lys Arg Gu Ala Phe Asp Pro Lys Gly Phe Met  
65 70 80  
ctc gct tac aac gct tac caa acc gct ttc aac gtt gtg gtg ctg gga 288  
Leu Ala Tyr Asn Ala Tyr Gn Thr Ala Phe Asn Val Val Val Leu Gly  
85 90 95  
atg ttc gct aga gag atc tct gga ttg gga caa cct gtt tgg gga tct 336  
Met Phe Ala Arg Gu Ile Ser Gly Leu Gly Gn Pro Val Trp Gly Ser  
100 105 110  
act atg cct tgg agc gat agg aag tcc ttc aag att ttg ttg gga gtg 384  
Thr Met Pro Trp Ser Asp Arg Lys Ser Phe Lys Ile Leu Leu Gly Val  
115 120 125  
tgg ctg cat tac aac aat aag tac ctc gag ttg ttg gat act gtg ttc 432  
Trp Leu His Tyr Asn Asn Lys Tyr Leu Gu Leu Leu Asp Thr Val Phe  
130 135 140  
atg gtg gct agg aaa aag acc aag cag ctc tct ttc ttg cat gtg tac 480  
Met Val Ala Arg Lys Lys Thr Lys Gn Leu Ser Phe Leu His Val Tyr  
145 150 160  
cat cat gct ttg ttg att tgg gct tgg tgg ctt gtt tgt cat ctg atg 528  
His His Ala Leu Leu Ile Trp Ala Trp Trp Leu Val Cys His Leu Met  
165 170 175  
gct acc aac gat tgc atc gat gct tat ttc gga gct gct tgc aac tct 576  
Ala Thr Asn Asp Cys Ile Asp Ala Tyr Phe Gly Ala Ala Cys Asn Ser  
180 185 190  
ttc atc cac atc gtg atg tac tcc tac tac ctc atg tct gct ttg gga 624  
Phe Ile His Ile Val Met Tyr Ser Tyr Tyr Leu Met Ser Ala Leu Gly  
195 200 205  
att aga tgc cct tgg aag aga tat atc acc cag gct cag atg ttg caa 672  
Ile Arg Cys Pro Trp Lys Arg Tyr Ile Thr Gn Ala Gn Met Leu Gn  
210 215 220  
ttc gtg atc gtg ttc gct cat gct gtt ttc gtg ctg aga caa aag cac 720  
Phe Val Ile Val Phe Ala His Ala Val Phe Val Leu Arg Gn Lys His  
225 230 235 240  
tgc cct gtt act ttg cct tgg gca caa atg ttc gtg atg aca aat atg 768  
Cys Pro Val Thr Leu Pro Trp Ala Gn Met Phe Val Met Thr Asn Met  
245 250 255  
ttg gtg ctg ttc gga aac ttc tac ctc aag gct tac tct aac aag tct 816  
Leu Val Leu Phe Gly Asn Phe Tyr Leu Lys Ala Tyr Ser Asn Lys Ser  
260 265 270 275  
agg gga gat gga gct tct tct gtt aag cct gct gag act act aga gca 864  
Arg Gly Asp Gly Ala Ser Ser Val Lys Pro Ala Gu Thr Thr Arg Ala  
280 285  
cct tct gtg aga aga acc agg tcc agg aag atc gat tga 903  
Pro Ser Val Arg Arg Thr Arg Ser Arg Lys Ile Asp  
290 295 300

<210> 65  
<211> 300  
<212> PRT  
<213> *Ostreococcus tauri*

<400> 65  
Met Ser Ala Ser Gly Ala Leu Leu Pro Ala Ile Ala Phe Ala Ala Tyr  
1 5 10 15  
Ala Tyr Ala Thr Tyr Ala Tyr Ala Phe Gu Trp Ser His Ala Asn Gly  
20 25 30  
Ile Asp Asn Val Asp Ala Arg Gu Trp Ile Gly Ala Leu Ser Leu Arg  
35 40 45  
Leu Pro Ala Ile Ala Thr Thr Met Tyr Leu Leu Phe Cys Leu Val Gly  
Sei te 86

PF58307. txt

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

<210> 66  
 <211> 834  
 <212> DNA  
 <213> Pavlova sp

<220>  
 <221> CDS  
 <222> (1)..(834)

<400>	66																	
atg	atg	ttg	gcc	gca	ggc	tat	ctt	cta	gtg	ctc	tcg	gcc	gct	cgc	cag			
Met	Met	Leu	Ala	Ala	Gly	Tyr	Leu	Leu	Val	Leu	Ser	Ala	Ala	Arg	Gln			48
1				5					10					15				
agc	ttc	cag	cag	gac	att	gac	aac	ccc	aac	ggg	gcc	tac	tcg	acc	tcg			
Ser	Phe	Gln	Gln	Asp	Ile	Asp	Asn	Pro	Asn	Gly	Ala	Tyr	Ser	Thr	Ser			96
			20					25					30					
tgg	act	ggc	ctg	ccc	att	gtg	atg	tct	gtg	gtc	tat	ctc	agc	ggg	gtg			
Trp	Thr	Gly	Leu	Pro	Ile	Val	Met	Ser	Val	Val	Tyr	Leu	Ser	Gly	Val			144
			35				40					45						
ttt	ggg	ctc	aca	aag	tac	ttc	gag	aac	cgg	aag	ccc	atg	acg	ggg	ctg			
Phe	Gly	Leu	Thr	Lys	Tyr	Phe	Glu	Asn	Arg	Lys	Pro	Met	Thr	Gly	Leu			192
	50				55						60							
aag	gac	tac	atg	ttc	act	tac	aat	ctc	tac	cag	gtg	atc	atc	aac	gtg			
Lys	Asp	Tyr	Met	Phe	Thr	Tyr	Asn	Leu	Tyr	Gln	Val	Ile	Ile	Asn	Val			240
65					70				75						80			
tgg	tgc	gtg	gtg	gcc	ttt	ctc	ctg	gag	gtg	cgg	cgt	gcg	ggc	atg	tca			
Trp	Cys	Val	Val	Ala	Phe	Leu	Leu	Glu	Val	Arg	Arg	Ala	Gly	Met	Ser			288
				85					90					95				
ctc	atc	ggc	aat	aag	gtg	gac	ctt	ggg	ccc	aac	tcc	ttc	agg	ctc	ggc			
Leu	Ile	Gly	Asn	Lys	Val	Asp	Leu	Gly	Pro	Asn	Ser	Phe	Arg	Leu	Gly			336
			100					105					110					
ttc	gtc	acg	tgg	gtg	cac	tac	aac	aag	tac	gtg	gag	ctc	ctc	gac				
Phe	Val	Thr	Trp	Val	His	Tyr	Asn	Asn	Lys	Tyr	Val	Glu	Leu	Leu	Asp			384
			115				120					125						
acc	cta	tgg	atg	gtg	ctg	cgc	aag	aag	acg	cag	cag	gtc	tcc	ttc	ctc			
Thr	Leu	Trp	Met	Val	Leu	Arg	Lys	Lys	Thr	Gln	Gln	Val	Ser	Phe	Leu			432



Al a Val Lys Ala Gl u  
275

<210> 68  
<211> 1077  
<212> DNA  
<213> Thal assi osi r a pseudonana

<220>  
<221> CDS  
<222> (1).. (1077)

<400> 68  
at g t gc t ca t ca ccg ccg t ca caa t cc aaa aca aca t cc ct c ct a gca 48  
Met Oys Ser Ser Pro Pro Ser Gl n Ser Lys Thr Thr Ser Leu Leu Ala  
1 5 10 15  
cgg t ac acc acc gcc gcc ct c ct c ct c acc ct c aca aca t gg t gc 96  
Arg Tyr Thr Thr Ala Ala Leu Leu Leu Thr Leu Thr Thr Trp Cys  
20 25 30  
cac t tc gcc t tc cca gcc gcc acc gcc aca ccc ggc ct c acc gcc gaa 144  
His Phe Ala Phe Pro Ala Ala Thr Ala Thr Pro Gly Leu Thr Ala Gl u  
35 40 45  
at g cac t cc t ac aaa gt c cca ct c ggt ct c acc gt a t tc t ac ct g ct g 192  
Met His Ser Tyr Lys Val Pro Leu Gly Leu Thr Val Phe Tyr Leu Leu  
50 55 60  
agt ct a ccg t ca ct a aag t ac gt t acg gac aac t ac ct t gcc aaa aag 240  
Ser Leu Pro Ser Leu Lys Tyr Val Thr Asp Asn Tyr Leu Ala Lys Lys  
65 70 75 80  
t at gat at g aag t ca ct c ct a acg gaa t ca at g gt g t tg t ac aat gt g 288  
Tyr Asp Met Lys Ser Leu Leu Thr Gl u Ser Met Val Leu Tyr Asn Val  
85 90 95  
gcg caa gt g ct g ct c aat ggg t gg acg gt g t at gcg at t gt g gat gcg 336  
Ala Gl n Val Leu Leu Asn Gly Trp Thr Val Tyr Ala Ile Val Asp Ala  
100 105 110  
gt g at g aat aga gac cat ccg t tt at t gga agt aga agt t tg gt t ggg 384  
Val Met Asn Arg Asp His Pro Phe Ile Gly Ser Arg Ser Leu Val Gly  
115 120 125  
gcg gcg t tg cat agt ggg agc t cg t at gcg gt g t gg gt t cat t at t gt 432  
Ala Ala Leu His Ser Gly Ser Ser Tyr Ala Val Trp Val His Tyr Cys  
130 140  
gat aag t at t tg gag t tc t tt gat acg t at t tt at g gt g t tg agg ggg 480  
Asp Lys Tyr Leu Gl u Phe Phe Asp Thr Tyr Phe Met Val Leu Arg Gly  
145 150 155 160  
aaa at g gac cag gt c t cc t tc ct c cac at c t ac cac cac acg acc at a 528  
Lys Met Asp Gl n Val Ser Phe Leu His Ile Tyr His His Thr Thr Ile  
165 170 175  
gcg t gg gca t gg t gg at c gcc ct c cgc t tc t cc ccc ggt gga gac at t 576  
Ala Trp Ala Trp Trp Ile Ala Leu Arg Phe Ser Pro Gly Gly Asp Ile  
180 185 190  
t ac t tc ggg gca ct c ct c aac t cc at c at c cac gt c ct c at g t at t cc 624  
Tyr Phe Gly Ala Leu Leu Asn Ser Ile Ile His Val Leu Met Tyr Ser  
195 200 205  
t ac t ac gcc ct t gcc ct a ct c aag gt c agt t gt cca t gg aaa cga t ac 672  
Tyr Tyr Ala Leu Ala Leu Lys Val Ser Cys Val Ser Pro Trp Lys Arg Tyr  
210 215 220  
ct g act caa gct caa t ta t tg caa t tc aca agt gt g gt g gt t t at acg 720  
Leu Thr Gl n Ala Gl n Leu Leu Gl n Phe Thr Ser Val Val Val Tyr Thr  
225 230 235 240  
ggg t gt acg ggt t at act cat t ac t at cat acg aag cat gga gcg gat 768  
Gly Cys Thr Gly Tyr Thr His Tyr Tyr His Thr Lys His Gly Ala Asp  
245 250 255  
gag aca cag cct agt t ta gga acg t at t at t tc t gt t gt gga gt g cag 816  
Gl u Thr Gl n Pro Ser Leu Gly Thr Tyr Tyr Phe Cys Cys Gly Val Gl n  
260 265 270  
gt g t tt gag at g gt t agt t tg t tt gt a ct c t tt t cc at c t tt t at aaa 864  
Val Phe Gl u Met Val Ser Leu Phe Val Leu Phe Ser Ile Phe Tyr Lys  
275 280 285

PF58307. txt

cga	t cc	t at	t cg	aag	aag	aac	aag	t ca	gga	gga	aag	gat	agc	aag	aag	912
Arg	Ser	Tyr	Ser	Lys	Lys	Asn	Lys	Ser	G y	G y	Lys	Asp	Ser	Lys	Lys	
	290					295					300					
aat	gat	gat	ggg	aat	aat	gag	gat	caa	t gt	cac	aag	gct	at g	aag	gat	960
Asn	Asp	Asp	G y	Asn	Asn	G u	Asp	G n	Cys	Hi s	Lys	Al a	Mët	Lys	Asp	
305				310						315					320	
at a	t cg	gag	ggt	gcg	aag	gag	gt t	gt g	ggg	cat	gca	gcg	aag	gat	gct	1008
I l e	Ser	G u	G y	Al a	Lys	G u	Val	Val	G y	Hi s	Al a	Al a	Lys	Asp	Al a	
				325					330					335		
gga	aag	t t g	gt g	gct	acg	gcg	agt	aag	gct	gt a	aag	agg	aag	gat	act	1056
G y	Lys	Leu	Val	Al a	Thr	Al a	Ser	Lys	Al a	Val	Lys	Arg	Lys	G y	Thr	
			340					345					350			
cgt	gt t	act	ggt	gcc	at g	t ag										1077
Arg	Val	Thr	G y	Al a	Mët											
		355														

<210> 69

<211> 358

<212> PRT

<213> Thal assi osi r a pseudonana

<400> 69

Mët	Cys	Ser	Ser	Pro	Pro	Ser	G n	Ser	Lys	Thr	Thr	Ser	Leu	Leu	Al a	
1				5					10					15		
Arg	Tyr	Thr	Thr	Al a	Al a	Leu	Leu	Leu	Leu	Thr	Leu	Thr	Thr	Tr p	Cys	
			20					25					30			
Hi s	Phe	Al a	Phe	Pro	Al a	Al a	Thr	Al a	Thr	Pro	G y	Leu	Thr	Al a	G u	
	35						40					45				
Mët	Hi s	Ser	Tyr	Lys	Val	Pro	Leu	G y	Leu	Thr	Val	Phe	Tyr	Leu	Leu	
	50				55						60					
Ser	Leu	Pro	Ser	Leu	Lys	Tyr	Val	Thr	Asp	Asn	Tyr	Leu	Al a	Lys	Lys	
65					70					75					80	
Tyr	Asp	Mët	Lys	Ser	Leu	Leu	Thr	G u	Ser	Mët	Val	Leu	Tyr	Asn	Val	
				85				90						95		
Al a	G n	Val	Leu	Leu	Asn	G y	Tr p	Thr	Val	Tyr	Al a	I l e	Val	Asp	Al a	
			100					105					110			
Val	Mët	Asn	Arg	Asp	Hi s	Pro	Phe	I l e	G y	Ser	Arg	Ser	Leu	Val	G y	
		115					120					125				
Al a	Al a	Leu	Hi s	Ser	G y	Ser	Ser	Tyr	Al a	Val	Tr p	Val	Hi s	Tyr	Cys	
	130					135					140					
Asp	Lys	Tyr	Leu	G u	Phe	Asp	Thr	Tyr	Phe	Mët	Val	Leu	Arg	G y		
145					150				155					160		
Lys	Mët	Asp	G n	Val	Ser	Phe	Leu	Hi s	I l e	Tyr	Hi s	Hi s	Thr	Thr	I l e	
				165					170					175		
Al a	Tr p	Al a	Tr p	Tr p	I l e	Al a	Leu	Arg	Phe	Ser	Pro	G y	G y	Asp	I l e	
			180					185					190			
Tyr	Phe	G y	Al a	Leu	Leu	Asn	Ser	I l e	I l e	Hi s	Val	Leu	Mët	Tyr	Ser	
	195						200					205				
Tyr	Tyr	Al a	Leu	Al a	Leu	Leu	Lys	Val	Ser	Cys	Pro	Tr p	Lys	Arg	Tyr	
	210					215					220					
Leu	Thr	G n	Al a	G n	Leu	Leu	G n	Phe	Thr	Ser	Val	Val	Tyr	Thr		
225					230				235					240		
G y	Cys	Thr	G y	Tyr	Thr	Hi s	Tyr	Tyr	Hi s	Thr	Lys	Hi s	G y	Al a	Asp	
				245					250					255		
G u	Thr	G n	Pro	Ser	Leu	G y	Thr	Tyr	Tyr	Phe	Cys	Cys	G y	Val	G n	
			260					265					270			
Val	Phe	G u	Mët	Val	Ser	Leu	Phe	Val	Leu	Phe	Ser	I l e	Phe	Tyr	Lys	
	275						280					285				
Arg	Ser	Tyr	Ser	Lys	Lys	Asn	Lys	Ser	G y	G y	Lys	Asp	Ser	Lys	Lys	
	290					295					300					
Asn	Asp	Asp	G y	Asn	Asn	G u	Asp	G n	Cys	Hi s	Lys	Al a	Mët	Lys	Asp	
305				310						315					320	
I l e	Ser	G u	G y	Al a	Lys	G u	Val	Val	G y	Hi s	Al a	Al a	Lys	Asp	Al a	
				325					330					335		
G y	Lys	Leu	Val	Al a	Thr	Al a	Ser	Lys	Al a	Val	Lys	Arg	Lys	G y	Thr	
			340					345					350			
Arg	Val	Thr	G y	Al a	Mët											
		355														



<210> 70  
 <211> 903  
 <212> DNA  
 <213> *Ostreococcus tauri*

<220>  
 <221> CDS  
 <222> (1)..(903)

<400> 70  
 at g agc gcc t cc ggt gcg ct g ct g ccc gcg at c gcg t cc gcc gcg t ac 48  
 Met Ser Ala Ser Gly Ala Leu Leu Pro Ala Ile Ala Ser Ala Ala Tyr  
 1 5 10 15  
 gcg t ac gcg acg t ac gcc t ac gcc t t t gag t gg t cg cac gcg aat ggc 96  
 Ala Tyr Ala Thr Tyr Ala Tyr Ala Phe Gu Trp Ser His Ala Asn Gly  
 20 25 30  
 at c gac aac gt c gac gcg cgc gag t gg at c ggt gcg ct g t cg t t g agg 144  
 Ile Asp Asn Val Asp Ala Arg Gu Trp Ile Gly Ala Leu Ser Leu Arg  
 35 40 45  
 ct c ccg gcg at c gcg acg acg at g t ac ct g t t g t t c t gc ct g gt c gga 192  
 Leu Pro Ala Ile Ala Thr Thr Met Tyr Leu Leu Phe Cys Leu Val Gly  
 50 55 60  
 ccg agg t t g at g gcg aag cgc gag gcg t t c gac ccg aag ggg t t c at g 240  
 Pro Arg Leu Met Ala Lys Arg Gu Ala Phe Asp Pro Lys Gly Phe Met  
 65 70 75 80  
 ct g gcg t ac aat gcg t at cag acg gcg t t c aac gt c gt c gt g ct c ggg 288  
 Leu Ala Tyr Asn Ala Tyr Gn Thr Ala Phe Asn Val Val Val Leu Gly  
 85 90 95  
 at g t t c gcg cga gag at c t cg ggg ct g ggg cag ccc gt g t gg ggg t ca 336  
 Met Phe Ala Arg Gu Ile Ser Gly Leu Gly Gn Pro Val Trp Gly Ser  
 100 105 110  
 acc at g ccg t gg agc gat aga aaa t cg t t t aag at c ct c ct c ggg gt g 384  
 Thr Met Pro Trp Ser Asp Arg Lys Ser Phe Lys Ile Leu Leu Gly Val  
 115 120 125  
 t gg t t g cac t ac aac aac aaa t at t t g gag ct a t t g gac act gt g t t c 432  
 Trp Leu His Tyr Asn Asn Lys Tyr Leu Gu Leu Leu Asp Thr Val Phe  
 130 135 140  
 at g gt t gcg cgc aag aag acg aag cag t t g agc t t c t t g cac gt t t at 480  
 Met Val Ala Arg Lys Lys Thr Lys Gn Leu Ser Phe Leu His Val Tyr  
 145 150 155 160  
 cat cac gcc ct g t t g at c t gg gcg t gg t gg t t g gt g t gt cac t t g at g 528  
 His His Ala Leu Leu Ile Trp Ala Trp Trp Leu Val Cys His Leu Met  
 165 170 175  
 gcc acg aac gat t gt at c gat gcc t ac t t c ggc gcg gcg t gc aac t cg 576  
 Ala Thr Asn Asp Cys Ile Asp Ala Tyr Phe Gly Ala Ala Cys Asn Ser  
 180 185 190  
 t t c att cac at c gt g at g t ac t cg t at t at ct c at g t cg gcg ct c ggc 624  
 Phe Ile His Ile Val Met Tyr Ser Tyr Tyr Leu Met Ser Ala Leu Gly  
 195 200 205  
 at t cga t gc ccg t gg aag cga t ac at c acc cag gct caa at g ct c caa 672  
 Ile Arg Cys Pro Trp Lys Arg Tyr Ile Thr Gn Ala Gn Met Leu Gn  
 210 215 220  
 t t c gt c att gt c t t c gcg cac gcc gt g t t c gt g ct g cgt cag aag cac 720  
 Phe Val Ile Val Phe Ala His Ala Val Phe Val Leu Arg Gn Lys His  
 225 230 235 240  
 t gc ccg gt c acc ct t cct t gg gcg caa at g t t c gt c at g acg aac at g 768  
 Cys Pro Val Thr Leu Pro Trp Ala Gn Met Phe Val Met Thr Asn Met  
 245 250 255  
 ct c gt g ct c t t c ggg aac t t c t ac ct c aag gcg t ac t cg aac aag t cg 816  
 Leu Val Leu Phe Gly Asn Phe Tyr Leu Lys Ala Tyr Ser Asn Lys Ser  
 260 265 270  
 cgc ggc gac ggc gcg agt t cc gt g aaa cca gcc gag acc acg cgc gcg 864  
 Arg Gly Asp Gly Ala Ser Ser Val Lys Pro Ala Gu Thr Thr Arg Ala  
 275 280 285  
 ccc agc gt g cga cgc acg cga t ct cga aaa at t gac t aa 903  
 Pro Ser Val Arg Arg Thr Arg Ser Arg Lys Ile Asp  
 290 295 300

<210> 71  
 <211> 300  
 <212> PRT  
 <213> *Ostreococcus tauri*

<400> 71  
 Met Ser Ala Ser Gly Ala Leu Leu Pro Ala Ile Ala Ser Ala Ala Tyr  
 1 5 10 15  
 Ala Tyr Ala Thr Tyr Ala Tyr Ala Phe Gu Trp Ser His Ala Asn Gly  
 20 25 30  
 Ile Asp Asn Val Asp Ala Arg Gu Trp Ile Gly Ala Leu Ser Leu Arg  
 35 40 45  
 Leu Pro Ala Ile Ala Thr Thr Met Tyr Leu Leu Phe Cys Leu Val Gly  
 50 55 60  
 Pro Arg Leu Met Ala Lys Arg Gu Ala Phe Asp Pro Lys Gly Phe Met  
 65 70 75 80  
 Leu Ala Tyr Asn Ala Tyr Gn Thr Ala Phe Asn Val Val Val Leu Gly  
 85 90 95  
 Met Phe Ala Arg Gu Ile Ser Gly Leu Gly Gn Pro Val Trp Gly Ser  
 100 105 110  
 Thr Met Pro Trp Ser Asp Arg Lys Ser Phe Lys Ile Leu Leu Gly Val  
 115 120 125  
 Trp Leu His Tyr Asn Asn Lys Tyr Leu Gu Leu Leu Asp Thr Val Phe  
 130 135 140  
 Met Val Ala Arg Lys Lys Thr Lys Gn Leu Ser Phe Leu His Val Tyr  
 145 150 155 160  
 His His Ala Leu Leu Ile Trp Ala Trp Trp Leu Val Cys His Leu Met  
 165 170 175  
 Ala Thr Asn Asp Cys Ile Asp Ala Tyr Phe Gly Ala Ala Cys Asn Ser  
 180 185 190  
 Phe Ile His Ile Val Met Tyr Ser Tyr Tyr Leu Met Ser Ala Leu Gly  
 195 200 205  
 Ile Arg Cys Pro Trp Lys Arg Tyr Ile Thr Gn Ala Gn Met Leu Gn  
 210 215 220  
 Phe Val Ile Val Phe Ala His Ala Val Phe Val Leu Arg Gn Lys His  
 225 230 235 240  
 Cys Pro Val Thr Leu Pro Trp Ala Gn Met Phe Val Met Thr Asn Met  
 245 250 255  
 Leu Val Leu Phe Gly Asn Phe Tyr Leu Lys Ala Tyr Ser Asn Lys Ser  
 260 265 270  
 Arg Gly Asp Gly Ala Ser Ser Val Lys Pro Ala Gu Thr Thr Arg Ala  
 275 280 285  
 Pro Ser Val Arg Arg Thr Arg Ser Arg Lys Ile Asp  
 290 295 300

<210> 72  
 <211> 339  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (3)..(29)  
 <223> xaa in position 3 to 29 is any amino acid  
 <220>  
 <221> Variant  
 <222> (30)..(43)  
 <223> xaa in position 30 to 43 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (45)..(50)  
 <223> xaa in position 45 to 50 is any amino acid  
 <220>  
 <221> Variant  
 <222> (53)..(73)  
 <223> xaa in position 53 to 73 is any amino acid

<220>  
<221> Variant  
<222> (74)..(79)  
<223> xaa in position 74 to 79 is any or no amino acid  
<220>  
<221> Variant  
<222> (81)..(82)  
<223> xaa in position 81 to 82 is any amino acid  
<220>  
<221> Variant  
<222> (85)..(86)  
<223> xaa in position 85 to 86 is any amino acid  
<220>  
<221> Variant  
<222> (88)..(90)  
<223> xaa in position 88 to 90 is any amino acid  
<220>  
<221> Variant  
<222> (92)..(110)  
<223> xaa in position 92 to 110 is any amino acid  
<220>  
<221> Variant  
<222> (112)..(127)  
<223> xaa in position 112 to 127 is any amino acid  
<220>  
<221> Variant  
<222> (128)..(128)  
<223> xaa in position 128 is any or no amino acid  
<220>  
<221> Variant  
<222> (130)..(130)  
<223> xaa in position 130 is any amino acid  
<220>  
<221> Variant  
<222> (133)..(134)  
<223> xaa in position 133 to 134 is any amino acid  
<220>  
<221> Variant  
<222> (137)..(137)  
<223> xaa in position 137 is any amino acid  
<220>  
<221> Variant  
<222> (139)..(140)  
<223> xaa in position 139 to 140 is any amino acid  
<220>  
<221> Variant  
<222> (143)..(144)  
<223> xaa in position 143 to 144 is any amino acid  
<220>  
<221> Variant  
<222> (147)..(147)  
<223> xaa in position 147 is any amino acid  
<220>  
<221> Variant  
<222> (149)..(149)  
<223> xaa in position 149 is any amino acid  
<220>  
<221> Variant  
<222> (151)..(152)  
<223> xaa in position 151 to 152 is any amino acid  
<220>  
<221> Variant  
<222> (154)..(154)  
<223> xaa in position 154 is any amino acid  
<220>  
<221> Variant  
<222> (159)..(159)  
<223> xaa in position 159 is any amino acid

<220>  
<221> Variant  
<222> (163)..(166)  
<223> xaa in position 163 to 166 is any amino acid  
<220>  
<221> Variant  
<222> (170)..(179)  
<223> xaa in position 170 to 179 is any amino acid  
<220>  
<221> Variant  
<222> (180)..(182)  
<223> xaa in position 180 to 182 is any or no amino acid  
<220>  
<221> Variant  
<222> (184)..(184)  
<223> xaa in position 184 is any amino acid  
<220>  
<221> Variant  
<222> (188)..(190)  
<223> xaa in position 188 to 190 is any amino acid  
<220>  
<221> Variant  
<222> (193)..(193)  
<223> xaa in position 193 is any amino acid  
<220>  
<221> Variant  
<222> (196)..(197)  
<223> xaa in position 196 to 197 is any amino acid  
<220>  
<221> Variant  
<222> (203)..(206)  
<223> xaa in position 203 to 206 is any amino acid  
<220>  
<221> Variant  
<222> (208)..(210)  
<223> xaa in position 208 to 210 is any amino acid  
<220>  
<221> Variant  
<222> (217)..(217)  
<223> xaa in position 217 is any amino acid  
<220>  
<221> Variant  
<222> (222)..(223)  
<223> xaa in position 222 to 223 is any amino acid  
<220>  
<221> Variant  
<222> (226)..(252)  
<223> xaa in position 226 to 252 is any amino acid  
<220>  
<221> Variant  
<222> (253)..(267)  
<223> xaa in position 253 to 267 is any or no amino acid  
<220>  
<221> Variant  
<222> (269)..(272)  
<223> xaa in position 269 to 272 is any amino acid  
<220>  
<221> Variant  
<222> (276)..(277)  
<223> xaa in position 276 to 277 is any amino acid  
<220>  
<221> Variant  
<222> (280)..(282)  
<223> xaa in position 280 to 282 is any amino acid  
<220>  
<221> Variant  
<222> (284)..(291)  
<223> xaa in position 284 to 291 is any amino acid  
Seite 94

<220>  
 <221> Variant  
 <222> (292)..(324)  
 <223> xaa in position 292 to 324 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (326)..(337)  
 <223> xaa in position 326 to 337 is any amino acid  
 <220>  
 <221> Variant  
 <222> (338)..(338)  
 <223> xaa in position 338 is any or no amino acid

<400> 72  
 Leu Leu xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 20 25 30  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Pro xaa xaa xaa xaa  
 35 40 45  
 xaa xaa Tyr Leu xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 50 55 60  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Met  
 65 70 75 80  
 xaa xaa Tyr Asn xaa xaa G n xaa xaa xaa Asn xaa xaa xaa xaa xaa  
 85 90 95  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa  
 100 105 110  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 115 120 125  
 Trp xaa His Tyr xaa xaa Lys Tyr xaa G u xaa xaa Asp Thr xaa xaa  
 130 135 140  
 Met Val xaa Arg xaa Lys xaa xaa G n xaa Ser Phe Leu His xaa Tyr  
 145 150 155 160  
 His His xaa xaa xaa xaa Trp Ala Trp xaa xaa xaa xaa xaa xaa xaa  
 165 170 175  
 xaa xaa xaa xaa xaa xaa Asp xaa Tyr Phe Gly xaa xaa xaa xaa Asn Ser  
 180 185 190  
 xaa Ile His xaa xaa Met Tyr Ser Tyr Tyr xaa xaa xaa xaa Leu xaa  
 195 200 205  
 xaa xaa Cys Pro Trp Lys Arg Tyr xaa Thr G n Ala G n xaa xaa G n  
 210 215 220  
 Phe xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 225 230 235 240  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 245 250 255  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Met xaa xaa xaa xaa  
 260 265 270  
 Val Leu Phe xaa xaa Phe Tyr xaa xaa xaa Tyr xaa xaa xaa xaa xaa  
 275 280 285  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 290 295 300  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 305 310 315 320  
 xaa xaa xaa xaa Ala xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 325 330 335  
 xaa xaa Val

<210> 73  
 <211> 45  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (4)..(4)

<223> xaa in position 4 is Ala or Gly  
<220>  
<221> Variant  
<222> (5)..(5)  
<223> xaa in position 5 is Ala or Leu  
<220>  
<221> Variant  
<222> (6)..(6)  
<223> xaa in position 6 is any amino acid  
<220>  
<221> Variant  
<222> (9)..(9)  
<223> xaa in position 9 is Phe or Ile  
<220>  
<221> Variant  
<222> (12)..(12)  
<223> xaa in position 12 is any or no amino acid  
<220>  
<221> Variant  
<222> (14)..(14)  
<223> xaa in position 14 is any or no amino acid  
<220>  
<221> Variant  
<222> (20)..(20)  
<223> xaa in position 20 is any amino acid  
<220>  
<221> Variant  
<222> (21)..(21)  
<223> xaa in position 21 is Leu or Met  
<220>  
<221> Variant  
<222> (22)..(22)  
<223> xaa in position 22 is Ala or Ser  
<220>  
<221> Variant  
<222> (23)..(23)  
<223> xaa in position 23 is Ala or Leu  
<220>  
<221> Variant  
<222> (25)..(26)  
<223> xaa in position 25 to 26 is any amino acid  
<220>  
<221> Variant  
<222> (27)..(27)  
<223> xaa in position 27 is Arg or Ser  
<220>  
<221> Variant  
<222> (34)..(34)  
<223> xaa in position 34 is Ile or Leu  
<220>  
<221> Variant  
<222> (39)..(39)  
<223> xaa in position 39 is Leu or Met  
<220>  
<221> Variant  
<222> (40)..(40)  
<223> xaa in position 40 is Leu or Val  
<220>  
<221> Variant  
<222> (43)..(43)  
<223> xaa in position 43 is Cys, Thr or Val  
<220>  
<221> Variant  
<222> (44)..(44)  
<223> xaa in position 44 is any amino acid  
<220>  
<221> Variant  
<222> (45)..(45)

<223> xaa in position 45 is Cys or Val

<400> 73

```

Tyr Phe Gly xaa xaa xaa Asn Ser xaa Ile His xaa Val xaa Met Tyr
 1           5           10
Ser Tyr Tyr xaa xaa xaa xaa Leu xaa xaa xaa Cys Pro Trp Lys Arg
           20           25           30
Tyr xaa Thr Gln Ala Gln xaa xaa Gln Phe xaa xaa xaa
           35           40           45

```

<210> 74

<211> 43

<212> PRT

<213> Artificial sequence

<220>

<221> Variant

<222> (2)..(2)

<223> xaa in position 2 is Leu or Val

<220>

<221> Variant

<222> (5)..(5)

<223> xaa in position 5 is Cys or Asn

<220>

<221> Variant

<222> (6)..(6)

<223> xaa in position 6 is Asp or Asn

<220>

<221> Variant

<222> (9)..(10)

<223> xaa in position 9 to 10 is any or no amino acid

<220>

<221> Variant

<222> (12)..(12)

<223> xaa in position 12 is any amino acid

<220>

<221> Variant

<222> (13)..(14)

<223> xaa in position 13 to 14 is any or no amino acid

<220>

<221> Variant

<222> (17)..(17)

<223> xaa in position 17 is any amino acid

<220>

<221> Variant

<222> (18)..(18)

<223> xaa in position 18 is Phe or Trp

<220>

<221> Variant

<222> (21)..(21)

<223> xaa in position 21 is Ala or Leu

<220>

<221> Variant

<222> (23)..(23)

<223> xaa in position 23 is any amino acid

<220>

<221> Variant

<222> (25)..(26)

<223> xaa in position 25 to 26 is any amino acid

<220>

<221> Variant

<222> (28)..(28)

<223> xaa in position 28 is Leu or Val

<220>

<221> Variant

<222> (33)..(33)

<223> xaa in position 33 is Ile or Val

<220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Ala, Thr or Val  
 <220>  
 <221> Variant  
 <222> (38)..(38)  
 <223> xaa in position 38 is any amino acid  
 <220>  
 <221> Variant  
 <222> (39)..(39)  
 <223> xaa in position 39 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (40)..(40)  
 <223> xaa in position 40 is any amino acid  
  
 <400> 74  
 Trp xaa His Tyr xaa xaa Lys Tyr xaa xaa Leu xaa xaa xaa Asp Thr  
   1          5          10          15  
 xaa xaa Met Val xaa Arg xaa Lys xaa xaa Gln xaa Ser Phe Leu His  
           20          25          30  
 xaa Tyr His His xaa xaa xaa xaa Trp Ala Trp  
           35          40

<210> 75  
 <211> 37  
 <212> PRT  
 <213> Artificial sequence  
  
 <220>  
 <221> Variant  
 <222> (2)..(3)  
 <223> xaa in position 2 to 3 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is Gly, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(14)  
 <223> xaa in position 13 to 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is Gu or Lys  
 <220>



<221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Ala, Asn or Ser  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Lys or Arg  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Ala, Gly or Ser  
 <220>  
 <221> Variant  
 <222> (23)..(23)  
 <223> xaa in position 23 is any amino acid  
 <220>  
 <221> Variant  
 <222> (24)..(24)  
 <223> xaa in position 24 is Gu or Gly  
 <220>  
 <221> Variant  
 <222> (25)..(29)  
 <223> xaa in position 25 to 29 is any amino acid  
 <220>  
 <221> Variant  
 <222> (30)..(30)  
 <223> xaa in position 30 is Asn, Pro or Gn  
 <220>  
 <221> Variant  
 <222> (31)..(32)  
 <223> xaa in position 31 to 32 is Ala, Asp or Gu  
 <220>  
 <221> Variant  
 <222> (33)..(33)  
 <223> xaa in position 33 is Gly, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (34)..(34)  
 <223> xaa in position 34 is Asn, Gn or Thr  
 <220>  
 <221> Variant  
 <222> (35)..(36)  
 <223> xaa in position 35 to 36 is any amino acid  
 <220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Ala, Asp or Pro

<400> 75  
 Met xaa xaa Leu xaa Val Leu Phe xaa xaa Phe Tyr xaa xaa xaa Tyr  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 20 25 30  
 xaa xaa xaa xaa xaa  
 35

<210> 76  
 <211> 19  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(3)  
 <223> xaa in position 2 to 3 is any amino acid  
 <220>  
 <221> Variant

<222> (4)..(7)  
 <223> xaa in position 4 to 7 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is Phe, Leu or Val  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Ala, Leu or Val  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Thr or Val  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Ala, Ile or Leu  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Phe, Ile or Leu

<400> 76  
 Lys xaa xaa xaa xaa xaa xaa Met xaa xaa Tyr Asn xaa xaa Gln xaa  
 1 5 10 15  
 xaa xaa Asn

<210> 77  
 <211> 1560  
 <212> DNA  
 <213> Traustochytrium sp.

<220>  
 <221> CDS  
 <222> (1)..(1560)

<400> 77  
 atg acg gtc ggc tac gac gag gag atc ccg ttc gag cag gtc cgc gcg 48  
 Met Thr Val Gly Tyr Asp Gln Gln Ile Pro Phe Gln Gln Val Arg Ala  
 1 5 10 15  
 cac aac aag ccg gat gac gcc tgg tgc gcg atc cac ggg cac gtg tac 96  
 His Asn Lys Pro Asp Asp Ala Trp Cys Ala Ile His Gly His Val Tyr  
 20 25 30  
 gat gtg acc aag ttc gcg agc gtg cac ccg ggc ggc gac att atc ctg 144  
 Asp Val Thr Lys Phe Ala Ser Val His Pro Gly Gly Asp Ile Ile Leu  
 35 40 45  
 ctg gcc gca ggc aag gag gcc acc gtg ctg tac gag act tac cat gtg 192  
 Leu Ala Ala Gly Lys Gln Ala Thr Val Leu Tyr Gln Thr Tyr His Val  
 50 55 60  
 cgg ggc gtc tcc gac gcg gtg ctg cgc aag tac ccg atc ggc aag ctg 240  
 Arg Gly Val Ser Asp Ala Val Leu Arg Lys Tyr Arg Ile Gly Lys Leu  
 65 70 75 80  
 ccg gac ggc caa ggc ggc gcg aac gag aag gaa aag ccg acg ctg tcc 288  
 Pro Asp Gly Gln Gly Ala Asn Gln Lys Gln Lys Arg Thr Leu Ser  
 85 90 95  
 ggc ctg tcc tcc gcc tcc tac tac acg tgg aac agc gac ttt tac agg 336  
 Gly Leu Ser Ser Ala Ser Tyr Tyr Thr Trp Asn Ser Asp Phe Tyr Arg  
 100 105 110

PF58307. txt

gt a	at g	cgc	gag	cgc	gt c	gt g	gct	cgg	ct c	aag	gag	cgc	ggc	aag	gcc	384
Val	Met	Arg	Gl u	Arg	Val	Val	Al a	Arg	Leu	Lys	Gl u	Arg	Gl y	Lys	Al a	
		115					120					125				
cgc	cgc	gga	ggc	t ac	gag	ct c	t gg	at c	aag	gcg	t t c	ct g	ct g	ct c	gt c	432
Arg	Arg	Gl y	Gl y	Tyr	Gl u	Leu	Tr p	I l e	Lys	Al a	Phe	Leu	Leu	Leu	Val	
		130					135				140					
ggc	t t c	t gg	agc	t cg	ct g	t ac	t gg	at g	t gc	acg	ct g	gac	ccc	t cg	t t c	480
Gl y	Phe	Tr p	Ser	Ser	Leu	Tyr	Tr p	Met	Cys	Thr	Leu	Asp	Pro	Ser	Phe	
		145			150					155					160	
ggg	gcc	at c	ct g	ggc	ggc	at g	t cg	ct g	ggc	gt c	t t t	gcc	gcc	t t t	gt g	528
Gl y	Al a	I l e	Leu	Al a	Al a	Met	Ser	Leu	Gl y	Val	Phe	Al a	Al a	Phe	Val	
				165				170						175		
ggc	acg	t gc	at c	cag	cac	gac	ggc	aac	cac	ggc	gcc	t t t	gcc	cag	t cg	576
Gl y	Thr	Cys	I l e	Gl n	Hi s	Asp	Gl y	Asn	Hi s	Gl y	Al a	Phe	Al a	Gl n	Ser	
			180					185						190		
cga	t gg	gt c	aac	aag	gt t	gcc	ggg	t gg	acg	ct c	gac	at g	at c	ggc	gcc	624
Arg	Tr p	Val	Asn	Lys	Val	Al a	Gl y	Tr p	Thr	Leu	Asp	Met	I l e	Gl y	Al a	
			195				200					205				
agc	ggc	at g	acg	t gg	gag	t t c	cag	cac	gt c	ct g	ggc	cac	cat	ccg	t ac	672
Ser	Gl y	Met	Thr	Tr p	Gl u	Phe	Gl n	Hi s	Val	Leu	Gl y	Hi s	Hi s	Pro	Tyr	
		210				215					220					
acg	aac	ct g	at c	gag	gag	gag	aac	ggc	ct g	caa	aag	gt g	agc	ggc	aag	720
Thr	Asn	Leu	I l e	Gl u	Gl u	Gl u	Asn	Gl y	Leu	Gl n	Lys	Val	Ser	Gl y	Lys	
		225			230					235					240	
aag	at g	gac	acc	aag	ct g	gcc	gac	cag	gag	agc	gat	ccg	gac	gt c	t t t	768
Lys	Met	Asp	Thr	Lys	Leu	Al a	Asp	Gl n	Gl u	Ser	Asp	Pro	Asp	Val	Phe	
			245					250						255		
t cc	acg	t ac	ccg	at g	cgc	ct g	cac	ccg	t gg	cac	cag	aag	ccg	t gg		816
Ser	Thr	Tyr	Pro	Met	Met	Arg	Leu	Hi s	Pro	Tr p	Hi s	Gl n	Lys	Arg	Tr p	
			260					265					270			
t ac	cac	cgt	t t c	cag	cac	at t	t ac	ggc	ccc	t t c	at c	t t t	ggc	t t c	at g	864
Tyr	Hi s	Arg	Phe	Gl n	Hi s	I l e	Tyr	Gl y	Pro	Phe	I l e	Phe	Gl y	Phe	Met	
		275					280					285				
acc	at c	aac	aag	gt g	gt c	acg	cag	gac	gt c	ggt	gt g	gt g	ct c	cgc	aag	912
Thr	I l e	Asn	Lys	Val	Val	Thr	Gl n	Asp	Val	Gl y	Val	Val	Leu	Arg	Lys	
		290				295					300					
cgg	ct c	t t c	cag	at t	gac	ggc	gag	t gc	cgg	t ac	gcg	agc	cca	at g	t ac	960
Arg	Leu	Phe	Gl n	I l e	Asp	Al a	Gl u	Cys	Arg	Tyr	Al a	Ser	Pro	Met	Tyr	
		305			310					315					320	
gt g	gcg	cgt	t t c	t gg	at c	at g	aag	gcg	ct c	acg	gt g	ct c	t ac	at g	gt g	1008
Val	Al a	Arg	Phe	Tr p	I l e	Met	Lys	Al a	Leu	Thr	Val	Leu	Tyr	Met	Val	
				325					330					335		
gcc	ct g	ccg	t gc	t ac	at g	cag	ggc	ccg	t gg	cac	ggc	ct c	aag	ct g	t t c	1056
Al a	Leu	Pro	Cys	Tyr	Met	Gl n	Gl y	Pro	Tr p	Hi s	Gl y	Leu	Lys	Leu	Phe	
			340					345					350			
gcg	at c	gcg	cac	t t t	acg	t gc	ggc	gag	gt g	ct c	gca	acc	at g	t t c	at t	1104
Al a	I l e	Al a	Hi s	Phe	Thr	Cys	Gl y	Gl u	Val	Leu	Al a	Thr	Met	Phe	I l e	
		355				360						365				
gt g	aac	cac	at c	at c	gag	ggc	gt c	t cg	t ac	gct	t cc	aag	gac	gcg	gt c	1152
Val	Asn	Hi s	I l e	I l e	Gl u	Gl y	Val	Ser	Tyr	Al a	Ser	Lys	Asp	Al a	Val	
		370				375					380					
aag	ggc	acg	at g	gcg	ccg	ccg	aag	acg	at g	cac	ggc	gt g	acg	ccc	at g	1200
Lys	Gl y	Thr	Met	Al a	Pro	Pro	Lys	Thr	Met	Hi s	Gl y	Val	Thr	Pro	Met	
		385			390					395				400		
aac	aac	acg	cgc	aag	gag	gt g	gag	gcg	gag	cgc	t cc	aag	t ct	ggc	gcc	1248
Asn	Asn	Thr	Arg	Lys	Gl u	Val	Gl u	Al a	Gl u	Al a	Ser	Lys	Ser	Gl y	Al a	
				405					410					415		
gt g	gt c	aag	t ca	gt c	ccg	ct c	gac	gac	t gg	gcc	gcc	gt c	cag	t gc	cag	1296
Val	Val	Lys	Ser	Val	Pro	Leu	Asp	Asp	Tr p	Al a	Al a	Val	Gl n	Cys	Gl n	
			420					425					430			
acc	t cg	gt g	aac	t gg	agc	gt c	ggc	t cg	t gg	t t c	t gg	aat	cac	t t t	t cc	1344
Thr	Ser	Val	Asn	Tr p	Ser	Val	Gl y	Ser	Tr p	Phe	Tr p	Asn	Hi s	Phe	Ser	
		435					440					445				
ggc	ggc	ct c	aac	cac	cag	at t	gag	cac	cac	ct g	t t c	ccc	ggg	ct c	agc	1392
Gl y	Gl y	Leu	Asn	Hi s	Gl n	I l e	Gl u	Hi s	Hi s	Leu	Phe	Pro	Gl y	Leu	Ser	
		450				455					460					
cac	gag	acg	t ac	t ac	cac	at c	cag	gac	gt c	gt t	cag	t cc	acc	t gc	gcc	1440
Hi s	Gl u	Thr	Tyr	Tyr	Hi s	I l e	Gl n	Asp	Val	Val	Gl n	Ser	Thr	Cys	Al a	

PF58307.txt

465					470					475					480	
gag	t ac	ggc	gt c	ccg	t ac	cag	cac	gag	cct	t cg	ct c	t gg	acc	gcg	t ac	1488
Gu	Tyr	Gly	Val	Pro	Tyr	Gln	His	Gu	Pro	Ser	Leu	Trp	Thr	Ala	Tyr	
				485					490					495		
t gg	aag	at g	ct c	gag	cac	ct c	cgt	cag	ct c	ggc	aat	gag	gag	acc	cac	1536
Trp	Lys	Met	Leu	Gu	His	Leu	Arg	Gln	Leu	Gly	Asn	Gu	Gu	Thr	His	
			500					505					510			
gag	t cc	t gg	cag	cgc	gct	gcc	t ga									1560
Gu	Ser	Trp	Gln	Arg	Ala	Ala										
		515					520									

<210> 78  
 <211> 519  
 <212> PRT  
 <213> *Tr aust ochyt ri um sp.*

<400> 78

Met	Thr	Val	Gly	Tyr	Asp	Gu	Gu	Ile	Pro	Phe	Gu	Gln	Val	Arg	Ala	
1				5					10					15		
His	Asn	Lys	Pro	Asp	Asp	Ala	Trp	Cys	Ala	Ile	His	Gly	His	Val	Tyr	
			20					25					30			
Asp	Val	Thr	Lys	Phe	Ala	Ser	Val	His	Pro	Gly	Gly	Asp	Ile	Ile	Leu	
		35					40					45				
Leu	Ala	Ala	Gly	Lys	Gu	Ala	Thr	Val	Leu	Tyr	Gu	Thr	Tyr	His	Val	
	50					55					60					
Arg	Gly	Val	Ser	Asp	Ala	Val	Leu	Arg	Lys	Tyr	Arg	Ile	Gly	Lys	Leu	
65				70					75					80		
Pro	Asp	Gly	Gln	Gly	Ala	Asn	Gu	Lys	Gu	Lys	Arg	Thr	Leu	Ser		
			85					90					95			
Gly	Leu	Ser	Ser	Ala	Ser	Tyr	Tyr	Thr	Trp	Asn	Ser	Asp	Phe	Tyr	Arg	
		100						105					110			
Val	Met	Arg	Gu	Arg	Val	Val	Ala	Arg	Leu	Lys	Gu	Arg	Gly	Lys	Ala	
		115					120					125				
Arg	Arg	Gly	Gly	Tyr	Gu	Leu	Trp	Ile	Lys	Ala	Phe	Leu	Leu	Leu	Val	
	130					135					140					
Gly	Phe	Trp	Ser	Ser	Leu	Tyr	Trp	Met	Cys	Thr	Leu	Asp	Pro	Ser	Phe	
145				150					155					160		
Gly	Ala	Ile	Leu	Ala	Ala	Met	Ser	Leu	Gly	Val	Phe	Ala	Ala	Phe	Val	
			165						170					175		
Gly	Thr	Cys	Ile	Gln	His	Asp	Gly	Asn	His	Gly	Ala	Phe	Ala	Gln	Ser	
		180						185					190			
Arg	Trp	Val	Asn	Lys	Val	Ala	Gly	Trp	Thr	Leu	Asp	Met	Ile	Gly	Ala	
		195					200					205				
Ser	Gly	Met	Thr	Trp	Gu	Phe	Gln	His	Val	Leu	Gly	His	His	Pro	Tyr	
	210					215					220					
Thr	Asn	Leu	Ile	Gu	Gu	Asn	Gly	Leu	Gln	Lys	Val	Ser	Gly	Lys		
225				230					235					240		
Lys	Met	Asp	Thr	Lys	Leu	Ala	Asp	Gln	Gu	Ser	Asp	Pro	Asp	Val	Phe	
			245					250					255			
Ser	Thr	Tyr	Pro	Met	Met	Arg	Leu	His	Pro	Trp	His	Gln	Lys	Arg	Trp	
		260						265					270			
Tyr	His	Arg	Phe	Gln	His	Ile	Tyr	Gly	Pro	Phe	Ile	Phe	Gly	Phe	Met	
		275					280					285				
Thr	Ile	Asn	Lys	Val	Val	Thr	Gln	Asp	Val	Gly	Val	Val	Leu	Arg	Lys	
		290				295					300					
Arg	Leu	Phe	Gln	Ile	Asp	Ala	Gu	Cys	Arg	Tyr	Ala	Ser	Pro	Met	Tyr	
305				310					315					320		
Val	Ala	Arg	Phe	Trp	Ile	Met	Lys	Ala	Leu	Thr	Val	Leu	Tyr	Met	Val	
			325						330					335		
Ala	Leu	Pro	Cys	Tyr	Met	Gln	Gly	Pro	Trp	His	Gly	Leu	Lys	Leu	Phe	
		340						345					350			
Ala	Ile	Ala	His	Phe	Thr	Cys	Gly	Gu	Val	Leu	Ala	Thr	Met	Phe	Ile	
		355					360					365				
Val	Asn	His	Ile	Ile	Gu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val	
	370					375					380					
Lys	Gly	Thr	Met	Ala	Pro	Pro	Lys	Thr	Met	His	Gly	Val	Thr	Pro	Met	
385				390					395					400		
Asn	Asn	Thr	Arg	Lys	Gu	Val	Gu	Ala	Gu	Ala	Ser	Lys	Ser	Gly	Ala	

PF58307. txt

Val Val Lys Ser Val Pro Leu Asp Asp Trp Ala Ala Val Gln Cys Gln  
 405 410 415  
 Thr Ser Val Asn Trp Ser Val Gly Ser Trp Phe Trp Asn His Phe Ser  
 420 425 430  
 Gly Gly Leu Asn His Gln Ile Gu His His Leu Phe Pro Gly Leu Ser  
 435 440 445  
 450 455 460  
 His Gu Thr Tyr Tyr His Ile Gln Asp Val Val Gln Ser Thr Cys Ala  
 465 470 475 480  
 Gu Tyr Gly Val Pro Tyr Gln His Gu Pro Ser Leu Trp Thr Ala Tyr  
 485 490 495  
 Trp Lys Met Leu Gu His Leu Arg Gln Leu Gly Asn Gu Gu Thr His  
 500 505 510  
 Gu Ser Trp Arg Ala Ala  
 515

<210> 79  
 <211> 1560  
 <212> DNA  
 <213> Thr aust ochyt r i u m sp

<220>  
 <221> CDS  
 <222> (1).. (1560)

<400> 79  
 at g acg gt c ggc t ac gac ggg gag at c ccg t t c gag cag gt c cgc gcg 48  
 Met Thr Val Gly Tyr Asp Gly Gu Ile Pro Phe Gu Gln Val Arg Ala  
 1 5 10 15  
 cac aac aag ccg gat gac gcc t gg t gc gcg at c cac ggg cac gt g t ac 96  
 His Asn Lys Pro Asp Asp Ala Trp Cys Ala Ile His Gly His Val Tyr  
 20 25 30  
 gat gt g acc aag t t c gcg agc gt g cac ccg ggc ggc gac at t at c ct g 144  
 Asp Val Thr Lys Phe Ala Ser Val His Pro Gly Gly Asp Ile Ile Leu  
 35 40 45  
 ct g gcc gca ggc aag gat gcc acc gt g ct g t ac gag act t ac cat gt g 192  
 Leu Ala Ala Gly Lys Asp Ala Thr Val Leu Tyr Gln Thr Tyr His Val  
 50 55 60  
 cgg ggc gt c t cg gac gcg gt g ct g cgc aag t ac cgc at c ggc aag ct g 240  
 Arg Gly Val Ser Asp Ala Val Leu Arg Lys Tyr Arg Ile Gly Lys Leu  
 65 70 75 80  
 ccg gac ggc caa ggc ggc gcg aac gag aag gaa aag ccg acg ct c t cg 288  
 Pro Asp Gly Gln Gly Gly Ala Asn Gu Lys Gu Lys Arg Thr Leu Ser  
 85 90 95  
 ggc ct c t cg t cg gcc t cg t ac t ac acg t gg aac agc gac t t t t ac agg 336  
 Gly Leu Ser Ala Ser Tyr Tyr Thr Trp Asn Ser Asp Phe Tyr Arg  
 100 105 110  
 gt a at g cgc gag cgc gt c gt g gct cgg ct c aag gag cgc ggc aag gcc 384  
 Val Met Arg Gu Arg Val Val Ala Arg Leu Lys Gu Arg Gly Lys Ala  
 115 120 125  
 cgc cgc gga ggc t ac gag ct c t gg at c aag gcg ct c ct g ct c gt c 432  
 Arg Arg Gly Gly Tyr Gu Leu Trp Ile Lys Ala Leu Leu Leu Leu Val  
 130 135 140  
 ggc t t c t gg agc t cg ct g t gc t gg at g t gc acg ct g gac ccc t cg t t c 480  
 Gly Phe Trp Ser Ser Leu Cys Trp Met Cys Thr Leu Asp Pro Ser Phe  
 145 150 155 160  
 ggg gcc at c ct g gcc gcc at g t cg ct g ggc gt c t t t gcc gcc t t t gt g 528  
 Gly Ala Ile Leu Ala Ala Met Ser Leu Gly Val Phe Ala Ala Phe Val  
 165 170 175  
 ggc acg t gc at c cag cac gac ggc aac cac ggc gcc t t t gcc cag t cg 576  
 Gly Thr Cys Ile Gln His Asp Gly Asn His Gly Ala Phe Ala Gln Ser  
 180 185 190  
 cga t gg gt c aac aag gt t gcc ggg t gg acg ct c gac at g at c ggc gcc 624  
 Arg Trp Val Asn Lys Val Ala Gly Trp Thr Leu Asp Met Ile Gly Ala  
 195 200 205  
 agc ggc at g acg t gg gag t t c cag cac gcc ct g ggc cac cat ccg t ac 672  
 Ser Gly Met Thr Trp Gu Phe Gln His Ala Leu Gly His His Pro Tyr  
 210 215 220

PF58307. txt

acg aac ct g at c gag gag gag aac ggc ct g caa aag gt g agc ggc aag 720  
 Thr Asn Leu Ile Glu Glu Glu Asn Gly Leu Gl n Lys Val Ser Gly Lys  
 225 230 235 240  
 aag at g gac acc aag ct g gcc gac cag gag agc gat ccg gac gt c ttt 768  
 Lys Met Asp Thr Lys Leu Ala Asp Gl n Gl u Ser Asp Pro Asp Val ttt  
 245 250 255  
 t cc acg t ac ccg at g at g cgc ct g cac ccg t gg cac cag aag cgc t gg 816  
 Ser Thr Tyr Pro Met Met Arg Leu His Pro Trp His Gl n Lys Arg Trp  
 260 265 270  
 t ac cac cgt ttc cag cac att t ac ggc ccc ttc at c ttt ggc ttc at g 864  
 Tyr His Arg Phe Gl n His Ile Tyr Gly Pro Phe Ile Phe Gly Phe Met  
 275 280 285  
 acc at c aac aag gt g gt c acg cag gac gt c ggt gt g gt g ttc cgc aag 912  
 Thr Ile Asn Lys Val Val Thr Gl n Asp Val Gly Val Phe Arg Lys  
 290 295 300  
 cgg ct c ttc cag att gac gcc gag tgc cgg t ac gcg agc cca at g t ac 960  
 Arg Leu Phe Gl n Ile Asp Ala Gl u Cys Arg Tyr Ala Ser Pro Met Tyr  
 305 310 315 320  
 gt g gcg cgt ttc tgg at c at g aag gcg ct c acg gt g ct c t ac at g gt g 1008  
 Val Ala Arg Phe Trp Ile Met Lys Ala Leu Thr Val Leu Tyr Met Val  
 325 330 335  
 gcc ct g ccg tgc t ac at g cag ggc ccg t gg cac ggc ct c aag ct g ttc 1056  
 Ala Leu Pro Cys Tyr Met Gl n Gly Pro Trp His Gly Leu Lys Leu Phe  
 340 345 350  
 gcg at c gcg cac ttt acg tgc ggc gag gt g ct c gca acc at g ttc att 1104  
 Ala Ile Ala His Phe Thr Cys Gly Gl u Val Leu Ala Thr Met Phe Ile  
 355 360 365  
 gt g aac cac gt c at c gag ggc gt c t cg t ac gct t cc aag gac gcg gt c 1152  
 Val Asn His Val Ile Gl u Gly Val Ser Tyr Ala Ser Lys Asp Ala Val  
 370 375 380  
 aag ggc acg at g gcg ccg ccg aag acg at g cac ggc gt g acg ccc at g 1200  
 Lys Gly Thr Met Ala Pro Pro Lys Thr Met His Gly Val Thr Pro Met  
 385 390 395 400  
 aac aac acg cgc aag gag gt g gag gcg gag gcg t cc aag t ct ggc gcc 1248  
 Asn Asn Thr Arg Lys Gl u Val Gl u Ala Gl u Ala Ser Lys Ser Gly Ala  
 405 410 415  
 gt g gt c aag t ca gt c ccg ct c gac gac t gg gcc gcc gt c cag t gc cag 1296  
 Val Val Lys Ser Val Pro Leu Asp Asp Trp Ala Ala Val Gl n Cys Gl n  
 420 425 430  
 acc t cg gcg aac t gg agc gt c ggc t cg t gg ttc t gg aac cac ttt t cc 1344  
 Thr Ser Ala Asn Trp Ser Val Gly Ser Trp Phe Trp Asn His Phe Ser  
 435 440 445  
 ggc ggt ct c aac cac cag att gag cac cac ct g ttc ccc ggg ct c agc 1392  
 Gly Gly Leu Asn His Gl n Ile Gl u His His Leu Phe Pro Gly Leu Ser  
 450 455 460  
 cac gag acg t ac t ac cac at c cag gac gt c gt t cag t cc acc t gc gcc 1440  
 His Gl u Thr Tyr Tyr His Ile Gl n Asp Val Val Gl n Ser Thr Cys Ala  
 465 470 475 480  
 gag t ac ggc gt c ccg t ac cag cac gag cct t cg ct c t gg acc gcg t ac 1488  
 Gl u Tyr Gly Val Pro Tyr Gl n His Gl u Pro Ser Leu Trp Thr Ala Tyr  
 485 490 495  
 t gg aag at g ct c gag cac ct c cgt cgg ct c ggc aat gag gag acc cac 1536  
 Trp Lys Met Leu Gl u His Leu Arg Arg Leu Gly Asn Gl u Gl u Thr His  
 500 505 510  
 gag t cc t gg cag cgc gct gcc t ga 1560  
 Gl u Ser Trp Gl n Arg Ala Ala

<210> 80  
 <211> 519  
 <212> PRT  
 <213> Thr aust ochyt ri um sp

<400> 80  
 Met Thr Val Gly Tyr Asp Gly Gl u Ile Pro Phe Gl u Gl n Val Arg Ala  
 1 5 10 15  
 His Asn Lys Pro Asp Asp Ala Trp Cys Ala Ile His Gly His Val Tyr  
 20 25 30

PF58307.txt

Asp Val Thr Lys Phe Ala Ser Val His Pro Gly Gly Asp Ile Ile Leu  
 35 40 45  
 Leu Ala Ala Gly Lys Asp Ala Thr Val Leu Tyr Gu Thr Tyr His Val  
 50 55 60  
 Arg Gly Val Ser Asp Ala Val Leu Arg Lys Tyr Arg Ile Gly Lys Leu  
 65 70 75 80  
 Pro Asp Gly Gn Gly Gly Ala Asn Gu Lys Gu Lys Arg Thr Leu Ser  
 85 90 95  
 Gly Leu Ser Ser Ala Ser Tyr Tyr Thr Trp Asn Ser Asp Phe Tyr Arg  
 100 105 110  
 Val Met Arg Gu Arg Val Val Ala Arg Leu Lys Gu Arg Gly Lys Ala  
 115 120 125  
 Arg Arg Gly Gly Tyr Gu Leu Trp Ile Lys Ala Leu Leu Leu Val  
 130 135 140  
 Gly Phe Trp Ser Ser Leu Cys Trp Met Cys Thr Leu Asp Pro Ser Phe  
 145 150 155 160  
 Gly Ala Ile Leu Ala Ala Met Ser Leu Gly Val Phe Ala Ala Phe Val  
 165 170 175  
 Gly Thr Cys Ile Gn His Asp Gly Asn His Gly Ala Phe Ala Gn Ser  
 180 185 190  
 Arg Trp Val Asn Lys Val Ala Gly Trp Thr Leu Asp Met Ile Gly Ala  
 195 200 205  
 Ser Gly Met Thr Trp Gu Phe Gn His Ala Leu Gly His His Pro Tyr  
 210 215 220  
 Thr Asn Leu Ile Gu Gu Gu Asn Gly Leu Gn Lys Val Ser Gly Lys  
 225 230 235 240  
 Lys Met Asp Thr Lys Leu Ala Asp Gn Gu Ser Asp Pro Asp Val Phe  
 245 250 255  
 Ser Thr Tyr Pro Met Met Arg Leu His Pro Trp His Gn Lys Arg Trp  
 260 265 270  
 Tyr His Arg Phe Gn His Ile Tyr Gly Pro Phe Ile Phe Gly Phe Met  
 275 280 285  
 Thr Ile Asn Lys Val Val Thr Gn Asp Val Gly Val Val Phe Arg Lys  
 290 295 300  
 Arg Leu Phe Gn Ile Asp Ala Gu Cys Arg Tyr Ala Ser Pro Met Tyr  
 305 310 315 320  
 Val Ala Arg Phe Trp Ile Met Lys Ala Leu Thr Val Leu Tyr Met Val  
 325 330 335  
 Ala Leu Pro Cys Tyr Met Gn Gly Pro Trp His Gly Leu Lys Leu Phe  
 340 345 350  
 Ala Ile Ala His Phe Thr Cys Gly Gu Val Leu Ala Thr Met Phe Ile  
 355 360 365  
 Val Asn His Val Ile Gu Gly Val Ser Tyr Ala Ser Lys Asp Ala Val  
 370 375 380  
 Lys Gly Thr Met Ala Pro Lys Thr Met His Gly Val Thr Pro Met  
 385 390 395 400  
 Asn Asn Thr Arg Lys Gu Val Gu Ala Gu Ala Ser Lys Ser Gly Ala  
 405 410 415  
 Val Val Lys Ser Val Pro Leu Asp Asp Trp Ala Ala Val Gn Cys Gn  
 420 425 430  
 Thr Ser Ala Asn Trp Ser Val Gly Ser Trp Phe Trp Asn His Phe Ser  
 435 440 445  
 Gly Gly Leu Asn His Gn Ile Gu His His Leu Phe Pro Gly Leu Ser  
 450 455 460  
 His Gu Thr Tyr Tyr His Ile Gn Asp Val Val Gn Ser Thr Cys Ala  
 465 470 475 480  
 Gu Tyr Gly Val Pro Tyr Gn His Gu Pro Ser Leu Trp Thr Ala Tyr  
 485 490 495  
 Trp Lys Met Leu Gu His Leu Arg Arg Leu Gly Asn Gu Gu Thr His  
 500 505 510  
 Gu Ser Trp Gn Arg Ala Ala  
 515

<210> 81

<211> 1653

<212> DNA

<213> Thal assi osi r a pseudonana



<220>  
 <221> CDS  
 <222> (1).. (1653)

```

<400> 81
at g ggc aac ggc aac ct c cca gca t cc acc gca cag ct c aag t cc acc      48
Met Gly Asn Gly Asn Leu Pro Ala Ser Thr Ala Gln Leu Lys Ser Thr
1
t cg aag ccc cag cag caa cat gag cat cgc acc at c t cc aag t cc gag      96
Ser Lys Pro Gln Gln Gln His Gu His Arg Thr Ile Ser Lys Ser Gu
20
ct c gcc caa cac aac acg ccc aaa t ca gca t gg t gt gcc gt c cac t cc      144
Leu Ala Gln His Asn Thr Pro Lys Ser Ala Trp Cys Ala Val His Ser
35
act ccc gcc acc gac cca t cc cac t cc aac aac aaa caa cac gca cac      192
Thr Pro Ala Thr Asp Pro Ser His Ser Asn Asn Lys Gln His Ala His
50
ct a gt c ct c gac att acc gac ttt gcg t cc cgc cat cca ggg gga gac      240
Leu Val Leu Asp Ile Thr Asp Phe Ala Ser Arg His Pro Gly Gly Asp
65
ct c at c ct c ct c gct t cc ggc aaa gac gcc t cg gt g ct g ttt gaa aca      288
Leu Ile Leu Leu Ala Ser Gly Lys Asp Ala Ser Val Leu Phe Gu Thr
85
t ac cat cca cgt gga gtt ccg acg t ct ct c att caa aag ct g cag att      336
Tyr His Pro Arg Gly Val Pro Thr Ser Leu Ile Gln Lys Leu Gln Ile
100
gga gt g at g gag gag gag gcg ttt cgg gat t cg ttt t ac agt t gg act      384
Gly Val Met Gu Gu Gu Ala Phe Arg Asp Ser Phe Tyr Ser Trp Thr
115
gat t ct gac ttt t at act gt g tt g aag agg agg gt t gt g gag cgg tt g      432
Asp Ser Asp Phe Tyr Thr Val Leu Lys Arg Arg Val Val Gu Arg Leu
130
gag gag agg ggg tt g gac agg agg gga t cg aaa gag att t gg at c aag      480
Gu Gu Arg Gly Leu Asp Arg Arg Gly Ser Lys Gu Ile Trp Ile Lys
145
gct tt g ttc tt g tt g gtt gga ttt t gg t ac t gt tt g t ac aag at g t at      528
Ala Leu Phe Leu Val Gly Phe Trp Tyr Cys Leu Tyr Lys Met Tyr
165
act acg t cg gat att gat cag t ac ggt att gcc att gcc t at t ct att      576
Thr Thr Ser Asp Ile Asp Gln Tyr Gly Ile Ala Ile Ala Tyr Ser Ile
180
gga at g gga acc ttt gcg gca ttc at c ggc acg t gt att caa cac gat      624
Gly Met Gly Thr Phe Ala Ala Phe Ile Gly Thr Cys Ile Gln His Asp
195
gga aat cac ggt gca ttc gct cag aac aag tta ct c aac aag tt g gct      672
Gly Asn His Gly Ala Phe Ala Gln Asn Lys Leu Asn Lys Leu Ala
210
ggg t gg acg tt g gat at g att ggt gcg agt gcg ttt acg t gg gag ct t      720
Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Phe Thr Trp Gu Leu
225
cag cac at g ct g ggg cat cat cca t at acg aat gt g tt g gat ggg gt g      768
Gln His Met Leu Gly His His Pro Tyr Thr Asn Val Leu Asp Gly Val
245
gag gag gag agg aag gag agg ggg gag gat gt t gct tt g gaa gaa aag      816
Gu Gu Gu Arg Lys Gu Arg Gly Gu Asp Val Ala Leu Gu Gu Lys
260
gat cag gaa t ca gat cca gac gt a ttc t cc t cc ttc cct ct c at g aga      864
Asp Gln Glu Ser Asp Pro Asp Val Phe Ser Ser Phe Pro Leu Met Arg
275
at g cat ccc cac cat aca acc t ca t gg t at cat aaa t ac caa cac ct c      912
Met His Pro His His Thr Thr Ser Trp Tyr His Lys Tyr Gln His Leu
290
t ac gct cca ccc ct c ttt gca tt g at g aca ct t gcc aaa gt a ttc caa      960
Tyr Ala Pro Pro Leu Phe Ala Leu Met Thr Leu Ala Lys Val Phe Gln
305
cag gat ttt gaa gtt gcc aca t cc gga cga tta t at cat att gat gcc      1008
Gln Asp Phe Glu Val Ala Thr Ser Gly Arg Leu Tyr His Ile Asp Ala
325
    
```



PF58307. txt

aat gta cgt tat ggt t c g gta tgg aat gtc atg agg ttt tgg gct atg 1056  
 Asn Val Arg Tyr Gly Ser Val Trp Asn Val Met Arg Phe Trp Ala Met  
 340 345 350  
 aag gtc att acg atg gga tat atg gga tta cca atc tac ttt cat 1104  
 Lys Val Ile Thr Met Gly Tyr Met Gly Leu Pro Ile Tyr Phe His  
 355 360 365  
 gga gta ctg agg gga gtt gga ttg ttt gtt att ggg cat ttg gcg tgt 1152  
 Gly Val Leu Arg Gly Val Gly Leu Phe Val Ile Gly His Leu Ala Cys  
 370 375 380  
 gga gag ttg ttg gcg acg atg ttt att gtg aat cac gtc att gag ggt 1200  
 Gly Gu Leu Leu Ala Thr Met Phe Ile Val Asn His Val Ile Gu Gly  
 385 390 395 400  
 gtg agt tat gga acg aag gat ttg gtt ggt ggt gcg agt cat gga gat 1248  
 Val Ser Tyr Gly Thr Lys Asp Leu Val Gly Ala Ser His Gly Asp  
 405 410 415  
 gag aag aag att gtc aag cca acg act gta ttg gga gat aca cca atg 1296  
 Gu Lys Lys Ile Val Lys Pro Thr Thr Val Leu Gly Asp Thr Pro Met  
 420 425 430  
 gaa aag act cgc gag gag gca ttg aaa agc aac agc aat aac aag 1344  
 Gu Lys Thr Arg Gu Gu Ala Leu Lys Ser Asn Ser Asn Asn Asn Lys  
 435 440 445  
 aag aag gga gag aag aac t c g gta cca t c c gtt cca t t c aac gac t g g 1392  
 Lys Lys Gly Gu Lys Asn Ser Val Pro Ser Val Pro Phe Asn Asp Trp  
 450 455 460  
 gca gca gtc caa tgc cag acc t c c gtg aat tgg t c t cca ggc t c a t g g 1440  
 Ala Ala Val Gn Cys Gn Thr Ser Val Asn Trp Ser Pro Gly Ser Trp  
 465 470 475 480  
 t t c t g g a a t c a c t t t t c t g g g g a c t c t c t c a g a t t g a g c a t c a c 1488  
 Phe Trp Asn His Phe Ser Gly Gly Leu Ser His Gn Ile Gu His His  
 485 490 495  
 t t g t t c c c c a g c a t t t g t c a t a c a a c t a c t g t c a t a t c c a g g a t g t t 1536  
 Leu Phe Pro Ser Ile Cys His Thr Asn Tyr Cys His Ile Gn Asp Val  
 500 505 510  
 gtg gag agt acg tgt gct gag tac gga gtt ccg tat cag agt gag agt 1584  
 Val Gu Ser Thr Cys Ala Gu Tyr Gly Val Pro Tyr Gn Ser Gu Ser  
 515 520 525  
 aat ttg ttt gtt gct tat gga aag atg att agt cat ttg aag ttt ttg 1632  
 Asn Leu Phe Val Ala Tyr Gly Lys Met Ile Ser His Leu Lys Phe Leu  
 530 535 540  
 ggt aaa gcc aag tgt gag t ag 1653  
 Gly Lys Ala Lys Cys Gu 550

<210> 82  
 <211> 550  
 <212> PRT  
 <213> Thal assi osi r a pseudonana

<400> 82  
 Met Gly Asn Gly Asn Leu Pro Ala Ser Thr Ala Gn Leu Lys Ser Thr  
 1 5 10 15  
 Ser Lys Pro Gn Gn Gn His Gu His Arg Thr Ile Ser Lys Ser Gu  
 20 25 30  
 Leu Ala Gn His Asn Thr Pro Lys Ser Ala Trp Cys Ala Val His Ser  
 35 40 45  
 Thr Pro Ala Thr Asp Pro Ser His Ser Asn Asn Lys Gn His Ala His  
 50 55 60  
 Leu Val Leu Asp Ile Thr Asp Phe Ala Ser Arg His Pro Gly Gly Asp  
 65 70 75 80  
 Leu Ile Leu Leu Ala Ser Gly Lys Asp Ala Ser Val Leu Phe Gu Thr  
 85 90 95  
 Tyr His Pro Arg Gly Val Pro Thr Ser Leu Ile Gn Lys Leu Gn Ile  
 100 105 110  
 Gly Val Met Gu Gu Gu Ala Phe Arg Asp Ser Phe Tyr Ser Trp Thr  
 115 120 125  
 Asp Ser Asp Phe Tyr Thr Val Leu Lys Arg Arg Val Val Gu Arg Leu  
 130 135 140  
 Gu Gu Arg Gly Leu Asp Arg Arg Gly Ser Lys Gu Ile Trp Ile Lys  
 Seite 107

PF58307.txt

145 Ala Leu Phe Leu Leu Val Gly Phe Trp Tyr Cys Leu Tyr Lys Met Tyr 160  
 165 Thr Thr Ser Asp Ile Asp Gln Tyr Gly Ile Ala Ile Ala Tyr Ser Ile  
 180 Gly Met Gly Thr Phe Ala Ala Phe Ile Gly Thr Cys Ile Gln His Asp  
 195 Gly Asn His Gly Ala Phe Ala Gln Asn Lys Leu Leu Asn Lys Leu Ala  
 210 Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Phe Thr Trp Gu Leu  
 225 Gln His Met Leu Gly His His Pro Tyr Thr Asn Val Leu Asp Gly Val  
 245 Gu Gu Gu Arg Lys Gu Arg Gly Gu Asp Val Ala Leu Gu Gu Lys  
 260 Asp Gln Gu Ser Asp Pro Asp Val Phe Ser Ser Phe Pro Leu Met Arg  
 275 Met His Pro His His Thr Thr Ser Trp Tyr His Lys Tyr Gln His Leu  
 290 Tyr Ala Pro Pro Leu Phe Ala Leu Met Thr Leu Ala Lys Val Phe Gln  
 305 Gln Asp Phe Gu Val Ala Thr Ser Gly Arg Leu Tyr His Ile Asp Ala  
 325 Asn Val Arg Tyr Gly Ser Val Trp Asn Val Met Arg Phe Trp Ala Met  
 340 Lys Val Ile Thr Met Gly Tyr Met Met Gly Leu Pro Ile Tyr Phe His  
 355 Gly Val Leu Arg Gly Val Gly Leu Phe Val Ile Gly His Leu Ala Cys  
 370 Gly Gu Leu Leu Ala Thr Met Phe Ile Val Asn His Val Ile Gu Gly  
 385 Val Ser Tyr Gly Thr Lys Asp Leu Val Gly Ala Ser His Gly Asp  
 405 Gu Lys Lys Ile Val Lys Pro Thr Thr Val Leu Gly Asp Thr Pro Met  
 420 Gu Lys Thr Arg Gu Gu Ala Leu Lys Ser Asn Ser Asn Asn Lys  
 435 Lys Lys Gly Gu Lys Asn Ser Val Pro Ser Val Pro Phe Asn Asp Trp  
 450 Ala Ala Val Gln Cys Gln Thr Ser Val Asn Trp Ser Pro Gly Ser Trp  
 465 Phe Trp Asn His Phe Ser Gly Gly Leu Ser His Gln Ile Gu His His  
 485 Leu Phe Pro Ser Ile Cys His Thr Asn Tyr Cys His Ile Gln Asp Val  
 500 Val Gu Ser Thr Cys Ala Gu Tyr Gly Lys Met Ile Ser His Leu Lys Phe Leu  
 515 Asn Leu Phe Val Ala Tyr Gly Lys Met Ile Ser His Leu Lys Phe Leu  
 530 Gly Lys Ala Lys Cys Gu 550  
 545

<210> 83  
 <211> 1626  
 <212> DNA  
 <213> Euglena gracilis

<220>  
 <221> CDS  
 <222> (1)..(1626)

<400> 83  
 atg ttg gtg ctg ttt ggc aat ttc tat gtc aag caa tac tcc caa aag 48  
 Met Leu Val Leu Phe Gly Asn Phe Tyr Val Lys Gln Tyr Ser Gln Lys  
 1 aac ggc aag ccg gag aac gga gcc acc cct gag aac gga gcg aag ccg 96  
 Asn Gly Lys Pro Gu Asn Gly Ala Thr Pro Gu Asn Gly Ala Lys Pro  
 20 25 30

PF58307. txt

caa	cct	tgc	gag	aac	ggc	acg	gtg	gaa	aag	cga	gag	aat	gac	acc	gcc	144
Gn	Pro	Oys	Glu	Asn	Gly	Thr	Val	Glu	Lys	Arg	Glu	Asn	Asp	Thr	Ala	
		35				40						45				
aac	gtt	cgg	ccc	acc	cgt	cca	gct	gga	ccc	ccg	ccg	gcc	acg	tac	tac	192
Asn	Val	Arg	Pro	Thr	Arg	Pro	Ala	Gly	Pro	Pro	Pro	Ala	Thr	Tyr	Tyr	
	50					55					60					
gac	tcc	ctg	gca	gtg	tcg	ggg	cag	ggc	aag	gag	cgg	ctg	ttc	acc	acc	240
Asp	Ser	Leu	Ala	Val	Ser	Gly	Gln	Gly	Lys	Glu	Arg	Leu	Phe	Thr	Thr	
65					70					75				80		
gat	gag	gtg	agg	cgg	cac	atc	ctc	ccc	acc	gat	ggc	tgg	ctg	acg	tgc	288
Asp	Glu	Val	Arg	Arg	His	Ile	Leu	Pro	Thr	Asp	Gly	Trp	Leu	Thr	Cys	
			85					90					95			
cac	gaa	gga	gtc	tac	gat	gtc	act	gat	ttc	ctt	gcc	aag	cac	cct	ggt	336
His	Glu	Gly	Val	Tyr	Asp	Val	Thr	Asp	Phe	Leu	Ala	Lys	His	Pro	Gly	
			100					105					110			
ggc	ggt	gtc	atc	acg	ctg	ggc	ctt	gga	agg	gac	tgc	aca	atc	ctc	atc	384
Gly	Gly	Val	Ile	Thr	Leu	Gly	Leu	Gly	Arg	Asp	Cys	Thr	Ile	Leu	Ile	
		115				120					125					
gag	tca	tac	cac	cct	gct	ggg	cgc	ccg	gac	aag	gtg	atg	gag	aag	tac	432
Glu	Ser	Tyr	His	Pro	Ala	Gly	Arg	Pro	Asp	Lys	Val	Met	Glu	Lys	Tyr	
	130					135					140					
cgc	att	ggt	acg	ctg	cag	gac	ccc	aag	acg	ttc	tat	gct	tgg	gga	gag	480
Arg	Ile	Gly	Thr	Leu	Gln	Asp	Pro	Lys	Thr	Phe	Tyr	Ala	Trp	Gly	Glu	
145					150					155				160		
tcc	gat	ttc	tac	cct	gag	tig	aag	cgc	cgg	gcc	ctt	gca	agg	ctg	aag	528
Ser	Asp	Phe	Tyr	Pro	Glu	Leu	Lys	Arg	Arg	Ala	Leu	Ala	Arg	Leu	Lys	
				165						170				175		
gag	gct	ggt	cag	gcg	cgg	cgc	ggc	ggc	ctt	ggg	gtg	aag	gcc	ctc	ctg	576
Glu	Ala	Gly	Gln	Ala	Arg	Arg	Gly	Gly	Leu	Gly	Val	Lys	Ala	Leu	Leu	
			180					185					190			
gtg	ctc	acc	ctc	ttc	ttc	gtg	tcg	tgg	tac	atg	tgg	gtg	gcc	cac	aag	624
Val	Leu	Thr	Leu	Phe	Phe	Val	Ser	Trp	Tyr	Met	Trp	Trp	Val	Ala	His	Lys
		195					200						205			
tcc	ttc	ctc	tgg	gcc	gcc	gtc	tgg	ggc	ttc	gcc	ggc	tcc	cac	gtc	ggg	672
Ser	Phe	Leu	Trp	Ala	Ala	Val	Trp	Gly	Phe	Ala	Gly	Ser	His	Val	Gly	
		210				215					220					
ctg	agc	atc	cag	cac	gat	ggc	aac	cac	ggc	gcg	ttc	agc	cgc	aac	aca	720
Leu	Ser	Ile	Gln	His	Asp	Gly	Asn	His	Gly	Ala	Phe	Ser	Arg	Asn	Thr	
225					230					235				240		
ctg	gtg	aac	cgc	ctg	gcg	ggg	tgg	ggc	atg	gac	tig	atc	ggc	gcg	tcg	768
Leu	Val	Asn	Arg	Leu	Ala	Gly	Trp	Gly	Met	Asp	Leu	Ile	Gly	Ala	Ser	
			245					250						255		
tcc	acg	gtg	tgg	gag	tac	cag	cac	gtc	atc	ggc	cac	cac	cag	tac	acc	816
Ser	Thr	Val	Trp	Glu	Tyr	Gln	His	Val	Ile	Gly	His	His	Gln	Tyr	Thr	
			260					265					270			
aac	ctc	gtg	tcg	gac	acg	cta	ttc	agt	ctg	cct	gag	aac	gat	ccg	gac	864
Asn	Leu	Val	Ser	Asp	Thr	Leu	Phe	Ser	Leu	Pro	Glu	Asn	Asp	Pro	Asp	
			275					280					285			
gtc	ttc	tcc	agc	tac	ccg	ctg	atg	cgc	atg	cac	ccg	gat	acg	gcg	tgg	912
Val	Phe	Ser	Ser	Tyr	Pro	Leu	Met	Arg	Met	His	Pro	Asp	Thr	Ala	Trp	
		290				295						300				
cag	ccg	cac	cac	cgc	ttc	cag	cac	ctg	ttc	gcg	ttc	cca	ctg	ttc	gcc	960
Gln	Pro	His	His	Arg	Phe	Gln	His	Leu	Phe	Ala	Phe	Pro	Leu	Phe	Ala	
					310					315					320	
ctg	atg	aca	atc	agc	aag	gtg	ctg	acc	agc	gat	ttc	gct	gtc	tgc	ctc	1008
Leu	Met	Thr	Ile	Ser	Lys	Val	Leu	Thr	Ser	Asp	Phe	Ala	Val	Cys	Leu	
				325						330				335		
agc	atg	aag	aag	ggg	tcc	atc	gac	tgc	tcc	tcc	agg	ctc	gtc	cca	ctg	1056
Ser	Met	Lys	Lys	Gly	Ser	Ile	Asp	Cys	Ser	Ser	Arg	Leu	Val	Pro	Leu	
				340				345					350			
gag	ggg	cag	ctg	ctg	ttc	tgg	ggg	gcc	aag	ctg	gcg	aac	ttc	ctg	tig	1104
Glu	Gly	Gln	Leu	Leu	Phe	Trp	Gly	Ala	Lys	Leu	Ala	Asn	Phe	Leu	Leu	
			355					360					365			
cag	att	gtg	tig	cca	tgc	tac	ctc	cac	ggg	aca	gct	atg	ggc	ctg	gcc	1152
Gln	Ile	Val	Leu	Pro	Cys	Tyr	Leu	His	Gly	Thr	Ala	Met	Gly	Leu	Ala	
			370					375				380				
ctc	ttc	tct	gtt	gct	cac	ctt	gtg	tcg	ggg	gag	tac	ctc	gcg	atc	tgc	1200
Leu	Phe	Ser	Val	Ala	His	Leu	Val	Ser	Gly	Glu	Tyr	Leu	Ala	Ile	Cys	

PF58307. txt

```

385          390          395          400
t t c at c at c aac cac at c agc gag t c t t g t gag t t t at g aat aca agc      1248
Phe Ile Ile Asn His Ile Ser Gu Ser Cys Gu Phe Met Asn Thr Ser
405
t t t caa acc gcc gcc cgg agg aca gag at g c t t cag gca gca cat cag      1296
Phe Gl n Thr Ala Ala Arg Arg Thr Gu Met Leu Gl n Ala Ala His Gl n
420          425          430
gca gcg gag gcc aag aag gt g aag ccc acc cct cca ccg aac gat t gg      1344
Al a Al a Gl u Ala Lys Lys Val Lys Pro Thr Pro Pro Ala Asn Asp Tr p
435          440          445
gct gt g aca cag gt c caa t gc t gc gt g aat t gg aga t ca ggt ggc gt g      1392
Al a Val Thr Gl n Val Gl n Cys Cys Val Asn Tr p Arg Ser Gl y Gl y Val
450          455          460
t t g gcc aat cac ct c t ct t g g ggc t t g aac cac cag at c gag cat cat      1440
Leu Ala Asn His Leu Ser Gl y Gl y Leu Asn His Gl n Ile Gl u His His
465          470          475
ct g t t c ccc agc at c t cg cat gcc aac t ac ccc acc at c gcc cct gt t      1488
Leu Phe Pro Ser Ile Ser His Ala Asn Tyr Pro Thr Ile Ala Pro Val
485          490          495
gt g aag gag gt g t gc gag gag t ac ggg t t g ccg t ac aag aat t ac gt c      1536
Val Lys Gl u Val Cys Gl u Gl u Tyr Gl y Leu Pro Tyr Lys Asn Tyr Val
500          505          510
acg t t c t gg gat gca gt c t gt ggc at g gt t cag cac ct c cgg t t g at g      1584
Thr Phe Tr p Asp Ala Val Cys Gl y Met Val Gl n His Leu Arg Leu Met
515          520          525
ggt gct cca ccg gt g cca acg aac ggg gac aaa aag t ca t aa      1626
Gl y Ala Pro Pro Val Pro Thr Asn Gl y Asp Lys Lys Ser
530          535          540

```

<210> 84  
 <211> 541  
 <212> PRT  
 <213> Euglena gracilis

```

<400> 84
Met Leu Val Leu Phe Gl y Asn Phe Tyr Val Lys Gl n Tyr Ser Gl n Lys
1      5      10      15
Asn Gl y Lys Pro Gl u Asn Gl y Ala Thr Pro Gl u Asn Gl y Ala Lys Pro
20      25      30
Gl n Pro Cys Gl u Asn Gl y Thr Val Gl u Lys Arg Gl u Asn Asp Thr Ala
35      40      45
Asn Val Arg Pro Thr Arg Pro Ala Gl y Pro Pro Pro Ala Thr Tyr Tyr
50      55      60
Asp Ser Leu Ala Val Ser Gl y Gl n Gl y Lys Gl u Arg Leu Phe Thr Thr
65      70      75      80
Asp Gl u Val Arg Arg His Ile Leu Pro Thr Asp Gl y Tr p Leu Thr Cys
85      90      95
His Gl u Gl y Val Tyr Asp Val Thr Asp Phe Leu Ala Lys His Pro Gl y
100      105      110
Gl y Gl y Val Ile Thr Leu Gl y Leu Gl y Arg Asp Cys Thr Ile Leu Ile
115      120      125
Gl u Ser Tyr His Pro Ala Gl y Arg Pro Asp Lys Val Met Gl u Lys Tyr
130      135      140
Arg Ile Gl y Thr Leu Gl n Asp Pro Lys Thr Phe Tyr Ala Tr p Gl y Gl u
145      150      155      160
Ser Asp Phe Tyr Pro Gl u Leu Lys Arg Arg Ala Leu Ala Arg Leu Lys
165      170      175
Gl u Ala Gl y Gl n Ala Arg Arg Gl y Gl y Leu Gl y Val Lys Ala Leu Leu
180      185      190
Val Leu Thr Leu Phe Phe Val Ser Tr p Tyr Met Tr p Val Ala His Lys
195      200      205
Ser Phe Leu Tr p Ala Ala Val Tr p Gl y Phe Ala Gl y Ser His Val Gl y
210      215      220
Leu Ser Ile Gl n His Asp Gl y Asn His Gl y Ala Phe Ser Arg Asn Thr
225      230      235      240
Leu Val Asn Arg Leu Ala Gl y Tr p Gl y Met Asp Leu Ile Gl y Ala Ser
245      250      255
Ser Thr Val Tr p Gl u Tyr Gl n His Val Ile Gl y His His Gl n Tyr Thr
Sei te 110

```

PF58307. txt

260  
 Asn Leu Val Ser Asp Thr Leu Phe Ser Leu Pro Glu Asn Asp Pro Asp  
 275  
 Val Phe Ser Ser Tyr Pro Leu Met Arg Met His Pro Asp Thr Ala Trp  
 290  
 Gln Pro His His Arg Phe Gln His Leu Phe Ala Phe Pro Leu Phe Ala  
 305  
 Leu Met Thr Ile Ser Lys Val Leu Thr Ser Asp Phe Ala Val Cys Leu  
 325  
 Ser Met Lys Lys Gly Ser Ile Asp Cys Ser Ser Arg Leu Val Pro Leu  
 340  
 Glu Gly Gln Leu Leu Phe Trp Gly Ala Lys Leu Ala Asn Phe Leu Leu  
 355  
 Gln Ile Val Leu Pro Cys Tyr Leu His Gly Thr Ala Met Gly Leu Ala  
 370  
 Leu Phe Ser Val Ala His Leu Val Ser Gly Glu Tyr Leu Ala Ile Cys  
 385  
 Phe Ile Ile Asn His Ile Ser Glu Ser Cys Glu Phe Met Asn Thr Ser  
 405  
 Phe Gln Thr Ala Ala Arg Arg Thr Glu Met Leu Gln Ala Ala His Gln  
 420  
 Ala Ala Glu Ala Lys Lys Val Lys Pro Thr Pro Pro Pro Asn Asp Trp  
 435  
 Ala Val Thr Gln Val Gln Cys Cys Val Asn Trp Arg Ser Gly Gly Val  
 450  
 Leu Ala Asn His Leu Ser Gly Gly Leu Asn His Gln Ile Glu His His  
 465  
 Leu Phe Pro Ser Ile Ser His Ala Asn Tyr Pro Thr Ile Ala Pro Val  
 485  
 Val Lys Glu Val Cys Glu Glu Tyr Gly Leu Pro Tyr Lys Asn Tyr Val  
 500  
 Thr Phe Trp Asp Ala Val Cys Gly Met Val Gln His Leu Arg Leu Met  
 515  
 Gly Ala Pro Pro Val Pro Thr Asn Gly Asp Lys Lys Ser  
 530

<210> 85  
 <211> 1548  
 <212> DNA  
 <213> Thr aust ochyt r i um aur eum

<220>  
 <221> CDS  
 <222> (1) . . (1548)

<400> 85  
 at g acg gt c ggg t t t gac gaa acg gt g act at g gac acg gt c cgc aac 48  
 Met Thr Val Gly Phe Asp Glu Thr Val Thr Met Asp Thr Val Arg Asn  
 5  
 cac aac at g ccg gac gac gcc t gg t gc gcg at c cac ggc acc gt g t ac 96  
 His Asn Met Pro Asp Asp Ala Trp Cys Ala Ile His Gly Thr Val Tyr  
 20  
 gac at c acc aag t t c agc aag gt g cac ccc ggc ggg gac at c at c at g 144  
 Asp Ile Thr Lys Phe Ser Lys Val His Pro Gly Gly Asp Ile Ile Met  
 35  
 ct g gcc gct ggc aag gag gcc acc at c ct g t t c gag acc t ac cac at c 192  
 Leu Ala Ala Gly Lys Glu Ala Thr Ile Leu Phe Glu Thr Tyr His Ile  
 50  
 aag ggc gt c ccg gac gcg gt g ct g cgc aag t ac aag gt c ggc aag ct c 240  
 Lys Gly Val Pro Asp Ala Val Leu Arg Lys Tyr Lys Val Gly Lys Leu  
 65  
 ccc cag ggc aag aag ggc gaa acg agc cac at g ccc acc ggg ct c gac 288  
 Pro Gln Gly Lys Lys Gly Glu Thr Ser His Met Pro Thr Gly Leu Asp  
 85  
 t cg gcc t cc t ac t ac t cg t gg gac agc gag t t t t ac agg gt g ct c cgc 336  
 Ser Ala Ser Tyr Tyr Ser Trp Asp Ser Glu Phe Tyr Arg Val Leu Arg  
 100  
 gag cgc gt c gcc aag aag ct g gcc gag ccc ggc ct c at g cag cgc gcg 384  
 110

PF58307.txt

Gu	Arg	Val	Ala	Lys	Lys	Leu	Ala	Gu	Pro	Gly	Leu	Met	Gln	Arg	Ala		
cgc	atg	gag	ctc	tgg	gcc	aag	gcg	atc	ttc	ctc	ctg	gca	ggt	ttc	tgg		432
Arg	Met	Glu	Leu	Trp	Ala	Lys	Ala	ile	Phe	Leu	Leu	Ala	Gly	Phe	Trp		
ggc	tcc	ctt	tac	gcc	atg	tgc	gtg	cta	gac	ccg	cac	ggc	ggt	gcc	atg		480
Gly	Ser	Leu	Tyr	Ala	Met	Cys	Val	Leu	Asp	Pro	His	Gly	Gly	Ala	Met		
gt a	gcc	gcc	gtt	acg	ctc	ggc	gtg	ttc	gct	gcc	ttt	gtc	gga	act	tgc		528
Val	Ala	Ala	Val	Thr	Leu	Gly	Val	Phe	Ala	Ala	Phe	Val	Gly	Thr	Cys		
atc	cag	cac	gac	ggc	agc	cac	ggc	gcc	ttc	tcc	aag	tcg	cga	ttc	atg		576
ile	Gln	His	Asp	Gly	Ser	His	Gly	Ala	Phe	Ser	Lys	Ser	Arg	Phe	Met		
aac	aag	gcg	gcg	ggc	tgg	acc	ctc	gac	atg	atc	ggc	gcg	agt	gcg	atg		624
Asn	Lys	Ala	Ala	Gly	Trp	Thr	Leu	Asp	Met	ile	Gly	Ala	Ser	Ala	Met		
acc	tgg	gag	atg	cag	cac	gtt	ctt	ggc	cac	cac	ccg	tac	acc	aac	ctc		672
Thr	Trp	Glu	Met	Gln	His	Val	Leu	Gly	His	His	Pro	Tyr	Thr	Asn	Leu		
atc	gag	atg	gag	aac	ggt	ttg	gcc	aag	gtc	aag	ggc	gcc	gac	gtc	gac		720
ile	Glu	Met	Glu	Asn	Gly	Leu	Ala	Lys	Val	Lys	Gly	Ala	Asp	Val	Asp		
ccg	aag	aag	gtc	gac	cag	gag	agc	gac	ccg	gac	gtc	ttc	agt	acg	tac		768
Pro	Lys	Lys	Val	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe	Ser	Thr	Tyr		
ccg	atg	ctt	cgc	ctg	cac	ccg	tgg	cac	cgc	cag	cgg	ttt	tac	cac	aag		816
Pro	Met	Leu	Arg	Leu	His	Pro	Trp	His	Arg	Gln	Arg	Phe	Tyr	His	Lys		
ttc	cag	cac	ctg	tac	gcc	ccg	ttt	atc	ttt	ggg	tct	atg	acg	att	aac		864
Phe	Gln	His	Leu	Tyr	Ala	Pro	Phe	ile	Phe	Gly	Ser	Met	Thr	ile	Asn		
aag	gtg	att	tcc	cag	gat	gtc	ggg	gtt	gtg	ctg	cgc	aag	cgc	ctg	ttc		912
Lys	Val	ile	Ser	Gln	Asp	Val	Gly	Val	Val	Leu	Arg	Lys	Arg	Leu	Phe		
cag	atc	gac	gcc	aac	tgc	cgg	tat	ggc	agc	ccc	tgg	tac	gtg	gcc	cgc		960
Gln	ile	Asp	Ala	Asn	Cys	Arg	Tyr	Gly	Ser	Pro	Trp	Tyr	Val	Ala	Arg		
ttc	tgg	atc	atg	aag	ctc	ctc	acc	acg	ctc	tac	atg	gtg	gcg	ctt	ccc		1008
Phe	Trp	ile	Met	Lys	Leu	Leu	Thr	Thr	Leu	Tyr	Met	Val	Ala	Leu	Pro		
atg	tac	atg	cag	ggg	cct	gct	cag	ggc	ttg	aag	ctt	ttc	ttc	atg	gcc		1056
Met	Tyr	Met	Gln	Gly	Pro	Ala	Gln	Gly	Leu	Lys	Leu	Phe	Phe	Met	Ala		
cac	ttc	acc	tgc	gga	gag	gtc	ctc	gcc	acc	atg	ttt	att	gtc	aac	cac		1104
His	Phe	Thr	Cys	Gly	Glu	Val	Leu	Ala	Thr	Met	Phe	ile	Val	Asn	His		
atc	atc	gag	ggc	gtc	agc	tac	gct	tcc	aag	gac	gcg	gtc	aag	ggc	gtc		1152
ile	ile	Glu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val	Lys	Gly	Val		
atg	gct	ccg	ccg	cgc	act	gtg	cac	ggt	gtc	acc	ccg	atg	cag	gtg	acg		1200
Met	Ala	Pro	Pro	Arg	Thr	Val	His	Gly	Val	Thr	Pro	Met	Gln	Val	Thr		
caa	aag	gcg	ctc	agt	gcg	gcc	gag	tcg	gcc	aag	tcg	gac	gcc	gac	aag		1248
Gln	Lys	Ala	Leu	Ser	Ala	Ala	Glu	Ser	Ala	Lys	Ser	Asp	Ala	Asp	Lys		
acg	acc	atg	atc	ccc	ctc	aac	gac	tgg	gcc	gct	gtg	cag	tgc	cag	acc		1296
Thr	Thr	Met	ile	Pro	Leu	Asn	Asp	Trp	Ala	Ala	Val	Gln	Cys	Gln	Thr		
tct	gtg	aac	tgg	gct	gtc	ggg	tcg	ttt	tgg	aac	cac	ttt	tcg	ggc			1344
Ser	Val	Asn	Trp	Ala	Val	Gly	Ser	Trp	Phe	Trp	Asn	His	Phe	Ser	Gly		
ggc	ctc	aac	cac	cag	att	gag	cac	cac	tgc	ttc	ccc	caa	aac	ccc	cac		1392
Gly	Leu	Asn	His	Gln	ile	Glu	His	His	Cys	Phe	Pro	Gln	Asn	Pro	His		
acg	gtc	aac	gtc	tac	atc	tcg	ggc	atc	gtc	aag	gag	acc	tgc	gaa	gaa		1440
Thr	Val	Asn	Val	Tyr	ile	Ser	Gly	ile	Val	Lys	Glu	Thr	Cys	Glu	Glu		

PF58307. txt

t ac	ggc	gt g	ccg	t ac	cag	gct	gag	at c	agc	ct c	t t c	t ct	gcc	t at	t t c	1488
Tyr	G y	Val	Pro	Tyr	G n	Al a	G u	I le	Ser	Leu	Phe	Ser	Al a	Tyr	Phe	
				485					490					495		
aag	at g	ct g	t cg	cac	ct c	cgc	acg	ct c	ggc	aac	gag	gac	ct c	acg	gcc	1536
Lys	Mæt	Leu	Ser	Hi s	Leu	Arg	Thr	Leu	G y	Asn	G u	Asp	Leu	Thr	Al a	
			500					505					510			
t gg	t cc	acg	t ga													1548
Tr p	Ser	Thr														
		515														

<210> 86  
 <211> 515  
 <212> PRT  
 <213> Thr aust ochyt r i um aur eum

<400> 86

Mæt	Thr	Val	G y	Phe	Asp	G u	Thr	Val	Thr	Mæt	Asp	Thr	Val	Arg	Asn	
1				5					10					15		
Hi s	Asn	Mæt	Pro	Asp	Asp	Al a	Tr p	Cys	Al a	I le	Hi s	G y	Thr	Val	Tyr	
			20					25					30			
Asp	I le	Thr	Lys	Phe	Ser	Lys	Val	Hi s	Pro	G y	G y	Asp	I le	I le	Mæt	
		35					40					45				
Leu	Al a	Al a	G y	Lys	G u	Al a	Thr	I le	Leu	Phe	G u	Thr	Tyr	Hi s	I le	
	50					55					60					
Lys	G y	Val	Pro	Asp	Al a	Val	Leu	Arg	Lys	Tyr	Lys	Val	G y	Lys	Leu	
65					70					75				80		
Pro	G n	G y	Lys	Lys	G y	G u	Thr	Ser	Hi s	Mæt	Pro	Thr	G y	Leu	Asp	
				85					90					95		
Ser	Al a	Ser	Tyr	Tyr	Ser	Tr p	Asp	Ser	G u	Phe	Tyr	Arg	Val	Leu	Arg	
			100					105					110			
G u	Arg	Val	Al a	Lys	Lys	Leu	Al a	G u	Pro	G y	Leu	Mæt	G n	Arg	Al a	
		115					120					125				
Arg	Mæt	G u	Leu	Tr p	Al a	Lys	Al a	I le	Phe	Leu	Leu	Al a	G y	Phe	Tr p	
	130				135						140					
G y	Ser	Leu	Tyr	Al a	Mæt	Cys	Val	Leu	Asp	Pro	Hi s	G y	G y	Al a	Mæt	
145					150				155					160		
Val	Al a	Al a	Val	Thr	Leu	G y	Val	Phe	Al a	Al a	Phe	Val	G y	Thr	Cys	
				165				170						175		
I le	G n	Hi s	Asp	G y	Ser	Hi s	G y	Al a	Phe	Ser	Lys	Ser	Arg	Phe	Mæt	
			180					185					190			
Asn	Lys	Al a	Al a	G y	Tr p	Thr	Leu	Asp	Mæt	I le	G y	Al a	Ser	Al a	Mæt	
		195					200					205				
Thr	Tr p	G u	Mæt	G n	Hi s	Val	Leu	G y	Hi s	Hi s	Pro	Tyr	Thr	Asn	Leu	
	210					215					220					
I le	G u	Mæt	G u	Asn	G y	Leu	Al a	Lys	Val	Lys	G y	Al a	Asp	Val	Asp	
225					230						235				240	
Pro	Lys	Lys	Val	Asp	G n	G u	Ser	Asp	Pro	Asp	Val	Phe	Ser	Thr	Tyr	
				245					250					255		
Pro	Mæt	Leu	Arg	Leu	Hi s	Pro	Tr p	Hi s	Arg	G n	Arg	Phe	Tyr	Hi s	Lys	
			260					265					270			
Phe	G n	Hi s	Leu	Tyr	Al a	Pro	Phe	I le	Phe	G y	Ser	Mæt	Thr	I le	Asn	
		275					280					285				
Lys	Val	I le	Ser	G n	Asp	Val	G y	Val	Val	Leu	Arg	Lys	Arg	Leu	Phe	
	290					295					300					
G n	I le	Asp	Al a	Asn	Cys	Arg	Tyr	G y	Ser	Pro	Tr p	Tyr	Val	Al a	Arg	
305					310					315					320	
Phe	Tr p	I le	Mæt	Lys	Leu	Leu	Thr	Thr	Leu	Tyr	Mæt	Val	Al a	Leu	Pro	
				325					330					335		
Mæt	Tyr	Mæt	G n	G y	Pro	Al a	G n	G y	Leu	Lys	Leu	Phe	Phe	Mæt	Al a	
			340					345					350			
Hi s	Phe	Thr	Cys	G y	G u	Val	Leu	Al a	Thr	Mæt	Phe	I le	Val	Asn	Hi s	
		355					360					365				
I le	I le	G u	G y	Val	Ser	Tyr	Al a	Ser	Lys	Asp	Al a	Val	Lys	G y	Val	
	370					375					380					
Mæt	Al a	Pro	Pro	Arg	Thr	Val	Hi s	G y	Val	Thr	Pro	Mæt	G n	Val	Thr	
385					390					395					400	
G n	Lys	Al a	Leu	Ser	Al a	Al a	G u	Ser	Al a	Lys	Ser	Asp	Al a	Asp	Lys	
				405					410					415		



PF58307. txt

Thr Thr Met Ile Pro Leu Asn Asp Trp Ala Ala Val Gln Cys Gln Thr  
 420 425 430  
 Ser Val Asn Trp Ala Val Gly Ser Trp Phe Trp Asn His Phe Ser Gly  
 435 440 445  
 Gly Leu Asn His Gln Ile Gu His His Cys Phe Pro Gln Asn Pro His  
 450 455 460  
 Thr Val Asn Val Tyr Ile Ser Gly Ile Val Lys Gu Thr Cys Gu Gu  
 465 470 475 480  
 Tyr Gly Val Pro Tyr Gln Ala Gu Ile Ser Leu Phe Ser Ala Tyr Phe  
 485 490 495  
 Lys Met Leu Ser His Leu Arg Thr Leu Gly Asn Gu Asp Leu Thr Ala  
 500 505 510  
 Trp Ser Thr  
 515

<210> 87  
 <211> 1548  
 <212> DNA  
 <213> Thraustochytrium aurum

<220>  
 <221> CDS  
 <222> (1)..(1548)

<400> 87  
 atg acg gtc ggg ttt gac gaa acg gtg act atg gac acg gtc cgc aac 48  
 Met Thr Val Gly Phe Asp Gu Thr Val Thr Met Asp Thr Val Arg Asn  
 1 5 10  
 cac aac atg ccg gac gac gcc tgg tgc gcg atc cac ggc acc gtg tac 96  
 His Asn Met Pro Asp Asp Ala Trp Cys Ala Ile His Gly Thr Val Tyr  
 20 25 30  
 gac atc acc aag ttc agc aag gtg cac ccc ggc ggg gac atc atc atg 144  
 Asp Ile Thr Lys Phe Ser Lys Val His Pro Gly Gly Asp Ile Ile Met  
 35 40 45  
 ctg gcc gct ggc aag gag gcc acc atc ctg ttc gag acg tac cac atc 192  
 Leu Ala Ala Gly Lys Gu Ala Thr Ile Leu Phe Gu Thr Tyr His Ile  
 50 55 60  
 aag ggc gtt ccg gac gcg gtg ctg cgc aag tac aag gtc ggc aag ctg 240  
 Lys Gly Val Pro Asp Ala Val Leu Arg Lys Tyr Lys Val Gly Lys Leu  
 65 70 75 80  
 ccc cag ggc aag aag ggc gaa acg agc cac atg ccc acc ggc ctg gac 288  
 Pro Gln Gly Lys Lys Gly Gu Thr Ser His Met Pro Thr Gly Leu Asp  
 85 90 95  
 tcg gcc ttc tac tac tcg tgg gac agc gag ttt tac agg gtg ctg cgc 336  
 Ser Ala Phe Tyr Tyr Ser Trp Asp Ser Gu Phe Tyr Arg Val Leu Arg  
 100 105 110  
 gag cgc gtc gcc aag aag ctg gcc gag ccc ggc ctg atg cag cgc gcg 384  
 Gu Arg Val Ala Lys Lys Leu Ala Gu Pro Gly Leu Met Gln Arg Ala  
 115 120 125  
 cgc atg gag ctg tgg gcc aag gcg atc ttc ctg ctg gca ggt ttc tgg 432  
 Arg Met Gu Leu Trp Ala Lys Ala Ile Phe Leu Ala Gly Phe Trp  
 130 135 140  
 ggc tcc ctt tac gcc atg tgc gtg cta gac ccg cac ggc ggt gcc atg 480  
 Gly Ser Leu Tyr Ala Met Cys Val Leu Asp Pro His Gly Gly Ala Met  
 145 150 155 160  
 gta gcc gcc gtt acg ctg gcc gtg ttc gct gcc ttt gtc gga act tgc 528  
 Val Ala Ala Val Thr Leu Gly Val Phe Ala Ala Phe Val Gly Thr Cys  
 165 170 175  
 atc cag cac gac ggc agc cac ggc gcc ttc tcc aag tcg cga ttc atg 576  
 Ile Gln His Asp Gly Ser His Gly Ala Phe Ser Lys Ser Arg Phe Met  
 180 185 190  
 aac aag gcg gcg ggc tgg acc ctg gac atg atc ggc gcg agc gcg atg 624  
 Asn Lys Ala Ala Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Met  
 195 200 205  
 acc tgg gag atg cag cac gtt ctt ggc cac cac ccg tac acc aac ctg 672  
 Thr Trp Gu Met Gln His Val Leu Gly His His Pro Tyr Thr Asn Leu  
 210 215 220  
 atc gag atg gag aac ggt ttg gcc aag gtc aag ggc gcc gac gtc gac 720



PF58307. txt

I l e	G u	M e t	G u	A s n	G y	L e u	A l a	L y s	V a l	L y s	G y	A l a	A s p	V a l	A s p	
225					230					235					240	
c c g	a a g	a a g	g t c	g a c	c a g	g a g	a g c	g a c	c c g	g a c	g t c	t t c	a g t	a c g	t a c	768
P r o	L y s	L y s	V a l	A s p	G n	G u	S e r	A s p	P r o	A s p	V a l	P h e	S e r	T h r	T y r	
				245					250					255		
c c g	a t g	c t t	c g c	c t g	c a c	c c g	t g g	c a c	c g c	c a g	c g g	t t t	t a c	c a c	a a g	816
P r o	M e t	L e u	A r g	L e u	H i s	P r o	T r p	H i s	A r g	G n	A r g	P h e	T y r	H i s	L y s	
			260					265					270			
t t c	c a g	c a c	c t g	t a c	g c c	c c g	t t t	a t c	t t t	g g g	t t t	a t g	a c g	a t t	a a c	864
P h e	G n	H i s	L e u	T y r	A l a	P r o	P h e	I l e	P h e	G y	P h e	M e t	T h r	I l e	A s n	
			275				280					285				
a a g	g t g	a t t	t c c	c a g	g a t	g t c	g g g	g t t	g t g	c t g	c g c	a a g	c g c	c t g	t t c	912
L y s	V a l	I l e	S e r	G n	A s p	V a l	G y	V a l	V a l	L e u	A r g	L y s	A r g	L e u	P h e	
			290			295					300					
c a g	a t c	g a c	g c c	a a c	t g c	c g g	t a t	g g c	a g c	c c c	t g g	t a c	g t g	g c c	c g c	960
G n	I l e	A s p	A l a	A s n	C y s	A r g	T y r	G y	S e r	P r o	T r p	T y r	V a l	A l a	A r g	
305				310					315						320	
t t c	t g g	a t c	a t g	a a g	c t c	c t c	a c c	a c g	c t c	t a c	a t g	g t g	g c g	c t t	c c c	1008
P h e	T r p	I l e	M e t	L y s	L e u	L e u	T h r	T h r	L e u	T y r	M e t	V a l	A l a	c t t	c c c	
				325				330						335		
a t g	t a c	a t g	c a g	g g g	c c t	g c t	c a g	g g c	t t g	a a g	c t t	t t c	t t c	a t g	g c c	1056
M e t	T y r	M e t	G n	G y	P r o	A l a	G n	G y	L e u	L y s	L e u	P h e	P h e	M e t	A l a	
			340				345						350			
c a c	t t c	a c c	t g c	g g a	g a g	g t c	c t c	g c c	a c c	a t g	t t t	a t t	g t c	a a c	c a c	1104
H i s	P h e	T h r	C y s	G y	G u	V a l	L e u	A l a	T h r	M e t	P h e	I l e	V a l	A s n	H i s	
			355				360					365				
a t c	a t c	g a g	g g c	g t c	a g c	t a c	g c t	t c c	a a g	g a c	g c g	g t c	a a g	g g c	g t c	1152
I l e	I l e	G u	G y	V a l	S e r	T y r	A l a	S e r	L y s	A s p	A l a	V a l	L y s	G y	V a l	
			370			375					380					
a t g	g c t	c c g	c c g	c g c	a c t	g t g	c a c	g g t	g t c	a c c	c c g	a t g	c a g	g t g	a c g	1200
M e t	A l a	P r o	P r o	A r g	T h r	V a l	H i s	G y	V a l	T h r	P r o	M e t	G n	V a l	T h r	
				385						395					400	
c a a	a a g	g c g	c t c	a g t	g c g	g c c	g a g	t c g	a c c	a a g	t c g	g a c	g c c	g a c	a a g	1248
G n	L y s	A l a	L e u	S e r	A l a	A l a	G u	S e r	T h r	L y s	S e r	A s p	A l a	A s p	L y s	
				405					410					415		
a c g	a c c	a t g	a t c	c c c	c t c	a a c	g a c	t g g	g c c	g c t	g t g	c a g	t g c	c a g	a c c	1296
T h r	T h r	M e t	I l e	P r o	L e u	A s n	A s p	T r p	A l a	A l a	V a l	G n	C y s	G n	T h r	
			420					425					430			
t c t	g t g	a a c	t g g	g c t	g t c	g g g	t c g	t g g	t t t	t g g	a a c	c a c	t t t	t c g	g g c	1344
S e r	V a l	A s n	T r p	A l a	V a l	G y	S e r	T r p	P h e	T r p	A s n	H i s	P h e	S e r	G y	
			435				440						445			
g g c	c t c	a a c	c a c	c a g	a t t	g a g	c a c	c a c	t g c	t t c	c c c	c a a	a a c	c c c	c a c	1392
G y	L e u	A s n	H i s	G n	I l e	G u	H i s	H i s	C y s	P h e	P r o	G n	A s n	P r o	H i s	
				450		455					460					
a c g	g t c	a a c	g t c	t a c	a t c	t c a	g g c	a t c	g t c	a a g	g a g	a c c	t g c	g a a	g a a	1440
T h r	V a l	A s n	V a l	T y r	I l e	S e r	G y	I l e	V a l	L y s	G u	T h r	C y s	G u	G u	
				465		470				475					480	
t a c	g g c	g t g	c c g	t a c	c a g	g c t	g a g	a t c	a g c	c t c	t t c	t c t	g c c	t a t	t t c	1488
T y r	G y	V a l	P r o	T y r	G n	A l a	G u	I l e	S e r	L e u	P h e	S e r	A l a	T y r	P h e	
				485					490					495		
a a g	a t g	c t g	t c g	c a c	c t c	c g c	a c g	c t c	g g c	a a c	g a g	g a c	c t c	a c g	g c c	1536
L y s	M e t	L e u	S e r	H i s	L e u	A r g	T h r	L e u	G y	A s n	G u	A s p	L e u	T h r	A l a	
			500					505					510			
t g g	t c c	a c g	t g a													1548
T r p	S e r	T h r														
			515													

<210> 88  
 <211> 515  
 <212> PRT  
 <213> Thr aust ochyt r i u m a u r e u m

<400> 88  
 M e t T h r V a l G y P h e A s p G u T h r V a l T h r M e t A s p T h r V a l A r g A s n  
 1 5 10 15  
 H i s A s n M e t P r o A s p A s p A l a T r p C y s A l a I l e H i s G y T h r V a l T y r  
 20 25 30  
 A s p I l e T h r L y s P h e S e r L y s V a l H i s P r o G y G y A s p I l e I l e M e t  
 S e i t e 115

PF58307. txt

Leu Ala Ala Gly Lys Gu Ala Thr Ile Leu Phe Gu Thr Tyr His Ile  
 35 40 45  
 50 55 60  
 Lys Gly Val Pro Asp Ala Val Leu Arg Lys Tyr Lys Val Gly Lys Leu  
 65 70 75 80  
 Pro Gn Gly Lys Lys Gly Gu Thr Ser His Met Pro Thr Gly Leu Asp  
 85 90 95  
 Ser Ala Phe Tyr Tyr Ser Trp Asp Ser Gu Phe Tyr Arg Val Leu Arg  
 100 110  
 Gu Arg Val Ala Lys Lys Leu Ala Gu Pro Gly Leu Met Gn Arg Ala  
 115 120 125  
 Arg Met Gu Leu Trp Ala Lys Ala Ile Phe Leu Leu Ala Gly Phe Trp  
 130 135 140  
 Gly Ser Leu Tyr Ala Met Cys Val Leu Asp Pro His Gly Gly Ala Met  
 145 150 155 160  
 Val Ala Ala Val Thr Leu Gly Val Phe Ala Ala Phe Val Gly Thr Cys  
 165 170 175  
 Ile Gn His Asp Gly Ser His Gly Ala Phe Ser Lys Ser Arg Phe Met  
 180 185 190  
 Asn Lys Ala Ala Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Met  
 195 200 205  
 Thr Trp Gu Met Gn His Val Leu Gly His His Pro Tyr Thr Asn Leu  
 210 215 220  
 Ile Gu Met Gu Asn Gly Leu Ala Lys Val Lys Gly Ala Asp Val Asp  
 225 230 235 240  
 Pro Lys Lys Val Asp Gn Gu Ser Asp Pro Asp Val Phe Ser Thr Tyr  
 245 250 255  
 Pro Met Leu Arg Leu His Pro Trp His Arg Gn Arg Phe Tyr His Lys  
 260 265 270  
 Phe Gn His Leu Tyr Ala Pro Phe Ile Phe Gly Phe Met Thr Ile Asn  
 275 280 285  
 Lys Val Ile Ser Gn Asp Val Gly Val Val Leu Arg Lys Arg Leu Phe  
 290 295 300  
 Gn Ile Asp Ala Asn Cys Arg Tyr Gly Ser Pro Trp Tyr Val Ala Arg  
 305 310 315 320  
 Phe Trp Ile Met Lys Leu Leu Thr Thr Leu Tyr Met Val Ala Leu Pro  
 325 330 335  
 Met Tyr Met Gn Gly Pro Ala Gn Gly Leu Lys Leu Phe Phe Met Ala  
 340 345 350  
 His Phe Thr Cys Gly Gu Val Leu Ala Thr Met Phe Ile Val Asn His  
 355 360 365  
 Ile Ile Gu Gly Val Ser Tyr Ala Ser Lys Asp Ala Val Lys Gly Val  
 370 375 380  
 Met Ala Pro Pro Arg Thr Val His Gly Val Thr Pro Met Gn Val Thr  
 385 390 395 400  
 Gn Lys Ala Leu Ser Ala Ala Gu Ser Thr Lys Ser Asp Ala Asp Lys  
 405 410 415  
 Thr Thr Met Ile Pro Leu Asn Asp Trp Ala Ala Val Gn Cys Gn Thr  
 420 425 430  
 Ser Val Asn Trp Ala Val Gly Ser Trp Phe Trp Asn His Phe Ser Gly  
 435 440 445  
 Gly Leu Asn His Gn Ile Gu His His Cys Phe Pro Gn Asn Pro His  
 450 455 460  
 Thr Val Asn Val Tyr Ile Ser Gly Ile Val Lys Gu Thr Cys Gu Gu  
 465 470 475 480  
 Tyr Gly Val Pro Tyr Gn Ala Gu Ile Ser Leu Phe Ser Ala Tyr Phe  
 485 490 495  
 Lys Met Leu Ser His Leu Arg Thr Leu Gly Asn Gu Asp Leu Thr Ala  
 500 505 510  
 Trp Ser Thr  
 515

<210> 89  
 <211> 1548  
 <212> DNA  
 <213> Thr aust ochyt r i um aur eum

<220>

<221> CDS

<222> (1).. (1548)

<400> 89

at g	acg	gt c	ggg	t t t	gac	gaa	acg	gt g	act	at g	gac	acg	gt c	cgc	aac	48
Met	Thr	Val	Gly	Phe	Asp	Glu	Thr	Val	Thr	Met	Asp	Thr	Val	Arg	Asn	
1			5						10					15		
cac	aac	at g	ccg	gac	gac	gcc	t gg	t gc	gcg	at c	cac	ggc	acc	gt g	t ac	96
His	Asn	Met	Pro	Asp	Asp	Ala	Trp	Cys	Ala	Ile	His	Gly	Thr	Val	Tyr	
			20					25					30			
gac	at c	acc	aag	t t c	agc	aag	gt g	cac	ccc	ggc	ggg	gac	at c	at c	at g	144
Asp	Ile	Thr	Lys	Phe	Ser	Lys	Val	His	Pro	Gly	Gly	Asp	Ile	Ile	Met	
			35				40					45				
ct g	gcc	gct	ggc	aag	gag	gcc	acc	at c	ct g	t t c	gag	acg	t ac	cac	at c	192
Leu	Ala	Ala	Gly	Lys	Glu	Ala	Thr	Ile	Leu	Phe	Glu	Thr	Tyr	His	Ile	
	50					55					60					
aag	ggc	gt c	ccg	gac	gcg	gt g	ct g	cgc	aag	t ac	aag	gt c	ggc	aag	ct c	240
Lys	Gly	Val	Pro	Asp	Ala	Val	Leu	Arg	Lys	Tyr	Lys	Val	Gly	Lys	Leu	
	65				70					75					80	
ccc	cag	ggc	aag	aag	ggc	gaa	acg	agc	cac	at g	ccc	acc	ggg	ct c	gac	288
Pro	Gln	Gly	Lys	Lys	Gly	Glu	Thr	Ser	His	Met	Pro	Thr	Gly	Leu	Asp	
				85					90					95		
t cg	gcc	t cc	t ac	t ac	t cg	t gg	gac	agc	gag	t t t	t ac	agg	gt g	ct c	cgc	336
Ser	Ala	Ser	Tyr	Tyr	Ser	Trp	Asp	Ser	Glu	Phe	Tyr	Arg	Val	Leu	Arg	
			100					105					110			
gag	cgc	gt c	gcc	aag	aag	ct g	gcc	gag	ccc	ggc	ct c	at g	cag	cgc	gcg	384
Glu	Arg	Val	Ala	Lys	Lys	Leu	Ala	Glu	Pro	Gly	Leu	Met	Gln	Arg	Ala	
			115				120					125				
cgc	at g	gag	ct c	t gg	gcc	aag	gcg	at c	t t c	ct c	ct g	gca	ggg	t t c	t gg	432
Arg	Met	Glu	Leu	Trp	Ala	Lys	Ala	Ile	Phe	Leu	Leu	Ala	Gly	Phe	Trp	
			130			135					140					
ggc	t cc	ct t	t ac	gcc	at g	t gc	gt g	ct a	gac	ccg	cac	ggc	ggg	gcc	at g	480
Gly	Ser	Leu	Tyr	Ala	Met	Cys	Val	Leu	Asp	Pro	His	Gly	Gly	Ala	Met	
					150					155					160	
gt a	gcc	gcc	gt t	acg	ct c	ggc	gt g	t t c	gct	gcc	t t t	gt c	gga	act	t gc	528
Val	Ala	Ala	Val	Thr	Leu	Gly	Val	Phe	Ala	Ala	Phe	Val	Gly	Thr	Cys	
				165					170					175		
at c	cag	cac	gac	ggc	agc	cac	ggc	gcc	t t c	t cc	aag	t cg	cga	t t c	at g	576
Ile	Gln	His	Asp	Gly	Ser	His	Gly	Ala	Phe	Ser	Lys	Ser	Arg	Phe	Met	
			180					185					190			
aac	aag	gcg	gcg	ggc	t gg	acc	ct c	gac	at g	at c	ggc	gcg	agc	gcg	at g	624
Asn	Lys	Ala	Ala	Gly	Trp	Thr	Leu	Asp	Met	Ile	Gly	Ala	Ser	Ala	Met	
							200					205				
acc	t gg	gag	at g	cag	cac	gt t	ct t	ggc	cac	cac	ccg	t ac	acc	aac	ct c	672
Thr	Trp	Glu	Met	Gln	His	Val	Leu	Gly	His	His	Pro	Tyr	Thr	Asn	Leu	
			210			215					220					
at c	gag	at g	gag	aac	ggg	t t g	gcc	aag	gt c	aag	ggc	gcc	gac	gt c	gac	720
Ile	Glu	Met	Glu	Asn	Gly	Leu	Ala	Lys	Val	Lys	Gly	Ala	Asp	Val	Asp	
				230				235							240	
ccg	aag	aag	gt c	gac	cag	gag	agc	gac	ccg	gac	gt c	t t c	agt	acg	t ac	768
Pro	Lys	Lys	Val	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe	Ser	Thr	Tyr	
				245					250					255		
ccg	at g	ct t	cgc	ct g	cac	ccg	t gg	cac	cgc	cag	cgg	t t t	t ac	cac	aag	816
Pro	Met	Leu	Arg	Leu	His	Pro	Trp	His	Arg	Gln	Arg	Phe	Tyr	His	Lys	
			260					265					270			
t t c	cag	cac	ct g	t ac	gcc	ccg	t t t	at c	t t t	ggg	t t t	at g	acg	at t	aac	864
Phe	Gln	His	Leu	Tyr	Ala	Pro	Phe	Ile	Phe	Gly	Phe	Met	Thr	Ile	Asn	
							280					285				
aag	gt g	at t	t cc	cag	gat	gt c	ggg	gt t	gt g	ct g	cgc	aag	cgc	ct g	t t c	912
Lys	Val	Ile	Ser	Gln	Asp	Val	Gly	Val	Val	Leu	Arg	Lys	Arg	Leu	Phe	
						295					300					
cag	at c	gac	gcc	aac	t gc	ccg	t at	ggc	agc	ccc	t gg	t ac	gt g	gcc	cgc	960
Gln	Ile	Asp	Ala	Asn	Cys	Arg	Tyr	Gly	Ser	Pro	Trp	Tyr	Val	Ala	Arg	
					310					315					320	
t t c	t gg	at c	at g	aag	ct c	ct c	acc	acg	ct c	t ac	acg	gt g	gcg	ct t	ccc	1008
Phe	Trp	Ile	Met	Lys	Leu	Leu	Thr	Thr	Leu	Tyr	Thr	Val	Ala	Leu	Pro	
				325					330					335		
at g	t ac	at g	cag	ggg	cct	gct	cag	ggc	t t g	aag	ct t	t t c	t t c	at g	gcc	1056

PF58307. txt

Met	Tyr	Met	Gln	Gly	Pro	Ala	Gln	Gly	Leu	Lys	Leu	Phe	Phe	Met	Ala		
			340					345					350				
cac	ttc	acc	tgc	gga	gag	gtc	ctc	gcc	acc	atg	ttt	att	gtc	aac	cac		1104
His	Phe	Thr	Cys	Gly	Gu	Val	Leu	Ala	Thr	Met	Phe	Ile	Val	Asn	His		
			355				360					365					
atc	atc	gag	ggc	gtc	agc	tac	gct	tcc	aag	gac	gcg	gtc	aag	ggc	gtc		1152
Ile	Ile	Gu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val	Lys	Gly	Val		
			370			375				380							
atg	gct	ccg	ccg	cgc	act	gtg	cac	ggt	gtc	acc	ccg	atg	cag	gtg	acg		1200
Met	Ala	Pro	Pro	Arg	Thr	Val	His	Gly	Val	Thr	Pro	Met	Gln	Val	Thr		
				390						395					400		
caa	aag	gcg	ctc	agt	gcg	gcc	gag	tcg	acc	aag	tcg	gac	gcc	gac	aag		1248
Gln	Lys	Ala	Leu	Ser	Ala	Ala	Gu	Ser	Thr	Lys	Ser	Asp	Ala	Asp	Lys		
				405				410					415				
acg	acc	atg	atc	ccc	ctc	aac	gac	tgg	gcc	gct	gtg	cag	tgc	cag	acc		1296
Thr	Thr	Met	Ile	Pro	Leu	Asn	Asp	Trp	Ala	Ala	Val	Gln	Cys	Gln	Thr		
			420					425					430				
tct	gtg	aac	tgg	gct	gtc	ggg	tcg	tgg	ttt	tgg	aac	cac	ttt	tcg	ggc		1344
Ser	Val	Asn	Trp	Ala	Val	Gly	Ser	Trp	Phe	Trp	Asn	His	Phe	Ser	Gly		
			435			440						445					
ggc	ctc	aac	cac	cag	att	gag	cac	cac	tgc	ttc	ccc	caa	aac	ccc	cac		1392
Gly	Leu	Asn	His	Gln	Ile	Gu	His	His	Cys	Phe	Pro	Gln	Asn	Pro	His		
			450			455					460						
acg	gtc	aac	gtc	tac	atc	tcg	ggc	atc	gtc	aag	gag	acc	tgc	gaa	gaa		1440
Thr	Val	Asn	Val	Tyr	Ile	Ser	Gly	Ile	Val	Lys	Gu	Thr	Cys	Gu	Gu		
				470				475						480			
tac	ggc	gtg	ccg	tac	cag	gct	gag	atc	agc	ctc	ttc	tct	gcc	tat	ttc		1488
Tyr	Gly	Val	Pro	Tyr	Gln	Ala	Gu	Ile	Ser	Leu	Phe	Ser	Ala	Tyr	Phe		
				485					490					495			
aag	atg	ctg	tcg	cac	ctc	cgc	acg	ctc	ggc	aac	gag	gac	ctc	acg	gcc		1536
Lys	Met	Leu	Ser	His	Leu	Arg	Thr	Leu	Gly	Asn	Gu	Asp	Leu	Thr	Ala		
			500					505					510				
agg	tcc	acg	tga														1548
Arg	Ser	Thr															
			515														

<210> 90  
 <211> 515  
 <212> PRT  
 <213> Thr aust ochyt r i u m aur eum

<400> 90

Met	Thr	Val	Gly	Phe	Asp	Gu	Thr	Val	Thr	Met	Asp	Thr	Val	Arg	Asn		
1			5					10					15				
His	Asn	Met	Pro	Asp	Asp	Ala	Trp	Cys	Ala	Ile	His	Gly	Thr	Val	Tyr		
			20					25					30				
Asp	Ile	Thr	Lys	Phe	Ser	Lys	Val	His	Pro	Gly	Gly	Asp	Ile	Ile	Met		
			35				40					45					
Leu	Ala	Ala	Gly	Lys	Gu	Ala	Thr	Ile	Leu	Phe	Gu	Thr	Tyr	His	Ile		
			50			55					60						
Lys	Gly	Val	Pro	Asp	Ala	Val	Leu	Arg	Lys	Tyr	Lys	Val	Gly	Lys	Leu		
65				70						75				80			
Pro	Gln	Gly	Lys	Lys	Gly	Gu	Thr	Ser	His	Met	Pro	Thr	Gly	Leu	Asp		
			85					90					95				
Ser	Ala	Ser	Tyr	Ser	Trp	Asp	Ser	Gu	Phe	Tyr	Arg	Val	Leu	Arg			
			100				105					110					
Gu	Arg	Val	Ala	Lys	Lys	Leu	Ala	Gu	Pro	Gly	Leu	Met	Gln	Arg	Ala		
			115			120						125					
Arg	Met	Gu	Leu	Trp	Ala	Lys	Ala	Ile	Phe	Leu	Leu	Ala	Gly	Phe	Trp		
			130			135					140						
Gly	Ser	Leu	Tyr	Ala	Met	Cys	Val	Leu	Asp	Pro	His	Gly	Gly	Ala	Met		
145				150						155				160			
Val	Ala	Ala	Val	Thr	Leu	Gly	Val	Phe	Ala	Ala	Phe	Val	Gly	Thr	Cys		
				165				170						175			
Ile	Gln	His	Asp	Gly	Ser	His	Gly	Ala	Phe	Ser	Lys	Ser	Arg	Phe	Met		
			180					185					190				
Asn	Lys	Ala	Ala	Gly	Trp	Thr	Leu	Asp	Met	Ile	Gly	Ala	Ser	Ala	Met		
			195				200					205					

PF58307.txt

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

<210> 91  
 <211> 1548  
 <212> DNA  
 <213> Thr aust ochyt r i u m a u r e u m

<220>  
 <221> CDS  
 <222> (1)..(1548)

<400> 91  
 at g acg gt c ggg t t t gac gaa acg gt g act at g gac acg gt c cgc aac 48  
 Met Thr Val Gly Phe Asp Glu Thr Val Thr Met Asp Thr Val Arg Asn  
 1 5 10 15  
 cac aac at g ccg gac gac gcc tgg tgc gcg at c cac ggc acc gt g t ac 96  
 His Asn Met Pro Asp Asp Ala Trp Cys Ala Ile His Gly Thr Val Tyr  
 20 25 30  
 gac at c acc aag t t c agc aag gt g cac ccc ggc ggg gac at c at c at g 144  
 Asp Ile Thr Lys Phe Ser Lys Val His Pro Gly Gly Asp Ile Ile Met  
 35 40 45  
 ct g gcc gct ggc aag gag gcc acc at c ct g t t c gag acc t ac cac at c 192  
 Leu Ala Ala Gly Lys Glu Ala Thr Ile Leu Phe Glu Thr Tyr His Ile  
 50 55 60  
 aag ggc gt c ccg gac gcg gt g ct g cgc aag t ac aag gt c ggc aag ct c 240  
 Lys Gly Val Pro Asp Ala Val Leu Arg Lys Tyr Lys Val Gly Lys Leu  
 65 70 75 80  
 ccc cag ggc aag aag ggc gaa acg agc cac at g ccc acc ggg ct c gac 288  
 Pro Gln Gly Lys Lys Gly Glu Thr Ser His Met Pro Thr Gly Leu Asp

PF58307. txt

																	85																		90																		95
t cg	gcc	t cc	t ac	t ac	t cg	t gg	gac	agc	gag	t t t	t ac	agg	gt g	ct c	cgc		336	Ser	Al a	Ser	Tyr	Tyr	Ser	Tr p	Asp	Ser	Al a	Phe	Tyr	Arg	Val	Leu	Arg																				
gag	cgc	gt c	gcc	aag	aag	ct g	gcc	gag	ccc	ggc	ct c	at g	cag	cgc	gcg		384	Gu	Arg	Val	Al a	Lys	Lys	Leu	Al a	Gu	Pro	Gly	Leu	Met	G n	Arg	Al a																				
cgc	at g	gag	ct c	t gg	gcc	aag	gcg	at c	t t c	ct c	ct g	gca	ggg	t t c	t gg		432	Arg	Met	Gu	Leu	Tr p	Al a	Lys	Al a	l l e	Phe	Leu	Al a	Gly	Phe	Tr p																					
ggc	t cc	ct t	t ac	gcc	at g	t gc	gt g	ct a	gac	ccg	cac	ggc	ggg	gcc	at g		480	Gly	Ser	Leu	Tyr	Al a	Met	Cys	Val	Leu	Asp	Pro	His	Gly	Gly	Al a	Met																				
gt a	gcc	gcc	gt t	acg	ct c	ggc	gt g	t t c	gct	gcc	t t t	gt c	gga	act	t gc		528	Val	Al a	Al a	Val	Thr	Leu	Gly	Val	Phe	Al a	Al a	Phe	Val	Gly	Thr	Cys																				
at c	cag	cac	gac	ggc	agc	cac	ggc	gcc	t t c	t cc	aag	t cg	cga	t t c	at g		576	l l e	G n	His	Asp	Gly	Ser	His	Gly	Al a	Phe	Ser	Lys	Ser	Arg	Phe	Met																				
aac	aag	gcg	gcg	ggc	t gg	acc	ct c	gac	at g	at c	ggc	gcg	agc	gcg	at g		624	Asn	Lys	Al a	Al a	Gly	Tr p	Thr	Leu	Asp	Met	l l e	Gly	Al a	Ser	Al a	Met																				
acc	t gg	gag	at g	cag	cac	gt t	ct t	ggg	cac	cac	ccg	t ac	acc	aac	ct c		672	Thr	Tr p	Gu	Met	G n	His	Val	Leu	Gly	His	His	Pro	Tyr	Thr	Asn	Leu																				
at c	gag	at g	gag	aac	ggg	t t g	gcc	aag	gt c	aag	ggc	gcc	gac	gt c	gac		720	l l e	Gu	Met	Gu	Asn	Gly	Leu	Al a	Lys	Val	Lys	Gly	Al a	Asp	Val	Asp																				
ccg	aag	aag	gt c	gac	cag	gag	agc	gac	ccg	gac	gt c	t t c	agt	acg	t ac		768	Pro	Lys	Lys	Val	Asp	G n	Gu	Ser	Asp	Pro	Asp	Val	Phe	Ser	Thr	Tyr																				
ccg	at g	ct t	cgc	ct g	cac	ccg	t gg	cac	cgc	cag	ccg	t t t	t ac	cac	aag		816	Pro	Met	Leu	Arg	Leu	His	Pro	Tr p	His	Arg	G n	Arg	Phe	Tyr	His	Lys																				
t t c	cag	cac	ct g	t ac	gcc	ccg	ct t	at c	t t t	ggg	t t t	at g	acg	at t	aac		864	Phe	G n	His	Leu	Tyr	Al a	Pro	Leu	l l e	Phe	Gly	Phe	Met	Thr	l l e	Asn																				
aag	gt g	at t	t cc	cag	gat	gt c	ggg	gt t	gt g	ct g	cgc	aag	cgc	ct g	t t c		912	Lys	Val	l l e	Ser	G n	Asp	Val	Gly	Val	Val	Leu	Arg	Lys	Arg	Leu	Phe																				
cag	at c	gac	gcc	aac	t gc	cgg	t at	ggc	agc	ccc	t gg	aac	gt g	gcc	cgc		960	G n	l l e	Asp	Al a	Asn	Cys	Arg	Tyr	Gly	Ser	Pro	Tr p	Asn	Val	Al a	Arg																				
t t c	t gg	at c	at g	aag	ct c	ct c	acc	acg	ct c	t ac	at g	gt g	gcg	ct t	ccc		1008	Phe	Tr p	l l e	Met	Lys	Leu	Leu	Thr	Thr	Leu	Tyr	Met	Val	Al a	Leu	Pro																				
at g	t ac	at g	cag	ggg	cct	gct	cag	ggc	t t g	aag	ct t	t t c	t t c	at g	gcc		1056	Met	Tyr	Met	G n	Gly	Pro	Al a	G n	Gly	Leu	Lys	Leu	Phe	Phe	Met	Al a																				
cac	t t c	acc	t gc	gga	gag	gt c	ct c	gcc	acc	at g	t t t	at t	gt c	aac	cac		1104	His	Phe	Thr	Cys	Gly	Gu	Val	Leu	Al a	Thr	Met	Phe	l l e	Val	Asn	His																				
at c	at c	gag	ggc	gt c	agc	t ac	gct	t cc	aag	gac	gcg	gt c	aag	ggc	gt c		1152	l l e	l l e	Gu	Gly	Val	Ser	Tyr	Al a	Ser	Lys	Asp	Al a	Val	Lys	Gly	Val																				
at g	gct	ccg	ccg	cgc	act	gt g	cac	ggg	gt c	acc	ccg	at g	cag	gt g	acg		1200	Met	Al a	Pro	Pro	Arg	Thr	Val	His	Gly	Val	Thr	Pro	Met	G n	Val	Thr																				
caa	aag	gcg	ct c	agt	gcg	gcc	gag	t cg	acc	aag	t cg	gac	gcc	gac	aag		1248	G n	Lys	Al a	Leu	Ser	Al a	Al a	Gu	Ser	Thr	Lys	Ser	Asp	Al a	Asp	Lys																				
acg	acc	at g	at c	ccc	ct c	aac	gac	t gg	gcc	gct	gt g	cag	t gc	cag	acc		1296	Thr	Thr	Met	l l e	Pro	Leu	Asn	Asp	Tr p	Al a	Al a	Val	G n	Cys	G n	Thr																				
t ct	gt g	aac	t gg	gct	gt c	ggg	t cg	t gg	t t t	t gg	aac	cac	t t t	t cg	ggc		1344	Ser	Val	Asn	Tr p	Al a	Val	Gly	Ser	Tr p	Phe	Tr p	Asn	His	Phe	Ser	Gly																				
ggc	ct a	aac	cac	cag	at t	gag	cac	cac	t gc	t t c	ccc	caa	aac	ccc	cac		1392																																				

PF58307.txt

Gly	Leu	Asn	His	Gln	Ile	Glu	His	His	Cys	Phe	Pro	Gln	Asn	Pro	His		
	450					455					460						
acg	gtc	aac	gtc	tac	atc	tcg	ggc	atc	gtc	aag	gag	acc	tgc	gaa	gaa	1440	
Thr	Val	Asn	Val	Tyr	Ile	Ser	Gly	Ile	Val	Lys	Glu	Thr	Cys	Glu	Glu		
	465				470					475				480			
tac	ggc	gtg	ccg	tac	cag	gct	gag	atc	agc	ctc	ttc	tct	gcc	tat	ttc	1488	
Tyr	Gly	Val	Pro	Tyr	Gln	Ala	Glu	Ile	Ser	Leu	Phe	Ser	Ala	Tyr	Phe		
				485					490					495			
aag	atg	ctg	tcg	cac	ctc	cgc	acg	ctc	ggc	aac	gag	gac	ctc	acg	gcc	1536	
Lys	Met	Leu	Ser	His	Leu	Arg	Thr	Leu	Gly	Asn	Glu	Asp	Leu	Thr	Ala		
			500					505					510				
tgg	tcc	acg	tga													1548	
Trp	Ser	Thr															
		515															

<210> 92  
 <211> 515  
 <212> PRT  
 <213> Thr aust ochyt r i um aur eum

<400> 92

Met	Thr	Val	Gly	Phe	Asp	Glu	Thr	Val	Thr	Met	Asp	Thr	Val	Arg	Asn		
1				5					10					15			
His	Asn	Met	Pro	Asp	Asp	Ala	Trp	Cys	Ala	Ile	His	Gly	Thr	Val	Tyr		
			20					25					30				
Asp	Ile	Thr	Lys	Phe	Ser	Lys	Val	His	Pro	Gly	Gly	Asp	Ile	Ile	Met		
		35					40					45					
Leu	Ala	Ala	Gly	Lys	Glu	Ala	Thr	Ile	Leu	Phe	Glu	Thr	Tyr	His	Ile		
	50					55					60						
Lys	Gly	Val	Pro	Asp	Ala	Val	Leu	Arg	Lys	Tyr	Lys	Val	Gly	Lys	Leu		
65				70					75					80			
Pro	Gln	Gly	Lys	Lys	Gly	Glu	Thr	Ser	His	Met	Pro	Thr	Gly	Leu	Asp		
			85						90					95			
Ser	Ala	Ser	Tyr	Tyr	Ser	Trp	Asp	Ser	Glu	Phe	Tyr	Arg	Val	Leu	Arg		
		100						105					110				
Glu	Arg	Val	Ala	Lys	Lys	Leu	Ala	Glu	Pro	Gly	Leu	Met	Gln	Arg	Ala		
		115					120					125					
Arg	Met	Glu	Leu	Trp	Ala	Lys	Ala	Ile	Phe	Leu	Leu	Ala	Gly	Phe	Trp		
	130					135					140						
Gly	Ser	Leu	Tyr	Ala	Met	Cys	Val	Leu	Asp	Pro	His	Gly	Gly	Ala	Met		
145					150					155					160		
Val	Ala	Ala	Val	Thr	Leu	Gly	Val	Phe	Ala	Ala	Phe	Val	Gly	Thr	Cys		
				165					170					175			
Ile	Gln	His	Asp	Gly	Ser	His	Gly	Ala	Phe	Ser	Lys	Ser	Arg	Phe	Met		
			180					185					190				
Asn	Lys	Ala	Ala	Gly	Trp	Thr	Leu	Asp	Met	Ile	Gly	Ala	Ser	Ala	Met		
		195					200					205					
Thr	Trp	Glu	Met	Gln	His	Val	Leu	Gly	His	His	Pro	Tyr	Thr	Asn	Leu		
	210					215					220						
Ile	Glu	Met	Glu	Asn	Gly	Leu	Ala	Lys	Val	Lys	Gly	Ala	Asp	Val	Asp		
225				230						235					240		
Pro	Lys	Lys	Val	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe	Ser	Thr	Tyr		
				245					250					255			
Pro	Met	Leu	Arg	Leu	His	Pro	Trp	His	Arg	Gln	Arg	Phe	Tyr	His	Lys		
		260						265					270				
Phe	Gln	His	Leu	Tyr	Ala	Pro	Leu	Ile	Phe	Gly	Phe	Met	Thr	Ile	Asn		
		275					280					285					
Lys	Val	Ile	Ser	Gln	Asp	Val	Gly	Val	Val	Leu	Arg	Lys	Arg	Leu	Phe		
	290					295					300						
Gln	Ile	Asp	Ala	Asn	Cys	Arg	Tyr	Gly	Ser	Pro	Trp	Asn	Val	Ala	Arg		
305					310					315					320		
Phe	Trp	Ile	Met	Lys	Leu	Leu	Thr	Thr	Leu	Tyr	Met	Val	Ala	Leu	Pro		
				325					330					335			
Met	Tyr	Met	Gln	Gly	Pro	Ala	Gln	Gly	Leu	Lys	Leu	Phe	Phe	Met	Ala		
			340					345					350				
His	Phe	Thr	Cys	Gly	Glu	Val	Leu	Ala	Thr	Met	Phe	Ile	Val	Asn	His		
		355					360					365					
Ile	Ile	Glu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val	Lys	Gly	Val		



PF58307. txt

370 375 380  
Met Ala Pro Pro Arg Thr Val His Gly Val Thr Pro Met Gln Val Thr  
385 390 395 400  
Gln Lys Ala Leu Ser Ala Ala Glu Ser Thr Lys Ser Asp Ala Asp Lys  
405 410 415  
Thr Thr Met Ile Pro Leu Asn Asp Trp Ala Ala Val Gln Cys Gln Thr  
420 425 430  
Ser Val Asn Trp Ala Val Gly Ser Trp Phe Trp Asn His Phe Ser Gly  
435 440 445  
Gly Leu Asn His Gln Ile Glu His His Cys Phe Pro Gln Asn Pro His  
450 455 460  
Thr Val Asn Val Tyr Ile Ser Gly Ile Val Lys Glu Thr Cys Glu Glu  
465 470 475 480  
Tyr Gly Val Pro Tyr Gln Ala Glu Ile Ser Leu Phe Ser Ala Tyr Phe  
485 490 495  
Lys Met Leu Ser His Leu Arg Thr Leu Gly Asn Glu Asp Leu Thr Ala  
500 505 510  
Trp Ser Thr  
515

<210> 93  
<211> 1560  
<212> DNA  
<213> Thraustochytrium sp

<220>  
<221> CDS  
<222> (1).. (1560)

<400> 93  
atg acg gtc ggc tac gac gag gag atc ccg ttc gag cag gtc cgc gcg 48  
Met Thr Val Gly Tyr Asp Glu Glu Ile Pro Phe Glu Gln Val Arg Ala  
1 5 10 15  
cac aac aag ccg gat gac gcc tgg tgc gcg atc cac ggg cac gtg tac 96  
His Asn Lys Pro Asp Asp Ala Trp Cys Ala Ile His Gly His Val Tyr  
20 25 30  
gat gtg acc aag ttc gcg agc gtg cac ccg ggc ggc gac att atc ctg 144  
Asp Val Thr Lys Phe Ala Ser Val His Pro Gly Gly Asp Ile Ile Leu  
35 40 45  
ctg gcc gca ggc aag gag gcc acc gtg ctg tac gag act tac cat gtg 192  
Leu Ala Ala Gly Lys Glu Ala Thr Val Leu Tyr Glu Thr Tyr His Val  
50 55 60  
cgg ggc gtc tcg gac gcg gtg ctg cgc aag tac cgc atc ggc aag ctg 240  
Arg Gly Val Ser Asp Ala Val Leu Arg Lys Tyr Arg Ile Gly Lys Leu  
65 70 75 80  
ccg gac ggc caa ggc gcc gcg aac gag aag gaa aag cgg acg ctg tcg 288  
Pro Asp Gly Gln Gly Gly Ala Asn Glu Lys Glu Lys Arg Thr Leu Ser  
85 90 95  
ggc ctg tcg tcg gcc tcg tac tac acg tgg aac agc gac ttt tac agg 336  
Gly Leu Ser Ser Ala Ser Tyr Tyr Thr Trp Asn Ser Asp Phe Tyr Arg  
100 105 110  
gta atg cgc gag cgc gtc gtg gct cgg ctg aag gag cgc ggc aag gcc 384  
Val Met Arg Glu Arg Val Val Ala Arg Leu Lys Glu Arg Gly Lys Ala  
115 120 125  
cgc cgc gga ggc tac gag ctg tgg atc aag gcg ttc ctg ctg ctg gtc 432  
Arg Arg Gly Gly Tyr Glu Leu Trp Ile Lys Ala Phe Leu Leu Leu Val  
130 135 140  
ggc ttc tgg agc tcg ctg tac tgg atg tgc acg ctg gac ccc tcg ttc 480  
Gly Phe Trp Ser Ser Leu Tyr Trp Met Cys Thr Leu Asp Pro Ser Phe  
145 150 155 160  
ggg gcc atc ctg gcc gcc atg tcg ctg ggc gtc ttt gcc gcc ttt gtg 528  
Gly Ala Ile Leu Ala Ala Met Ser Leu Gly Val Phe Ala Ala Phe Val  
165 170 175  
ggc acg tgc atc cag cac gac ggc aac cac gcc gcc ttt gcc cag tcg 576  
Gly Thr Cys Ile Gln His Asp Gly Asn His Gly Ala Phe Ala Gln Ser  
180 185 190  
cga tgg gtc aac aag gtt gcc ggg tgg acg ctg gac atg atc ggc gcc 624  
Arg Trp Val Asn Lys Val Ala Gly Trp Thr Leu Asp Met Ile Gly Ala



PF58307. txt

	195					200						205							
agc	ggc	at g	acg	t gg	gag	t t c	cag	cac	gt c	ct g	ggc	cac	cat	ccg	t ac				672
Ser	G y	Met	Thr	Tr p	G u	Phe	G n	Hi s	Val	Leu	G y	Hi s	Hi s	Pro	Tyr				
	210					215					220								
acg	aac	ct g	at c	gag	gag	gag	aac	ggc	ct g	caa	aag	gt g	agc	ggc	aag				720
Thr	Asn	Leu	I le	G u	G u	G u	Asn	G y	Leu	G n	Lys	Val	Ser	G y	Lys				
	225					230					235				240				
aag	at g	gac	acc	aag	ct g	gcc	gac	cag	gag	agc	gat	ccg	gac	gt c	t t t				768
Lys	Met	Asp	Thr	Lys	Leu	Al a	Asp	G n	G u	Ser	Asp	Pro	Asp	Val	Phe				
				245					250					255					
t cc	acg	t ac	ccg	at g	at g	cgc	ct g	cac	ccg	t gg	cac	cag	aag	cgc	t gg				816
Ser	Thr	Tyr	Pro	Met	Met	Arg	Leu	Hi s	Pro	Tr p	Hi s	G n	Lys	Arg	Tr p				
			260					265					270						
t ac	cac	cgt	t t c	cag	cac	at t	t ac	ggc	ccc	t t c	at c	t t t	ggc	t t c	at g				864
Tyr	Hi s	Arg	Phe	G n	Hi s	I le	Tyr	G y	Pro	Phe	I le	Phe	G y	Phe	Met				
		275					280					285							
acc	at c	aac	aag	gt g	gt c	acg	cag	gac	gt c	ggt	gt g	gt g	ct c	cgc	aag				912
Thr	I le	Asn	Lys	Val	Val	Thr	G n	Asp	Val	G y	Val	Val	Leu	Arg	Lys				
		290				295					300								
cgg	ct c	t t c	cag	at t	gac	gcc	gag	t gc	cgg	t ac	gcg	agc	cca	at g	t ac				960
Arg	Leu	Phe	G n	I le	Asp	Al a	G u	Cys	Arg	Tyr	Al a	Ser	Pro	Met	Tyr				
				310					315					320					
gt g	gcg	cgt	t t c	t gg	at c	at g	aag	gcg	ct c	acg	gt g	ct c	t ac	at g	gt g				1008
Val	Al a	Arg	Phe	Tr p	I le	Met	Lys	Al a	Leu	Thr	Val	Leu	Tyr	Met	Val				
				325					330					335					
gcc	ct g	ccg	t gc	t ac	at g	cag	ggc	ccg	t gg	cac	ggc	ct c	aag	ct g	t t c				1056
Al a	Leu	Pro	Cys	Tyr	Met	G n	G y	Pro	Tr p	Hi s	G y	Leu	Lys	Leu	Phe				
			340					345					350						
gcg	at c	gcg	cac	t t t	acg	t gc	ggc	gag	gt g	ct c	gca	acc	at g	t t c	at t				1104
Al a	I le	Al a	Hi s	Phe	Thr	Cys	G y	G u	Val	Leu	Al a	Thr	Met	Phe	I le				
		355				360					365								
gt g	aac	cac	at c	at c	gag	ggc	gt c	t cg	t ac	gct	t cc	aag	gac	gcg	gt c				1152
Val	Asn	Hi s	I le	I le	G u	G y	Val	Ser	Tyr	Al a	Ser	Lys	Asp	Al a	Val				
	370					375					380								
aag	ggc	acg	at g	gcg	ccg	ccg	aag	acg	at g	cac	ggc	gt g	acg	ccc	at g				1200
Lys	G y	Thr	Met	Al a	Pro	Pro	Lys	Thr	Met	Hi s	G y	Val	Thr	Pro	Met				
					390					395				400					
aac	aac	acg	cgc	aag	gag	gt g	gag	gcg	gag	gcg	t cc	aag	t ct	ggc	gcc				1248
Asn	Asn	Thr	Arg	Lys	G u	Val	G u	Al a	G u	Al a	Ser	Lys	Ser	G y	Al a				
				405				410					415						
gt g	gt c	aag	t ca	gt c	ccg	ct c	gac	gac	t gg	gcc	gt c	gt c	cag	t gc	cag				1296
Val	Val	Lys	Ser	Val	Pro	Leu	Asp	Asp	Tr p	Al a	Val	Val	G n	Cys	G n				
			420					425					430						
acc	t cg	gt g	aac	t gg	agc	gt c	ggc	t cg	t gg	t t c	t gg	aat	cac	t t t	t cc				1344
Thr	Ser	Val	Asn	Tr p	Ser	Val	G y	Ser	Tr p	Phe	Tr p	Asn	Hi s	Phe	Ser				
		435					440					445							
ggc	ggc	ct c	aac	cac	cag	at t	gag	cac	cac	ct g	t t c	ccc	ggr	ct c	agc				1392
G y	G y	Leu	Asn	Hi s	G n	I le	G u	Hi s	Hi s	Leu	Phe	Pro	G y	Leu	Ser				
		450				455					460								
cac	gag	acg	t ac	t ac	cac	at t	cag	gac	gt c	t t t	cag	t cc	acc	t gc	gcc				1440
Hi s	G u	Thr	Tyr	Tyr	Hi s	I le	G n	Asp	Val	Phe	G n	Ser	Thr	Cys	Al a				
		465			470				475					480					
gag	t ac	ggc	gt c	ccg	t ac	cag	cac	gag	cct	t cg	ct c	t gg	acc	gcg	t ac				1488
G u	Tyr	G y	Val	Pro	Tyr	G n	Hi s	G u	Pro	Ser	Leu	Tr p	Thr	Al a	Tyr				
				485					490					495					
t gg	aag	at g	ct c	gag	cac	ct c	cgt	cag	ct c	ggc	aat	gag	gag	acc	cac				1536
Tr p	Lys	Met	Leu	G u	Hi s	Leu	Arg	G n	Leu	G y	Asn	G u	G u	Thr	Hi s				
				500				505					510						
gag	t cc	t gg	cag	cgc	gct	gcc	t ga												1560
G u	Ser	Tr p	G n	Arg	Al a	Al a													
		515																	

<210> 94  
 <211> 519  
 <212> PRT  
 <213> Thr aust ochyt r i u m sp

<400> 94

PF58307.txt

Met Thr Val Gly Tyr Asp Glu Glu Ile Pro Phe Glu Gln Val Arg Ala  
1 5 10 15  
His Asn Lys Pro Asp Asp Ala Trp Cys Ala Ile His Gly His Val Tyr  
20 25 30  
Asp Val Thr Lys Phe Ala Ser Val His Pro Gly Gly Asp Ile Ile Leu  
35 40 45  
Leu Ala Ala Gly Lys Glu Ala Thr Val Leu Tyr Glu Thr Tyr His Val  
50 55 60  
Arg Gly Val Ser Asp Ala Val Leu Arg Lys Tyr Arg Ile Gly Lys Leu  
65 70 75  
Pro Asp Gly Gln Gly Gly Ala Asn Glu Lys Glu Lys Arg Thr Leu Ser  
85 90 95  
Gly Leu Ser Ser Ala Ser Tyr Tyr Thr Trp Asn Ser Asp Phe Tyr Arg  
100 105 110  
Val Met Arg Glu Arg Val Val Ala Arg Leu Lys Glu Arg Gly Lys Ala  
115 120 125  
Arg Arg Gly Gly Tyr Glu Leu Trp Ile Lys Ala Phe Leu Leu Leu Val  
130 135 140  
Gly Phe Trp Ser Ser Leu Tyr Trp Met Cys Thr Leu Asp Pro Ser Phe  
145 150 155 160  
Gly Ala Ile Leu Ala Ala Met Ser Leu Gly Val Phe Ala Ala Phe Val  
165 170 175  
Gly Thr Cys Ile Gln His Asp Gly Asn His Gly Ala Phe Ala Gln Ser  
180 185 190  
Arg Trp Val Asn Lys Val Ala Gly Trp Thr Leu Asp Met Ile Gly Ala  
195 200 205  
Ser Gly Met Thr Trp Glu Phe Gln His Val Leu Gly His His Pro Tyr  
210 215 220  
Thr Asn Leu Ile Glu Glu Glu Asn Gly Leu Gln Lys Val Ser Gly Lys  
225 230 235 240  
Lys Met Asp Thr Lys Leu Ala Asp Gln Glu Ser Asp Pro Asp Val Phe  
245 250 255  
Ser Thr Tyr Pro Met Met Arg Leu His Pro Trp His Gln Lys Arg Trp  
260 265 270  
Tyr His Arg Phe Gln His Ile Tyr Gly Pro Phe Ile Phe Gly Phe Met  
275 280 285  
Thr Ile Asn Lys Val Val Thr Gln Asp Val Gly Val Val Leu Arg Lys  
290 295 300  
Arg Leu Phe Gln Ile Asp Ala Glu Cys Arg Tyr Ala Ser Pro Met Tyr  
305 310 315 320  
Val Ala Arg Phe Trp Ile Met Lys Ala Leu Thr Val Leu Tyr Met Val  
325 330 335  
Ala Leu Pro Cys Tyr Met Gln Gly Pro Trp His Gly Leu Lys Leu Phe  
340 345 350  
Ala Ile Ala His Phe Thr Cys Gly Glu Val Leu Ala Thr Met Phe Ile  
355 360 365  
Val Asn His Ile Ile Glu Gly Val Ser Tyr Ala Ser Lys Asp Ala Val  
370 375 380  
Lys Gly Thr Met Ala Pro Pro Lys Thr Met His Gly Val Thr Pro Met  
385 390 395 400  
Asn Asn Thr Arg Lys Glu Val Glu Ala Glu Ala Ser Lys Ser Gly Ala  
405 410 415  
Val Val Lys Ser Val Pro Leu Asp Asp Trp Ala Val Val Gln Cys Gln  
420 425 430  
Thr Ser Val Asn Trp Ser Val Gly Ser Trp Phe Trp Asn His Phe Ser  
435 440 445  
Gly Gly Leu Asn His Gln Ile Glu His His Leu Phe Pro Gly Leu Ser  
450 455 460  
His Glu Thr Tyr Tyr His Ile Gln Asp Val Phe Gln Ser Thr Cys Ala  
465 470 475 480  
Glu Tyr Gly Val Pro Tyr Gln His Glu Pro Ser Leu Trp Thr Ala Tyr  
485 490 495  
Trp Lys Met Leu Glu His Leu Arg Gln Leu Gly Asn Glu Glu Thr His  
500 505 510  
Glu Ser Trp Gln Arg Ala Ala  
515

<210> 95  
<211> 575  
<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (2)..(3)  
<223> xaa in position 2 to 3 is any amino acid  
<220>  
<221> Variant  
<222> (5)..(13)  
<223> xaa in position 5 to 13 is any amino acid  
<220>  
<221> Variant  
<222> (14)..(48)  
<223> xaa in position 14 to 48 is any or no amino acid  
<220>  
<221> Variant  
<222> (51)..(51)  
<223> xaa in position 51 is any amino acid  
<220>  
<221> Variant  
<222> (54)..(54)  
<223> xaa in position 54 is any amino acid  
<220>  
<221> Variant  
<222> (56)..(56)  
<223> xaa in position 56 is any amino acid  
<220>  
<221> Variant  
<222> (62)..(62)  
<223> xaa in position 62 is any amino acid  
<220>  
<221> Variant  
<222> (64)..(65)  
<223> xaa in position 64 to 65 is any amino acid  
<220>  
<221> Variant  
<222> (66)..(81)  
<223> xaa in position 66 to 81 is any or no amino acid  
<220>  
<221> Variant  
<222> (85)..(85)  
<223> xaa in position 85 is any amino acid  
<220>  
<221> Variant  
<222> (87)..(87)  
<223> xaa in position 87 is any amino acid  
<220>  
<221> Variant  
<222> (89)..(91)  
<223> xaa in position 89 to 91 is any amino acid  
<220>  
<221> Variant  
<222> (97)..(97)  
<223> xaa in position 97 is any amino acid  
<220>  
<221> Variant  
<222> (99)..(99)  
<223> xaa in position 99 is any amino acid  
<220>  
<221> Variant  
<222> (102)..(102)  
<223> xaa in position 102 is any amino acid  
<220>  
<221> Variant  
<222> (105)..(105)

<223> xaa in position 105 is any amino acid  
<220>  
<221> Variant  
<222> (108)..(108)  
<223> xaa in position 108 is any amino acid  
<220>  
<221> Variant  
<222> (110)..(110)  
<223> xaa in position 110 is any amino acid  
<220>  
<221> Variant  
<222> (115)..(116)  
<223> xaa in position 115 to 116 is any amino acid  
<220>  
<221> Variant  
<222> (119)..(119)  
<223> xaa in position 119 is any amino acid  
<220>  
<221> Variant  
<222> (121)..(121)  
<223> xaa in position 121 is any amino acid  
<220>  
<221> Variant  
<222> (123)..(124)  
<223> xaa in position 123 to 124 is any amino acid  
<220>  
<221> Variant  
<222> (127)..(128)  
<223> xaa in position 127 to 128 is any amino acid  
<220>  
<221> Variant  
<222> (130)..(130)  
<223> xaa in position 130 is any amino acid  
<220>  
<221> Variant  
<222> (132)..(137)  
<223> xaa in position 132 to 137 is any amino acid  
<220>  
<221> Variant  
<222> (138)..(154)  
<223> xaa in position 138 to 154 is any or no amino acid  
<220>  
<221> Variant  
<222> (156)..(156)  
<223> xaa in position 156 is any amino acid  
<220>  
<221> Variant  
<222> (158)..(158)  
<223> xaa in position 158 is any amino acid  
<220>  
<221> Variant  
<222> (159)..(159)  
<223> xaa in position 159 is any or no amino acid  
<220>  
<221> Variant  
<222> (161)..(161)  
<223> xaa in position 161 is any amino acid  
<220>  
<221> Variant  
<222> (164)..(164)  
<223> xaa in position 164 is any amino acid  
<220>  
<221> Variant  
<222> (166)..(168)  
<223> xaa in position 166 to 168 is any amino acid  
<220>  
<221> Variant  
<222> (171)..(173)

<223> xaa in position 171 to 173 is any amino acid  
<220>  
<221> Variant  
<222> (175)..(175)  
<223> xaa in position 175 is any amino acid  
<220>  
<221> Variant  
<222> (177)..(177)  
<223> xaa in position 177 is any amino acid  
<220>  
<221> Variant  
<222> (179)..(181)  
<223> xaa in position 179 to 181 is any amino acid  
<220>  
<221> Variant  
<222> (183)..(184)  
<223> xaa in position 183 to 184 is any amino acid  
<220>  
<221> Variant  
<222> (185)..(185)  
<223> xaa in position 185 is any or no amino acid  
<220>  
<221> Variant  
<222> (189)..(189)  
<223> xaa in position 189 is any amino acid  
<220>  
<221> Variant  
<222> (192)..(193)  
<223> xaa in position 192 to 193 is any amino acid  
<220>  
<221> Variant  
<222> (196)..(196)  
<223> xaa in position 196 is any amino acid  
<220>  
<221> Variant  
<222> (200)..(200)  
<223> xaa in position 200 is any amino acid  
<220>  
<221> Variant  
<222> (204)..(204)  
<223> xaa in position 204 is any amino acid  
<220>  
<221> Variant  
<222> (206)..(207)  
<223> xaa in position 206 to 207 is any amino acid  
<220>  
<221> Variant  
<222> (208)..(211)  
<223> xaa in position 208 to 211 is any or no amino acid  
<220>  
<221> Variant  
<222> (213)..(214)  
<223> xaa in position 213 to 214 is any amino acid  
<220>  
<221> Variant  
<222> (215)..(216)  
<223> xaa in position 215 to 216 is any or no amino acid  
<220>  
<221> Variant  
<222> (218)..(220)  
<223> xaa in position 218 to 220 is any amino acid  
<220>  
<221> Variant  
<222> (222)..(224)  
<223> xaa in position 222 to 224 is any amino acid  
<220>  
<221> Variant  
<222> (226)..(226)

<223> xaa in position 226 is any amino acid  
<220>  
<221> Variant  
<222> (240)..(240)  
<223> xaa in position 240 is any amino acid  
<220>  
<221> Variant  
<222> (245)..(250)  
<223> xaa in position 245 to 250 is any amino acid  
<220>  
<221> Variant  
<222> (253)..(253)  
<223> xaa in position 253 is any amino acid  
<220>  
<221> Variant  
<222> (265)..(266)  
<223> xaa in position 265 to 266 is any amino acid  
<220>  
<221> Variant  
<222> (270)..(270)  
<223> xaa in position 270 is any amino acid  
<220>  
<221> Variant  
<222> (273)..(273)  
<223> xaa in position 273 is any amino acid  
<220>  
<221> Variant  
<222> (283)..(284)  
<223> xaa in position 283 to 284 is any amino acid  
<220>  
<221> Variant  
<222> (285)..(290)  
<223> xaa in position 285 to 290 is any or no amino acid  
<220>  
<221> Variant  
<222> (292)..(293)  
<223> xaa in position 292 to 293 is any amino acid  
<220>  
<221> Variant  
<222> (295)..(297)  
<223> xaa in position 295 to 297 is any amino acid  
<220>  
<221> Variant  
<222> (299)..(302)  
<223> xaa in position 299 to 302 is any amino acid  
<220>  
<221> Variant  
<222> (313)..(313)  
<223> xaa in position 313 is any amino acid  
<220>  
<221> Variant  
<222> (316)..(317)  
<223> xaa in position 316 to 317 is any amino acid  
<220>  
<221> Variant  
<222> (319)..(319)  
<223> xaa in position 319 is any amino acid  
<220>  
<221> Variant  
<222> (322)..(322)  
<223> xaa in position 322 is any amino acid  
<220>  
<221> Variant  
<222> (324)..(327)  
<223> xaa in position 324 to 327 is any amino acid  
<220>  
<221> Variant  
<222> (330)..(330)

<223> xaa in position 330 is any amino acid  
<220>  
<221> Variant  
<222> (334)..(334)  
<223> xaa in position 334 is any amino acid  
<220>  
<221> Variant  
<222> (336)..(336)  
<223> xaa in position 336 is any amino acid  
<220>  
<221> Variant  
<222> (338)..(339)  
<223> xaa in position 338 to 339 is any amino acid  
<220>  
<221> Variant  
<222> (341)..(342)  
<223> xaa in position 341 to 342 is any amino acid  
<220>  
<221> Variant  
<222> (346)..(346)  
<223> xaa in position 346 is any amino acid  
<220>  
<221> Variant  
<222> (349)..(350)  
<223> xaa in position 349 to 350 is any amino acid  
<220>  
<221> Variant  
<222> (353)..(354)  
<223> xaa in position 353 to 354 is any amino acid  
<220>  
<221> Variant  
<222> (356)..(359)  
<223> xaa in position 356 to 359 is any amino acid  
<220>  
<221> Variant  
<222> (362)..(363)  
<223> xaa in position 362 to 363 is any amino acid  
<220>  
<221> Variant  
<222> (367)..(368)  
<223> xaa in position 367 to 368 is any amino acid  
<220>  
<221> Variant  
<222> (371)..(371)  
<223> xaa in position 371 is any amino acid  
<220>  
<221> Variant  
<222> (373)..(375)  
<223> xaa in position 373 to 375 is any amino acid  
<220>  
<221> Variant  
<222> (377)..(377)  
<223> xaa in position 377 is any amino acid  
<220>  
<221> Variant  
<222> (381)..(381)  
<223> xaa in position 381 is any amino acid  
<220>  
<221> Variant  
<222> (384)..(385)  
<223> xaa in position 384 to 385 is any amino acid  
<220>  
<221> Variant  
<222> (387)..(387)  
<223> xaa in position 387 is any amino acid  
<220>  
<221> Variant  
<222> (390)..(392)

<223> xaa in position 390 to 392 is any amino acid  
<220>  
<221> Variant  
<222> (395)..(395)  
<223> xaa in position 395 is any amino acid  
<220>  
<221> Variant  
<222> (397)..(398)  
<223> xaa in position 397 to 398 is any amino acid  
<220>  
<221> Variant  
<222> (400)..(402)  
<223> xaa in position 400 to 402 is any amino acid  
<220>  
<221> Variant  
<222> (405)..(405)  
<223> xaa in position 405 is any amino acid  
<220>  
<221> Variant  
<222> (408)..(409)  
<223> xaa in position 408 to 409 is any amino acid  
<220>  
<221> Variant  
<222> (412)..(413)  
<223> xaa in position 412 to 413 is any amino acid  
<220>  
<221> Variant  
<222> (417)..(417)  
<223> xaa in position 417 is any amino acid  
<220>  
<221> Variant  
<222> (427)..(427)  
<223> xaa in position 427 is any amino acid  
<220>  
<221> Variant  
<222> (434)..(435)  
<223> xaa in position 434 to 435 is any amino acid  
<220>  
<221> Variant  
<222> (438)..(438)  
<223> xaa in position 438 is any amino acid  
<220>  
<221> Variant  
<222> (440)..(440)  
<223> xaa in position 440 is any amino acid  
<220>  
<221> Variant  
<222> (442)..(443)  
<223> xaa in position 442 to 443 is any amino acid  
<220>  
<221> Variant  
<222> (444)..(452)  
<223> xaa in position 444 to 452 is any or no amino acid  
<220>  
<221> Variant  
<222> (454)..(454)  
<223> xaa in position 454 is any amino acid  
<220>  
<221> Variant  
<222> (456)..(457)  
<223> xaa in position 456 to 457 is any amino acid  
<220>  
<221> Variant  
<222> (459)..(459)  
<223> xaa in position 459 is any amino acid  
<220>  
<221> Variant  
<222> (463)..(464)



<223> xaa in position 463 to 464 is any amino acid  
<220>  
<221> Variant  
<222> (466)..(466)  
<223> xaa in position 466 is any amino acid  
<220>  
<221> Variant  
<222> (468)..(472)  
<223> xaa in position 468 to 472 is any amino acid  
<220>  
<221> Variant  
<222> (473)..(473)  
<223> xaa in position 473 is any or no amino acid  
<220>  
<221> Variant  
<222> (475)..(483)  
<223> xaa in position 475 to 483 is any amino acid  
<220>  
<221> Variant  
<222> (484)..(489)  
<223> xaa in position 484 to 489 is any or no amino acid  
<220>  
<221> Variant  
<222> (491)..(492)  
<223> xaa in position 491 to 492 is any amino acid  
<220>  
<221> Variant  
<222> (496)..(496)  
<223> xaa in position 496 is any amino acid  
<220>  
<221> Variant  
<222> (506)..(507)  
<223> xaa in position 506 to 507 is any amino acid  
<220>  
<221> Variant  
<222> (527)..(527)  
<223> xaa in position 527 is any amino acid  
<220>  
<221> Variant  
<222> (530)..(532)  
<223> xaa in position 530 to 532 is any amino acid  
<220>  
<221> Variant  
<222> (534)..(538)  
<223> xaa in position 534 to 538 is any amino acid  
<220>  
<221> Variant  
<222> (540)..(542)  
<223> xaa in position 540 to 542 is any amino acid  
<220>  
<221> Variant  
<222> (544)..(545)  
<223> xaa in position 544 to 545 is any amino acid  
<220>  
<221> Variant  
<222> (548)..(548)  
<223> xaa in position 548 is any amino acid  
<220>  
<221> Variant  
<222> (556)..(556)  
<223> xaa in position 556 is any amino acid  
<220>  
<221> Variant  
<222> (558)..(559)  
<223> xaa in position 558 to 559 is any amino acid  
<220>  
<221> Variant  
<222> (561)..(562)

<223> xaa in position 561 to 562 is any amino acid  
 <220>  
 <221> Variant  
 <222> (565)..(565)  
 <223> xaa in position 565 is any amino acid  
 <220>  
 <221> Variant  
 <222> (568)..(569)  
 <223> xaa in position 568 to 569 is any amino acid  
 <220>  
 <221> Variant  
 <222> (573)..(573)  
 <223> xaa in position 573 is any amino acid

<400> 95  
 Met xaa xaa Gly xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 20 25 30  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 35 40 45  
 Val Arg xaa His Asn xaa Pro xaa Asp Ala Trp Cys Ala xaa His xaa  
 50 55 60  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 65 70 75 80  
 xaa Val Tyr Asp xaa Thr xaa Phe xaa xaa xaa His Pro Gly Gly Asp  
 85 90 95  
 xaa Ile xaa Leu Ala xaa Gly Lys xaa Ala Thr xaa Leu xaa Gu Thr  
 100 105 110  
 Tyr His xaa xaa Gly Val xaa Asp xaa Val xaa xaa Lys Tyr xaa xaa  
 115 120 125  
 Gly xaa Leu xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 130 135 140  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Tyr xaa Trp xaa xaa Ser  
 145 150 155 160  
 xaa Phe Tyr xaa Val xaa xaa xaa Arg Val xaa xaa xaa Leu xaa Gu  
 165 170 175  
 xaa Gly xaa xaa xaa Arg xaa xaa xaa Gu Leu Trp xaa Lys Ala xaa  
 180 185 190  
 xaa Leu Leu xaa Gly Phe Trp xaa Ser Leu Tyr xaa Met xaa xaa xaa  
 195 200 205  
 xaa xaa xaa Asp xaa xaa xaa xaa Ala xaa xaa xaa Ala xaa xaa xaa  
 210 215 220  
 Gly xaa Phe Ala Ala Phe Val Gly Thr Cys Ile Gn His Asp Gly xaa  
 225 230 235 240  
 His Gly Ala Phe xaa xaa xaa xaa xaa Asn Lys xaa Ala Gly Trp  
 245 250 255  
 Thr Leu Asp Met Ile Gly Ala Ser xaa xaa Thr Trp Gu xaa Gn His  
 260 265 270  
 xaa Leu Gly His His Pro Tyr Thr Asn Leu xaa xaa xaa xaa xaa  
 275 280 285  
 xaa xaa Lys xaa xaa Gly xaa xaa xaa Asp xaa xaa xaa xaa Asp Gn  
 290 295 300  
 Gu Ser Asp Pro Asp Val Phe Ser xaa Tyr Pro xaa xaa Arg xaa His  
 305 310 315 320  
 Pro xaa His xaa xaa xaa xaa Tyr His xaa Phe Gn His xaa Tyr xaa  
 325 330 335  
 Pro xaa xaa Phe xaa xaa Met Thr Ile xaa Lys Val xaa xaa Gn Asp  
 340 345 350  
 xaa xaa Val xaa xaa xaa xaa Arg Leu xaa xaa Ile Asp Ala xaa xaa  
 355 360 365  
 Arg Tyr xaa Ser xaa xaa xaa Val xaa Arg Phe Trp xaa Met Lys xaa  
 370 375 380  
 xaa Thr xaa Leu Tyr xaa xaa xaa Leu Pro xaa Tyr xaa xaa Gly xaa  
 385 390 395 400  
 xaa xaa Gly Leu xaa Leu Phe xaa xaa Ala His xaa xaa Cys Gly Gu  
 405 410 415  
 xaa Leu Ala Thr Met Phe Ile Val Asn His xaa Ile Gu Gly Val Ser

PF58307.txt

```

          420          425          430
Tyr xaa xaa Lys Asp xaa Val xaa Gly xaa xaa xaa xaa xaa xaa xaa
          435          440          445
xaa xaa xaa xaa Pro xaa Thr xaa xaa Gly xaa Thr Pro Met xaa xaa
          450          455          460
Thr xaa Lys xaa xaa xaa xaa xaa xaa xaa Ser xaa xaa xaa xaa xaa xaa
465          470          475          480
xaa xaa xaa xaa xaa xaa xaa xaa xaa Pro xaa xaa Asp Trp Ala xaa
          485          490          495
Val G n Cys G n Thr Ser Val Asn Trp xaa Gly Ser Trp Phe Trp
          500          505          510
Asn His Phe Ser Gly Gly Leu Asn His G n Ile G u His His xaa Phe
          515          520          525
Pro xaa xaa xaa His xaa xaa xaa xaa xaa Ile xaa xaa xaa Val xaa
          530          535          540
xaa Thr Cys xaa G u Tyr Gly Val Pro Tyr G n xaa G u xaa xaa Leu
545          550          555          560
xaa xaa Ala Tyr xaa Lys Met xaa xaa His Leu Arg xaa Leu Gly
          565          570          575

```

<210> 96  
 <211> 58  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(4)  
 <223> xaa in position 2 to 4 is any amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is Ala or Val  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Pro or Val  
 <220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is Asn or Ser  
 <220>  
 <221> Variant  
 <222> (34)..(34)  
 <223> xaa in position 34 is any amino acid  
 <220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Gly, G n or Ser  
 <220>  
 <221> Variant  
 <222> (38)..(39)  
 <223> xaa in position 38 to 39 is any amino acid  
 <220>  
 <221> Variant  
 <222> (41)..(41)  
 <223> xaa in position 41 is G u or Thr  
 <220>  
 <221> Variant  
 <222> (42)..(42)  
 <223> xaa in position 42 is Asn, Thr or Val  
 <220>  
 <221> Variant

<222> (43)..(44)  
 <223> xaa in position 43 to 44 is any amino acid  
 <220>  
 <221> Variant  
 <222> (45)..(45)  
 <223> xaa in position 45 is His or Tyr  
 <220>  
 <221> Variant  
 <222> (47)..(47)  
 <223> xaa in position 47 is Gln or Ser  
 <220>  
 <221> Variant  
 <222> (48)..(48)  
 <223> xaa in position 48 is Asp or Gly  
 <220>  
 <221> Variant  
 <222> (49)..(49)  
 <223> xaa in position 49 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (50)..(50)  
 <223> xaa in position 50 is Phe or Val  
 <220>  
 <221> Variant  
 <222> (51)..(51)  
 <223> xaa in position 51 is any amino acid  
 <220>  
 <221> Variant  
 <222> (52)..(52)  
 <223> xaa in position 52 is Gln or Ser  
 <220>  
 <221> Variant  
 <222> (55)..(55)  
 <223> xaa in position 55 is Ala or Gln

<400> 96  
 Trp xaa xaa xaa Gln Cys Gln Thr Ser xaa Asn Trp xaa xaa Gly Ser  
 1 5 10 15  
 Trp Phe Trp Asn His Phe Ser Gly Gly Leu xaa His Gln Ile Gu His  
 20 25 30  
 His xaa Phe Pro xaa xaa xaa His xaa xaa xaa xaa xaa Ile xaa xaa  
 35 40 45  
 xaa xaa xaa xaa Thr Cys xaa Gu Tyr Gly  
 50 55

<210> 97  
 <211> 61  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is any amino acid  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Asn or Ser  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Ala or Ser  
 <220>

```

<221> Variant
<222> (22)..(22)
<223> xaa in position 22 is any amino acid
<220>
<221> Variant
<222> (23)..(23)
<223> xaa in position 23 is Asn or Ser
<220>
<221> Variant
<222> (24)..(24)
<223> xaa in position 24 is Lys or Arg
<220>
<221> Variant
<222> (25)..(25)
<223> xaa in position 25 is any amino acid
<220>
<221> Variant
<222> (26)..(26)
<223> xaa in position 26 is Leu, Met or Val
<220>
<221> Variant
<222> (29)..(29)
<223> xaa in position 29 is any or no amino acid
<220>
<221> Variant
<222> (31)..(31)
<223> xaa in position 31 is any or no amino acid
<220>
<221> Variant
<222> (42)..(42)
<223> xaa in position 42 is Ala or Gly
<220>
<221> Variant
<222> (43)..(43)
<223> xaa in position 43 is Phe or Met
<220>
<221> Variant
<222> (47)..(47)
<223> xaa in position 47 is Phe, Leu or Met
<220>
<221> Variant
<222> (50)..(50)
<223> xaa in position 50 is any amino acid
<220>
<221> Variant
<222> (59)..(59)
<223> xaa in position 59 is Leu or Val
<220>
<221> Variant
<222> (60)..(60)
<223> xaa in position 60 is Ile or Leu
<220>
<221> Variant
<222> (61)..(61)
<223> xaa in position 61 is Asp or Glu

<400> 97
Gly xaa Phe Ala Ala Phe xaa Gly Thr Cys Ile Gln His Asp Gly xaa
 1          5          10          15
His Gly Ala Phe xaa xaa xaa xaa xaa Asn Lys xaa Ala xaa Gly
          20          25          30
Trp Thr Leu Asp Met Ile Gly Ala Ser xaa xaa Thr Trp Glu xaa Gln
          35          40          45
His xaa Leu Gly His His Pro Tyr Thr Asn xaa xaa xaa
 50          55          60

```

<210> 98

<211> 40  
<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (3)..(3)  
<223> xaa in position 3 is Lys or Arg  
<220>  
<221> Variant  
<222> (4)..(4)  
<223> xaa in position 4 is Phe or Tyr  
<220>  
<221> Variant  
<222> (7)..(7)  
<223> xaa in position 7 is Ile or Leu  
<220>  
<221> Variant  
<222> (9)..(9)  
<223> xaa in position 9 is Ala or Gly  
<220>  
<221> Variant  
<222> (11)..(12)  
<223> xaa in position 11 to 12 is any or no amino acid  
<220>  
<221> Variant  
<222> (14)..(15)  
<223> xaa in position 14 to 15 is any amino acid  
<220>  
<221> Variant  
<222> (16)..(17)  
<223> xaa in position 16 to 17 is any or no amino acid  
<220>  
<221> Variant  
<222> (20)..(20)  
<223> xaa in position 20 is Ile or Leu  
<220>  
<221> Variant  
<222> (21)..(21)  
<223> xaa in position 21 is Ala or Asn  
<220>  
<221> Variant  
<222> (24)..(24)  
<223> xaa in position 24 is Phe, Ile or Val  
<220>  
<221> Variant  
<222> (25)..(25)  
<223> xaa in position 25 is Gln, Ser or Thr  
<220>  
<221> Variant  
<222> (28)..(28)  
<223> xaa in position 28 is Phe or Val  
<220>  
<221> Variant  
<222> (29)..(29)  
<223> xaa in position 29 is Glu or Gly  
<220>  
<221> Variant  
<222> (31)..(31)  
<223> xaa in position 31 is Ala or Val  
<220>  
<221> Variant  
<222> (32)..(32)  
<223> xaa in position 32 is any amino acid  
<220>  
<221> Variant  
<222> (33)..(33)  
<223> xaa in position 33 is Arg or Ser

<220>  
 <221> Variant  
 <222> (34)..(34)  
 <223> xaa in position 34 is any amino acid  
 <220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Phe or Tyr  
 <220>  
 <221> Variant  
 <222> (38)..(38)  
 <223> xaa in position 38 is any amino acid  
  
 <400> 98  
 Tyr His xaa xaa Gln His xaa Tyr xaa Pro xaa xaa Phe xaa xaa xaa  
 1 5 10 15  
 xaa Met Thr xaa xaa Lys Val xaa xaa Gln Asp xaa xaa Val xaa xaa  
 20 25 30  
 xaa xaa Arg Leu xaa xaa Ile Asp  
 35 40

<210> 99  
 <211> 59  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is Ala or Ile  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Ala, Leu or Val  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is any amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is Gly or Leu  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Met or Val  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Ala or Gly  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is any amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is Phe or Met  
 <220>

<221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is any amino acid  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Pro or Val  
 <220>  
 <221> Variant  
 <222> (23)..(24)  
 <223> xaa in position 23 to 24 is any amino acid  
 <220>  
 <221> Variant  
 <222> (26)..(26)  
 <223> xaa in position 26 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is any amino acid  
 <220>  
 <221> Variant  
 <222> (30)..(30)  
 <223> xaa in position 30 is any amino acid  
 <220>  
 <221> Variant  
 <222> (31)..(31)  
 <223> xaa in position 31 is Ile or Met  
 <220>  
 <221> Variant  
 <222> (32)..(32)  
 <223> xaa in position 32 is Ala or Gly  
 <220>  
 <221> Variant  
 <222> (34)..(34)  
 <223> xaa in position 34 is Phe or Leu  
 <220>  
 <221> Variant  
 <222> (35)..(35)  
 <223> xaa in position 35 is Ala or Thr  
 <220>  
 <221> Variant  
 <222> (39)..(39)  
 <223> xaa in position 39 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (49)..(49)  
 <223> xaa in position 49 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (56)..(56)  
 <223> xaa in position 56 is Ala or Gly  
 <220>  
 <221> Variant  
 <222> (57)..(57)  
 <223> xaa in position 57 is Ser or Thr

<400> 99  
 Phe Trp xaa Met Lys xaa xaa Thr xaa xaa Tyr xaa xaa xaa Leu Pro  
 1 5 10 15  
 xaa Tyr xaa xaa Gly xaa xaa xaa Gly xaa xaa Leu Phe xaa xaa xaa  
 20 25 30  
 His xaa xaa Cys Gly Gu xaa Leu Ala Thr Met Phe Ile Val Asn His  
 35 40 45  
 xaa Ile Gu Gly Val Ser Tyr xaa xaa Lys Asp  
 50 55

<210> 100



<211> 45  
<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (3)..(3)  
<223> xaa in position 3 is Ile or Val  
<220>  
<221> Variant  
<222> (5)..(5)  
<223> xaa in position 5 is Asp or Lys  
<220>  
<221> Variant  
<222> (7)..(9)  
<223> xaa in position 7 to 9 is any amino acid  
<220>  
<221> Variant  
<222> (14)..(14)  
<223> xaa in position 14 is Asp or Gly  
<220>  
<221> Variant  
<222> (15)..(15)  
<223> xaa in position 15 is Ile or Val  
<220>  
<221> Variant  
<222> (17)..(17)  
<223> xaa in position 17 is any amino acid  
<220>  
<221> Variant  
<222> (19)..(19)  
<223> xaa in position 19 is Ala or Gly  
<220>  
<221> Variant  
<222> (20)..(20)  
<223> xaa in position 20 is Ala or Leu  
<220>  
<221> Variant  
<222> (22)..(22)  
<223> xaa in position 22 is Lys or Arg  
<220>  
<221> Variant  
<222> (23)..(23)  
<223> xaa in position 23 is Asp or Glu  
<220>  
<221> Variant  
<222> (24)..(24)  
<223> xaa in position 24 is Ala or Cys  
<220>  
<221> Variant  
<222> (26)..(26)  
<223> xaa in position 26 is Ile or Val  
<220>  
<221> Variant  
<222> (28)..(28)  
<223> xaa in position 28 is any amino acid  
<220>  
<221> Variant  
<222> (30)..(30)  
<223> xaa in position 30 is Ser or Thr  
<220>  
<221> Variant  
<222> (33)..(33)  
<223> xaa in position 33 is Ile, Pro or Val  
<220>  
<221> Variant  
<222> (34)..(34)  
<223> xaa in position 34 is any amino acid

<220>  
 <221> Variant  
 <222> (36)..(36)  
 <223> xaa in position 36 is any amino acid  
 <220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Pro or Ser  
 <220>  
 <221> Variant  
 <222> (39)..(39)  
 <223> xaa in position 39 is any amino acid  
 <220>  
 <221> Variant  
 <222> (41)..(41)  
 <223> xaa in position 41 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (42)..(42)  
 <223> xaa in position 42 is Glu or Arg  
 <220>  
 <221> Variant  
 <222> (45)..(45)  
 <223> xaa in position 45 is Lys or Arg  
  
 <400> 100  
 Tyr Asp xaa Thr xaa Phe xaa xaa xaa His Pro Gly Gly xaa xaa Ile  
 1 5 10 15  
 xaa Leu xaa xaa Gly xaa xaa xaa Thr xaa Leu xaa Glu xaa Tyr His  
 20 25 30  
 xaa xaa Gly xaa xaa Asp xaa Val xaa xaa Lys Tyr xaa  
 35 40 45

<210> 101  
 <211> 24  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(6)  
 <223> xaa in position 2 to 6 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Ser or Thr  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Phe or Tyr  
 <220>  
 <221> Variant  
 <222> (19)..(20)  
 <223> xaa in position 19 to 20 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Leu or Met

<400> 101  
 Asp xaa xaa xaa xaa xaa Gln Glu Ser Asp Pro Asp Val Phe Ser xaa  
 1 5 10 15  
 xaa Pro xaa xaa Arg xaa His Pro  
 20

<210> 102

```

<211> 27
<212> PRT
<213> Artificial sequence

<220>
<221> Variant
<222> (2)..(2)
<223> xaa in position 2 is Asp or Gu
<220>
<221> Variant
<222> (5)..(5)
<223> xaa in position 5 is any amino acid
<220>
<221> Variant
<222> (7)..(7)
<223> xaa in position 7 is Leu or Met
<220>
<221> Variant
<222> (8)..(8)
<223> xaa in position 8 is Lys or Arg
<220>
<221> Variant
<222> (9)..(9)
<223> xaa in position 9 is Gu or Arg
<220>
<221> Variant
<222> (12)..(12)
<223> xaa in position 12 is Ala or Val
<220>
<221> Variant
<222> (13)..(13)
<223> xaa in position 13 is any amino acid
<220>
<221> Variant
<222> (14)..(14)
<223> xaa in position 14 is Lys or Arg
<220>
<221> Variant
<222> (16)..(16)
<223> xaa in position 16 is any amino acid
<220>
<221> Variant
<222> (18)..(18)
<223> xaa in position 18 is any amino acid
<220>
<221> Variant
<222> (20)..(21)
<223> xaa in position 20 to 21 is any amino acid
<220>
<221> Variant
<222> (22)..(22)
<223> xaa in position 22 is Gn or Arg
<220>
<221> Variant
<222> (24)..(24)
<223> xaa in position 24 is Ala or Gly
<220>
<221> Variant
<222> (25)..(26)
<223> xaa in position 25 to 26 is any amino acid

<400> 102
Ser xaa Phe Tyr xaa Val xaa xaa xaa Arg Val xaa xaa xaa Leu xaa
 1           5           10           15
Gu xaa Gly xaa xaa xaa Arg xaa xaa xaa Gu
 20           25

```

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

<220>  
 <221> Variant  
 <222> (2)..(3)  
 <223> xaa in position 2 to 3 is any amino acid  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(10)  
 <223> xaa in position 6 to 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(12)  
 <223> xaa in position 11 to 12 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (17)..(18)  
 <223> xaa in position 17 to 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Asp or Thr  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Ala or Gly  
 <220>  
 <221> Variant  
 <222> (24)..(24)  
 <223> xaa in position 24 is any amino acid  
 <220>  
 <221> Variant  
 <222> (25)..(25)  
 <223> xaa in position 25 is Ala or Thr  
 <220>  
 <221> Variant  
 <222> (26)..(26)  
 <223> xaa in position 26 is any amino acid  
 <220>  
 <221> Variant  
 <222> (28)..(28)  
 <223> xaa in position 28 is Gu or Gly

<400> 103  
 G y xaa xaa xaa G u xaa xaa xaa xaa xaa xaa Val Arg xaa Hi s  
 1 5 10 15  
 xaa xaa Pro xaa Asp xaa Trp xaa xaa xaa Hi s xaa  
 20 25

<210> 104  
 <211> 28  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant

<222> (5)..(5)  
 <223> xaa in position 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is Phe or Trp  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is Ser, Thr or Val  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is Gu or Ser  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Lys or Arg  
 <220>  
 <221> Variant  
 <222> (23)..(23)  
 <223> xaa in position 23 is any amino acid  
 <220>  
 <221> Variant  
 <222> (26)..(26)  
 <223> xaa in position 26 is any amino acid  
 <220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is Ala or Gu  
 <220>  
 <221> Variant  
 <222> (28)..(28)  
 <223> xaa in position 28 is Asp, Gu or Lys  
  
 <400> 104  
 Val Pro Tyr G n xaa G u xaa Ser xaa Leu xaa xaa Ala Tyr xaa Lys  
   1                  5                  10                  15  
 Met xaa xaa His Leu xaa xaa Leu G y xaa xaa xaa  
                   20                  25

<210> 105  
 <211> 15  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is any amino acid  
 <220>

<221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Cys or Ser  
 <220>  
 <221> Variant  
 <222> (7)..(8)  
 <223> xaa in position 7 to 8 is any amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(12)  
 <223> xaa in position 10 to 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Ile or Pro  
 <400> 105  
 G y Phe Trp xaa xaa Leu xaa xaa Met xaa xaa xaa xaa Asp xaa  
 1 5 10 15

<210> 106  
 <211> 29  
 <212> PRT  
 <213> Artificial sequence  
 <220>  
 <221> Variant  
 <222> (2)..(7)  
 <223> xaa in position 2 to 7 is any amino acid  
 <220>  
 <221> Variant  
 <222> (8)..(12)  
 <223> xaa in position 8 to 12 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Asp or Val  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Gu, Asn or Gn  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is any amino acid  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Gn or Arg  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is Gu or Lys  
 <220>  
 <221> Variant  
 <222> (23)..(23)  
 <223> xaa in position 23 is Ala or Gu  
 <220>  
 <221> Variant  
 <222> (24)..(24)  
 <223> xaa in position 24 is Ala, Leu or Val  
 <220>  
 <221> Variant

<222> (25)..(26)  
 <223> xaa in position 25 to 26 is any amino acid  
 <220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is Ala, Glu or Ser  
 <220>  
 <221> Variant  
 <222> (28)..(28)  
 <223> xaa in position 28 is Ala, Glu or Asn  
  
 <400> 106  
 Val xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa Thr Pro  
   1          5          10          15  
 Met xaa xaa Thr xaa xaa xaa xaa xaa xaa xaa xaa Ser  
           20          25

<210> 107  
 <211> 11  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Glu or Asn  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is Cys or Val  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Ala or Gly  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Pro or Val  
 <220>  
 <221> Variant  
 <222> (9)..(10)  
 <223> xaa in position 9 to 10 is any amino acid

<400> 107  
 Ala xaa xaa Arg Tyr xaa Ser xaa xaa xaa Val  
   1          5          10

<210> 108  
 <211> 8  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Ala or Ile  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Phe, Ile or Leu  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Phe or Leu

<400> 108

Trp xaa Lys Ala xaa xaa Leu Leu  
 1 5

<210> 109  
 <211> 903  
 <212> DNA  
 <213> *Ostreococcus tauri*

<220>  
 <221> CDS  
 <222> (1)..(903)

<400> 109  
 at g agc gcc t cc ggt gcg ct g ct g ccc gcg at c gcg t t c gcc gcg t ac 48  
 Met Ser Ala Ser Gly Ala Leu Leu Pro Ala Ile Ala Phe Ala Ala Tyr  
 1 5 10 15  
 gcg t ac gcg acg t ac gcc t ac gcc t t t gag t gg t cg cac gcg aat ggc 96  
 Ala Tyr Ala Thr Tyr Ala Tyr Ala Phe Gu Trp Ser His Ala Asn Gly  
 20 25 30  
 at c gac aac gt c gac gcg cgc gag t gg at c ggt gcg ct g t cg t t g agg 144  
 Ile Asp Asn Val Asp Ala Arg Gu Trp Ile Gly Ala Leu Ser Leu Arg  
 35 40 45  
 ct c ccg gcg at c gcg acg acg at g t ac ct g t t g t t c t gc ct g gt c gga 192  
 Leu Pro Ala Ile Ala Thr Thr Met Tyr Leu Leu Phe Cys Leu Val Gly  
 50 55 60  
 ccg agg t t g at g gcg aag cgc gag gcg t t c gac ccg aag ggg t t c at g 240  
 Pro Arg Leu Met Ala Lys Arg Gu Ala Phe Asp Pro Lys Gly Phe Met  
 65 70 75  
 ct g gcg t ac aat gcg t at cag acg gcg t t c aac gt c gt c gt g ct c ggg 288  
 Leu Ala Tyr Asn Ala Tyr Gn Thr Ala Phe Asn Val Val Val Leu Gly  
 85 90 95  
 at g t t c gcg cga gag at c t cg ggg ct g ggg cag ccc gt g t gg ggg t ca 336  
 Met Phe Ala Arg Gu Ile Ser Gly Leu Gly Gn Pro Val Trp Gly Ser  
 100 105 110  
 acc at g ccg t gg agc gat aga aaa t cg t t t aag at c ct c ct c ggg gt g 384  
 Thr Met Pro Trp Ser Asp Arg Lys Ser Phe Lys Ile Leu Leu Gly Val  
 115 120 125  
 t gg t t g cac t ac aac aac aaa t at t t g gag ct a t t g gac act gt g t t c 432  
 Trp Leu His Tyr Asn Asn Lys Tyr Leu Gu Leu Leu Asp Thr Val Phe  
 130 135 140  
 at g gt t gcg cgc aag aag acg aag cag t t g agc t t c t t g cac gt t t at 480  
 Met Val Ala Arg Lys Lys Thr Lys Gn Leu Ser Phe Leu His Val Tyr  
 145 150 155 160  
 cat cac gcc ct g t t g at c t gg gcg t gg t gg t t g gt g t gt cac t t g at g 528  
 His His Ala Leu Leu Ile Trp Ala Trp Trp Leu Val Cys His Leu Met  
 165 170 175  
 gcc acg aac gat t gt at c gat gcc t ac t t c ggc gcg gcg t gc aac t cg 576  
 Ala Thr Asn Asp Cys Ile Asp Ala Tyr Phe Gly Ala Ala Cys Asn Ser  
 180 185 190  
 t t c att cac at c gt g at g t ac t cg t at t at ct c at g t cg gcg ct c ggc 624  
 Phe Ile His Ile Val Met Tyr Ser Tyr Tyr Leu Met Ser Ala Leu Gly  
 195 200 205  
 at t cga t gc ccg t gg aag cga t ac at c acc cag gct caa at g ct c caa 672  
 Ile Arg Cys Pro Trp Lys Arg Tyr Ile Thr Gn Ala Gn Met Leu Gn  
 210 215 220  
 t t c gt c att gt c t t c gcg cac gcc gt g t t c gt g ct g cgt cag aag cac 720  
 Phe Val Ile Val Phe Ala His Ala Val Phe Val Leu Arg Gn Lys His  
 225 230 235 240  
 t gc ccg gt c acc ct t cct t gg gcg caa at g t t c gt c at g acg aac at g 768  
 Cys Pro Val Thr Leu Pro Trp Ala Gn Met Phe Val Met Thr Asn Met  
 245 250 255  
 ct c gt g ct c t t c ggg aac t t c t ac ct c aag gcg t ac t cg aac aag t cg 816  
 Leu Val Leu Phe Gly Asn Phe Tyr Leu Lys Ala Tyr Ser Asn Lys Ser  
 260 265 270  
 gcg ggc gac ggc gcg agt t cc gt g aaa cca gcc gag acc acg cgc gcg 864  
 Arg Gly Asp Gly Ala Ser Ser Val Lys Pro Ala Gu Thr Thr Arg Ala  
 275 280 285  
 ccc agc gt g cga cgc acg cga t ct cga aaa at t gac t aa 903



Pro Ser Val Arg Arg Thr Arg Ser Arg Lys Ile Asp  
 290 295 300

<210> 110  
 <211> 300  
 <212> PRT  
 <213> *Ostreococcus tauri*

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

<210> 111  
 <211> 834  
 <212> DNA  
 <213> *Pavlova sp*

<220>  
 <221> CDS  
 <222> (1)..(834)

<400> 111  
 atg atg ttg gcc gca ggc tat ct t cta gtg ct c t cg gcc gct cgc cag 48  
 Met Met Leu Ala Ala Gly Tyr Leu Leu Val 10 Leu Ser Ala Ala 15 Arg Gln  
 1  
 agc ttc cag cag gac att gac aac ccc aac ggg gcc tac t cg acc t cg 96  
 Ser Phe Gln Gln Asp Ile Asp Asn Pro Asn Gly Ala Tyr Ser Thr Ser  
 20 25 30  
 tgg act ggc ctg ccc att gtg atg t ct gtg gt c t at ct c agc ggt gt g 144  
 Trp Thr Gly Leu Pro Ile Val Met Ser Val Val Tyr Leu Ser Gly Val  
 35 40 45  
 ttt ggg ct c aca aag tac ttc gag aac cgg aag ccc atg acg ggg ct g 192

PF58307. txt

Phe	Gly	Leu	Thr	Lys	Tyr	Phe	Glu	Asn	Arg	Lys	Pro	Met	Thr	Gly	Leu		
50						55					60						
aag	gac	tac	atg	ttc	act	tac	aat	ctc	tac	cag	gtg	atc	atc	aac	gtg		240
Lys	Asp	Tyr	Met	Phe	Thr	Tyr	Asn	Leu	Tyr	Gln	Val	Ile	Ile	Asn	Val		
65					70					75					80		
tgg	tgc	gtg	gtg	gcc	ttt	ctc	ctg	gag	gtg	cgg	cgt	gcg	ggc	atg	tca		288
Trp	Cys	Val	Val	Ala	Phe	Leu	Leu	Glu	Val	Arg	Arg	Ala	Gly	Met	Ser		
				85					90				95				
ctc	atc	ggc	aat	aag	gtg	gac	ctt	ggg	ccc	aac	tcc	ttc	agg	ctc	ggc		336
Leu	Ile	Gly	Asn	Lys	Val	Asp	Leu	Gly	Pro	Asn	Ser	Phe	Arg	Leu	Gly		
			100					105					110				
ttc	gtc	acg	tgg	gtg	cac	tac	aac	aac	aag	tac	gtg	gag	ctc	ctc	gac		384
Phe	Val	Thr	Trp	Val	His	Tyr	Asn	Asn	Lys	Tyr	Val	Glu	Leu	Leu	Asp		
			115					120				125					
acc	cta	tgg	atg	gtg	ctg	cgc	aag	aag	acg	cag	cag	gtc	tcc	ttc	ctc		432
Thr	Leu	Trp	Met	Val	Leu	Arg	Lys	Lys	Thr	Gln	Gln	Val	Ser	Phe	Leu		
			130					135				140					
cac	gtc	tat	cat	cac	gtg	ctt	ctg	atg	tgg	gcc	tgg	ttc	gtt	gtc	gtc		480
His	Val	Tyr	His	His	Val	Leu	Leu	Met	Trp	Ala	Trp	Phe	Val	Val	Val		
					150					155					160		
aag	ctc	ggc	aat	ggt	ggt	gac	gca	tat	ttt	ggc	ggt	ctc	atg	aac	tcg		528
Lys	Leu	Gly	Asn	Gly	Gly	Asp	Ala	Tyr	Phe	Gly	Gly	Leu	Met	Asn	Ser		
				165					170					175			
atc	atc	cac	gtg	atg	atg	tat	tcc	tac	tac	acc	atg	gcg	ctc	ctg	ggc		576
Ile	Ile	His	Val	Met	Met	Tyr	Ser	Tyr	Tyr	Thr	Met	Ala	Leu	Leu	Gly		
				180				185					190				
tgg	tca	tgc	ccc	tgg	aag	cgc	tac	ctc	acg	cag	gca	cag	ctc	gtg	cag		624
Trp	Ser	Cys	Pro	Trp	Lys	Arg	Tyr	Leu	Thr	Gln	Ala	Gln	Leu	Val	Gln		
			195				200					205					
ttt	tgc	atc	tgc	ctc	gcc	cac	tcc	aca	tgg	gcg	gca	gta	acg	ggt	gcc		672
Phe	Cys	Ile	Cys	Leu	Ala	His	Ser	Thr	Trp	Ala	Ala	Val	Thr	Gly	Ala		
						215						220					
tac	ccg	tgg	cga	att	tgc	ttg	gtg	gag	gtg	tgg	gtg	atg	gtg	tcc	atg		720
Tyr	Pro	Trp	Arg	Ile	Cys	Leu	Val	Glu	Val	Trp	Val	Met	Val	Ser	Met		
					230					235					240		
ctg	gtg	ctc	ttc	aca	cgc	ttc	tac	cgc	cag	gcc	tat	gcc	aag	gag	gcg		768
Leu	Val	Leu	Phe	Thr	Arg	Phe	Tyr	Arg	Gln	Ala	Tyr	Ala	Lys	Glu	Ala		
				245					250					255			
aag	gcc	aag	gag	gcg	aaa	aag	ctc	gca	cag	gag	gca	tca	cag	gcc	aag		816
Lys	Ala	Lys	Glu	Ala	Lys	Lys	Leu	Ala	Gln	Glu	Ala	Ser	Gln	Ala	Lys		
			260					265					270				
gcg	gtc	aag	gcg	gag	t aa												834
Ala	Val	Lys	Ala	Glu													
			275														

<210> 112  
 <211> 277  
 <212> PRT  
 <213> Pavlova sp

<400> 112  
 Met Met Leu Ala Ala Gly Tyr Leu Leu Val Leu Ser Ala Ala Arg Gln  
 1 5 10  
 Ser Phe Gln Gln Asp Ile Asp Asn Pro Asn Gly Ala Tyr Ser Thr Ser  
 20 25 30  
 Trp Thr Gly Leu Pro Ile Val Met Ser Val Val Tyr Leu Ser Gly Val  
 35 40 45  
 Phe Gly Leu Thr Lys Tyr Phe Glu Asn Arg Lys Pro Met Thr Gly Leu  
 50 55 60  
 Lys Asp Tyr Met Phe Thr Tyr Asn Leu Tyr Gln Val Ile Ile Asn Val  
 65 70 75 80  
 Trp Cys Val Val Ala Phe Leu Leu Glu Val Arg Arg Ala Gly Met Ser  
 85 90 95  
 Leu Ile Gly Asn Lys Val Asp Leu Gly Pro Asn Ser Phe Arg Leu Gly  
 100 105 110  
 Phe Val Thr Trp Val His Tyr Asn Asn Lys Tyr Val Glu Leu Leu Asp  
 115 120 125  
 Thr Leu Trp Met Val Leu Arg Lys Lys Thr Gln Gln Val Ser Phe Leu  
 Sei te 148

PF58307. txt

	130					135					140					
His	Val	Tyr	His	His	Val	Leu	Leu	Met	Trp	Ala	Trp	Phe	Val	Val	Val	
145					150					155					160	
Lys	Leu	Gly	Asn	Gly	Gly	Asp	Ala	Tyr	Phe	Gly	Gly	Leu	Met	Asn	Ser	
				165					170					175		
Ile	Ile	His	Val	Met	Met	Tyr	Ser	Tyr	Tyr	Thr	Met	Ala	Leu	Leu	Gly	
			180					185					190			
Trp	Ser	Cys	Pro	Trp	Lys	Arg	Tyr	Leu	Thr	Gln	Ala	Gln	Leu	Val	Gln	
		195					200									
Phe	Cys	Ile	Cys	Leu	Ala	His	Ser	Thr	Trp	Ala	Ala	Val	Thr	Gly	Ala	
	210					215				220						
Tyr	Pro	Trp	Arg	Ile	Cys	Leu	Val	Glu	Val	Trp	Val	Met	Val	Ser	Met	
225					230					235					240	
Leu	Val	Leu	Phe	Thr	Arg	Phe	Tyr	Arg	Gln	Ala	Tyr	Ala	Lys	Glu	Ala	
				245					250					255		
Lys	Ala	Lys	Glu	Ala	Lys	Lys	Leu	Ala	Gln	Glu	Ala	Ser	Gln	Ala	Lys	
			260					265					270			
Ala	Val	Lys	Ala	Glu												
		275														

<210> 113  
 <211> 1077  
 <212> DNA  
 <213> Thalassiopsis pseudonana

<220>  
 <221> CDS  
 <222> (1)..(1077)

<400> 113	
atg tgc tca tca ccg ccg tca caa tcc aaa aca aca tcc ct c ct a gca	48
Met Cys Ser Ser Pro Pro Ser Gln Ser Lys Thr Thr Ser Leu Leu Ala	
1 5	
cgg tac acc acc gcc gcc ct c ct c ct c ct c acc ct c aca aca tgg tgc	96
Arg Tyr Thr Thr Ala Ala Leu Leu Leu Leu Thr Leu Thr Thr Trp Cys	
20 25 30	
cac ttc gcc ttc cca gcc gcc acc gcc aca ccc ggc ct c acc gcc gaa	144
His Phe Ala Phe Pro Ala Ala Thr Ala Thr Pro Gly Leu Thr Ala Glu	
35 40 45	
atg cac tcc tac aaa gtc cca ct c ggt ct c acc gta ttc tac ct g ct g	192
Met His Ser Tyr Lys Val Pro Leu Gly Leu Thr Val Phe Tyr Leu Leu	
50 55 60	
agt cta ccg tca cta aag tac gtt acg gac aac tac ct t gcc aaa aag	240
Ser Leu Pro Ser Leu Lys Tyr Val Thr Asp Asn Tyr Leu Ala Lys Lys	
65 70 75	
tat gat atg aag tca ct c ct a acg gaa tca atg gt g tt g tac aat gt g	288
Tyr Asp Met Lys Ser Leu Leu Thr Gu Ser Met Val Leu Tyr Asn Val	
85 90 95	
gcg caa gt g ct g ct c aat ggg tgg acg gt g tat gcg at t gt g gat gcg	336
Ala Gln Val Leu Leu Asn Gly Trp Thr Val Tyr Ala Ile Val Asp Ala	
100 105 110	
gt g atg aat aga gac cat ccg ttt at t gga agt aga agt tt g gt t ggg	384
Val Met Asn Arg Asp His Pro Phe Ile Gly Ser Arg Ser Leu Val Gly	
115 120 125	
gcg gcg ttg cat agt ggg agc tcg Ser Tyr Ala Val Trp Val His Tyr Cys	432
Ala Ala Leu His Ser Gly Ser Ser Tyr Ala Val Trp Val His Tyr Cys	
130 135 140	
gat aag tat ttg gag ttc ttt gat acg tat ttt atg gt g tt g agg ggg	480
Asp Lys Tyr Leu Glu Phe Phe Asp Thr Tyr Phe Met Val Leu Arg Gly	
145 150 155 160	
aaa atg gac cag gtc tcc ttc ct c cac at c tac cac cac acg acc ata	528
Lys Met Asp Gln Val Ser Phe Leu His Ile Tyr His His Thr Thr Ile	
165 170 175	
gcg tgg gca tgg ttg at c gcc ct c cgc ttc tcc ccc ggt gga gac at t	576
Ala Trp Ala Trp Ile Ala Leu Arg Phe Ser Pro Gly Gly Asp Ile	
180 185 190	
tac ttc ggg gca ct c ct c aac tcc at c at c cac gt c ct c at g tat tcc	624
Tyr Phe Gly Ala Leu Leu Asn Ser Ile Ile His Val Leu Met Tyr Ser	

PF58307. txt

t ac	t ac	gcc	ct t	gcc	ct a	ct c	aag	gt c	agt	t gt	cca	t gg	aaa	cga	t ac	672
Tyr	Tyr	Al a	Leu	Al a	Leu	Leu	Lys	Val	Ser	Cys	Pro	Tr p	Lys	Arg	Tyr	
	210					215					220					
ct g	act	caa	gct	caa	t t a	t t g	caa	t t c	aca	agt	gt g	gt t	t at	acg	720	
Leu	Thr	Gl n	Al a	Gl n	Leu	Leu	Gl n	Phe	Thr	Ser	Val	Val	Val	Tyr	Thr	
225					230					235					240	
ggg	t gt	acg	ggt	t at	act	cat	t ac	t at	cat	acg	aag	cat	gga	gcg	gat	768
Gl y	Cys	Thr	Gl y	Tyr	Thr	Hi s	Tyr	Tyr	Hi s	Thr	Lys	Hi s	Gl y	Al a	Asp	
				245					250					255		
gag	aca	cag	cct	agt	t t a	gga	acg	t at	t at	t t c	t gt	t gt	gga	gt g	cag	816
Gl u	Thr	Gl n	Pro	Ser	Leu	Gl y	Thr	Tyr	Tyr	Phe	Cys	Cys	Gl y	Val	Gl n	
			260					265					270			
gt g	t t t	gag	at g	gt t	agt	t t g	t t t	gt a	ct c	t t t	t cc	at c	t t t	t at	aaa	864
Val	Phe	Gl u	Met	Val	Ser	Leu	Phe	Val	Leu	Phe	Ser	Il e	Phe	Tyr	Lys	
			275				280						285			
cga	t cc	t at	t cg	aag	aag	aac	aag	t ca	gga	gga	aag	gat	agc	aag	aag	912
Arg	Ser	Tyr	Ser	Lys	Lys	Asn	Lys	Ser	Gl y	Gl y	Lys	Asp	Ser	Lys	Lys	
	290					295					300					
aat	gat	gat	ggg	aat	aat	gag	gat	caa	t gt	cac	aag	gct	at g	aag	gat	960
Asn	Asp	Asp	Gl y	Asn	Asn	Gl u	Asp	Gl n	Cys	Hi s	Lys	Al a	Met	Lys	Asp	
305				310						315					320	
at a	t cg	gag	ggt	gcg	aag	gag	gt t	gt g	ggg	cat	gca	gcg	aag	gat	gct	1008
Il e	Ser	Gl u	Gl y	Al a	Lys	Gl u	Val	Val	Gl y	Hi s	Al a	Al a	Lys	Asp	Al a	
				325					330					335		
gga	aag	t t g	gt g	gct	acg	gcg	agt	aag	gct	gt a	aag	agg	aag	gga	act	1056
Gl y	Lys	Leu	Val	Al a	Thr	Al a	Ser	Lys	Al a	Val	Lys	Arg	Lys	Gl y	Thr	
			340					345					350			
cgt	gt t	act	ggt	gcc	at g	t ag										1077
Arg	Val	Thr	Gl y	Al a	Met											
			355													

<210> 114

<211> 358

<212> PRT

<213> Thal assi osi r a pseudonana

<400> 114

Met	Cys	Ser	Ser	Pro	Pro	Ser	Gl n	Ser	Lys	Thr	Thr	Ser	Leu	Leu	Al a	
1				5					10						15	
Arg	Tyr	Thr	Thr	Al a	Al a	Leu	Leu	Leu	Leu	Thr	Leu	Thr	Thr	Tr p	Cys	
			20						25					30		
Hi s	Phe	Al a	Phe	Pro	Al a	Al a	Thr	Al a	Thr	Pro	Gl y	Leu	Thr	Al a	Gl u	
			35				40						45			
Met	Hi s	Ser	Tyr	Lys	Val	Pro	Leu	Gl y	Leu	Thr	Val	Phe	Tyr	Leu	Leu	
	50					55					60					
Ser	Leu	Pro	Ser	Leu	Lys	Tyr	Val	Thr	Asp	Asn	Tyr	Leu	Al a	Lys	Lys	
65					70					75				80		
Tyr	Asp	Met	Lys	Ser	Leu	Leu	Thr	Gl u	Ser	Met	Val	Leu	Tyr	Asn	Val	
				85					90					95		
Al a	Gl n	Val	Leu	Leu	Asn	Gl y	Tr p	Thr	Val	Tyr	Al a	Il e	Val	Asp	Al a	
			100					105					110			
Val	Met	Asn	Arg	Asp	Hi s	Pro	Phe	Il e	Gl y	Ser	Arg	Ser	Leu	Val	Gl y	
		115					120					125				
Al a	Al a	Leu	Hi s	Ser	Gl y	Ser	Tyr	Al a	Val	Tr p	Val	Hi s	Tyr	Cys		
	130					135				140						
Asp	Lys	Tyr	Leu	Gl u	Phe	Phe	Asp	Thr	Tyr	Phe	Met	Val	Leu	Arg	Gl y	
145					150					155				160		
Lys	Met	Asp	Gl n	Val	Ser	Phe	Leu	Hi s	Il e	Tyr	Hi s	Hi s	Thr	Thr	Il e	
				165						170				175		
Al a	Tr p	Al a	Tr p	Tr p	Il e	Al a	Leu	Arg	Phe	Ser	Pro	Gl y	Gl y	Asp	Il e	
			180					185					190			
Tyr	Phe	Gl y	Al a	Leu	Leu	Asn	Ser	Il e	Il e	Hi s	Val	Leu	Met	Tyr	Ser	
			195				200						205			
Tyr	Tyr	Al a	Leu	Al a	Leu	Leu	Lys	Val	Ser	Cys	Pro	Tr p	Lys	Arg	Tyr	
	210					215					220					
Leu	Thr	Gl n	Al a	Gl n	Leu	Leu	Gl n	Phe	Thr	Ser	Val	Val	Val	Tyr	Thr	
225					230					235					240	

PF58307.txt

G y Cys Thr G y Tyr Thr H i s Tyr Tyr H i s Thr Lys H i s G y A l a Asp  
 245 250  
 G u Thr G l n P r o Ser Leu G y Thr Tyr Tyr Phe Cys Cys G y Val G n  
 260 265  
 Val Phe G u M e t Val Ser Leu Phe Val Leu Phe Ser I l e Phe Tyr Lys  
 275 280 285  
 Arg Ser Tyr Ser Lys Lys Asn Lys Ser G y G y Lys Asp Ser Lys Lys  
 290 295 300  
 Asn Asp Asp G y Asn Asn G u Asp G n Cys H i s Lys A l a M e t Lys Asp  
 305 310 315  
 I l e Ser G u G y A l a Lys G u Val Val G y H i s A l a A l a Lys Asp A l a  
 325 330 335  
 G y Lys Leu Val A l a Thr A l a Ser Lys A l a Val Lys Arg Lys G y Thr  
 340 345 350  
 Arg Val Thr G y A l a M e t  
 355

<210> 115  
 <211> 903  
 <212> DNA  
 <213> *Ostreococcus tauri*

<220>  
 <221> CDS  
 <222> (1)..(903)

<400> 115  
 at g agc gcc t cc ggt gcg ct g ct g ccc gcg at c gcg t cc gcc gcg t ac 48  
 Met Ser A l a Ser G y A l a Leu Leu P r o A l a I l e A l a Ser A l a A l a Tyr  
 1 5 10 15  
 gcg t ac gcg acg t ac gcc t ac gcc t t t gag t gg t cg cac gcg aat ggc 96  
 A l a Tyr A l a Thr Tyr A l a Tyr A l a Phe G u Trp Ser H i s A l a Asn G y  
 20 25 30  
 at c gac aac gt c gac gcg cgc gag t gg at c ggt gcg ct g t cg t t g agg 144  
 I l e Asp Asn Val Asp A l a Arg G u Trp I l e G y A l a Leu Ser Leu Arg  
 35 40 45  
 ct c ccg gcg at c gcg acg acg at g t ac ct g t t g t t c t gc ct g gt c gga 192  
 Leu P r o A l a I l e A l a Thr Thr Met Tyr Leu Leu Phe Cys Leu Val G y  
 50 55 60  
 ccg agg t t g at g gcg aag cgc gag gcg t t c gac ccg aag ggg t t c at g 240  
 P r o Arg Leu M e t A l a Lys Arg G u A l a Phe Asp P r o Lys G y Phe M e t  
 65 70 75 80  
 ct g gcg t ac aat gcg t at cag acg gcg t t c aac gt c gt c gt g ct c ggg 288  
 Leu A l a Tyr Asn A l a Tyr G n Thr A l a Phe Asn Val Val Val Leu G y  
 85 90 95  
 at g t t c gcg cga gag at c t cg ggg ct g ggg cag ccc gt g t gg ggg t ca 336  
 Met Phe A l a Arg G u I l e Ser G y Leu G y G n P r o Val Trp G y Ser  
 100 105 110  
 acc at g ccg t gg agc gat aga aaa t cg t t t aag at c ct c ct c ggg gt g 384  
 Thr Met P r o Trp Ser Asp Arg Lys Ser Phe Lys I l e Leu Leu G y Val  
 115 120 125  
 t gg t t g cac t ac aac aac aaa t at t t g gag ct a t t g gac act gt g t t c 432  
 Trp Leu H i s Tyr Asn Asn Lys Tyr Leu G u Leu Leu Asp Thr Val Phe  
 130 135 140  
 at g g t t gcg cgc aag aag acg aag cag t t g agc t t c t t g cac gt t t at 480  
 Met Val A l a Arg Lys Lys Thr Lys G n Leu Ser Phe Leu H i s Val Tyr  
 145 150 155 160  
 cat cac gcc ct g t t g at c t gg gcg t gg t gg t t g gt g t gt cac t t g at g 528  
 H i s H i s A l a Leu Leu I l e Trp A l a Trp Trp Leu Val Cys H i s Leu M e t  
 165 170 175  
 gcc acg aac gat t gt at c gat gcc t ac t t c ggc gcg gcg t gc aac t cg 576  
 A l a Thr Asn Asp Cys I l e Asp A l a Tyr Phe G y A l a A l a Cys Asn Ser  
 180 185 190  
 t t c att cac at c gt g at g t ac t cg t at t at ct c at g t cg gcg ct c ggc 624  
 Phe I l e H i s I l e Val M e t Tyr Ser Tyr Tyr Leu M e t Ser A l a Leu G y  
 195 200 205  
 att cga t gc ccg t gg aag cga t ac at c acc cag gct caa at g ct c caa 672  
 I l e Arg Cys P r o Trp Lys Arg Tyr I l e Thr G n A l a G n M e t Leu G n

Seite 151

PF58307. txt

210	215	220	
t t c g t c a t t g t c t t c g c g c a c g c c g t g t t c g t g c t g c g t c a g a a g c a c			720
P h e V a l I l e V a l P h e A l a H i s A l a V a l P h e V a l L e u A r g G n L y s H i s			
225	230	235	240
t g c c c g g t c a c c c t t c c t t g g g c g c a a a t g t t c g t c a t g a c g a a c a t g			768
C y s P r o V a l T h r L e u P r o T r p A l a G n M e t P h e V a l M e t T h r A s n M e t			
245	250	255	
c t c g t g c t c t t c g g g a a c t t c t a c c t c a a g g c g t a c t c g a a c a a g t c g			816
L e u V a l L e u P h e G y A s n P h e T y r L e u L y s A l a T y r S e r A s n L y s S e r			
260	265	270	
c g c g g c g a c g g c g c g a g t t c c g t g a a a c c a g c c g a g a c c a c g c g c g c g			864
A r g G y A s p G y A l a S e r S e r V a l L y s P r o A l a G u T h r T h r A r g A l a			
275	280	285	
c c c a g c g t g c g a c g c a c g c g a t c t c g a a a a t t g a c t a a			903
P r o S e r V a l A r g A r g T h r A r g S e r A r g L y s I l e A s p			
290	295	300	

<210> 116  
 <211> 300  
 <212> PRT  
 <213> *Ostreococcus tauri*

<400> 116

M e t S e r A l a S e r G y A l a L e u L e u P r o A l a I l e A l a S e r A l a A l a T y r	
1 5 10 15	
A l a T y r A l a T h r T y r A l a T y r A l a P h e G u T r p S e r H i s A l a A s n G y	
20 25 30	
I l e A s p A s n V a l A s p A l a A r g G u T r p I l e G y A l a L e u S e r L e u A r g	
35 40 45	
L e u P r o A l a I l e A l a T h r T h r M e t T y r L e u L e u P h e C y s L e u V a l G y	
50 55 60	
P r o A r g L e u M e t A l a L y s A r g G u A l a P h e A s p P r o L y s G y P h e M e t	
65 70 75 80	
L e u A l a T y r A s n A l a T y r G n T h r A l a P h e A s n V a l V a l V a l L e u G y	
85 90 95	
M e t P h e A l a A r g G u I l e S e r G y L e u G y G n P r o V a l T r p G y S e r	
100 105 110	
T h r M e t P r o T r p S e r A s p A r g L y s S e r P h e L y s I l e L e u L e u G y V a l	
115 120 125	
T r p L e u H i s T y r A s n A s n L y s T y r L e u G u L e u L e u A s p T h r V a l P h e	
130 135 140	
M e t V a l A l a A r g L y s L y s T h r L y s G n L e u S e r P h e L e u H i s V a l T y r	
145 150 155 160	
H i s H i s A l a L e u L e u I l e T r p A l a T r p T r p L e u V a l C y s H i s L e u M e t	
165 170 175	
A l a T h r A s n A s p C y s I l e A s p A l a T y r P h e G y A l a A l a C y s A s n S e r	
180 185 190	
P h e I l e H i s I l e V a l M e t T y r S e r T y r T y r L e u M e t S e r A l a L e u G y	
195 200 205	
I l e A r g C y s P r o T r p L y s A r g T y r I l e T h r G n A l a G n M e t L e u G n	
210 215 220	
P h e V a l I l e V a l P h e A l a H i s A l a V a l P h e V a l L e u A r g G n L y s H i s	
225 230 235 240	
C y s P r o V a l T h r L e u P r o T r p A l a G n M e t P h e V a l M e t T h r A s n M e t	
245 250 255	
L e u V a l L e u P h e G y A s n P h e T y r L e u L y s A l a T y r S e r A s n L y s S e r	
260 265 270	
A r g G y A s p G y A l a S e r S e r V a l L y s P r o A l a G u T h r T h r A r g A l a	
275 280 285	
P r o S e r V a l A r g A r g T h r A r g S e r A r g L y s I l e A s p	
290 295 300	

<210> 117  
 <211> 339  
 <212> PRT  
 <213> Artificial sequence

<220>  
<221> Variant  
<222> (3)..(29)  
<223> xaa in position 3 to 29 is any amino acid  
<220>  
<221> Variant  
<222> (30)..(43)  
<223> xaa in position 30 to 43 is any or no amino acid  
<220>  
<221> Variant  
<222> (45)..(50)  
<223> xaa in position 45 to 50 is any amino acid  
<220>  
<221> Variant  
<222> (53)..(73)  
<223> xaa in position 53 to 73 is any amino acid  
<220>  
<221> Variant  
<222> (74)..(79)  
<223> xaa in position 74 to 79 is any or no amino acid  
<220>  
<221> Variant  
<222> (81)..(82)  
<223> xaa in position 81 to 82 is any amino acid  
<220>  
<221> Variant  
<222> (85)..(86)  
<223> xaa in position 85 to 86 is any amino acid  
<220>  
<221> Variant  
<222> (88)..(90)  
<223> xaa in position 88 to 90 is any amino acid  
<220>  
<221> Variant  
<222> (92)..(110)  
<223> xaa in position 92 to 110 is any amino acid  
<220>  
<221> Variant  
<222> (112)..(127)  
<223> xaa in position 112 to 127 is any amino acid  
<220>  
<221> Variant  
<222> (128)..(128)  
<223> xaa in position 128 is any or no amino acid  
<220>  
<221> Variant  
<222> (130)..(130)  
<223> xaa in position 130 is any amino acid  
<220>  
<221> Variant  
<222> (133)..(134)  
<223> xaa in position 133 to 134 is any amino acid  
<220>  
<221> Variant  
<222> (137)..(137)  
<223> xaa in position 137 is any amino acid  
<220>  
<221> Variant  
<222> (139)..(140)  
<223> xaa in position 139 to 140 is any amino acid  
<220>  
<221> Variant  
<222> (143)..(144)  
<223> xaa in position 143 to 144 is any amino acid  
<220>  
<221> Variant  
<222> (147)..(147)  
<223> xaa in position 147 is any amino acid

<220>  
 <221> Vari ant  
 <222> ( 149).. ( 149)  
 <223> xaa in position 149 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 151).. ( 152)  
 <223> xaa in position 151 to 152 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 154).. ( 154)  
 <223> xaa in position 154 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 159).. ( 159)  
 <223> xaa in position 159 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 163).. ( 166)  
 <223> xaa in position 163 to 166 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 170).. ( 179)  
 <223> xaa in position 170 to 179 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 180).. ( 182)  
 <223> xaa in position 180 to 182 is any or no ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 184).. ( 184)  
 <223> xaa in position 184 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 188).. ( 190)  
 <223> xaa in position 188 to 190 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 193).. ( 193)  
 <223> xaa in position 193 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 196).. ( 197)  
 <223> xaa in position 196 to 197 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 203).. ( 206)  
 <223> xaa in position 203 to 206 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 208).. ( 210)  
 <223> xaa in position 208 to 210 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 217).. ( 217)  
 <223> xaa in position 217 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 222).. ( 223)  
 <223> xaa in position 222 to 223 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 226).. ( 252)  
 <223> xaa in position 226 to 252 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 253).. ( 267)  
 <223> xaa in position 253 to 267 is any or no ami no aci d



<220>  
 <221> Vari ant  
 <222> ( 269) .. ( 272)  
 <223> xaa in position 269 to 272 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 276) .. ( 277)  
 <223> xaa in position 276 to 277 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 280) .. ( 282)  
 <223> xaa in position 280 to 282 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 284) .. ( 291)  
 <223> xaa in position 284 to 291 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 292) .. ( 324)  
 <223> xaa in position 292 to 324 is any or no ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 326) .. ( 337)  
 <223> xaa in position 326 to 337 is any ami no aci d  
 <220>  
 <221> Vari ant  
 <222> ( 338) .. ( 338)  
 <223> xaa in position 338 is any or no ami no aci d

<400> 117  
 Leu Leu xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 20 25 30  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Pro xaa xaa xaa xaa  
 35 40 45  
 xaa xaa Tyr Leu xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 50 55 60  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Met  
 65 70 75 80  
 xaa xaa Tyr Asn xaa xaa G n xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 85 90 95  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Gly xaa  
 100 105 110  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 115 120 125  
 Trp xaa Hi s Tyr xaa xaa Lys Tyr xaa G u xaa xaa Asp Thr xaa xaa  
 130 135 140  
 Met Val xaa Arg xaa Lys xaa xaa G n xaa Ser Phe Leu Hi s xaa Tyr  
 145 150 155 160  
 Hi s Hi s xaa xaa xaa xaa Trp Al a Trp xaa xaa xaa xaa xaa xaa  
 165 170 175  
 xaa xaa xaa xaa xaa xaa Asp xaa Tyr Phe Gl y xaa xaa xaa Asn Ser  
 180 185 190  
 xaa Ile Hi s xaa xaa Met Tyr Ser Tyr Tyr xaa xaa xaa xaa Leu xaa  
 195 200 205  
 xaa xaa Cys Pro Trp Lys Arg Tyr xaa Thr G n Al a G n xaa xaa G n  
 210 215 220  
 Phe xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 225 230 235 240  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 245 250 255  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Met xaa xaa xaa xaa  
 260 265 270  
 Val Leu Phe xaa xaa Phe Tyr xaa xaa xaa Tyr xaa xaa xaa xaa xaa  
 275 280 285  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 290 295 300



<223> xaa in position 39 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (40)..(40)  
 <223> xaa in position 40 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (43)..(43)  
 <223> xaa in position 43 is Cys, Thr or Val  
 <220>  
 <221> Variant  
 <222> (44)..(44)  
 <223> xaa in position 44 is any amino acid  
 <220>  
 <221> Variant  
 <222> (45)..(45)  
 <223> xaa in position 45 is Cys or Val

<400> 118  
 Tyr Phe Gly xaa xaa xaa Asn Ser xaa Ile His xaa Val xaa Met Tyr  
 1 5 10 15  
 Ser Tyr Tyr xaa xaa xaa Leu xaa xaa xaa Cys Pro Trp Lys Arg  
 20 25 30  
 Tyr xaa Thr Gln Ala Gln xaa xaa Gln Phe xaa xaa xaa  
 35 40 45

<210> 119  
 <211> 43  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Cys or Asn  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Asp or Asn  
 <220>  
 <221> Variant  
 <222> (9)..(10)  
 <223> xaa in position 9 to 10 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(14)  
 <223> xaa in position 13 to 14 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Phe or Trp  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Ala or Leu

<220>  
 <221> Variant  
 <222> (23)..(23)  
 <223> xaa in position 23 is any amino acid  
 <220>  
 <221> Variant  
 <222> (25)..(26)  
 <223> xaa in position 25 to 26 is any amino acid  
 <220>  
 <221> Variant  
 <222> (28)..(28)  
 <223> xaa in position 28 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (33)..(33)  
 <223> xaa in position 33 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Ala, Thr or Val  
 <220>  
 <221> Variant  
 <222> (38)..(38)  
 <223> xaa in position 38 is any amino acid  
 <220>  
 <221> Variant  
 <222> (39)..(39)  
 <223> xaa in position 39 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (40)..(40)  
 <223> xaa in position 40 is any amino acid  
  
 <400> 119  
 Trp xaa His Tyr xaa xaa Lys Tyr xaa xaa Leu xaa xaa xaa Asp Thr  
   1          5          10          15  
 xaa xaa Met Val xaa Arg xaa Lys xaa xaa Gln xaa Ser Phe Leu His  
           20          25          30  
 xaa Tyr His His xaa xaa xaa xaa Trp Ala Trp  
           35          40

<210> 120  
 <211> 37  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(3)  
 <223> xaa in position 2 to 3 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is Gly, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(14)  
 <223> xaa in position 13 to 14 is any amino acid  
 <220>

```

<221> Variant
<222> (15)..(15)
<223> xaa in position 15 is Ala or Ser
<220>
<221> Variant
<222> (17)..(17)
<223> xaa in position 17 is Ala or Ser
<220>
<221> Variant
<222> (18)..(18)
<223> xaa in position 18 is any amino acid
<220>
<221> Variant
<222> (19)..(19)
<223> xaa in position 19 is Gu or Lys
<220>
<221> Variant
<222> (20)..(20)
<223> xaa in position 20 is Ala, Asn or Ser
<220>
<221> Variant
<222> (21)..(21)
<223> xaa in position 21 is Lys or Arg
<220>
<221> Variant
<222> (22)..(22)
<223> xaa in position 22 is Ala, Gly or Ser
<220>
<221> Variant
<222> (23)..(23)
<223> xaa in position 23 is any amino acid
<220>
<221> Variant
<222> (24)..(24)
<223> xaa in position 24 is Gu or Gly
<220>
<221> Variant
<222> (25)..(29)
<223> xaa in position 25 to 29 is any amino acid
<220>
<221> Variant
<222> (30)..(30)
<223> xaa in position 30 is Asn, Pro or Gn
<220>
<221> Variant
<222> (31)..(32)
<223> xaa in position 31 to 32 is Ala, Asp or Gu
<220>
<221> Variant
<222> (33)..(33)
<223> xaa in position 33 is Gly, Ser or Thr
<220>
<221> Variant
<222> (34)..(34)
<223> xaa in position 34 is Asn, Gn or Thr
<220>
<221> Variant
<222> (35)..(36)
<223> xaa in position 35 to 36 is any amino acid
<220>
<221> Variant
<222> (37)..(37)
<223> xaa in position 37 is Ala, Asp or Pro

<400> 120
Met xaa xaa Leu xaa Val Leu Phe xaa xaa Phe Tyr xaa xaa xaa Tyr
  1           5           10           15
xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa

```

20  
xaa xaa xaa xaa xaa  
35

<210> 121  
<211> 19  
<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (2)..(3)  
<223> xaa in position 2 to 3 is any amino acid  
<220>  
<221> Variant  
<222> (4)..(7)  
<223> xaa in position 4 to 7 is any or no amino acid  
<220>  
<221> Variant  
<222> (9)..(9)  
<223> xaa in position 9 is Phe, Leu or Val  
<220>  
<221> Variant  
<222> (10)..(10)  
<223> xaa in position 10 is any amino acid  
<220>  
<221> Variant  
<222> (13)..(13)  
<223> xaa in position 13 is Ala, Leu or Val  
<220>  
<221> Variant  
<222> (14)..(14)  
<223> xaa in position 14 is any amino acid  
<220>  
<221> Variant  
<222> (16)..(16)  
<223> xaa in position 16 is Thr or Val  
<220>  
<221> Variant  
<222> (17)..(17)  
<223> xaa in position 17 is Ala, Ile or Leu  
<220>  
<221> Variant  
<222> (18)..(18)  
<223> xaa in position 18 is Phe, Ile or Leu

<400> 121  
Lys xaa xaa xaa xaa xaa Met xaa xaa Tyr Asn xaa xaa Gln xaa  
1 5 10 15  
xaa xaa Asn

<210> 122  
<211> 879  
<212> DNA  
<213> *Ostreococcus tauri*

<220>  
<221> CDS  
<222> (1)..(879)

<400> 122  
atg tct gga ttg agg gct cct aac ttc ttg cat agg ttc tgg acc aag 48  
Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys  
1 5 10 15  
tgg gat tac gct atc tct aag gtg gtg ttc act tgc gct gat tct ttc 96  
Trp Asp Tyr Ala Ile Ser Lys Val Val Phe Thr Cys Ala Asp Ser Phe  
Seite 160

PF58307. txt

20 25 30  
 cag tgg gat atc gga cct gtt tct tct tct acc gct cat ttg cct gct 144  
 G n Trp Asp lle Gly Pro Val Ser Ser Ser Thr Ala His Leu Pro Ala  
 35  
 att gag tct cct act cct ttg gtg acc tct ttg ctg ttc tac ttg gtg 192  
 lle Gu Ser Pro Thr Pro Leu Val Thr Ser Leu Leu Phe Tyr Leu Val  
 50  
 act gtg ttc ttg tgg tac gga aga ttg acc aga tcc tcc gat aag aag 240  
 Thr Val Phe Leu Trp Tyr Gly Arg Leu Thr Arg Ser Ser Asp Lys Lys  
 65  
 atc aga gag cct acc tgg ttg agg aga ttc atc atc tgc cac aac gct 288  
 lle Arg Glu Pro Thr Trp Leu Arg Arg Phe lle lle Cys His Asn Ala  
 85  
 ttc ttg att gtg ctg tcc ttg tac atg tgt ttg gga tgc gtt gct caa 336  
 Phe Leu lle Val Leu Ser Leu Tyr Met Cys Leu Gly Cys Val Ala G n  
 100  
 gct tac caa aac gga tac acc ttg tgg gga aac gag ttc aag gct act 384  
 Ala Tyr G n Asn Gly Tyr Thr Leu Trp Gly Asn Gu Phe Lys Ala Thr  
 115  
 gag acc caa ttg gct ctg tac atc tac atc ttc tac gtg tcc aag atc 432  
 Gu Thr G n Leu Ala Leu Tyr lle Tyr lle Phe Tyr Val Ser Lys lle  
 130  
 tac gag ttc gtg gat acc tac atc atg ctg ctg aag aac ctg agg 480  
 Tyr Gu Phe Val Asp Thr Tyr lle Met Leu Leu Lys Asn Asn Leu Arg  
 145  
 caa gtg tct ttc ttg cac atc tac cac cac tct acc atc tct ttc atc 528  
 G n Val Ser Phe Leu His lle Tyr His His Ser Thr lle Ser Phe lle  
 165  
 tgg tgg atc atc gct aga aga gca cct gga gga gat gct tat ttc tcc 576  
 Trp Trp lle lle Ala Arg Arg Ala Pro Gly Gly Asp Ala Tyr Phe Ser  
 180  
 gct gct ctg aac tct tgg gtt cat gtg tgc atg tac act tac tac ctg 624  
 Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr Tyr Leu  
 195  
 ctg tct acc ttg att gga aag gaa gat cct aag agg tct aac tac ctg 672  
 Leu Ser Thr Leu lle Gly Lys Gu Asp Pro Lys Arg Ser Asn Tyr Leu  
 210  
 tgg tgg gga agg cat ttg acc caa atg caa atg ctg cag ttc ttc ttc 720  
 Trp Trp Gly Arg His Leu Thr G n Met G n Met Leu G n Phe Phe Phe  
 225  
 aac gtg ctg caa gct ctt tat tgc gct tcc ttc tcc act tac cct aag 768  
 Asn Val Leu G n Ala Leu Tyr Cys Ala Ser Phe Ser Thr Tyr Pro aag  
 245  
 ttc ctg tcc aag atc ttg ctg gtg tac atg atg tct ttg ctg gga ctt 816  
 Phe Leu Ser Lys lle Leu Leu Val Tyr Met Met Ser Leu Leu Gly Leu  
 260  
 ttc gga cac ttc tac tac tct aag cac atc gct gct gct aag ttg caa 864  
 Phe Gly His Phe Tyr Tyr Ser Lys His lle Ala Ala Ala Lys Leu G n  
 275  
 aag aag cag cag tga 879  
 Lys Lys G n G n  
 290

<210> 123  
 <211> 292  
 <212> PRT  
 <213> *Ostreococcus tauri*

<400> 123  
 Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys  
 1 5 10 15  
 Trp Asp Tyr Ala lle Ser Lys Val Val Phe Thr Cys Ala Asp Ser Phe  
 20 25 30  
 G n Trp Asp lle Gly Pro Val Ser Ser Thr Ala His Leu Pro Ala  
 35 40 45  
 lle Gu Ser Pro Thr Pro Leu Val Thr Ser Leu Leu Phe Tyr Leu Val  
 50 55 60  
 Thr Val Phe Leu Trp Tyr Gly Arg Leu Thr Arg Ser Ser Asp Lys Lys

PF58307. txt

65 Ile Arg Glu Pro Thr Trp Leu Arg Arg Phe Ile Ile Cys His Asn Ala  
 85 Phe Leu Ile Val Leu Ser Leu Tyr Met Cys Leu Gly Cys Val Ala Gn  
 100 Ala Tyr Gn Asn Gly Tyr Thr Leu Trp Gly Asn Gu Phe Lys Ala Thr  
 115 Gu Thr Gn Leu Ala Leu Tyr Ile Tyr Ile Phe Tyr Val Ser Lys Ile  
 130 Tyr Gu Phe Val Asp Thr Tyr Ile Met Leu Leu Lys Asn Asn Leu Arg  
 145 Gn Val Ser Phe Leu His Ile Tyr His His Ser Thr Ile Ser Phe Ile  
 165 Trp Trp Ile Ile Ala Arg Arg Ala Pro Gly Gly Asp Ala Tyr Phe Ser  
 180 Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr Tyr Leu  
 195 Leu Ser Thr Leu Ile Gly Lys Gu Asp Pro Lys Arg Ser Asn Tyr Leu  
 210 Trp Trp Gly Arg His Leu Thr Gn Met Gn Met Leu Gn Phe Phe Phe  
 225 Asn Val Leu Gn Ala Leu Tyr Cys Ala Ser Phe Ser Thr Tyr Pro Lys  
 245 Phe Leu Ser Lys Ile Leu Leu Val Tyr Met Met Ser Leu Leu Gly Leu  
 260 Phe Gly His Phe Tyr Tyr Ser Lys His Ile Ala Ala Ala Lys Leu Gn  
 275 Lys Lys Gn Gn  
 290

<210> 124  
 <211> 1047  
 <212> DNA  
 <213> Marchantia polymorpha

<220>  
 <221> CDS  
 <222> (1)..(1047)

<400> 124  
 atg gcg acg aag agc ggg agc gga ttg ctg gag tgg ata gca gta gcg 48  
 Met Ala Thr Lys Ser Gly Ser Gly Leu Leu Gu Trp Ile Ala Val Ala  
 1 5  
 gcg aag atg aag caa gct cgg agc agc ccc gag ggt gag atc gtg ggt 96  
 Ala Lys Met Lys Gn Ala Arg Ser Ser Pro Gu Gly Gu Ile Val Gly  
 20 25 30  
 ggg aat agg atg ggc tct gga aac gga gct gag tgg acc acg agt ctg 144  
 Gly Asn Arg Met Gly Ser Gly Asn Gly Ala Gu Trp Thr Thr Ser Leu  
 35 40 45  
 att cat gca ttt ttg aat gcc acg aat ggg aag agc ggc ggt gct tctg 192  
 Ile His Ala Phe Leu Asn Ala Thr Asn Gly Lys Ser Gly Gly Ala Ser  
 50 55 60  
 aaa gtg agg cct ctg gag gag aga atc ggg gag gcg gtg ttc aga gtt 240  
 Lys Val Arg Pro Leu Gu Gu Arg Ile Gly Gu Ala Val Phe Arg Val  
 65 70 75 80  
 ctt gaa gat gtc gtg ggc gtg gat att agg aag ccg aat cct gtc acg 288  
 Leu Gu Asp Val Val Gly Val Asp Ile Arg Lys Pro Asn Pro Val Thr  
 85 90 95  
 aag gac ctt ccg atg gtc gag agt ccc gtg ccc gtg ttg gcc tgc att 336  
 Lys Asp Leu Pro Met Val Gu Ser Pro Val Pro Val Leu Ala Cys Ile  
 100 105 110  
 tct ctg tac ttg ctg gtg gtg tgg ctt tgg tct tct cac att aag gcg 384  
 Ser Leu Tyr Leu Leu Val Val Trp Leu Trp Ser Ser His Ile Lys Ala  
 115 120 125  
 tct ggc caa aag ccc agg aag gag gac ccg ctg gcc ctg cgt tgc ctt 432  
 Ser Gly Gn Lys Pro Arg Lys Gu Asp Pro Leu Ala Leu Arg Cys Leu  
 130 135 140  
 gtg att gcc cac aat ctg ttc ctg tgt tgc ttg agc ttg ttc atg tgc 480



PF58307. txt

Val 145	Ile	Ala	His	Asn	Leu 150	Phe	Leu	Cys	Cys	Leu 155	Ser	Leu	Phe	Met	Cys 160	
gtc	ggt	ctc	att	gcc	gca	gct	cga	cat	tac	ggg	tat	agt	gta	tgg	ggg	528
Val	Gly	Leu	Ile	Ala 165	Ala	Ala	Arg	His	Tyr 170	Gly	Tyr	Ser	Val	Trp 175	Gly	
aac	tac	tac	aga	gaa	aga	gaa	ccc	gca	atg	aat	tgt	ctc	att	tac	gtg	576
Asn	Tyr	Tyr	Arg	Glu 180	Arg	Glu	Pro	Ala 185	Met	Asn	Leu	Leu	Ile	Tyr	Val	
ttc	tac	atg	tcg	aag	ctg	tac	gaa	ttt	atg	gac	acg	gcc	att	atg	tta	624
Phe	Tyr	Met 195	Ser	Lys	Leu	Tyr	Glu 200	Phe	Met	Asp	Thr	Ala	Ile	Met	Leu	
ttc	aga	aga	aat	ctg	cga	caa	gtc	acg	tac	tgt	cat	gta	tat	cac	cac	672
Phe	Arg 210	Arg	Asn	Leu	Arg	Gln 215	Val	Thr	Tyr	Leu	His 220	Val	Tyr	His	His	
gca	agc	atc	gca	atg	att	tgg	tgg	ata	att	tgc	tat	cgg	ttt	cca	gga	720
Ala	Ser	Ile	Ala	Met 230	Ile	Trp	Trp	Ile	Ile	Cys 235	Tyr	Arg	Phe	Pro	Gly	
gct	gat	tcg	tat	ttc	tcc	gca	gca	ttc	aat	tcc	tgt	atc	cat	gta	gcg	768
Ala	Asp	Ser	Tyr	Phe 245	Ser	Ala	Ala	Phe	Asn 250	Ser	Cys	Ile	His	Val 255	Ala	
atg	tac	ctg	tat	tat	cta	ctc	gcg	gca	acc	gtc	gcc	aga	gac	gaa	aag	816
Met	Tyr	Leu	Tyr 260	Tyr	Leu	Leu	Ala	Ala 265	Thr	Val	Ala	Arg	Asp 270	Glu	Lys	
cgg	aga	cgc	aaa	tat	ctc	ttc	tgg	gga	aag	tat	ctg	acc	atc	ata	caa	864
Arg	Arg	Arg	Lys 275	Tyr	Leu	Phe	Trp 280	Gly	Lys	Tyr	Leu	Thr 285	Ile	Ile	Gln	
atg	ctt	cag	ttt	ttg	tcc	ttc	att	ggg	cag	gcg	att	tat	gca	atg	tgg	912
Met	Leu 290	Gln	Phe	Leu	Ser	Phe 295	Ile	Gly	Gln	Ala	Ile 300	Tyr	Ala	Met	Trp	
aag	ttt	gaa	tac	tat	ccc	aag	ggc	ttt	ggc	agg	atg	ttg	ttc	ttt	tac	960
Lys	Phe	Glu	Tyr	Tyr	Pro 310	Lys	Gly	Phe	Gly	Arg 315	Met	Leu	Phe	Phe	Tyr 320	
tct	gta	tca	ttg	ttg	gca	ttt	ttc	ggc	aac	ttc	ttt	gtc	aaa	aag	tat	1008
Ser	Val	Ser	Leu	Leu 325	Ala	Phe	Phe	Gly	Asn 330	Phe	Phe	Val	Lys	Lys 335	Tyr	
tcg	aac	gct	tca	cag	cct	aag	aca	gtt	aaa	gtg	gag	tga				1047
Ser	Asn	Ala	Ser 340	Gln	Pro	Lys	Thr	Val 345	Lys	Val	Glu					

<210> 125  
 <211> 348  
 <212> PRT  
 <213> Marchantia polymorpha

<400> 125

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

PF58307. txt

180 185 190  
Phe Tyr Met Ser Lys Leu Tyr Glu Phe Met Asp Thr Ala Ile Met Leu  
195 200 205  
Phe Arg Arg Asn Leu Arg Gn Val Thr Tyr Leu His Val Tyr His His  
210 215 220  
Ala Ser Ile Ala Met Ile Trp Trp Ile Ile Cys Tyr Arg Phe Pro Gly  
225 230 235 240  
Ala Asp Ser Tyr Phe Ser Ala Ala Phe Asn Ser Cys Ile His Val Ala  
245 250 255  
Met Tyr Leu Tyr Tyr Leu Leu Ala Ala Thr Val Ala Arg Asp Glu Lys  
260 265 270  
Arg Arg Arg Lys Tyr Leu Phe Trp Gly Lys Tyr Leu Thr Ile Ile Gn  
275 280 285  
Met Leu Gn Phe Leu Ser Phe Ile Gly Gn Ala Ile Tyr Ala Met Trp  
290 295 300  
Lys Phe Glu Tyr Tyr Pro Lys Gly Phe Gly Arg Met Leu Phe Phe Tyr  
305 310 315 320  
Ser Val Ser Leu Leu Ala Phe Phe Gly Asn Phe Phe Val Lys Lys Tyr  
325 330 335  
Ser Asn Ala Ser Gn Pro Lys Thr Val Lys Val Glu  
340 345

<210> 126  
<211> 831  
<212> DNA  
<213> Thr aust ochyt r i u m sp

<220>  
<221> CDS  
<222> (1).. (831)

<400> 126  
at g gat gt c gt c gag cag caa tgg cgc cgc t t c gt g gac gcc gt g gac 48  
Met Asp Val Val Glu Gn Gn Trp Arg Arg Phe Val Asp Ala Val Asp  
1 5 10 15  
aac gga at c gt g gag t t c at g gag cat gag gag ccc aac aag ct g aac 96  
Asn Gly Ile Val Glu Phe Met Glu His Glu Glu Pro Asn Lys Leu Asn  
20 25 30  
gag ggc aag ct c t cc acc t cg acc gag gag at g at g gcg ct t at c gt c 144  
Glu Gly Lys Leu Ser Thr Ser Thr Glu Glu Met Met Ala Leu Ile Val  
35 40 45  
ggc t ac ct g gcg t t c gt g gt c ct c ggg t cc gcc t t c at g aag gcc t t t 192  
Gly Tyr Leu Ala Phe Val Val Leu Gly Ser Ala Phe Met Lys Ala Phe  
50 55 60  
gt c gat aag cct t t c gag ct c aag t t c ct c aag ct c gt g cac aac at c 240  
Val Asp Lys Pro Phe Glu Leu Lys Phe Leu Lys Leu Val His Asn Ile  
65 70 75 80  
t t c ct c acc ggt ct g t cc at g t ac at g gcc acc gag t gc gcg cgc cag 288  
Phe Leu Thr Gly Leu Ser Met Tyr Met Ala Thr Glu Cys Ala Arg Gn  
85 90 95  
gca t ac ct c ggc ggc t ac aag ct c t t t ggc aac ccg at g gag aag ggc 336  
Ala Tyr Leu Gly Gly Tyr Lys Leu Phe Gly Asn Pro Met Glu Lys Gly  
100 105 110  
acc gag t cg cac gcc ccg ggc at g gcc aac at c at c t ac at c t t c t ac 384  
Thr Glu Ser His Ala Pro Gly Met Ala Asn Ile Ile Tyr Ile Phe Tyr  
115 120 125  
gt g agc aag t t c ct c gaa t t c ct c gac acc gt c t t c at g at c ct c ggc 432  
Val Ser Lys Phe Leu Glu Phe Leu Asp Thr Val Phe Met Ile Leu Gly  
130 135 140  
aag aag t gg aag cag ct c agc t t t ct c cac gt c t ac cac cac gcg agc 480  
Lys Lys Trp Lys Gn Leu Ser Phe Leu His Val Tyr His His Ala Ser  
145 150 155 160  
at c agc t t c at c t gg ggc at c at c gcc cgc t t c gcg ccc ggt ggc gac 528  
Ile Ser Phe Ile Trp Gly Ile Ile Ala Arg Phe Ala Pro Gly Gly Asp  
165 170 175  
gcc t ac t t c t ct acc at c ct c aac agc agc gt g cat gt c gt g ct c t ac 576  
Ala Tyr Phe Ser Thr Ile Leu Asn Ser Ser Val His Val Val Leu Tyr  
180 185 190

PF58307. txt

ggc t ac t ac gcc t cg acc acc ct c ggc t ac acc t t c at g cg c ccg ct g 624  
 G y Tyr Tyr Al a Ser Thr Thr Leu G y Tyr Thr Phe Met Arg Pro Leu  
 195 200 205  
 cg c ccg t ac att acc acc att ct c acg cag t t c at g gcc at g gt c 672  
 Arg Pro Tyr Ile Thr Thr Ile G n Leu Thr G n Phe Met Al a Met Val  
 210 215 220  
 gt c cag t cc gt c t at gac t ac t ac aac ccc t gc gac t ac ccg cag ccc 720  
 Val G n Ser Val Tyr Asp Tyr Tyr Asn Pro Cys Asp Tyr Pro G n Pro  
 225 230 240  
 ct c gt c aag ct g ct c t t c t gg t ac at g ct c acc at g ct c ggc ct c t t c 768  
 Leu Val Lys Leu Leu Phe Trp Tyr Met Leu Thr Met Leu G y Leu Phe  
 245 250 255  
 ggc aac t t c t t c gt g cag cag t ac ct c aag ccc aag gcg ccc aag aag 816  
 G y Asn Phe Phe Val G n G n Tyr Leu Lys Pro Lys Al a Pro Lys Lys  
 260 265 270  
 cag aag acc at c t aa 831  
 G n Lys Thr Ile  
 275

<210> 127  
 <211> 276  
 <212> PRT  
 <213> Thr aust ochyt ri um sp

<400> 127  
 Met Asp Val Val G u G n G n Trp Arg Arg Phe Val Asp Al a Val Asp  
 1 5 10 15  
 Asn G y Ile Val G u Phe Met G u His G u G u Pro Asn Lys Leu Asn  
 20 25 30  
 G u G y Lys Leu Ser Thr Ser Thr G u G u Met Met Al a Leu Ile Val  
 35 40 45  
 G y Tyr Leu Al a Phe Val Val Leu G y Ser Al a Phe Met Lys Al a Phe  
 50 55 60  
 Val Asp Lys Pro Phe G u Leu Lys Phe Leu Lys Leu Val His Asn Ile  
 65 70 75 80  
 Phe Leu Thr G y Leu Ser Met Tyr Met Al a Thr G u Cys Al a Arg G n  
 85 90 95  
 Al a Tyr Leu G y G y Tyr Lys Leu Phe G y Asn Pro Met G u Lys G y  
 100 105 110  
 Thr G u Ser His Al a Pro G y Met Al a Asn Ile Ile Tyr Ile Phe Tyr  
 115 120 125  
 Val Ser Lys Phe Leu G u Phe Leu Asp Thr Val Phe Met Ile Leu G y  
 130 135 140  
 Lys Lys Trp Lys G n Leu Ser Phe Leu His Val Tyr His His Al a Ser  
 145 150 155 160  
 Ile Ser Phe Ile Trp G y Ile Ile Al a Arg Phe Al a Pro G y G y Asp  
 165 170 175  
 Al a Tyr Phe Ser Thr Ile Leu Asn Ser Ser Val His Val Val Leu Tyr  
 180 185 190  
 G y Tyr Tyr Al a Ser Thr Thr Leu G y Tyr Thr Phe Met Arg Pro Leu  
 195 200 205  
 Arg Pro Tyr Ile Thr Thr Ile G n Leu Thr G n Phe Met Al a Met Val  
 210 215 220  
 Val G n Ser Val Tyr Asp Tyr Tyr Asn Pro Cys Asp Tyr Pro G n Pro  
 225 230 235 240  
 Leu Val Lys Leu Leu Phe Trp Tyr Met Leu Thr Met Leu G y Leu Phe  
 245 250 255  
 G y Asn Phe Phe Val G n G n Tyr Leu Lys Pro Lys Al a Pro Lys Lys  
 260 265 270  
 G n Lys Thr Ile  
 275

<210> 128  
 <211> 1146  
 <212> DNA  
 <213> Lei shmani a maj or

<220>

<221> CDS

<222> (1)..(1146)

<400> 128

at g	gt c	t ct	ct g	gag	cag	gcc	gaa	cag	at c	gcg	gcc	gcg	at c	gag	gt a	48
Met	Val	Ser	Leu	Gl u	Gl n	Al a	Gl u	Gl n	I le	Al a	Al a	Al a	I le	Gl u	Val	
1			5						10					15		
cct	gac	t gg	gt c	tt g	aca	aag	t ct	gcg	gcg	ct g	gt g	t ac	agc	t gc	tt c	96
Pro	Asp	Trp	Val	Leu	Thr	Lys	Ser	Al a	Al a	Leu	Val	Tyr	Ser	Cys	Phe	
			20					25					30			
ggc	t cc	gcc	gcc	aat	gct	tt t	gaa	agc	agt	at c	aag	at c	aac	tt c	ccg	144
Gl y	Ser	Al a	Al a	Asn	Al a	Phe	Gl u	Ser	Ser	I le	Lys	I le	Asn	Phe	Pro	
		35				40						45				
gcg	cag	cat	gcg	tt c	gt g	gag	gcg	t gg	at g	cg c	gcg	cg c	t cc	cac	cct	192
Al a	Gl n	Hi s	Al a	Phe	Val	Gl u	Al a	Trp	Met	Arg	Al a	Arg	Ser	Hi s	Pro	
	50					55					60					
tt t	gcg	gag	cg c	ct g	ccg	t ac	ct g	aat	ccg	t gg	cac	gt t	at c	gcc	t cg	240
Phe	Al a	Gl u	Arg	Leu	Pro	Tyr	Leu	Asn	Pro	Trp	Hi s	Val	I le	Al a	Ser	
65				70					75					80		
at a	ct g	gcc	t ac	ct c	t cc	tt g	att	gt c	acc	tt g	cg c	ct g	tt g	cat	cgt	288
I le	Leu	Al a	Tyr	Leu	Ser	Leu	I le	Val	Thr	Leu	Arg	Leu	Leu	Hi s	Arg	
				85					90					95		
gt a	ct c	ggt	aag	tt c	t cg	t gc	cg c	act	ct c	gga	t tg	gt g	cac	aac	ct c	336
Val	Leu	Gl y	Lys	Phe	Ser	Cys	Arg	Thr	Leu	Gl y	Leu	Val	Hi s	Asn	Leu	
			100					105					110			
ggt	ct c	cat	ct t	ct c	t cg	tt g	t ac	at g	agc	ct t	ggt	ct c	at g	at c	agc	384
Gl y	Leu	Hi s	Leu	Leu	Ser	Leu	Tyr	Met	Ser	Leu	Gl y	Leu	Met	I le	Ser	
		115						120				125				
gcg	cg c	gcc	gcg	ggg	t ac	t cg	ct c	t gg	aac	aac	gcg	gt c	ggc	acc	t cc	432
Al a	Arg	Al a	Al a	Gl y	Tyr	Ser	Leu	Trp	Asn	Asn	Al a	Val	Gl y	Thr	Ser	
	130					135					140					
ccg	gct	gag	t gg	cg c	att	gcg	aag	ct g	at c	t gg	ct c	tt c	t at	gt c	t cg	480
Pro	Al a	Gl u	Trp	Arg	I le	Al a	Lys	Leu	I le	Trp	Leu	Phe	Tyr	Val	Ser	
145					150					155					160	
aag	gt g	gt g	gaa	t gg	gt g	gac	acg	gt a	att	at g	t t a	t t a	aag	cag	aac	528
Lys	Val	Val	Gl u	Trp	Val	Asp	Thr	Val	I le	Met	Leu	Leu	Lys	Gl n	Asn	
				165					170					175		
t ac	cac	cag	gt c	acc	tt c	ct g	cac	gt g	t at	cac	cac	acg	acg	gt t	tt t	576
Tyr	Hi s	Gl n	Val	Thr	Phe	Leu	Hi s	Val	Tyr	Hi s	Hi s	Thr	Thr	Val	Phe	
				180					185					190		
gt g	ct g	t gg	t gg	ct g	gcg	tt g	ct g	gt c	gct	cct	ggc	ggc	gag	t cg	t ac	624
Val	Leu	Trp	Trp	Leu	Al a	Leu	Leu	Val	Al a	Pro	Gl y	Gl y	Gl u	Ser	Tyr	
			195					200				205				
t ac	agc	gcc	at g	gt g	aac	t ct	ggc	gt c	cac	gt t	t t c	at g	t ac	ggg	t ac	672
Tyr	Ser	Al a	Met	Val	Asn	Ser	Gl y	Val	Hi s	Val	Phe	Met	Tyr	Gl y	Tyr	
		210				215						220				
t ac	tt t	ct c	acg	ct g	ct c	tt c	cca	t cc	ggc	at c	gt g	cg c	gac	gt c	tt g	720
Tyr	Phe	Leu	Thr	Leu	Leu	Phe	Pro	Ser	Gl y	I le	Val	Arg	Asp	Val	Leu	
225					230					235					240	
agc	aag	tt c	aag	tt t	gcc	att	acg	aag	ggc	cag	at g	t gg	cag	tt c	gt c	768
Ser	Lys	Phe	Lys	Phe	Al a	I le	Thr	Lys	Gl y	Gl n	Met	Trp	Gl n	Phe	Val	
				245					250					255		
tt c	aac	t gc	ct a	cag	t cc	gcg	t ac	gac	ct c	gt g	t gg	gt g	ccg	cg g	gaa	816
Phe	Asn	Cys	Leu	Gl n	Ser	Al a	Tyr	Asp	Leu	Val	Trp	Val	Pro	Arg	Gl u	
			260					265								
gag	ct c	aag	t ac	agc	gcg	gag	ct g	ct g	cag	at c	ct c	t t c	t gg	t ac	at g	864
Gl u	Leu	Lys	Tyr	Ser	Al a	Gl u	Leu	Leu	Gl n	I le	Leu	Phe	Trp	Tyr	Met	
			275					280					285			
at c	t cc	ct c	tt g	gcg	ct c	tt t	ggc	aac	t t c	tt g	gt g	aag	aac	aag	aag	912
I le	Ser	Leu	Leu	Al a	Leu	Phe	Gl y	Asn	Phe	Leu	Val	Lys	Asn	Lys	Lys	
		290				295						300				
tt c	t cg	cac	cg c	cg c	t gc	gt t	gat	gcc	gcg	act	gct	t cg	ggc	gcg	aag	960
Phe	Ser	Hi s	Arg	Arg	Cys	Val	Asp	Al a	Al a	Thr	Al a	Ser	Gl y	Al a	Lys	
305					310					315					320	
gag	gac	acg	gcg	gcg	agg	t cc	cac	ggc	gac	cg c	acc	cac	aga	acc	cgt	1008
Gl u	Asp	Thr	Al a	Al a	Arg	Ser	Hi s	Gl y	Asp	Arg	Thr	Hi s	Arg	Thr	Arg	
				325					330					335		
gt g	aag	gct	ggc	at g	acc	aac	at g	caa	ct g	gag	agg	ct g	aag	aat	gag	1056

PF58307.txt

Val	Lys	Ala	Gly	Met	Thr	Asn	Met	Gln	Leu	Glu	Arg	Leu	Lys	Asn	Glu		
			340					345					350				
aag	tcc	acg	gag	atg	aag	ctg	ctg	atg	cgc	aag	aac	ggc	aac	ggc	aac		1104
Lys	Ser	Thr	Glu	Met	Lys	Leu	Leu	Met	Arg	Lys	Asn	Gly	Asn	Gly	Asn		
			355				360					365					
gga	caa	aaa	gcg	tcg	ctc	cag	gcc	atg	gca	ggc	agt	cga	tga				1146
Gly	Gln	Lys	Ala	Ser	Leu	Gln	Ala	Met	Ala	Gly	Ser	Arg					
			370			375					380						

<210> 129  
 <211> 381  
 <212> PRT  
 <213> Leishmani amajor

<400> 129

Met	Val	Ser	Leu	Glu	Gln	Ala	Glu	Gln	Ile	Ala	Ala	Ala	Ile	Glu	Val		
1			5					10					15				
Pro	Asp	Trp	Val	Leu	Thr	Lys	Ser	Ala	Ala	Leu	Val	Tyr	Ser	Cys	Phe		
			20					25					30				
Gly	Ser	Ala	Ala	Asn	Ala	Phe	Glu	Ser	Ser	Ile	Lys	Ile	Asn	Phe	Pro		
			35			40						45					
Ala	Gln	His	Ala	Phe	Val	Glu	Ala	Trp	Met	Arg	Ala	Arg	Ser	His	Pro		
			50			55					60						
Phe	Ala	Glu	Arg	Leu	Pro	Tyr	Leu	Asn	Pro	Trp	His	Val	Ile	Ala	Ser		
65					70					75					80		
Ile	Leu	Ala	Tyr	Leu	Ser	Leu	Ile	Val	Thr	Leu	Arg	Leu	Leu	His	Arg		
			85						90					95			
Val	Leu	Gly	Lys	Phe	Ser	Cys	Arg	Thr	Leu	Gly	Leu	Val	His	Asn	Leu		
			100					105					110				
Gly	Leu	His	Leu	Leu	Ser	Leu	Tyr	Met	Ser	Leu	Gly	Leu	Met	Ile	Ser		
			115				120					125					
Ala	Arg	Ala	Ala	Gly	Tyr	Ser	Leu	Trp	Asn	Asn	Ala	Val	Gly	Thr	Ser		
			130			135					140						
Pro	Ala	Glu	Trp	Arg	Ile	Ala	Lys	Leu	Ile	Trp	Leu	Phe	Tyr	Val	Ser		
145					150					155					160		
Lys	Val	Val	Glu	Trp	Val	Asp	Thr	Val	Ile	Met	Leu	Leu	Lys	Gln	Asn		
			165						170					175			
Tyr	His	Gln	Val	Thr	Phe	Leu	His	Val	Tyr	His	His	Thr	Thr	Val	Phe		
			180					185						190			
Val	Leu	Trp	Trp	Leu	Ala	Leu	Leu	Val	Ala	Pro	Gly	Gly	Glu	Ser	Tyr		
			195				200					205					
Tyr	Ser	Ala	Met	Val	Asn	Ser	Gly	Val	His	Val	Phe	Met	Tyr	Gly	Tyr		
			210			215					220						
Tyr	Phe	Leu	Thr	Leu	Leu	Phe	Pro	Ser	Gly	Ile	Val	Arg	Asp	Val	Leu		
225					230					235					240		
Ser	Lys	Phe	Lys	Phe	Ala	Ile	Thr	Lys	Gly	Gln	Met	Trp	Gln	Phe	Val		
			245						250					255			
Phe	Asn	Cys	Leu	Gln	Ser	Ala	Tyr	Asp	Leu	Val	Trp	Val	Pro	Arg	Glu		
			260					265					270				
Glu	Leu	Lys	Tyr	Ser	Ala	Glu	Leu	Gln	Ile	Leu	Phe	Trp	Tyr	Met			
			275				280				285						
Ile	Ser	Leu	Leu	Ala	Leu	Phe	Gly	Asn	Phe	Leu	Val	Lys	Asn	Lys	Lys		
			290			295					300						
Phe	Ser	His	Arg	Arg	Cys	Val	Asp	Ala	Ala	Thr	Ala	Ser	Gly	Ala	Lys		
305					310					315					320		
Glu	Asp	Thr	Ala	Ala	Arg	Ser	His	Gly	Asp	Arg	Thr	His	Arg	Thr	Arg		
			325					330						335			
Val	Lys	Ala	Gly	Met	Thr	Asn	Met	Gln	Leu	Glu	Arg	Leu	Lys	Asn	Glu		
			340					345					350				
Lys	Ser	Thr	Glu	Met	Lys	Leu	Leu	Met	Arg	Lys	Asn	Gly	Asn	Gly	Asn		
			355				360					365					
Gly	Gln	Lys	Ala	Ser	Leu	Gln	Ala	Met	Ala	Gly	Ser	Arg					
			370			375					380						

<210> 130  
 <211> 879  
 <212> DNA  
 <213> Ostreococcus tauri

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1).. (879)

&lt;400&gt; 130

atg agt ggc tta cgt gca ccc aac ttt tta cac aga ttc tgg aca aag	48
Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys	
1 5 10 15	
tgg gac tac gcg att tcc aaa gtc gtc ttc acg tgt gcc gac agt ttt	96
Trp Asp Tyr Ala Ile Ser Lys Val Val Phe Thr Cys Ala Asp Ser Phe	
20 25 30	
cag tgg gac atc ggg cca gtg agt tcg agt acg gcg cat tta ccc gcc	144
Gln Trp Asp Ile Gly Pro Val Ser Ser Thr Ala His Leu Pro Ala	
35 40 45	
att gaa tcc cct acc cca ctg gtg act agc ctg ttt ttc tac tta gtc	192
Ile Glu Ser Pro Thr Pro Leu Val Thr Ser Leu Leu Phe Tyr Leu Val	
50 55 60	
aca gtt ttc ttg tgg tat ggt cgt tta acc agg agt tca gac aag aaa	240
Thr Val Phe Leu Trp Tyr Gly Arg Leu Thr Arg Ser Ser Asp Lys Lys	
65 70 75 80	
att aga gag cct acg tgg tta aga aga ttc ata ata tgt cat aat gcg	288
Ile Arg Glu Pro Thr Trp Leu Arg Arg Phe Ile Ile Cys His Asn Ala	
85 90 95	
ttc ttg ata gtc ctg agt ctt tac atg tgc ctt ggt tgt gtg gcc caa	336
Phe Leu Ile Val Leu Ser Leu Tyr Met Cys Leu Gly Cys Val Ala Gln	
100 105 110	
gcg tat cag aat gga tat act tta tgg ggt aat gaa ttc aag gcc acg	384
Ala Tyr Gln Asn Gly Tyr Thr Leu Trp Gly Asn Glu Phe Lys Ala Thr	
115 120 125	
gaa act cag ctt gct ctg tac att tac att ttt tac gta agt aaa ata	432
Glu Thr Gln Leu Ala Leu Tyr Ile Tyr Ile Phe Tyr Val Ser Lys Ile	
130 135 140 145	
tac gag ttt gta gat act tac att atg ctt ctg aag aat aac ttg cgg	480
Tyr Glu Phe Val Asp Thr Tyr Ile Met Leu Leu Lys Asn Asn Leu Arg	
150 155 160	
caa gta agt ttc ctg cac att tat cac cac agc acg att tcc ttt att	528
Gln Val Ser Phe Leu His Ile Tyr His His Ser Thr Ile Ser Phe Ile	
165 170 175	
tgg tgg atc att gct cgg agg gct cgg ggt ggt gat gct tac ttc agc	576
Trp Trp Ile Ile Ala Arg Arg Ala Pro Gly Gly Asp Ala Tyr Phe Ser	
180 185 190	
gcg gcc ttg aac tca tgg gta cac gtg tgc atg tac acc tat tat cta	624
Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr Tyr Leu	
195 200 205	
tta tca acc ctt att gga aaa gaa gat cct aag cgt tcc aac tac ctt	672
Leu Ser Thr Leu Ile Gly Lys Glu Asp Pro Lys Arg Ser Asn Tyr Leu	
210 215 220	
tgg tgg ggt cgc cac ctg acg caa atg cag atg ctt cag ttt ttc ttc	720
Trp Trp Gly Arg His Leu Thr Gln Met Gln Met Leu Gln Phe Phe Phe	
225 230 235 240	
aac gta ctt caa gcg ttg tac tgc gct tcg ttc tct acg tat ccc aag	768
Asn Val Leu Gln Ala Leu Tyr Cys Ala Ser Phe Ser Thr Tyr Pro Lys	
245 250 255	
ttt ttg tcc aaa att ctg ctg gtc tat atg atg agc ctt ctg ggc ttg	816
Phe Leu Ser Lys Ile Leu Leu Val Tyr Met Met Ser Leu Leu Gly Leu	
260 265 270 275	
ttt ggg cat ttc tac tat tcc aag cac ata gca gca gct aag ctg cag	864
Phe Gly His Phe Tyr Tyr Ser Lys His Ile Ala Ala Ala Lys Leu Gln	
280 285	
aaa aaa cag cag tga	879
Lys Lys Gln Gln	
290	

&lt;210&gt; 131

&lt;211&gt; 292

&lt;212&gt; PRT

<213> *Ostreococcus tauri*

<400> 131

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

<210> 132

<211> 873

<212> DNA

<213> Marchantia polymorpha

<220>

<221> CDS

<222> (1)..(873)

<400> 132

atg gag gcg tac gag atg gtg gat agt ttt gtg tcg aag acg gtt ttc 48  
 Met Glu Ala Tyr Glu Met Val Asp Ser Phe Val Ser Lys Thr Val Phe  
 1 5 10 15  
 gaa acg ctg cag aga ctg agg ggc gga gtc gtg ttg acg gaa tct gcg 96  
 Gu Thr Leu Gln Arg Leu Arg Gly Gy Val Val Leu Thr Glu Ser Ala  
 20 25 30  
 atc acc aaa ggt ttg cca tgc gtc gat agc ccg acg ccg atc gtt ctt 144  
 Ile Thr Lys Gly Leu Pro Cys Val Asp Ser Pro Thr Ile Val Leu  
 35 40 45  
 ggg ttg tcg tcc tac ttg aca ttc gtg ttt ctg ggg ctg att gtc atc 192  
 Gy Leu Ser Ser Tyr Leu Thr Phe Val Phe Leu Gly Leu Ile Val Ile  
 50 55 60  
 aag agc ctg gat ctt aag ccc cgc tcc aag gag ccc gcc att ttg aac 240  
 Lys Ser Leu Asp Leu Lys Pro Arg Ser Lys Glu Pro Ala Ile Leu Asn  
 65 70 75 80  
 ctg ttt gtg atc ttc cac aac ttc gtc tgc ttc gca ctg agt ctg tac 288  
 Leu Phe Val Ile Phe His Asn Phe Val Cys Phe Ala Leu Ser Leu Tyr



PF58307. txt

```

      85          90          95
at g t gc gt g gga at t gt c cgt caa gct at c ct c aac agg t ac t ct ct g      336
Met Cys Val Gly Ile Val Arg Gn Ala Ile Leu Asn Arg Tyr Ser Leu
      100
t gg ggc aat gcg t ac aat ccc aaa gaa gt t caa at g ggc cac ct g ct c      384
Trp Gly Asn Ala Tyr Asn Pro Lys Gu Val Gn Met Gly His Leu Leu
      115
t ac att ttc t ac at g t ca aag t ac at c gag ttt at g gac acg gt c at t      432
Tyr Ile Phe Tyr Met Ser Lys Tyr Ile Gu Phe Met Asp Thr Val Ile
      130
at g att ttg aag cgc aac acg cgc cag at c act gt g tt g cat gt g t ac      480
Met Ile Leu Lys Arg Asn Thr Arg Gn Ile Thr Val Leu His Val Tyr
      145
cac cac gca tcc at c tcc ttc at c tgg tgg at c at c gcc t ac cat gct      528
His His Ala Ser Ile Ser Phe Ile Trp Trp Ile Ile Ala Tyr His Ala
      165
cct ggc ggt gaa gct t at ttc tct gcc gca ttg aac tcc gga gt a cat      576
Pro Gly Gly Glu Ala Tyr Phe Ser Ala Ala Leu Asn Ser Gly Val His
      180
gt g ct c at g t ac ct c t ac t ac ctt ttg gca gca act ct g gga aag aac      624
Val Leu Met Tyr Leu Tyr Tyr Leu Leu Ala Ala Thr Leu Gly Lys Asn
      195
gag aaa gct cgc cgc aag t ac ct a tgg tgg gga aaa t ac ttg aca cag      672
Gu Lys Ala Arg Arg Lys Tyr Leu Trp Trp Gly Lys Tyr Leu Thr Gn
      210
ct g cag at g ttc cag ttt gt c ctt aac at g att cag gct t ac t ac gat      720
Leu Gn Met Phe Gn Phe Val Leu Asn Met Ile Gn Ala Tyr Tyr Asp
      225
att aag aac aac t cg cct t ac cca caa ttt ttg at c cag att ttg ttc      768
Ile Lys Asn Asn Ser Pro Tyr Pro Gn Phe Leu Ile Gn Ile Leu Phe
      245
t ac t ac at g at c t cg ctt tta gcg ct a ttt gga aac ttt t ac gt t cac      816
Tyr Tyr Met Ile Ser Leu Leu Ala Leu Phe Gly Asn Phe Tyr Val His
      260
aaa t ac gt a t ca gcg ccc gca aaa cct gcg aag at c aag agc aaa aag      864
Lys Tyr Val Ser Ala Pro Ala Lys Pro Ala Lys Ile Lys Ser Lys Lys
      275
gca gaa t aa
Ala Gu
      290

```

<210> 133

<211> 290

<212> PRT

<213> Marchantia polymorpha

<400> 133

```

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

```

Seite 170



PF58307. txt

165  
 Pro Gly Gly Glu Ala Tyr Phe Ser Ala Ala Leu Asn Ser Gly Val His  
 170  
 180  
 Val Leu Met Tyr Leu Tyr Tyr Leu Leu Ala Ala Thr Leu Gly Lys Asn  
 195  
 200  
 Gu Lys Ala Arg Arg Lys Tyr Leu Trp Trp Gly Lys Tyr Leu Thr Gn  
 210  
 215  
 Leu Gn Met Phe Gn Phe Val Leu Asn Met Ile Gn Ala Tyr Tyr Asp  
 225  
 230  
 Ile Lys Asn Asn Ser Pro Tyr Pro Gn Phe Leu Ile Gn Ile Leu Phe  
 245  
 250  
 Tyr Tyr Met Ile Ser Leu Leu Ala Leu Phe Gly Asn Phe Tyr Val His  
 260  
 265  
 Lys Tyr Val Ser Ala Pro Ala Lys Pro Ala Lys Ile Lys Ser Lys Lys  
 275  
 280  
 Ala Gu  
 290

<210> 134

<211> 873

<212> DNA

<213> Physcomitrella patens

<220>

<221> CDS

<222> (1).. (873)

<400> 134

at g	gag	gt c	gt g	gag	aga	t t c	t a c	ggt	gag	t t g	gat	ggg	aag	gt c	t c g	48
Met	Gu	Val	Val	Gu	Arg	Phe	Tyr	Gly	Gu	Leu	Asp	Gly	Lys	Val	Ser	
1				5				10						15		
cag	ggc	gt g	aat	gca	t t g	ct g	ggt	agt	t t t	ggg	gt g	gag	t t g	acg	gat	96
Gn	Gly	Val	Asn	Ala	Leu	Leu	Gly	Ser	Phe	Gly	Val	Gu	Leu	Thr	Asp	
			20					25					30			
acg	ccc	act	acc	aaa	ggc	t t g	ccc	ct c	gt t	gac	agt	ccc	aca	ccc	at c	144
Thr	Pro	Thr	Thr	Lys	Gly	Leu	Pro	Leu	Val	Asp	Ser	Pro	Thr	Pro	Ile	
			35				40					45				
gt c	ct c	ggt	gt t	t ct	gt a	t a c	t t g	act	att	gt c	att	gga	ggg	ct t	t t g	192
Val	Leu	Gly	Val	Ser	Val	Tyr	Leu	Thr	Ile	Val	Ile	Gly	Gly	Leu	Leu	
			50			55				60						
t gg	at a	aag	gcc	agg	gat	ct g	aaa	ccg	cg c	gcc	t c g	gag	cca	t t t	t t g	240
Trp	Ile	Lys	Ala	Arg	Asp	Leu	Lys	Pro	Arg	Ala	Ser	Gu	Pro	Phe	Leu	
65				70					75						80	
ct c	caa	gct	t t g	gt g	ct t	gt g	cac	aac	ct g	t t c	t gt	t t t	gcg	ct c	agt	288
Leu	Gn	Ala	Leu	Val	Leu	Val	His	Asn	Leu	Phe	Cys	Phe	Ala	Leu	Ser	
				85				90						95		
ct g	t at	at g	t gc	gt g	ggc	at c	gct	t at	cag	gct	att	acc	t gg	cg g	t a c	336
Leu	Tyr	Met	Cys	Val	Gly	Ile	Ala	Tyr	Gn	Ala	Ile	Thr	Trp	Arg	Tyr	
			100				105					110				
t ct	ct c	t gg	ggc	aat	gca	t a c	aat	cct	aaa	cat	aaa	gag	at g	gcg	at t	384
Ser	Leu	Trp	Gly	Asn	Ala	Tyr	Asn	Pro	Lys	His	Lys	Gu	Met	Ala	Ile	
			115				120					125				
ct g	gt a	t a c	t t g	t t c	t a c	at g	t ct	aag	t a c	gt g	gaa	t t c	at g	gat	acc	432
Leu	Val	Tyr	Leu	Phe	Tyr	Met	Ser	Lys	Tyr	Val	Gu	Phe	Met	Asp	Thr	
			130			135						140				
gt t	at c	at g	at a	ct g	aag	cg c	agc	acc	agg	caa	at a	agc	t t c	ct c	cac	480
Val	Ile	Met	Ile	Leu	Lys	Arg	Ser	Thr	Arg	Gn	Ile	Ser	Phe	Leu	His	
				145		150				155				160		
gt t	t at	cat	cat	t ct	t ca	att	t cc	ct c	att	t gg	t gg	gct	att	gct	cat	528
Val	Tyr	His	His	Ser	Ser	Ile	Ser	Leu	Ile	Trp	Trp	Ala	Ile	Ala	His	
				165				170						175		
cac	gct	cct	ggc	ggt	gaa	gca	t at	t gg	t ct	gcg	gct	ct g	aac	t ca	gga	576
His	Ala	Pro	Gly	Gly	Gu	Ala	Tyr	Trp	Ser	Ala	Ala	Leu	Asn	Ser	Gly	
			180					185					190			
gt g	cat	gt t	ct c	at g	t at	gcg	t at	t a c	t t c	t t g	gct	gcc	t gc	ct t	cga	624
Val	His	Val	Leu	Met	Tyr	Ala	Tyr	Tyr	Phe	Leu	Ala	Ala	Cys	Leu	Arg	
			195				200					205				
agt	agc	cca	aag	t t a	aaa	aat	aag	t a c	ct t	t t t	t gg	ggc	agg	t a c	t t g	672

PF58307.txt

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu  
 210 215 220  
 aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct tac 720  
 Thr Gn Phe Gn Met Phe Gn Phe Met Leu Asn Leu Val Gn Ala Tyr  
 225 230 240  
 tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att 768  
 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gn Trp Leu Ile Lys Ile  
 245 250 255  
 ttg ttc tac tac atg atc tcg ttg ctg ttt ttt ctg ggc aat ttt tac 816  
 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr  
 260 265 270  
 gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa 864  
 Val Gn Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gn Lys Gly Ala Lys  
 275 280 285  
 act gag tga 873  
 Thr Gu  
 290

<210> 135  
 <211> 290  
 <212> PRT  
 <213> Physcomitrella patens

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

<210> 136  
 <211> 957  
 <212> DNA  
 <213> Mbrtierella alpina

<220>  
 <221> CDS  
 <222> (1)..(957)

<400> 136  
 at g gag t c g att g c g c c a t t c c t c c c a t c a a a g at g c c g c a a gat ct g 48  
 Met Gl u Ser Ile Ala Pro Phe Leu Pro Ser Lys Met Pro Gl n Asp Leu  
 1 5 10 15  
 t t t at g gac c t t g c c acc g c t at c g g t gt c c g g gcc gcc gcg ccc t at gt c 96  
 Phe Met Asp Leu Ala Thr Ala Ile Gl y Val Arg Ala Ala Pro Tyr Val  
 20 25 30  
 gat c c t ct c gag gcc gcg ct g gt g gcc cag gcc gag aag t ac at c ccc 144  
 Asp Pro Leu Gl u Ala Ala Leu Val Ala Gl n Ala Gl u Lys Tyr Ile Pro  
 35 40 45  
 acg att gt c cat cac acg cgt ggg t t c ct g gt c gcg gt g gag t c g cct 192  
 Thr Ile Val His His Thr Arg Gl y Phe Leu Val Ala Val Gl u Ser Pro  
 50 55 60  
 t t g gcc cgt gag ct g ccg t t g at g aac ccg t t c cac gt g ct g t t g at c 240  
 Leu Ala Arg Gl u Leu Pro Leu Met Asn Pro Phe His Val Leu Leu Ile  
 65 70 75 80  
 gt g ct c gct t at t t g gt c acg gt c t t t gt g ggc at g cag at c at g aag 288  
 Val Leu Ala Tyr Leu Val Thr Val Phe Val Gl y Met Gl n Ile Met Lys  
 85 90 95  
 aac t t t gag c g g t t c gag gt c aag acg t t t t c g ct c ct g cac aac t t t 336  
 Asn Phe Gl u Arg Phe Gl u Val Lys Thr Phe Ser Leu Leu His Asn Phe  
 100 105 110  
 t gt ct g gt c t c g at c agc gcc t ac at g t gc ggt ggg at c ct g t ac gag 384  
 Cys Leu Val Ser Ile Ser Ala Tyr Met Cys Gl y Gl y Ile Leu Tyr Gl u  
 115 120 125  
 gct t at cag gcc aac t at gga ct g t t t gag aac gct gct gat cat acc 432  
 Ala Tyr Gl n Ala Asn Tyr Gl y Leu Phe Gl u Asn Ala Ala Asp His Thr  
 130 135 140  
 t t c aag ggt c t t c c t at g gcc aag at g at c t gg ct c t t c t ac t t c t cc 480  
 Phe Lys Gl y Leu Pro Met Ala Lys Met Ile Trp Leu Phe Tyr Phe Ser  
 145 150 155 160  
 aag at c at g gag t t t gt c gac acc at g at c at g gt c ct c aag aag aac 528  
 Lys Ile Met Gl u Phe Val Asp Thr Met Ile Met Val Leu Lys Lys Asn  
 165 170 175  
 aac cgc cag at c t cc t t c t t g cac gt t t ac cac cac agc t cc at c t t c 576  
 Asn Arg Gl n Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Phe  
 180 185 190  
 acc at c t gg t gg t t g gt c acc t t t gt t gca ccc aac ggt gaa gcc t ac 624  
 Thr Ile Trp Trp Leu Val Thr Phe Val Ala Pro Asn Gl y Gl u Ala Tyr  
 195 200 205  
 t t c t ct gct gcg t t g aac t c g t t c at c cat gt g at c at g t ac ggc t ac 672  
 Phe Ser Ala Ala Leu Asn Ser Phe Ile His Val Ile Met Tyr Gl y Tyr  
 210 215 220  
 t ac t t c t t g t c g gcc t t g ggc t t c aag cag gt g t c g t t c at c aag t t c 720  
 Tyr Phe Leu Ser Ala Leu Gl y Phe Lys Gl n Val Ser Phe Ile Lys Phe  
 225 230 235 240  
 t ac at c acg cgc t c g cag at g aca cag t t c t gc at g at g t c g gt c cag 768  
 Tyr Ile Thr Arg Ser Gl n Met Thr Gl n Phe Cys Met Met Ser Val Gl n  
 245 250 255  
 t ct t cc t gg gac at g t ac gcc at g aag gt c ct t ggc cgc ccc gga t ac 816  
 Ser Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gl y Arg Pro Gl y Tyr  
 260 265 270  
 ccc t t c t t c at c acg gct ct g ct t t gg t t c t ac at g t gg acc at g ct c 864  
 Pro Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp Thr Met Leu  
 275 280 285  
 ggt ct c t t c t ac aac t t t t ac aga aag aac gcc aag t t g gcc aag cag 912  
 Gl y Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu Ala Lys Gl n  
 290 295 300  
 gcc aag gcc gac gct gcc aag gag aag gca agg aag t t g cag t aa 957  
 Ala Lys Ala Asp Ala Lys Gl u Lys Ala Arg Lys Leu Gl n  
 305 310 315

<210> 137  
 <211> 318

&lt;212&gt; PRT

&lt;213&gt; Mbrtierella alpina

&lt;400&gt; 137

```

Met 1  Glu Ser Ile Ala Pro Phe Leu Pro Ser Lys Met Pro Gln Asp Leu
1      5      10      15
Phe 2  Met Asp Leu Ala Thr Ala Ile Gly Val Arg Ala Ala Pro Tyr Val
20     25     30
Asp 3  Pro Leu Glu Ala Ala Leu Val Ala Gln Ala Glu Lys Tyr Ile Pro
35     40     45
Thr 4  Ile Val His His Thr Arg Gly Phe Leu Val Ala Val Glu Ser Pro
50     55     60
Leu 5  Ala Arg Glu Leu Pro Leu Met Asn Pro Phe His Val Leu Leu Ile
65     70     75     80
Val 6  Leu Ala Tyr Leu Val Thr Val Phe Val Gly Met Gln Ile Met Lys
85     90     95
Asn 7  Phe Glu Arg Phe Glu Val Lys Thr Phe Ser Leu Leu His Asn Phe
100    105
Cys 8  Leu Val Ser Ile Ser Ala Tyr Met Cys Gly Gly Ile Leu Tyr Glu
115    120    125
Ala 9  Tyr Gln Ala Asn Tyr Gly Leu Phe Glu Asn Ala Ala Asp His Thr
130    135    140
Phe 10 Lys Gly Leu Pro Met Ala Lys Met Ile Trp Leu Phe Tyr Phe Ser
145    150    155    160
Lys 11 Ile Met Glu Phe Val Asp Thr Met Ile Met Val Leu Lys Lys Asn
165    170    175
Asn 12 Arg Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Phe
180    185    190
Thr 13 Ile Trp Trp Leu Val Thr Phe Val Ala Pro Asn Gly Glu Ala Tyr
195    200    205
Phe 14 Ser Ala Ala Leu Asn Ser Phe Ile His Val Ile Met Tyr Gly Tyr
210    215    220
Tyr 15 Phe Leu Ser Ala Leu Gly Phe Lys Gln Val Ser Phe Ile Lys Phe
225    230    235    240
Tyr 16 Ile Thr Arg Ser Gln Met Thr Gln Phe Cys Met Met Ser Val Gln
245    250    255
Ser 17 Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gly Arg Pro Gly Tyr
260    265    270
Pro 18 Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp Thr Met Leu
275    280    285
Gly 19 Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu Ala Lys Gln
290    295    300
Ala 20 Lys Ala Asp Ala Ala Lys Glu Lys Ala Arg Lys Leu Gln
305    310    315

```

&lt;210&gt; 138

&lt;211&gt; 242

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (3)..(5)

&lt;223&gt; xaa in position 3 to 5 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (6)..(6)

&lt;223&gt; xaa in position 6 is any or no amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (8)..(16)

&lt;223&gt; xaa in position 8 to 16 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (19)..(20)

&lt;223&gt; xaa in position 19 to 20 is any amino acid

&lt;220&gt;

<221> Variant  
<222> (22)..(42)  
<223> xaa in position 22 to 42 is any amino acid  
<220>  
<221> Variant  
<222> (43)..(49)  
<223> xaa in position 43 to 49 is any or no amino acid  
<220>  
<221> Variant  
<222> (52)..(56)  
<223> xaa in position 52 to 56 is any amino acid  
<220>  
<221> Variant  
<222> (59)..(59)  
<223> xaa in position 59 is any amino acid  
<220>  
<221> Variant  
<222> (62)..(63)  
<223> xaa in position 62 to 63 is any amino acid  
<220>  
<221> Variant  
<222> (65)..(68)  
<223> xaa in position 65 to 68 is any amino acid  
<220>  
<221> Variant  
<222> (70)..(73)  
<223> xaa in position 70 to 73 is any amino acid  
<220>  
<221> Variant  
<222> (75)..(75)  
<223> xaa in position 75 is any amino acid  
<220>  
<221> Variant  
<222> (77)..(78)  
<223> xaa in position 77 to 78 is any amino acid  
<220>  
<221> Variant  
<222> (80)..(94)  
<223> xaa in position 80 to 94 is any amino acid  
<220>  
<221> Variant  
<222> (95)..(98)  
<223> xaa in position 95 to 98 is any or no amino acid  
<220>  
<221> Variant  
<222> (101)..(101)  
<223> xaa in position 101 is any amino acid  
<220>  
<221> Variant  
<222> (104)..(105)  
<223> xaa in position 104 to 105 is any amino acid  
<220>  
<221> Variant  
<222> (108)..(108)  
<223> xaa in position 108 is any amino acid  
<220>  
<221> Variant  
<222> (111)..(111)  
<223> xaa in position 111 is any amino acid  
<220>  
<221> Variant  
<222> (114)..(114)  
<223> xaa in position 114 is any amino acid  
<220>  
<221> Variant  
<222> (116)..(120)  
<223> xaa in position 116 to 120 is any amino acid  
<220>

<221> Variant  
<222> (122)..(124)  
<223> xaa in position 122 to 124 is any amino acid  
<220>  
<221> Variant  
<222> (127)..(127)  
<223> xaa in position 127 is any amino acid  
<220>  
<221> Variant  
<222> (131)..(132)  
<223> xaa in position 131 to 132 is any amino acid  
<220>  
<221> Variant  
<222> (134)..(135)  
<223> xaa in position 134 to 135 is any amino acid  
<220>  
<221> Variant  
<222> (139)..(143)  
<223> xaa in position 139 to 143 is any amino acid  
<220>  
<221> Variant  
<222> (148)..(149)  
<223> xaa in position 148 to 149 is any amino acid  
<220>  
<221> Variant  
<222> (151)..(151)  
<223> xaa in position 151 is any amino acid  
<220>  
<221> Variant  
<222> (154)..(155)  
<223> xaa in position 154 to 155 is any amino acid  
<220>  
<221> Variant  
<222> (158)..(159)  
<223> xaa in position 158 to 159 is any amino acid  
<220>  
<221> Variant  
<222> (162)..(162)  
<223> xaa in position 162 is any amino acid  
<220>  
<221> Variant  
<222> (165)..(165)  
<223> xaa in position 165 is any amino acid  
<220>  
<221> Variant  
<222> (168)..(168)  
<223> xaa in position 168 is any amino acid  
<220>  
<221> Variant  
<222> (170)..(184)  
<223> xaa in position 170 to 184 is any amino acid  
<220>  
<221> Variant  
<222> (185)..(190)  
<223> xaa in position 185 to 190 is any or no amino acid  
<220>  
<221> Variant  
<222> (192)..(193)  
<223> xaa in position 192 to 193 is any amino acid  
<220>  
<221> Variant  
<222> (196)..(196)  
<223> xaa in position 196 is any amino acid  
<220>  
<221> Variant  
<222> (199)..(203)  
<223> xaa in position 199 to 203 is any amino acid  
<220>

<221> Variant  
 <222> (205)..(206)  
 <223> xaa in position 205 to 206 is any amino acid  
 <220>  
 <221> Variant  
 <222> (208)..(213)  
 <223> xaa in position 208 to 213 is any amino acid  
 <220>  
 <221> Variant  
 <222> (214)..(219)  
 <223> xaa in position 214 to 219 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (222)..(227)  
 <223> xaa in position 222 to 227 is any amino acid  
 <220>  
 <221> Variant  
 <222> (229)..(230)  
 <223> xaa in position 229 to 230 is any amino acid  
 <220>  
 <221> Variant  
 <222> (233)..(235)  
 <223> xaa in position 233 to 235 is any amino acid  
 <220>  
 <221> Variant  
 <222> (237)..(237)  
 <223> xaa in position 237 is any amino acid  
 <220>  
 <221> Variant  
 <222> (241)..(241)  
 <223> xaa in position 241 is any amino acid

<400> 138  
 Leu Pro xaa xaa xaa xaa Pro xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 1 5 10 15  
 Tyr Leu xaa xaa Val xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 20 25 30  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 35 40 45  
 xaa His Asn xaa xaa xaa xaa xaa Leu Ser xaa Tyr Met xaa xaa Gly  
 50 55 60  
 xaa xaa xaa xaa Ala xaa xaa xaa xaa Tyr xaa Leu xaa xaa Asn xaa  
 65 70 75 80  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 85 90 95  
 xaa xaa Phe Tyr xaa Ser Lys xaa xaa Gu Phe xaa Asp Thr xaa Ile  
 100 105 110  
 Met xaa Leu xaa xaa xaa xaa xaa Gn xaa xaa xaa Leu His xaa Tyr  
 115 120 125  
 His His xaa xaa Ile xaa xaa Ile Trp Trp xaa xaa xaa xaa Ala  
 130 135 140  
 Pro Gly Gly xaa xaa Tyr xaa Ser Ala xaa xaa Asn Ser xaa xaa His  
 145 150 155 160  
 Val xaa Met Tyr xaa Tyr Tyr xaa Leu xaa xaa xaa xaa xaa xaa  
 165 170 175  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Thr xaa  
 180 185 190  
 xaa Gn Met xaa Gn Phe xaa xaa xaa xaa xaa Gn xaa xaa Tyr xaa  
 195 200 205  
 xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Tyr Pro xaa xaa xaa  
 210 215 220  
 xaa xaa xaa Leu xaa xaa Tyr Met xaa xaa xaa Leu xaa Leu Phe Gly  
 225 230 235 240  
 xaa Phe

<210> 139

<211> 60  
<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (2)..(2)  
<223> xaa in position 2 is Phe or Trp  
<220>  
<221> Variant  
<222> (3)..(3)  
<223> xaa in position 3 is Leu, Met or Val  
<220>  
<221> Variant  
<222> (6)..(6)  
<223> xaa in position 6 is any amino acid  
<220>  
<221> Variant  
<222> (7)..(7)  
<223> xaa in position 7 is Phe or Ile  
<220>  
<221> Variant  
<222> (9)..(9)  
<223> xaa in position 9 is any or no amino acid  
<220>  
<221> Variant  
<222> (11)..(11)  
<223> xaa in position 11 is any or no amino acid  
<220>  
<221> Variant  
<222> (13)..(15)  
<223> xaa in position 13 to 15 is any amino acid  
<220>  
<221> Variant  
<222> (16)..(16)  
<223> xaa in position 16 is any or no amino acid  
<220>  
<221> Variant  
<222> (18)..(18)  
<223> xaa in position 18 is Ile, Leu or Val  
<220>  
<221> Variant  
<222> (19)..(19)  
<223> xaa in position 19 is Ser or Thr  
<220>  
<221> Variant  
<222> (20)..(20)  
<223> xaa in position 20 is Phe or Val  
<220>  
<221> Variant  
<222> (23)..(23)  
<223> xaa in position 23 is Ile or Val  
<220>  
<221> Variant  
<222> (27)..(27)  
<223> xaa in position 27 is Ala, Ser or Thr  
<220>  
<221> Variant  
<222> (28)..(28)  
<223> xaa in position 28 is Ser or Thr  
<220>  
<221> Variant  
<222> (29)..(29)  
<223> xaa in position 29 is Ile or Val  
<220>  
<221> Variant  
<222> (30)..(31)  
<223> xaa in position 30 to 31 is any amino acid



<220>  
 <221> Vari ant  
 <222> (32)..(32)  
 <223> xaa in position 32 is Ile or Leu  
 <220>  
 <221> Vari ant  
 <222> (34)..(34)  
 <223> xaa in position 34 is any amino acid  
 <220>  
 <221> Vari ant  
 <222> (35)..(35)  
 <223> xaa in position 35 is Ala, Ile or Leu  
 <220>  
 <221> Vari ant  
 <222> (36)..(36)  
 <223> xaa in position 36 is Ala, Ile or Val  
 <220>  
 <221> Vari ant  
 <222> (37)..(39)  
 <223> xaa in position 37 to 39 is any amino acid  
 <220>  
 <221> Vari ant  
 <222> (42)..(42)  
 <223> xaa in position 42 is Gly or Asn  
 <220>  
 <221> Vari ant  
 <222> (44)..(44)  
 <223> xaa in position 44 is Asp or Glu  
 <220>  
 <221> Vari ant  
 <222> (45)..(45)  
 <223> xaa in position 45 is Ala or Ser  
 <220>  
 <221> Vari ant  
 <222> (47)..(47)  
 <223> xaa in position 47 is Phe, Trp or Tyr  
 <220>  
 <221> Vari ant  
 <222> (49)..(49)  
 <223> xaa in position 49 is Ala or Thr  
 <220>  
 <221> Vari ant  
 <222> (50)..(50)  
 <223> xaa in position 50 is any amino acid  
 <220>  
 <221> Vari ant  
 <222> (51)..(51)  
 <223> xaa in position 51 is Leu or Val  
 <220>  
 <221> Vari ant  
 <222> (54)..(54)  
 <223> xaa in position 54 is any amino acid  
 <220>  
 <221> Vari ant  
 <222> (55)..(55)  
 <223> xaa in position 55 is Ile or Val  
 <220>  
 <221> Vari ant  
 <222> (58)..(58)  
 <223> xaa in position 58 is any amino acid  
 <220>  
 <221> Vari ant  
 <222> (59)..(59)  
 <223> xaa in position 59 is Leu or Met

<400> 139  
 Glu xaa xaa Asp Thr xaa xaa Met xaa Leu xaa Lys xaa xaa xaa xaa  
 1 5 10 15

PF58307.txt

G n xaa xaa xaa Leu H i s xaa Tyr H i s H i s xaa xaa xaa xaa xaa xaa xaa  
                           20                          25                          30  
 Trp xaa xaa xaa xaa xaa xaa Ala Pro xaa Gly xaa xaa Tyr xaa Ser  
                   35                          40                          45  
 xaa xaa xaa Asn Ser xaa xaa H i s Val xaa xaa Tyr  
           50                          55                          60

<210> 140  
 <211> 22  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Ala or Leu  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is Phe or Tyr  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Cys or Ser  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Gly, Leu or Val  
 <220>  
 <221> Variant  
 <222> (8)..(9)  
 <223> xaa in position 8 to 9 is any amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(11)  
 <223> xaa in position 10 to 11 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(14)  
 <223> xaa in position 13 to 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(16)  
 <223> xaa in position 15 to 16 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Gly, Ser or Thr  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Phe or Trp  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Glu, Gly or Asn

<400> 140  
 Ser xaa xaa Met xaa xaa Gly xaa xaa xaa xaa Ala xaa xaa xaa xaa  
   1                  5                          10                          15  
 Tyr xaa xaa xaa xaa Asn  
           20

<210> 141  
 <211> 18  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(3)  
 <223> xaa in position 2 to 3 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is any amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(12)  
 <223> xaa in position 9 to 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Gly, Ile, Leu or Val  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is any amino acid  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Trp or Tyr  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Ala, Cys or Asp

<400> 141  
 Thr xaa xaa Gln Met xaa Gln Phe xaa xaa xaa xaa xaa Gln xaa xaa  
 1 5 10 15  
 xaa xaa

<210> 142  
 <211> 25  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (4)..(5)  
 <223> xaa in position 4 to 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Phe, Ile or Leu  
 <220>  
 <221> Variant  
 <222> (7)..(8)  
 <223> xaa in position 7 to 8 is any amino acid  
 <220>

<221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is Ile, Leu or Met  
 <220>  
 <221> Variant  
 <222> (11)..(12)  
 <223> xaa in position 11 to 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(15)  
 <223> xaa in position 14 to 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Ser or Thr  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is any amino acid  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Phe or Leu  
 <220>  
 <221> Variant  
 <222> (22)..(23)  
 <223> xaa in position 22 to 23 is any amino acid  
 <220>  
 <221> Variant  
 <222> (25)..(25)  
 <223> xaa in position 25 is Phe or Tyr

<400> 142  
 Tyr xaa Pro xaa xaa xaa xaa xaa xaa Leu xaa xaa Tyr xaa xaa xaa  
 1 5 10 15  
 xaa Leu xaa xaa Phe xaa xaa Phe xaa  
 20 25

<210> 143  
 <211> 879  
 <212> DNA  
 <213> *Ostreococcus tauri*

<220>  
 <221> CDS  
 <222> (1)..(879)

<400> 143  
 atg agt ggc tta cgt gca ccc aac ttt tta cac aga ttc tgg aca aag 48  
 Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys  
 1 5 10 15  
 tgg gac tac gcg att tcc aaa gtc gtc ttc acg tgt gcc gac agt ttt 96  
 Trp Asp Tyr Ala Ile Ser Lys Val Val Phe Thr Cys Ala Asp Ser Phe  
 20 25 30  
 cag tgg gac atc ggg cca gtg agt tcg agt acg gcg cat tta ccc gcc 144  
 Gln Trp Asp Ile Gly Pro Val Ser Ser Thr Ala His Leu Pro Ala  
 35 40 45  
 att gaa tcc cct acc cca ctg gtg act agc ctc ttg ttc tac tta gtc 192  
 Ile Glu Ser Pro Thr Pro Leu Val Thr Ser Leu Leu Phe Tyr Leu Val  
 50 55 60  
 aca gtt ttc ttg tgg tat ggt cgt tta acc agg agt tca gac aag aaa 240  
 Thr Val Phe Leu Trp Tyr Gly Arg Leu Thr Arg Ser Ser Asp Lys Lys  
 65 70 75 80  
 att aga gag cct acg tgg tta aga aga ttc ata ata tgt cat aat gcg 288

PF58307.txt

I l e A r g G l u P r o T h r T r p L e u A r g A r g P h e I l e I l e C y s H i s A s n A l a  
85 90 95  
t t c t t g a t a g t c c t c a g t c t t t a c a t g t g c c t t g g t t g t g t g g c c c a a  
P h e L e u I l e V a l L e u S e r L e u T y r M e t C y s L e u G l y C y s V a l A l a G l n 336  
100  
g c g t a t c a g a a t g g a t a t a c t t t a t g g g g t a a t g a a t t t c a a g g c c a c g  
A l a T y r G l n A s n G l y T y r T h r L e u T r p G l y A s n G l u P h e L y s A l a T h r 384  
115  
g a a a c t c a g c t t g c t c t a c a t t t a c a t t t t t t a c g t a a g t a a a a t a  
G u T h r G l n L e u A l a L e u T y r I l e T y r I l e P h e T y r V a l S e r L y s I l e 432  
130  
t a c g a g t t t g t a g a t a c t t a c a t t a t g c t t c t c a a g a a t a a c t t g c g g  
T y r G l u P h e V a l A s p T h r T y r I l e M e t L e u L e u L y s A s n A s n L e u A r g 480  
145  
c a a g t a a g t t t c c t a c a c a t t t a t c a c c a g c a c g a t t t c c t t t a t t  
G n V a l S e r P h e L e u H i s I l e T y r H i s H i s S e r T h r I l e S e r P h e I l e 528  
165  
t g g t g g a t c a t t g c t c g g a g g g c t c c g g g t g a t g c t t a c t t c a g c  
T r p T r p I l e I l e A l a A r g A r g A l a P r o G l y G l y A s p A l a T y r P h e S e r 576  
180  
g c g g c c t t g a a c t c a t g g t a c a c g t g t g c a t g t a c a c c t a t t a t c t a  
A l a A l a L e u A s n S e r T r p V a l H i s V a l C y s M e t T y r T h r T y r T y r L e u 624  
195  
t t a t c a a c c c t t a t t g g a a a g a a g a t c c t a a g c g t t c c a a c t a c c t t  
L e u S e r T h r L e u I l e G l y L y s G l u A s p P r o L y s A r g S e r A s n T y r L e u 672  
210  
t g g t g g g g t c g c c a c c t a a c g c a a a t g c a g a t g c t t c a g t t t t t c t t c  
T r p T r p G l y A r g H i s L e u T h r G n M e t G n M e t L e u G n P h e P h e P h e 720  
225  
a a c g t a c t t c a a g c g t t g t a c t g c g c t t c g t t c t c t a c g t a t c c c a a g  
A s n V a l L e u G n A l a L e u T y r C y s A l a S e r P h e S e r T h r T y r P r o P r o L y s 768  
245  
t t t t t g t c c a a a a t t c t g c t c g t c t a t a t g a t g a g c c t t c t c g g c t t g  
P h e L e u S e r L y s I l e L e u L e u V a l T y r M e t M e t S e r L e u L e u G l y L e u 816  
260  
t t t g g g c a t t t c t a c t a t t c c a a g c a c a t a g c a g c a g t a a g c t c c a g  
P h e G l y H i s P h e T y r T y r S e r L y s H i s I l e A l a A l a G l y A l a L y s L e u G n 864  
275  
a a a a a a c a g c a g t g a  
L y s L y s G n G n 879  
290

<210> 144  
<211> 292  
<212> PRT  
<213> *Ostreococcus tauri*

<400> 144  
M e t S e r G l y L e u A r g A l a P r o A s n P h e L e u H i s A r g P h e T r p T h r L y s  
1 5 10  
T r p A s p T y r A l a I l e S e r L y s V a l V a l P h e T h r C y s A l a A s p S e r P h e  
20 25  
G n T r p A s p I l e G l y P r o V a l S e r S e r S e r T h r A l a H i s L e u P r o A l a  
35 40 45  
I l e G l u S e r P r o T h r P r o L e u V a l T h r S e r L e u L e u P h e T y r L e u V a l  
50 55 60  
T h r V a l P h e L e u T r p T y r G l y A r g L e u T h r A r g S e r S e r A s p L y s L y s  
65 70 75 80  
I l e A r g G l u P r o T h r T r p L e u A r g A r g P h e I l e I l e C y s H i s A s n A l a  
85 90 95  
P h e L e u I l e V a l L e u S e r L e u T y r M e t C y s L e u G l y C y s V a l A l a G n  
100 105 110  
A l a T y r G n A s n G l y T y r T h r L e u T r p G l y A s n G l u P h e L y s A l a T h r  
115 120 125  
G u T h r G n L e u A l a L e u T y r I l e T y r I l e P h e T y r V a l S e r L y s I l e  
130 135 140  
T y r G l u P h e V a l A s p T h r T y r I l e M e t L e u L e u L y s A s n A s n L e u A r g  
145 150 155 160

Seite 183

PF58307.txt

G n Val Ser Phe Leu His Ile Tyr His His Ser Thr Ile Ser Phe Ile  
 Trp Trp Ile Ile Ala Arg Arg Ala Pro Gly Gly Asp Ala Tyr Phe Ser  
 Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr Tyr Leu  
 Leu Ser Thr Leu Ile Gly Lys Gly Asp Pro Lys Arg Ser Asn Tyr Leu  
 Trp Trp Gly Arg His Leu Thr G n Met G n Met Leu G n Phe Phe Phe  
 Asn Val Leu G n Ala Leu Tyr Cys Ala Ser Phe Ser Thr Tyr Pro Lys  
 Phe Leu Ser Lys Ile Leu Leu Val Tyr Met Met Ser Leu Leu Gly Leu  
 Phe Gly His Phe Tyr Tyr Ser Lys His Ile Ala Ala Ala Lys Leu G n  
 Lys Lys G n G n  
 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290

<210> 145  
 <211> 831  
 <212> DNA  
 <213> Thr aust ochyt ri um sp

<220>  
 <221> CDS  
 <222> (1)..(831)

<400> 145  
 at g gat gt c gt c gag cag caa tgg cgc cgc t t c gt g gac gcc gt g gac 48  
 Met Asp Val Val Gl u G n G n Trp Arg Arg Phe Val Asp Ala Val Asp  
 1 5 10  
 aac gga at c gt g gag t t c at g gag cat gag gag ccc aac aag ct g aac 96  
 Asn Gly Ile Val Gl u Phe Met Gl u His Gl u Gl u Pro Asn Lys Leu Asn  
 20 25 30  
 gag ggc aag ct c t cc acc t cg acc gag gag at g at g gcg ct t at c gt c 144  
 Gl u Gly Lys Leu Ser Thr Ser Thr Gl u Gl u Met Met Ala Leu Ile Val  
 35 40 45  
 ggc t ac ct g gcg t t c gt g gt c ct c ggg t cc gcc t t c at g aag gcc t t t 192  
 Gly Tyr Leu Ala Phe Val Val Leu Gly Ser Ala Phe Met Lys Ala Phe  
 50 55 60  
 gt c gat aag cct t t c gag ct c aag t t c ct c aag ct c gt g cac aac at c 240  
 Val Asp Lys Pro Phe Gl u Leu Lys Phe Leu Lys Leu Val His Asn Ile  
 65 70 75 80  
 t t c ct c acc ggt ct g t cc at g t ac at g gcc acc gag t gc gcg cgc cag 288  
 Phe Leu Thr Gly Leu Ser Met Tyr Met Ala Thr Gl u Cys Ala Arg G n  
 85 90 95  
 gca t ac ct c ggc ggc t ac aag ct c t t t ggc aac ccg at g gag aag ggc 336  
 Ala Tyr Leu Gly Gly Tyr Lys Leu Phe Gly Asn Pro Met Gl u Lys Gly  
 100 110  
 acc gag t cg cac gcc ccg ggc at g gcc aac at c at c t ac at c t t c t ac 384  
 Thr Gl u Ser His Ala Pro Gly Met Ala Asn Ile Ile Tyr Ile Phe Tyr  
 115 120 125  
 gt g agc aag t t c ct c gaa t t c ct c gac acc gt c t t c at g at c ct c ggc 432  
 Val Ser Lys Phe Leu Gl u Phe Leu Asp Thr Val Phe Met Ile Leu Gly  
 130 135 140  
 aag aag t gg aag cag ct c agc t t t ct c cac gt c t ac cac cac gcg agc 480  
 Lys Lys Trp Lys G n Leu Ser Phe Leu His Val Tyr His His Ala Ser  
 145 150 155 160  
 at c agc t t c at c t gg ggc at c at c gcc cgc t t c gcg ccc ggt ggc gac 528  
 Ile Ser Phe Ile Trp Gly Ile Ile Ala Arg Phe Ala Pro Gly Gly Asp  
 165 170 175  
 gcc t ac t t c t ct acc at c ct c aac agc agc gt g cat gt c gt g ct c t ac 576  
 Ala Tyr Phe Ser Thr Ile Leu Asn Ser Ser Val His Val Val Leu Tyr  
 180 185 190  
 ggc t ac t ac gcc t cg acc acc ct c ggc t ac acc t t c at g cgc ccg ct g 624  
 Gly Tyr Tyr Ala Ser Thr Thr Leu Gly Tyr Thr Phe Met Arg Pro Leu  
 195 200 205

PF58307. txt

cgc	ccg	tac	att	acc	acc	att	cag	ctc	acg	cag	ttc	atg	gcc	atg	gtc	672
Arg	Pro	Tyr	Ile	Thr	Thr	Ile	Gln	Leu	Thr	Gln	Phe	Met	Ala	Met	Val	
	210					215					220					
gtc	cag	tcc	gtc	tat	gac	tac	tac	aac	ccc	tgc	gac	tac	ccg	cag	ccc	720
Val	Gln	Ser	Val	Tyr	Asp	Tyr	Tyr	Asn	Pro	Cys	Asp	Tyr	Pro	Gln	Pro	
	225				230					235					240	
ctc	gtc	aag	ctg	ctc	ttc	tgg	tac	atg	ctc	acc	atg	ctc	ggc	ctc	ttc	768
Leu	Val	Lys	Leu	Leu	Phe	Trp	Tyr	Met	Leu	Thr	Met	Leu	Gly	Leu	Phe	
				245					250					255		
ggc	aac	ttc	ttc	gtg	cag	cag	tac	ctc	aag	ccc	aag	gcg	ccc	aag	aag	816
Gly	Asn	Phe	Phe	Val	Gln	Gln	Tyr	Leu	Lys	Pro	Lys	Ala	Pro	Lys	Lys	
			260					265					270			
cag	aag	acc	atc	taa												831
Gln	Lys	Thr	Ile													
		275														

<210> 146  
 <211> 276  
 <212> PRT  
 <213> Thr aust ochyt r i u m sp

<400>	146																
Met	Asp	Val	Val	Glu	Gln	Gln	Trp	Arg	Arg	Phe	Val	Asp	Ala	Val	Asp		
1				5					10					15			
Asn	Gly	Ile	Val	Glu	Phe	Met	Glu	His	Glu	Glu	Pro	Asn	Lys	Leu	Asn		
			20					25					30				
Glu	Gly	Lys	Leu	Ser	Thr	Ser	Thr	Glu	Glu	Met	Met	Ala	Leu	Ile	Val		
		35					40					45					
Gly	Tyr	Leu	Ala	Phe	Val	Val	Leu	Gly	Ser	Ala	Phe	Met	Lys	Ala	Phe		
	50			55						60							
Val	Asp	Lys	Pro	Phe	Glu	Leu	Lys	Phe	Leu	Lys	Leu	Val	His	Asn	Ile		
65				70					75					80			
Phe	Leu	Thr	Gly	Leu	Ser	Met	Tyr	Met	Ala	Thr	Glu	Cys	Ala	Arg	Gln		
			85					90						95			
Ala	Tyr	Leu	Gly	Gly	Tyr	Lys	Leu	Phe	Gly	Asn	Pro	Met	Glu	Lys	Gly		
			100					105					110				
Thr	Glu	Ser	His	Ala	Pro	Gly	Met	Ala	Asn	Ile	Ile	Tyr	Ile	Phe	Tyr		
		115					120					125					
Val	Ser	Lys	Phe	Leu	Glu	Phe	Leu	Asp	Thr	Val	Phe	Met	Ile	Leu	Gly		
	130					135					140						
Lys	Lys	Trp	Lys	Gln	Leu	Ser	Phe	Leu	His	Val	Tyr	His	His	Ala	Ser		
145				150					155						160		
Ile	Ser	Phe	Ile	Trp	Gly	Ile	Ile	Ala	Arg	Phe	Ala	Pro	Gly	Gly	Asp		
			165					170						175			
Ala	Tyr	Phe	Ser	Thr	Ile	Leu	Asn	Ser	Ser	Val	His	Val	Val	Leu	Tyr		
			180					185					190				
Gly	Tyr	Tyr	Ala	Ser	Thr	Thr	Leu	Gly	Tyr	Thr	Phe	Met	Arg	Pro	Leu		
	195						200					205					
Arg	Pro	Tyr	Ile	Thr	Thr	Ile	Gln	Leu	Thr	Gln	Phe	Met	Ala	Met	Val		
	210					215					220						
Val	Gln	Ser	Val	Tyr	Asp	Tyr	Tyr	Asn	Pro	Cys	Asp	Tyr	Pro	Gln	Pro		
225					230					235					240		
Leu	Val	Lys	Leu	Leu	Phe	Trp	Tyr	Met	Leu	Thr	Met	Leu	Gly	Leu	Phe		
			245					250					255				
Gly	Asn	Phe	Phe	Val	Gln	Gln	Tyr	Leu	Lys	Pro	Lys	Ala	Pro	Lys	Lys		
			260					265					270				
Gln	Lys	Thr	Ile														
		275															

<210> 147  
 <211> 1146  
 <212> DNA  
 <213> Lei shmani a maj or

<220>  
 <221> CDS  
 <222> (1).. (1146)

PF58307. txt

<400> 147

at g	gt c	t ct	ct g	gag	cag	gcc	gaa	cag	at c	gcg	gcc	gcg	at c	gag	gt a	48
Met	Val	Ser	Leu	Gu	Gn	Ala	Gu	Gn	Ile	Ala	Ala	Ala	Ile	Gu	Val	
1			5						10					15		
cct	gac	tgg	gt c	t t g	aca	aag	t ct	gcg	gcg	ct g	gt g	t ac	agc	t gc	t t c	96
Pro	Asp	Trp	Val	Leu	Thr	Lys	Ser	Ala	Ala	Leu	Val	Tyr	Ser	Cys	Phe	
			20					25					30			
ggc	t cc	gcc	gcc	aat	gct	t t t	gaa	agc	agt	at c	aag	at c	aac	t t c	ccg	144
G y	Ser	Ala	Ala	Asn	Ala	Phe	Gu	Ser	Ser	Ile	Lys	Ile	Asn	Phe	Pro	
			35				40					45				
gcg	cag	cat	gcg	t t c	gt g	gag	gcg	t gg	at g	cgc	gcg	cgc	t cc	cac	cct	192
Ala	Gn	His	Ala	Phe	Val	Gu	Ala	Trp	Met	Arg	Ala	Arg	Ser	His	Pro	
			50			55				60						
t t t	gcg	gag	cgc	ct g	ccg	t ac	ct g	aat	ccg	t gg	cac	gt t	at c	gcc	t cg	240
Phe	Ala	Gu	Arg	Leu	Pro	Tyr	Leu	Asn	Pro	Trp	His	Val	Ile	Ala	Ser	
65					70				75						80	
at a	ct g	gcc	t ac	ct c	t cc	t t g	at t	gt c	acc	t t g	cgc	ct g	t t g	cat	cgt	288
Ile	Leu	Ala	Tyr	Leu	Ser	Leu	Ile	Val	Thr	Leu	Arg	Leu	Leu	His	Arg	
					85				90					95		
gt a	ct c	ggt	aag	t t c	t cg	t gc	cgc	act	ct c	gga	t t g	gt g	cac	aac	ct c	336
Val	Leu	G y	Lys	Phe	Ser	Cys	Arg	Thr	Leu	G y	Leu	Val	His	Asn	Leu	
			100					105					110			
ggt	ct c	cat	ct t	ct c	t cg	t t g	t ac	at g	agc	ct t	ggt	ct c	at g	at c	agc	384
G y	Leu	His	Leu	Leu	Ser	Leu	Tyr	Met	Ser	Leu	G y	Leu	Met	Ile	Ser	
			115					120				125				
gcg	cgc	gcc	gcg	ggg	t ac	t cg	ct c	t gg	aac	aac	gcg	gt c	ggc	acc	t cc	432
Ala	Arg	Ala	Ala	G y	Tyr	Ser	Leu	Trp	Asn	Asn	Ala	Val	G y	Thr	Ser	
			130			135					140					
ccg	gct	gag	t gg	cgc	at t	gcg	aag	ct g	at c	t gg	ct c	t t c	t at	gt c	t cg	480
Pro	Ala	Gu	Trp	Arg	Ile	Ala	Lys	Leu	Ile	Trp	Leu	Phe	Tyr	Val	Ser	
145					150				155					160		
aag	gt g	gt g	gaa	t gg	gt g	gac	acg	gt a	at t	at g	t t a	t t a	aag	cag	aac	528
Lys	Val	Val	Gu	Trp	Val	Asp	Thr	Val	Ile	Met	Leu	Leu	Lys	Gn	Asn	
				165				170					175			
t ac	cac	cag	gt c	acc	t t c	ct g	cac	gt g	t at	cac	cac	acg	acg	gt t	t t t	576
Tyr	His	Gn	Val	Thr	Phe	Leu	His	Val	Tyr	His	His	Thr	Thr	Val	Phe	
			180					185					190			
gt g	ct g	t gg	t gg	ct g	gcg	t t g	ct g	gt c	gct	cct	ggc	ggc	gag	t cg	t ac	624
Val	Leu	Trp	Trp	Leu	Ala	Leu	Leu	Val	Ala	Pro	G y	G y	Gu	Ser	Tyr	
			195			200					205					
t ac	agc	gcc	at g	gt g	aac	t ct	ggc	gt c	cac	gt t	t t c	at g	t ac	ggg	t ac	672
Tyr	Ser	Ala	Met	Val	Asn	Ser	G y	Val	His	Val	Phe	Met	Tyr	G y	Tyr	
			210			215				220						
t ac	t t t	ct c	acg	ct g	ct c	t t c	cca	t cc	ggc	at c	gt g	cgc	gac	gt c	t t g	720
Tyr	Phe	Leu	Thr	Leu	Leu	Phe	Pro	Ser	G y	Ile	Val	Arg	Asp	Val	Leu	
225					230				235					240		
agc	aag	t t c	aag	t t t	gcc	at t	acg	aag	ggc	cag	at g	t gg	cag	t t c	gt c	768
Ser	Lys	Phe	Lys	Phe	Ala	Ile	Thr	Lys	G y	Gn	Met	Trp	Gn	Phe	Val	
				245				250					255			
t t c	aac	t gc	ct a	cag	t cc	gcg	t ac	gac	ct c	gt g	t gg	gt g	ccg	cgg	gaa	816
Phe	Asn	Cys	Leu	Gn	Ser	Ala	Tyr	Asp	Leu	Val	Trp	Val	Pro	Arg	Gu	
			260					265					270			
gag	ct c	aag	t ac	agc	gcg	gag	ct g	ct g	cag	at c	ct c	t t c	t gg	t ac	at g	864
Gu	Leu	Lys	Tyr	Ser	Ala	Gu	Leu	Leu	Gn	Ile	Leu	Phe	Trp	Tyr	Met	
			275				280					285				
at c	t cc	ct c	t t g	gcg	ct c	t t t	ggc	aac	t t c	t t g	gt g	aag	aac	aag	aag	912
Ile	Ser	Leu	Leu	Ala	Leu	Phe	G y	Asn	Phe	Leu	Val	Lys	Asn	Lys	Lys	
			290			295				300						
t t c	t cg	cac	cgc	cgc	t gc	gt t	gat	gcc	gcg	act	gct	t cg	ggc	gcg	aag	960
Phe	Ser	His	Arg	Arg	Cys	Val	Asp	Ala	Ala	Thr	Ala	Ser	G y	Ala	Lys	
305					310				315						320	
gag	gac	acg	gcg	gcg	agg	t cc	cac	ggc	gac	cgc	acc	cac	aga	acc	cgt	1008
Gu	Asp	Thr	Ala	Ala	Arg	Ser	His	G y	Asp	Arg	Thr	His	Arg	Thr	Arg	
				325				330					335			
gt g	aag	gct	ggc	at g	acc	aac	at g	caa	ct g	gag	agg	ct g	aag	aat	gag	1056
Val	Lys	Ala	G y	Met	Thr	Asn	Met	Gn	Leu	Gu	Arg	Leu	Lys	Asn	Gu	
			340					345					350			
aag	t cc	acg	gag	at g	aag	ct g	ct g	at g	cgc	aag	aac	ggc	aac	ggc	aac	1104

Seite 186



PF58307. txt

Lys Ser Thr Glu Met Lys Leu Leu Met Arg Lys Asn Gly Asn Gly Asn  
 355 360 365  
 gga caa aaa gcg t cg ct c cag gcc at g gca ggc agt cga t ga  
 Gly Gln Lys Ala Ser Leu Gln Ala Met Ala Gly Ser Arg  
 370 375 380

1146

<210> 148  
 <211> 381  
 <212> PRT  
 <213> Lei shmani a maj or

<400> 148  
 Met Val Ser Leu Glu Gln Ala Glu Gln Ile Ala Ala Ala Ile Glu Val  
 1 5 10 15  
 Pro Asp Trp Val Leu Thr Lys Ser Ala Ala Leu Val Tyr Ser Cys Phe  
 20 25 30  
 Gly Ser Ala Ala Asn Ala Phe Glu Ser Ser Ile Lys Ile Asn Phe Pro  
 35 40 45  
 Ala Gln His Ala Phe Val Glu Ala Trp Met Arg Ala Arg Ser His Pro  
 50 55 60  
 Phe Ala Glu Arg Leu Pro Tyr Leu Asn Pro Trp His Val Ile Ala Ser  
 65 70 75  
 Ile Leu Ala Tyr Leu Ser Leu Ile Val Thr Leu Arg Leu Leu His Arg  
 85 90 95  
 Val Leu Gly Lys Phe Ser Cys Arg Thr Leu Gly Leu Val His Asn Leu  
 100 105 110  
 Gly Leu His Leu Leu Ser Leu Tyr Met Ser Leu Gly Leu Met Ile Ser  
 115 120 125  
 Ala Arg Ala Ala Gly Tyr Ser Leu Trp Asn Asn Ala Val Gly Thr Ser  
 130 135 140  
 Pro Ala Glu Trp Arg Ile Ala Lys Leu Ile Trp Leu Phe Tyr Val Ser  
 145 150 155  
 Lys Val Val Glu Trp Val Asp Thr Val Ile Met Leu Leu Lys Gln Asn  
 165 170 175  
 Tyr His Gln Val Thr Phe Leu His Val Tyr His His Thr Thr Val Phe  
 180 185 190  
 Val Leu Trp Trp Leu Ala Leu Leu Val Ala Pro Gly Gly Glu Ser Tyr  
 195 200 205  
 Tyr Ser Ala Met Val Asn Ser Gly Val His Val Phe Met Tyr Gly Tyr  
 210 215 220  
 Tyr Phe Leu Thr Leu Leu Phe Pro Ser Gly Ile Val Arg Asp Val Leu  
 225 230 235  
 Ser Lys Phe Lys Phe Ala Ile Thr Lys Gly Gln Met Trp Gln Phe Val  
 245 250 255  
 Phe Asn Cys Leu Gln Ser Ala Tyr Asp Leu Val Trp Val Pro Arg Glu  
 260 265 270  
 Glu Leu Lys Tyr Ser Ala Glu Leu Leu Gln Ile Leu Phe Trp Tyr Met  
 275 280 285  
 Ile Ser Leu Leu Ala Leu Phe Gly Asn Phe Leu Val Lys Asn Lys Lys  
 290 295 300  
 Phe Ser His Arg Arg Cys Val Asp Ala Ala Thr Ala Ser Gly Ala Lys  
 305 310 315 320  
 Glu Asp Thr Ala Ala Arg Ser His Gly Asp Arg Thr His Arg Thr Arg  
 325 330 335  
 Val Lys Ala Gly Met Thr Asn Met Gln Leu Glu Arg Leu Lys Asn Glu  
 340 345 350  
 Lys Ser Thr Glu Met Lys Leu Leu Met Arg Lys Asn Gly Asn Gly Asn  
 355 360 365  
 Gly Gln Lys Ala Ser Leu Gln Ala Met Ala Gly Ser Arg  
 370 375 380

<210> 149  
 <211> 879  
 <212> DNA  
 <213> Ostreococcus tauri

<220>  
 <221> CDS

<222> (1)..(879)

<400> 149

atg	agt	ggc	tta	cgt	gca	ccc	aac	ttt	tta	cac	aga	ttc	tgg	aca	aag	48
Met	Ser	Gly	Leu	Arg	Ala	Pro	Asn	Phe	Leu	His	Arg	Phe	Trp	Thr	Lys	
1				5				10					15			
tgg	gac	tac	gcg	att	tcc	aaa	gtc	gtc	ttc	acg	tgt	gcc	gac	agt	ttt	96
Trp	Asp	Tyr	Ala	Ile	Ser	Lys	Val	Val	Phe	Thr	Cys	Ala	Asp	Ser	Phe	
			20					25				30				
cag	tgg	gac	atc	ggg	cca	gtg	agt	tcg	agt	acg	gcg	cat	tta	ccc	gcc	144
Gln	Trp	Asp	Ile	Gly	Pro	Val	Ser	Ser	Ser	Thr	Ala	His	Leu	Pro	Ala	
			35				40					45				
att	gaa	tcc	cct	acc	cca	ctg	gtg	act	agc	ctc	ttg	ttc	tac	tta	gtc	192
Ile	Glu	Ser	Pro	Thr	Pro	Leu	Val	Thr	Ser	Leu	Leu	Phe	Tyr	Leu	Val	
	50					55					60					
aca	ggt	ttc	ttg	tgg	tat	ggt	cgt	tta	acc	agg	agt	tca	gac	aag	aaa	240
Thr	Val	Phe	Leu	Trp	Tyr	Gly	Arg	Leu	Thr	Arg	Ser	Ser	Asp	Lys	Lys	
65					70				75					80		
att	aga	gag	cct	acg	tgg	tta	aga	aga	ttc	ata	ata	tgt	cat	aat	gcg	288
Ile	Arg	Glu	Pro	Thr	Trp	Leu	Arg	Arg	Phe	Ile	Ile	Cys	His	Asn	Ala	
			85						90					95		
ttc	ttg	ata	gtc	ctc	agt	ctt	tac	atg	tgc	ctt	ggt	tgt	gtg	gcc	caa	336
Phe	Leu	Ile	Val	Leu	Ser	Leu	Tyr	Met	Cys	Leu	Gly	Cys	Val	Ala	Gln	
			100					105					110			
gcg	tat	cag	aat	gga	tat	act	tta	tgg	ggt	aat	gaa	ttc	aag	gcc	acg	384
Ala	Tyr	Gln	Asn	Gly	Tyr	Thr	Leu	Trp	Gly	Asn	Glu	Phe	Lys	Ala	Thr	
			115				120					125				
gaa	act	cag	ctt	gct	ctc	tac	att	tac	att	ttt	tac	gta	agt	aaa	ata	432
Glu	Thr	Gln	Leu	Ala	Leu	Tyr	Ile	Tyr	Ile	Phe	Tyr	Val	Ser	Lys	Ile	
	130				135						140					
tac	gag	ttt	gta	gat	act	tac	att	atg	ctt	ctc	aag	aat	aac	ttg	cgg	480
Tyr	Glu	Phe	Val	Asp	Thr	Tyr	Ile	Met	Leu	Leu	Lys	Asn	Asn	Leu	Arg	
145				150					155						160	
caa	gta	agt	ttc	cta	cac	att	tat	cac	cac	agc	acg	att	tcc	ttt	att	528
Gln	Val	Ser	Phe	Leu	His	Ile	Tyr	His	His	Ser	Thr	Ile	Ser	Phe	Ile	
				165					170					175		
tgg	tgg	atc	att	gct	cgg	agg	gct	ccg	ggt	gat	gct	tac	ttc	agc		576
Trp	Trp	Ile	Ile	Ala	Arg	Arg	Ala	Pro	Gly	Gly	Asp	Ala	Tyr	Phe	Ser	
				180				185					190			
gcg	gcc	ttg	aac	tca	tgg	gta	cac	gtg	tgc	atg	tac	acc	tat	tat	cta	624
Ala	Ala	Leu	Asn	Ser	Trp	Val	His	Val	Cys	Met	Tyr	Thr	Tyr	Tyr	Leu	
			195				200					205				
tta	tca	acc	ctt	att	gga	aaa	gaa	gat	cct	aag	cgt	tcc	aac	tac	ctt	672
Leu	Ser	Thr	Leu	Ile	Gly	Lys	Glu	Asp	Pro	Lys	Arg	Ser	Asn	Tyr	Leu	
			210			215					220					
tgg	tgg	ggt	cgc	cac	cta	acg	caa	atg	cag	atg	ctt	cag	ttt	ttc	ttc	720
Trp	Trp	Gly	Arg	His	Leu	Thr	Gln	Met	Gln	Met	Leu	Gln	Phe	Phe	Phe	
225					230				235						240	
aac	gta	ctt	caa	gcg	ttg	tac	tgc	gct	tcg	ttc	tct	acg	tat	ccc	aag	768
Asn	Val	Leu	Gln	Ala	Leu	Tyr	Cys	Ala	Ser	Phe	Ser	Thr	Tyr	Pro	Lys	
				245					250					255		
ttt	ttg	tcc	aaa	att	ctg	ctc	gtc	tat	atg	atg	agc	ctt	ctc	ggc	ttg	816
Phe	Leu	Ser	Lys	Ile	Leu	Leu	Val	Tyr	Met	Met	Ser	Leu	Leu	Gly	Leu	
			260					265					270			
ttt	ggg	cat	ttc	tac	tat	tcc	aag	cac	ata	gca	gca	gct	aag	ctc	cag	864
Phe	Gly	His	Phe	Tyr	Tyr	Ser	Lys	His	Ile	Ala	Ala	Ala	Lys	Leu	Gln	
			275				280					285				
aaa	aaa	cag	cag	tga												879
Lys	Lys	Gln	Gln													
			290													

<210> 150

<211> 292

<212> PRT

<213> *Ostreococcus tauri*

<400> 150

Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys  
Seite 188

PF58307. txt

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

<210> 151  
 <211> 873  
 <212> DNA  
 <213> Physcomitrella patens

<220>  
 <221> CDS  
 <222> (1)..(873)

<400> 151  
 atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg 48  
 Met Gu Val Val Gu Arg Phe Tyr Gy Gu Leu Asp Gy Lys Val Ser  
 1 5 10  
 cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg gat 96  
 Gn Gy Val Asn Ala Leu Leu Gy Ser Phe Gy Val Gu Leu Thr Asp  
 20 25 30  
 acg ccc act acc aaa ggc ttg ccc ctg gtt gac agt ccc aca ccc atc 144  
 Thr Pro Thr Thr Lys Gy Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
 35 40 45  
 gtc ctg ggt gtt tct gta tac ttg act att gtc att gga ggg ctt ttg 192  
 Val Leu Gy Val Ser Val Tyr Leu Thr Ile Val Ile Gy Gy Leu Leu  
 50 55 60  
 tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt ttg 240  
 Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Gu Pro Phe Leu  
 65 70 75 80  
 ctg caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctg agt 288  
 Leu Gn Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser  
 85 90 95  
 ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg tac 336  
 Leu Tyr Met Cys Val Gy Ile Ala Tyr Gn Ala Ile Thr Trp Arg Tyr

PF58307. txt

		100						105					110						
t ct	ct c	t gg	ggc	aat	gca	t ac	aat	c ct	aaa	cat	aaa	gag	at g	gcg	at t				384
Ser	Leu	Tr p	Gly	Asn	Al a	Tyr	Asn	Pro	Lys	His	Lys	Glu	Met	Al a	Ile				
		115						120				125							
ct g	gt a	t ac	ttg	ttc	t ac	at g	t ct	aag	t ac	gt g	gaa	ttc	at g	gat	acc				432
Leu	Val	Tyr	Leu	Phe	Tyr	Met	Ser	Lys	Tyr	Val	Glu	Phe	Met	Asp	Thr				
		130						135				140							
gt t	at c	at g	ata	ct g	aag	cgc	agc	acc	agg	caa	at a	agc	ttc	ct c	cac				480
Val	Ile	Met	Ile	Leu	Lys	Arg	Ser	Thr	Arg	Gln	Ile	Ser	Phe	Leu	His				160
		145						150				155							
gt t	t at	cat	cat	t ct	t ca	at t	t cc	ct c	at t	t gg	t gg	gct	at t	gct	cat				528
Val	Tyr	His	His	Ser	Ser	Ile	Ser	Leu	Ile	Trp	Trp	Al a	Ile	Al a	His				
		165						170				175							
cac	gct	cct	ggc	ggt	gaa	gca	t at	t gg	t ct	gcg	gct	ct g	aac	t ca	gga				576
His	Al a	Pro	Gly	Gly	Glu	Al a	Tyr	Trp	Ser	Al a	Al a	Leu	Asn	Ser	Gly				
		180						185				190							
gt g	cat	gt t	ct c	at g	t at	gcg	t at	t ac	ttc	ttg	gct	gcc	t gc	ct t	cga				624
Val	His	Val	Leu	Met	Tyr	Al a	Tyr	Tyr	Phe	Leu	Al a	Al a	Cys	Leu	Arg				
		195						200				205							
agt	agc	cca	aag	tta	aaa	aat	aag	t ac	ct t	ttt	t gg	ggc	agg	t ac	ttg				672
Ser	Ser	Pro	Lys	Leu	Lys	Asn	Lys	Tyr	Leu	Phe	Trp	Gly	Arg	Tyr	Leu				
		210						215				220							
aca	caa	ttc	caa	at g	ttc	cag	ttt	at g	ct g	aac	tta	gt g	cag	gct	t ac				720
Thr	Gln	Phe	Gln	Met	Phe	Gln	Phe	Met	Leu	Asn	Leu	Val	Gln	Al a	Tyr				240
		225						230				235							
t ac	gac	at g	aaa	acg	aat	gcg	cca	t at	cca	caa	t gg	ct g	at c	aag	at t				768
Tyr	Asp	Met	Lys	Thr	Asn	Al a	Pro	Tyr	Pro	Gln	Trp	Leu	Ile	Lys	Ile				
		245						250				255							
ttg	ttc	t ac	t ac	at g	at c	t cg	ttg	ct g	ttt	ct t	ttc	ggc	aat	ttt	t ac				816
Leu	Phe	Tyr	Tyr	Met	Ile	Ser	Leu	Leu	Phe	Leu	Phe	Gly	Asn	Phe	Tyr				
		260						265				270							
gt a	caa	aaa	t ac	at c	aaa	ccc	t ct	gac	gga	aag	caa	aag	gga	gct	aaa				864
Val	Gln	Lys	Tyr	Ile	Lys	Pro	Ser	Asp	Gly	Lys	Gln	Lys	Gly	Al a	Lys				
		275						280				285							
act	gag	t ga																	873
Thr	Glu																		
		290																	

<210> 152  
 <211> 290  
 <212> PRT  
 <213> Physcomitrella patens

<400> 152

Met	Glu	Val	Val	Glu	Arg	Phe	Tyr	Gly	Glu	Leu	Asp	Gly	Lys	Val	Ser				
1				5				10					15						
Gln	Gly	Val	Asn	Al a	Leu	Leu	Gly	Ser	Phe	Gly	Val	Glu	Leu	Thr	Asp				
			20					25				30							
Thr	Pro	Thr	Thr	Lys	Gly	Leu	Pro	Leu	Val	Asp	Ser	Pro	Thr	Pro	Ile				
		35					40					45							
Val	Leu	Gly	Val	Ser	Val	Tyr	Leu	Thr	Ile	Val	Ile	Gly	Gly	Leu	Leu				
		50				55					60								
Trp	Ile	Lys	Al a	Arg	Asp	Leu	Lys	Pro	Arg	Al a	Ser	Glu	Pro	Phe	Leu				
		65			70				75			80							
Leu	Gln	Al a	Leu	Val	Leu	Val	His	Asn	Leu	Phe	Cys	Phe	Al a	Leu	Ser				
				85				90						95					
Leu	Tyr	Met	Cys	Val	Gly	Ile	Al a	Tyr	Gln	Al a	Ile	Thr	Trp	Arg	Tyr				
		100						105				110							
Ser	Leu	Trp	Gly	Asn	Al a	Tyr	Asn	Pro	Lys	His	Lys	Glu	Met	Al a	Ile				
		115					120					125							
Leu	Val	Tyr	Leu	Phe	Tyr	Met	Ser	Lys	Tyr	Val	Glu	Phe	Met	Asp	Thr				
		130				135					140								
Val	Ile	Met	Ile	Leu	Lys	Arg	Ser	Thr	Arg	Gln	Ile	Ser	Phe	Leu	His				
		145			150				155						160				
Val	Tyr	His	His	Ser	Ser	Ile	Ser	Leu	Ile	Trp	Trp	Al a	Ile	Al a	His				
				165				170				175							
His	Al a	Pro	Gly	Gly	Glu	Al a	Tyr	Trp	Ser	Al a	Al a	Leu	Asn	Ser	Gly				
			180					185				190							

PF58307. txt

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg  
 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu  
 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr  
 225 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile  
 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr  
 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys  
 Thr Gu  
 290

<210> 153  
 <211> 957  
 <212> DNA  
 <213> Mbrtierella alpina

<220>  
 <221> CDS  
 <222> (1).. (957)

<400> 153  
 atg gag tgc att gcg cca ttc ctg cca tca aag atg ccg caa gat ctg 48  
 Met Gu Ser Ile Ala Pro Phe Leu Pro Ser Lys Met Pro Gln Asp Leu  
 1  
 ttt atg gac ctt gcc acc gct atc ggt gtc cgg gcc gcg ccc tat gtc 96  
 Phe Met Asp Leu Ala Thr Ala Ile Gly Val Arg Ala Ala Pro Tyr Val  
 20  
 gat cct ctg gag gcc gcg ctg gtg gcc cag gcc gag aag tac atc ccc 144  
 Asp Pro Leu Gu Ala Ala Leu Val Ala Gln Ala Gu Lys Tyr Ile Pro  
 35  
 acg att gtc cat cac acg cgt ggg ttc ctg gtc gcg gtg gag tgc cct 192  
 Thr Ile Val His His Thr Arg Gly Phe Leu Val Ala Val Gu Ser Pro  
 50  
 ttg gcc cgt gag ctg ccg ttg atg aac ccg ttc cac gtg ctg ttg atc 240  
 Leu Ala Arg Gu Leu Pro Leu Met Asn Pro Phe His Val Leu Leu Ile  
 65  
 gtg ctg gct tat ttg gtc acg gtc ttt gtg ggc atg cag atc atg aag 288  
 Val Leu Ala Tyr Leu Val Thr Val Phe Val Gly Met Gln Ile Met Lys  
 85  
 aac ttt gag ccg ttc gag gtc aag acg ttt tgc ctg ctg cac aac ttt 336  
 Asn Phe Gu Arg Phe Gu Val Lys Thr Phe Ser Leu Leu His Asn Phe  
 100  
 tgt ctg gtc tgc atc agc gcc tac atg tgc ggt ggg atc ctg tac gag 384  
 Cys Leu Val Ser Ile Ser Ala Tyr Met Cys Gly Gly Ile Leu Tyr Gu  
 115  
 gct tat cag gcc aac tat gga ctg ttt gag aac gct gat cat acc 432  
 Ala Tyr Gln Ala Asn Tyr Gly Leu Phe Gu Asn Ala Ala Asp His Thr  
 130  
 ttc aag ggt ctt cct atg gcc aag atg atc tgg ctg ttc tac ttc tcc 480  
 Phe Lys Gly Leu Pro Met Ala Lys Met Ile Trp Leu Phe Tyr Phe Ser  
 145  
 aag atc atg gag ttt gtc gac acc atg atc atg gtc ctg aag aag aac 528  
 Lys Ile Met Gu Phe Val Asp Thr Met Ile Met Val Leu Lys Lys Asn  
 165  
 aac cgc cag atc tcc ttc ttg cac gtt tac cac cac agc tcc atc ttc 576  
 Asn Arg Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Phe  
 180  
 acc atc tgg tgg ttg gtc acc ttt gtt gca ccc aac ggt gaa gcc tac 624  
 Thr Ile Trp Trp Leu Val Thr Phe Val Ala Pro Asn Gly Gu Ala Tyr  
 195  
 ttc tct gct gcg ttg aac tgc ttc atc cat gtg atc atg tac ggc tac 672  
 Phe Ser Ala Ala Leu Asn Ser Phe Ile His Val Ile Met Tyr Gly Tyr  
 210  
 tac ttc ttg tgc gcc ttg ggc ttc aag cag gtg tgc ttc atc aag ttc 720

PF58307.txt

Tyr 225	Phe	Leu	Ser	Ala	Leu 230	Gly	Phe	Lys	Gln	Val 235	Ser	Phe	Ile	Lys	Phe 240	
tac	atc	acg	cgc	tcg	cag	atg	aca	cag	ttc	tgc	atg	atg	tcg	gtc	cag	768
Tyr	Ile	Thr	Arg	Ser 245	Gln	Met	Thr	Gln	Phe 250	Cys	Met	Met	Ser	Val 255	Gln	
tct	tcc	tgg	gac	atg	tac	gcc	atg	aag	gtc	ctt	ggc	cgc	ccc	gga	tac	816
Ser	Ser	Trp	Asp 260	Met	Tyr	Ala	Met	Lys 265	Val	Leu	Gly	Arg	Pro 270	Gly	Tyr	
ccc	ttc	ttc	atc	acg	gct	ctg	ctt	tgg	ttc	tac	atg	tgg	acc	atg	ctc	864
Pro	Phe	Phe 275	Ile	Thr	Ala	Leu	Leu 280	Trp	Phe	Tyr	Met	Trp 285	Thr	Met	Leu	
ggt	ctc	ttc	tac	aac	ttt	tac	aga	aag	aac	gcc	aag	ttg	gcc	aag	cag	912
Gly	Leu	Phe 290	Tyr	Asn	Phe 295	Tyr	Arg	Lys	Asn	Ala 300	Lys	Leu	Ala	Lys	Gln	
gcc	aag	gcc	gac	gct	gcc	aag	gag	aag	gca	agg	aag	ttg	cag	t aa		957
Ala	Lys	Ala	Asp	Ala	Ala 310	Lys	Gu	Lys	Ala	Arg 315	Lys	Leu	Gln			

<210> 154

<211> 318

<212> PRT

<213> Mbrt i e r e l l a a l p i n a

<400> 154

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

<210> 155

<211> 236

<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (2)..(8)  
<223> xaa in position 2 to 8 is any amino acid

<220>  
<221> Variant  
<222> (10)..(10)  
<223> xaa in position 10 is any amino acid

<220>  
<221> Variant  
<222> (13)..(14)  
<223> xaa in position 13 to 14 is any amino acid

<220>  
<221> Variant  
<222> (16)..(36)  
<223> xaa in position 16 to 36 is any amino acid

<220>  
<221> Variant  
<222> (37)..(43)  
<223> xaa in position 37 to 43 is any or no amino acid

<220>  
<221> Variant  
<222> (46)..(47)  
<223> xaa in position 46 to 47 is any amino acid

<220>  
<221> Variant  
<222> (49)..(50)  
<223> xaa in position 49 to 50 is any amino acid

<220>  
<221> Variant  
<222> (53)..(53)  
<223> xaa in position 53 is any amino acid

<220>  
<221> Variant  
<222> (56)..(57)  
<223> xaa in position 56 to 57 is any amino acid

<220>  
<221> Variant  
<222> (59)..(62)  
<223> xaa in position 59 to 62 is any amino acid

<220>  
<221> Variant  
<222> (64)..(67)  
<223> xaa in position 64 to 67 is any amino acid

<220>  
<221> Variant  
<222> (69)..(69)  
<223> xaa in position 69 is any amino acid

<220>  
<221> Variant  
<222> (71)..(72)  
<223> xaa in position 71 to 72 is any amino acid

<220>  
<221> Variant  
<222> (74)..(82)  
<223> xaa in position 74 to 82 is any amino acid

<220>  
<221> Variant  
<222> (83)..(86)  
<223> xaa in position 83 to 86 is any or no amino acid

<220>  
<221> Variant  
<222> (88)..(89)  
<223> xaa in position 88 to 89 is any amino acid

<220>

<221> Variant  
<222> (91)..(92)  
<223> xaa in position 91 to 92 is any amino acid  
<220>  
<221> Variant  
<222> (95)..(95)  
<223> xaa in position 95 is any amino acid  
<220>  
<221> Variant  
<222> (98)..(99)  
<223> xaa in position 98 to 99 is any amino acid  
<220>  
<221> Variant  
<222> (102)..(102)  
<223> xaa in position 102 is any amino acid  
<220>  
<221> Variant  
<222> (105)..(105)  
<223> xaa in position 105 is any amino acid  
<220>  
<221> Variant  
<222> (108)..(108)  
<223> xaa in position 108 is any amino acid  
<220>  
<221> Variant  
<222> (111)..(114)  
<223> xaa in position 111 to 114 is any amino acid  
<220>  
<221> Variant  
<222> (116)..(116)  
<223> xaa in position 116 is any amino acid  
<220>  
<221> Variant  
<222> (121)..(121)  
<223> xaa in position 121 is any amino acid  
<220>  
<221> Variant  
<222> (125)..(126)  
<223> xaa in position 125 to 126 is any amino acid  
<220>  
<221> Variant  
<222> (128)..(129)  
<223> xaa in position 128 to 129 is any amino acid  
<220>  
<221> Variant  
<222> (133)..(137)  
<223> xaa in position 133 to 137 is any amino acid  
<220>  
<221> Variant  
<222> (142)..(142)  
<223> xaa in position 142 is any amino acid  
<220>  
<221> Variant  
<222> (145)..(145)  
<223> xaa in position 145 is any amino acid  
<220>  
<221> Variant  
<222> (148)..(148)  
<223> xaa in position 148 is any amino acid  
<220>  
<221> Variant  
<222> (152)..(152)  
<223> xaa in position 152 is any amino acid  
<220>  
<221> Variant  
<222> (156)..(156)  
<223> xaa in position 156 is any amino acid  
<220>



<221> Variant  
<222> (159)..(159)  
<223> xaa in position 159 is any amino acid  
<220>  
<221> Variant  
<222> (162)..(162)  
<223> xaa in position 162 is any amino acid  
<220>  
<221> Variant  
<222> (164)..(165)  
<223> xaa in position 164 to 165 is any amino acid  
<220>  
<221> Variant  
<222> (167)..(178)  
<223> xaa in position 167 to 178 is any amino acid  
<220>  
<221> Variant  
<222> (179)..(184)  
<223> xaa in position 179 to 184 is any or no amino acid  
<220>  
<221> Variant  
<222> (186)..(187)  
<223> xaa in position 186 to 187 is any amino acid  
<220>  
<221> Variant  
<222> (190)..(190)  
<223> xaa in position 190 is any amino acid  
<220>  
<221> Variant  
<222> (193)..(197)  
<223> xaa in position 193 to 197 is any amino acid  
<220>  
<221> Variant  
<222> (199)..(200)  
<223> xaa in position 199 to 200 is any amino acid  
<220>  
<221> Variant  
<222> (202)..(207)  
<223> xaa in position 202 to 207 is any amino acid  
<220>  
<221> Variant  
<222> (208)..(213)  
<223> xaa in position 208 to 213 is any or no amino acid  
<220>  
<221> Variant  
<222> (216)..(217)  
<223> xaa in position 216 to 217 is any amino acid  
<220>  
<221> Variant  
<222> (219)..(221)  
<223> xaa in position 219 to 221 is any amino acid  
<220>  
<221> Variant  
<222> (223)..(224)  
<223> xaa in position 223 to 224 is any amino acid  
<220>  
<221> Variant  
<222> (227)..(229)  
<223> xaa in position 227 to 229 is any amino acid  
<220>  
<221> Variant  
<222> (231)..(231)  
<223> xaa in position 231 is any amino acid  
<220>  
<221> Variant  
<222> (235)..(235)  
<223> xaa in position 235 is any amino acid

PF58307.txt

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

<210> 156  
 <211> 59  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is any amino acid  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Phe or Ile  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Ile, Leu or Val  
 <220>  
 <221> Variant  
 <222> (10)..(13)  
 <223> xaa in position 10 to 13 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Lys or Arg  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Ile, Leu or Val  
 <220>  
 <221> Variant

<222> (21)..(21)  
<223> xaa in position 21 is Ile or Val  
<220>  
<221> Variant  
<222> (25)..(25)  
<223> xaa in position 25 is Ala or Ser  
<220>  
<221> Variant  
<222> (26)..(26)  
<223> xaa in position 26 is Ser or Thr  
<220>  
<221> Variant  
<222> (28)..(29)  
<223> xaa in position 28 to 29 is any amino acid  
<220>  
<221> Variant  
<222> (32)..(32)  
<223> xaa in position 32 is any amino acid  
<220>  
<221> Variant  
<222> (33)..(33)  
<223> xaa in position 33 is Ala, Ile or Leu  
<220>  
<221> Variant  
<222> (34)..(34)  
<223> xaa in position 34 is Ile or Val  
<220>  
<221> Variant  
<222> (35)..(35)  
<223> xaa in position 35 is Ala or Thr  
<220>  
<221> Variant  
<222> (36)..(37)  
<223> xaa in position 36 to 37 is any amino acid  
<220>  
<221> Variant  
<222> (40)..(40)  
<223> xaa in position 40 is Gly or Asn  
<220>  
<221> Variant  
<222> (42)..(42)  
<223> xaa in position 42 is Asp or Glu  
<220>  
<221> Variant  
<222> (45)..(45)  
<223> xaa in position 45 is Phe or Trp  
<220>  
<221> Variant  
<222> (47)..(47)  
<223> xaa in position 47 is Ala or Thr  
<220>  
<221> Variant  
<222> (48)..(48)  
<223> xaa in position 48 is Ala or Ile  
<220>  
<221> Variant  
<222> (52)..(52)  
<223> xaa in position 52 is any amino acid  
<220>  
<221> Variant  
<222> (53)..(53)  
<223> xaa in position 53 is Ile or Val  
<220>  
<221> Variant  
<222> (56)..(56)  
<223> xaa in position 56 is any amino acid  
<220>  
<221> Variant

<222> (57)..(57)  
 <223> xaa in position 57 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (59)..(59)  
 <223> xaa in position 59 is Ala, Gly or Thr

<400> 156  
 Phe xaa Asp Thr xaa xaa Met xaa Leu xaa xaa xaa xaa xaa Gln xaa  
 1 5 10 15  
 Ser Phe Leu His xaa Tyr His His xaa xaa Ile xaa xaa Ile Trp xaa  
 20 25 30  
 xaa xaa xaa xaa xaa Ala Pro xaa Gly xaa Ala Tyr xaa Ser xaa xaa  
 35 40 45  
 Leu Asn Ser xaa xaa His Val xaa xaa Tyr xaa  
 50 55

<210> 157  
 <211> 22  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Ala, Cys or Ser  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is any amino acid  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Gu or Gly  
 <220>  
 <221> Variant  
 <222> (8)..(10)  
 <223> xaa in position 8 to 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is Gln or Ser  
 <220>  
 <221> Variant  
 <222> (13)..(16)  
 <223> xaa in position 13 to 16 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Phe or Trp  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Gly or Asn

<400> 157  
 Ser xaa Tyr Met xaa xaa xaa xaa xaa xaa xaa Ala xaa xaa xaa xaa  
 1 5 10 15  
 Tyr xaa Leu xaa xaa Asn

<210> 158  
 <211> 16  
 <212> PRT  
 <213> Artificial sequence  
  
 <220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(8)  
 <223> xaa in position 6 to 8 is any amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Trp or Tyr  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Cys or Asp  
  
 <400> 158  
 G n xaa xaa G n Phe xaa xaa xaa xaa Val xaa G n xaa xaa xaa xaa  
 1 5 10 15

<210> 159  
 <211> 24  
 <212> PRT  
 <213> Artificial sequence  
  
 <220>  
 <221> Variant  
 <222> (3)..(4)  
 <223> xaa in position 3 to 4 is any amino acid  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (6)..(7)  
 <223> xaa in position 6 to 7 is any amino acid  
 Seite 199

<220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (10)..(11)  
 <223> xaa in position 10 to 11 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Ser or Thr  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (21)..(22)  
 <223> xaa in position 21 to 22 is any amino acid  
 <220>  
 <221> Variant  
 <222> (24)..(24)  
 <223> xaa in position 24 is Phe or Tyr  
 <400> 159  
 Tyr Pro xaa xaa xaa xaa xaa xaa Leu xaa xaa Tyr Met xaa xaa xaa  
 1 5 10 15  
 Leu xaa Leu Phe xaa xaa Phe xaa  
 20

<210> 160  
 <211> 8  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Trp or Tyr  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Phe or Val

<400> 160  
 Ile xaa xaa Phe Tyr xaa Ser Lys  
 1 5

<210> 161  
 <211> 755  
 <212> DNA  
 <213> Cauliflower mosaic virus

<400> 161

PF58307. txt

```

cct gcaggt c aacat ggt gg agcacgacac act t gt ct ac t ccaaaaat a t caaagat ac 60
agt ct cagaa gaccaaagg caat t gagac t t t t caacaa aggggt aat at ccggaaacct 120
cct cggat t c cat t gccag ct at ct gt ca ct t t at t gt g aagat agt gg aaaaggaagg 180
t ggct cct ac aaat gccat c at t gcgat aa aggaaaggcc at cgt t gaag at gcct ct gc 240
cgacagt ggt cccaaagat g gacccccacc cacgaggagc at cgt ggaaa aagaagacgt 300
t ccaaccacg t ct t caaagc aagt ggat t g at gt gat aac at ggt ggagc acgacacact 360
t gt ct act cc aaaaat at ca aagat acagt ct cagaagac caaagggcaa t t gagact t t 420
t caacaagg gt aat at ccg gaaacct cct cggat t ccat t gccagct a t ct gt cact t 480
t at t gt gaag at agt ggaaa aggaaggt gg ct cct acaaa t gccat cat t gcgat aaagg 540
aaaggccat c gt t gaagat g cct ct gccga cagt ggt ccc aaagat ggac ccccaccac 600
gaggagcat c gt ggaaaaag aagacgt t cc aaccacgt ct t caaagcaag t ggat t gat g 660
t gat at ct cc act gacgt aa gggat gacgc acaat cccac t at cct t cgc aagaccct t c 720
ct ct at at aa ggaagt t cat t t cat t t gga gagga 755

```

<210> 162  
 <211> 211  
 <212> DNA  
 <213> Cauliflower mosaic virus

```

<400> 162
agt ccgcaa aat caccagt ct ct ct ct ac aat ct at ct ct ct ct at t t t ct ccagaa 60
t aat gt gt ga gt agt t ccca gat aagggaa t t aggggt t ct t at aggggt t t cgct cat gt g 120
t t gagcat at aagaaacct t agt at gt at t t gt at t t gt aaaat act t c t at caat aaa 180
at t t ct aat t cct aaaacca aaat ccagt g a 211

```

<210> 163  
 <211> 819  
 <212> DNA  
 <213> Thalassiosira pseudonana

<220>  
 <221> CDS  
 <222> (1)..(819)

```

<400> 163
at g gac gcc t ac aac gct gca at g gat aag at c ggt gcc gcc at c at c 48
Met Asp Ala Tyr Asn Ala Ala Met Asp Lys Ile Gly Ala Ala Ile Ile
1 5 10 15
gat t gg t ct gat ccc gat gga aag t t c cgt gcc gat aga gag gac t gg 96
Asp Trp Ser Asp Pro Asp Gly Lys Phe Arg Ala Asp Arg Glu Asp Trp
20 25 30
t gg ct c t gc gac t t c cgt agc gcc at c acc at c gcc ct c at c t ac at c 144
Trp Leu Cys Asp Phe Arg Ser Ala Ile Thr Ile Ala Leu Ile Tyr Ile
35 40 45
gcc t t c gt c at c ct c ggt t cc gcc gt c at g caa t cc ct c ccc gca at g 192
Ala Phe Val Ile Leu Gly Ser Ala Val Met Gl n Ser Leu Pro Ala Met
50 55 60
gat ccc t ac ccc at c aaa t t c ct c t ac aac gt c t cc caa at c t t c ct t 240
Asp Pro Tyr Pro Ile Lys Phe Leu Tyr Asn Val Ser Gl n Ile Phe Leu
65 70 75 80
t gt gcc t ac at g act gt c gag gcg gga t t t t t g gcc t ac cgc aat gga 288
Cys Ala Tyr Met Thr Val Gu Ala Gly Phe Leu Ala Tyr Arg Asn Gly
85 90 95
t at acc gt c at g cct t gc aat cat t t c aat gt g aat gat cct ccc gt g 336
Tyr Thr Val Met Pro Cys Asn His Phe Asn Val Asn Asp Pro Pro Val
100 105 110
gcg aat ct t ct t t gg t t g t t t t at at t t cc aag gt g t gg gac t t t t gg 384
Ala Asn Leu Leu Trp Leu Phe Tyr Ile Ser Lys Val Trp Asp Phe Trp
115 120 125
gat acc att t t c att gt g t t g gg aag aag t gg cgt caa t t a t ct t t c 432
Asp Thr Ile Phe Ile Val Leu Gly Lys Lys Trp Arg Gl n Leu Ser Phe
130 135 140
t t g cat gt a t ac cat cac acc acc at c t t t ct a t t c t at t gg ct g aat 480
Leu His Val Tyr His Thr Thr Ile Phe Leu Phe Tyr Trp Leu Asn
145 150 155 160
gcc aat gt c t t g t ac gat ggt gac at c t t c ct t acc at c t t g ct c aat 528
Ala Asn Val Leu Tyr Asp Gly Asp Ile Phe Leu Thr Ile Leu Leu Asn
165 170 175

```

PF58307. txt

gga ttc atc cac acg gtg atg tac acg tat tac ttc atc tgt atg cat	576
Gly Phe Ile His Thr Val Met Tyr Thr Tyr Tyr Phe Ile Cys Met His	
180 185 190	
acc aaa gat tcc aag acg ggc aag agt ctt cct ata tgg tgg aag tcc	624
Thr Lys Asp Ser Lys Thr Gly Lys Ser Leu Pro Ile Trp Trp Lys Ser	
195 200 205	
agt ttg acg gcg ttt cag ttg ttg caa ttc act atc atg atg agt cag	672
Ser Leu Thr Ala Phe Gln Leu Leu Gln Phe Thr Ile Met Met Ser Gln	
210 215 220	
gct acc tac ctt gtc ttc cac ggg tgt gat aag gtg tcc cgt atc	720
Ala Thr Tyr Leu Val Phe His Gly Cys Asp Lys Val Ser Leu Arg Ile	
225 230 235 240	
acg att gtg tac ttt gtg tcc ctt ttg agt ttg ttc ttc ctt ttt gct	768
Thr Ile Val Tyr Phe Val Ser Leu Leu Ser Leu Phe Phe Leu Phe Ala	
245 250 255	
cag ttc ttt gtg caa tca tac atg gca ccc aaa aag aag aag agt gct	816
Gln Phe Phe Val Gln Ser Tyr Met Ala Pro Lys Lys Lys Lys Ser Ala	
260 265 270	
t ag	819

<210> 164  
 <211> 272  
 <212> PRT  
 <213> Thal assi osi r a pseudonana

<400> 164	
Met Asp Ala Tyr Asn Ala Ala Met Asp Lys Ile Gly Ala Ala Ile Ile	
1 5 10 15	
Asp Trp Ser Asp Pro Asp Gly Lys Phe Arg Ala Asp Arg Glu Asp Trp	
20 25 30	
Trp Leu Cys Asp Phe Arg Ser Ala Ile Thr Ile Ala Leu Ile Tyr Ile	
35 40 45	
Ala Phe Val Ile Leu Gly Ser Ala Val Met Gln Ser Leu Pro Ala Met	
50 55 60	
Asp Pro Tyr Pro Ile Lys Phe Leu Tyr Asn Val Ser Gln Ile Phe Leu	
65 70 75 80	
Cys Ala Tyr Met Thr Val Glu Ala Gly Phe Leu Ala Tyr Arg Asn Gly	
85 90 95	
Tyr Thr Val Met Pro Cys Asn His Phe Asn Val Asn Asp Pro Pro Val	
100 105 110	
Ala Asn Leu Leu Trp Leu Phe Tyr Ile Ser Lys Val Trp Asp Phe Trp	
115 120 125	
Asp Thr Ile Phe Ile Val Leu Gly Lys Lys Trp Arg Gln Leu Ser Phe	
130 135 140	
Leu His Val Tyr His His Thr Thr Ile Phe Leu Phe Tyr Trp Leu Asn	
145 150 155 160	
Ala Asn Val Leu Tyr Asp Gly Asp Ile Phe Leu Thr Ile Leu Leu Asn	
165 170 175	
Gly Phe Ile His Thr Val Met Tyr Thr Tyr Tyr Phe Ile Cys Met His	
180 185 190	
Thr Lys Asp Ser Lys Thr Gly Lys Ser Leu Pro Ile Trp Trp Lys Ser	
195 200 205	
Ser Leu Thr Ala Phe Gln Leu Gln Phe Thr Ile Met Met Ser Gln	
210 215 220	
Ala Thr Tyr Leu Val Phe His Gly Cys Asp Lys Val Ser Leu Arg Ile	
225 230 235 240	
Thr Ile Val Tyr Phe Val Ser Leu Leu Ser Leu Phe Phe Leu Phe Ala	
245 250 255	
Gln Phe Phe Val Gln Ser Tyr Met Ala Pro Lys Lys Lys Lys Ser Ala	
260 265 270	

<210> 165  
 <211> 837  
 <212> DNA  
 <213> Phaeodactyl um tri cor nut um



<220>  
 <221> CDS  
 <222> (1).. (837)

<400> 165  
 at g at g gt a cct t ca agt t at gac gag t at at c gt c at g gt c aac gac 48  
 Met Met Val Pro Ser Ser Tyr Asp Gl u Tyr Ile Val Met Val Asn Asp  
 1 5 10 15  
 ct t ggc gac t ct att ct g agc t gg gcc gac cct gat cac t at cg t gga 96  
 Leu Gly Asp Ser Ile Leu Ser Trp Ala Asp Pro Asp His Tyr Arg Gly  
 20 25 30  
 cat acc gag gga t gg gag t t c act gac t t t t ct gct gct t t t agc at t 144  
 His Thr Gl u Gly Trp Gl u Phe Thr Asp Phe Ser Ala Ala Phe Ser Ile  
 35 40 45  
 gcc gt c gcg t ac ct c ct g t t t gt c t t t gt t gga t ct ct c at t at g agt 192  
 Ala Val Ala Tyr Leu Leu Phe Val Phe Val Gly Ser Leu Ile Met Ser  
 50 55 60  
 at g gga gt c ccc gca at t gac cct t at ccg ct c aag t t t gt c t ac aat 240  
 Met Gly Val Pro Ala Ile Asp Pro Tyr Pro Leu Lys Phe Val Tyr Asn  
 65 70 75 80  
 gt t t ca cag att at g ct t t gt gct t ac at g acc at t gaa gcc agt ct t 288  
 Val Ser Gl n Ile Met Leu Cys Ala Tyr Met Thr Ile Gl u Ala Ser Leu  
 85 90 95  
 ct a gct t at cg t aac ggc t ac aca t t c t gg cct t gc aac gat t gg gac 336  
 Leu Ala Tyr Arg Asn Gly Tyr Thr Phe Trp Pro Cys Asn Asp Trp Asp  
 100 105 110  
 t t t gaa aag ccg cct at c gct aag ct c ct c t gg ct c t t t t ac gt t t cc 384  
 Phe Gl u Lys Pro Pro Ile Ala Lys Leu Leu Trp Leu Phe Tyr Val Ser  
 115 120 125  
 aaa att t gg gat t t t t gg gac acc at c t t t at t gt t ct c ggg aag aag 432  
 Lys Ile Trp Asp Phe Trp Asp Thr Ile Phe Ile Val Leu Gly Lys Lys  
 130 135 140  
 t gg cg t caa ct t t cc t t c ct g cac gt c t ac cat cac acc acc at c t t t 480  
 Trp Arg Gl n Leu Ser Phe Leu His Val Tyr His His Thr Thr Ile Phe  
 145 150 155 160  
 ct c t t c t ac t gg t t g aat gca cat gt a aac t t t gat ggt gat at t t t c 528  
 Leu Phe Tyr Trp Leu Asn Ala His Val Asn Phe Asp Gly Asp Ile Phe  
 165 170 175  
 ct c acc at c gt c t t g aac ggt t t c at c cac acc gt c at g t ac acg t ac 576  
 Leu Thr Ile Val Leu Asn Gly Phe Ile His Thr Val Met Tyr Thr Tyr  
 180 185 190  
 t ac t t c att t gc at g cac acc aag gt c cca gag acc ggc aaa t cc t t g 624  
 Tyr Phe Ile Cys Met His Thr Lys Val Pro Gl u Thr Gly Lys Ser Leu  
 195 200 205  
 ccc att t gg t gg aaa t ct agt t t g aca agc at g cag ct g gt g cag t t c 672  
 Pro Ile Trp Trp Lys Ser Ser Leu Thr Ser Met Gl n Leu Val Gl n Phe  
 210 215 220  
 at c acg at g at g acg cag gct at c at g at c t t g t ac aag ggc t gt gct 720  
 Ile Thr Met Met Thr Gl n Ala Ile Met Ile Leu Tyr Lys Gly Cys Ala  
 225 230 235 240  
 gct ccc cat agc cgg gt g gt g aca t cg t ac t t g gt t t ac at t t t g t cg 768  
 Ala Pro His Ser Arg Val Val Thr Ser Tyr Leu Val Tyr Ile Leu Ser  
 245 250 255  
 ct c t t t att t t g t t c gcc cag t t c t t t gt c agc t ca t ac ct c aag ccg 816  
 Leu Phe Ile Leu Phe Ala Gl n Phe Phe Val Ser Ser Tyr Leu Lys Pro  
 260 265 270  
 aag aag aag aag aca gct t aa 837  
 Lys Lys Lys Lys Thr Ala

<210> 166  
 <211> 278  
 <212> PRT  
 <213> Phaeodactylum tricornutum

<400> 166  
 Met Met Val Pro Ser Ser Tyr Asp Gl u Tyr Ile Val Met Val Asn Asp  
 1 5 10 15

PF58307.txt

Leu Gly Asp Ser Ile Leu Ser Trp Ala Asp Pro Asp His Tyr Arg Gly  
 His Thr Glu Gly Trp Gu Phe Thr Asp Phe Ser Ala Ala Phe Ser Ile  
 Ala Val Ala Tyr Leu Leu Phe Val Phe Val Gly Ser Leu Ile Met Ser  
 Met Gly Val Pro Ala Ile Asp Pro Tyr Pro Leu Lys Phe Val Tyr Asn  
 Val Ser Gln Ile Met Leu Cys Ala Tyr Met Thr Ile Glu Ala Ser Leu  
 Leu Ala Tyr Arg Asn Gly Tyr Thr Phe Trp Pro Cys Asn Asp Trp Asp  
 Phe Gu Lys Pro Pro Ile Ala Lys Leu Leu Trp Leu Phe Tyr Val Ser  
 Lys Ile Trp Asp Phe Trp Asp Thr Ile Phe Ile Val Leu Gly Lys Lys  
 Trp Arg Gln Leu Ser Phe Leu His Val Tyr His His Thr Thr Ile Phe  
 Leu Phe Tyr Trp Leu Asn Ala His Val Asn Phe Asp Gly Asp Ile Phe  
 Leu Thr Ile Val Leu Asn Gly Phe Ile His Thr Val Met Tyr Thr Tyr  
 Tyr Phe Ile Cys Met His Thr Lys Val Pro Gu Thr Gly Lys Ser Leu  
 Pro Ile Trp Trp Lys Ser Ser Leu Thr Ser Met Gln Leu Val Gln Phe  
 Ile Thr Met Met Thr Gln Ala Ile Met Ile Leu Tyr Lys Gly Cys Ala  
 Ala Pro His Ser Arg Val Val Thr Ser Tyr Leu Val Tyr Ile Leu Ser  
 Leu Phe Ile Leu Phe Ala Gln Phe Phe Val Ser Ser Tyr Leu Lys Pro  
 Lys Lys Lys Lys Thr Ala  
 20  
 35  
 50  
 65  
 80  
 95  
 110  
 125  
 140  
 155  
 170  
 185  
 200  
 215  
 230  
 245  
 260  
 275

<210> 167  
 <211> 272  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is any amino acid  
 <220>  
 <221> Variant  
 <222> (4)..(10)  
 <223> xaa in position 4 to 10 is any amino acid  
 <220>  
 <221> Variant  
 <222> (12)..(13)  
 <223> xaa in position 12 to 13 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(16)  
 <223> xaa in position 15 to 16 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (22)..(23)  
 <223> xaa in position 22 to 23 is any amino acid  
 <220>  
 <221> Variant  
 <222> (24)..(24)

<223> xaa in position 24 is any or no amino acid  
<220>  
<221> Variant  
<222> (26)..(28)  
<223> xaa in position 26 to 28 is any amino acid  
<220>  
<221> Variant  
<222> (30)..(30)  
<223> xaa in position 30 is any amino acid  
<220>  
<221> Variant  
<222> (32)..(34)  
<223> xaa in position 32 to 34 is any amino acid  
<220>  
<221> Variant  
<222> (37)..(38)  
<223> xaa in position 37 to 38 is any amino acid  
<220>  
<221> Variant  
<222> (40)..(41)  
<223> xaa in position 40 to 41 is any amino acid  
<220>  
<221> Variant  
<222> (44)..(45)  
<223> xaa in position 44 to 45 is any amino acid  
<220>  
<221> Variant  
<222> (47)..(48)  
<223> xaa in position 47 to 48 is any amino acid  
<220>  
<221> Variant  
<222> (51)..(52)  
<223> xaa in position 51 to 52 is any amino acid  
<220>  
<221> Variant  
<222> (55)..(56)  
<223> xaa in position 55 to 56 is any amino acid  
<220>  
<221> Variant  
<222> (58)..(60)  
<223> xaa in position 58 to 60 is any amino acid  
<220>  
<221> Variant  
<222> (61)..(61)  
<223> xaa in position 61 is any or no amino acid  
<220>  
<221> Variant  
<222> (64)..(64)  
<223> xaa in position 64 is any amino acid  
<220>  
<221> Variant  
<222> (69)..(69)  
<223> xaa in position 69 is any amino acid  
<220>  
<221> Variant  
<222> (72)..(72)  
<223> xaa in position 72 is any amino acid  
<220>  
<221> Variant  
<222> (79)..(79)  
<223> xaa in position 79 is any amino acid  
<220>  
<221> Variant  
<222> (86)..(86)  
<223> xaa in position 86 is any amino acid  
<220>  
<221> Variant  
<222> (89)..(90)

<223> xaa in position 89 to 90 is any amino acid  
<220>  
<221> Variant  
<222> (99)..(100)  
<223> xaa in position 99 to 100 is any amino acid  
<220>  
<221> Variant  
<222> (104)..(109)  
<223> xaa in position 104 to 109 is any amino acid  
<220>  
<221> Variant  
<222> (112)..(112)  
<223> xaa in position 112 is any amino acid  
<220>  
<221> Variant  
<222> (114)..(114)  
<223> xaa in position 114 is any amino acid  
<220>  
<221> Variant  
<222> (121)..(121)  
<223> xaa in position 121 is any amino acid  
<220>  
<221> Variant  
<222> (124)..(124)  
<223> xaa in position 124 is any amino acid  
<220>  
<221> Variant  
<222> (162)..(162)  
<223> xaa in position 162 is any amino acid  
<220>  
<221> Variant  
<222> (164)..(165)  
<223> xaa in position 164 to 165 is any amino acid  
<220>  
<221> Variant  
<222> (174)..(174)  
<223> xaa in position 174 is any amino acid  
<220>  
<221> Variant  
<222> (195)..(197)  
<223> xaa in position 195 to 197 is any amino acid  
<220>  
<221> Variant  
<222> (212)..(213)  
<223> xaa in position 212 to 213 is any amino acid  
<220>  
<221> Variant  
<222> (216)..(216)  
<223> xaa in position 216 is any amino acid  
<220>  
<221> Variant  
<222> (219)..(220)  
<223> xaa in position 219 to 220 is any amino acid  
<220>  
<221> Variant  
<222> (223)..(223)  
<223> xaa in position 223 is any amino acid  
<220>  
<221> Variant  
<222> (226)..(231)  
<223> xaa in position 226 to 231 is any amino acid  
<220>  
<221> Variant  
<222> (234)..(238)  
<223> xaa in position 234 to 238 is any amino acid  
<220>  
<221> Variant  
<222> (240)..(243)

<223> xaa in position 240 to 243 is any amino acid  
 <220>  
 <221> Variant  
 <222> (245)..(245)  
 <223> xaa in position 245 is any amino acid  
 <220>  
 <221> Variant  
 <222> (247)..(248)  
 <223> xaa in position 247 to 248 is any amino acid  
 <220>  
 <221> Variant  
 <222> (253)..(253)  
 <223> xaa in position 253 is any amino acid  
 <220>  
 <221> Variant  
 <222> (261)..(261)  
 <223> xaa in position 261 is any amino acid  
 <220>  
 <221> Variant  
 <222> (264)..(265)  
 <223> xaa in position 264 to 265 is any amino acid  
 <220>  
 <221> Variant  
 <222> (271)..(271)  
 <223> xaa in position 271 is any amino acid

<400> 167  
 Asp xaa Tyr xaa xaa xaa xaa xaa xaa xaa Gly xaa xaa Ile xaa xaa  
 1 5 10 15  
 Trp xaa Asp Pro Asp xaa xaa xaa Arg xaa xaa xaa Gu xaa Trp xaa  
 20 25 30  
 xaa xaa Asp Phe xaa xaa Ala xaa xaa Ile Ala xaa xaa Tyr xaa xaa  
 35 40 45  
 Phe Val xaa xaa Gly Ser xaa xaa Met xaa xaa xaa xaa Pro Ala xaa  
 50 55 60  
 Asp Pro Tyr Pro xaa Lys Phe xaa Tyr Asn Val Ser G n Ile xaa Leu  
 65 70 75 80  
 Cys Ala Tyr Met Thr xaa Gu Ala xaa xaa Leu Ala Tyr Arg Asn Gly  
 85 90 95  
 Tyr Thr xaa xaa Pro Cys Asn xaa xaa xaa xaa xaa xaa Pro Pro xaa  
 100 105 110  
 Ala xaa Leu Leu Trp Leu Phe Tyr xaa Ser Lys xaa Trp Asp Phe Trp  
 115 120 125  
 Asp Thr Ile Phe Ile Val Leu Gly Lys Lys Trp Arg G n Leu Ser Phe  
 130 135 140  
 Leu Hi s Val Tyr Hi s Hi s Thr Thr Ile Phe Leu Phe Tyr Trp Leu Asn  
 145 150 155 160  
 Ala xaa Val xaa xaa Asp Gly Asp Ile Phe Leu Thr Ile xaa Leu Asn  
 165 170 175  
 Gly Phe Ile Hi s Thr Val Met Tyr Thr Tyr Tyr Phe Ile Cys Met Hi s  
 180 185 190  
 Thr Lys xaa xaa xaa Thr Gly Lys Ser Leu Pro Ile Trp Trp Lys Ser  
 195 200 205  
 Ser Leu Thr xaa xaa G n Leu xaa G n Phe xaa xaa Met Met xaa G n  
 210 215 220  
 Ala xaa xaa xaa xaa xaa xaa Gly Cys xaa xaa xaa xaa xaa Arg xaa  
 225 230 235 240  
 xaa xaa xaa Tyr xaa Val xaa xaa Leu Ser Leu Phe xaa Leu Phe Ala  
 245 250 255  
 G n Phe Phe Val xaa Ser Tyr xaa xaa Pro Lys Lys Lys Lys xaa Ala  
 260 265 270

<210> 168  
 <211> 59  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (27)..(27)  
 <223> xaa in position 27 is any amino acid  
 <220>  
 <221> Variant  
 <222> (29)..(29)  
 <223> xaa in position 29 is any amino acid  
 <220>  
 <221> Variant  
 <222> (30)..(34)  
 <223> xaa in position 30 to 34 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (39)..(39)  
 <223> xaa in position 39 is Leu or Val  
  
 <400> 168  
 G y Lys Lys Trp Arg G n Leu Ser Phe Leu H i s Val Tyr H i s H i s Thr  
   1                  5                  10                  15  
 Thr Ile Phe Leu Phe Tyr Trp Leu Asn Ala xaa Val xaa xaa xaa xaa  
           20                  25                  30  
 xaa xaa Phe Leu Thr Ile xaa Leu Asn G y Phe Ile H i s Thr Val Met  
           35                  40                  45  
 Tyr Thr Tyr Tyr Phe Ile Cys Met H i s Thr Lys  
       50                  55

<210> 169  
 <211> 60  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Phe or Met  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (28)..(28)  
 <223> xaa in position 28 is Phe or Val  
 <220>  
 <221> Variant  
 <222> (29)..(29)  
 <223> xaa in position 29 is any amino acid  
 <220>  
 <221> Variant  
 <222> (33)..(33)  
 <223> xaa in position 33 is Asp or H i s  
 <220>  
 <221> Variant

<222> (34)..(34)  
 <223> xaa in position 34 is Phe or Trp  
 <220>  
 <221> Variant  
 <222> (35)..(35)  
 <223> xaa in position 35 is Asp or Asn  
 <220>  
 <221> Variant  
 <222> (36)..(36)  
 <223> xaa in position 36 is Phe or Val  
 <220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is Gu or Asn  
 <220>  
 <221> Variant  
 <222> (38)..(38)  
 <223> xaa in position 38 is Asp or Lys  
 <220>  
 <221> Variant  
 <222> (41)..(41)  
 <223> xaa in position 41 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (43)..(43)  
 <223> xaa in position 43 is any amino acid  
 <220>  
 <221> Variant  
 <222> (50)..(50)  
 <223> xaa in position 50 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (53)..(53)  
 <223> xaa in position 53 is Ile or Val

<400> 169  
 Tyr Asn Val Ser G n Ile xaa Leu Cys Ala Tyr Met Thr xaa Glu Ala  
 1 5 10 15  
 xaa xaa Leu xaa Ala Tyr Arg Asn Gly Tyr Thr xaa xaa Pro Cys Asn  
 20 25 30  
 xaa xaa xaa xaa xaa xaa Pro Pro xaa Ala xaa Leu Leu Trp Leu Phe  
 35 40 45  
 Tyr xaa Ser Lys xaa Trp Asp Phe Trp Asp Thr Ile  
 50 55 60

<210> 170  
 <211> 39  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is Phe or Met  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any or no amino acid  
 <220>

<221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is any amino acid  
 <220>  
 <221> Variant  
 <222> (22)..(22)  
 <223> xaa in position 22 is any amino acid  
 <220>  
 <221> Variant  
 <222> (23)..(23)  
 <223> xaa in position 23 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (26)..(27)  
 <223> xaa in position 26 to 27 is any amino acid  
 <220>  
 <221> Variant  
 <222> (28)..(28)  
 <223> xaa in position 28 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (29)..(29)  
 <223> xaa in position 29 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (30)..(30)  
 <223> xaa in position 30 is Phe or Tyr  
 <220>  
 <221> Variant  
 <222> (31)..(31)  
 <223> xaa in position 31 is His or Lys  
 <220>  
 <221> Variant  
 <222> (34)..(34)  
 <223> xaa in position 34 is Ala or Asp  
 <220>  
 <221> Variant  
 <222> (35)..(35)  
 <223> xaa in position 35 is any amino acid  
 <220>  
 <221> Variant  
 <222> (36)..(36)  
 <223> xaa in position 36 is Pro or Val  
 <220>  
 <221> Variant  
 <222> (37)..(38)  
 <223> xaa in position 37 to 38 is any amino acid

<400> 170  
 Leu Pro Ile Trp Trp Lys Ser Ser Leu Thr xaa xaa Gln Leu xaa Gln  
 1 5 10 15  
 Phe xaa Ile xaa Met xaa xaa Gln Ala xaa xaa xaa xaa xaa xaa Gly  
 20 25 30  
 Cys xaa xaa xaa xaa xaa Arg  
 35

<210> 171  
 <211> 831  
 <212> DNA  
 <213> *Tr aust ochyt ri um sp.*

<220>  
 <221> CDS  
 <222> (1)..(831)

<400> 171  
 atg gac gtc gtc gag cag caa tgg cgc cgc ttc gtg gac gcc gtg gac  
 Met Asp Val Val Glu Gln Gln Trp Arg Arg Phe Val Asp Ala Val Asp  
 Seite 210



PF58307. txt

```

1          5          10          15
aac gga at c gt g gag t t c at g gag cat gag aag ccc aac aag ct g aac      96
Asn Gly Ile Val Gu Phe Met Gu His Gu Lys Pro Asn Lys Leu Asn
20
gag ggc aag ct c t t c acc t c g acc gag gag at g at g gcg ct t at c gt c      144
Gu Gy Lys Leu Phe Thr Ser Thr Gu Gu Met Met Ala Leu Ile Val
35
ggc t ac ct g gcg t t c gt g gt c ct c ggg t cc gcc t t c at g aag gcc t t t      192
Gy Tyr Leu Ala Phe Val Val Leu Gy Ser Ala Phe Met Lys Ala Phe
50
gt c gat aag cct t t c gag ct c aag t t c ct c aag ct c gt g cac aac at c      240
Val Asp Lys Pro Phe Gu Leu Lys Phe Leu Lys Leu Val His Asn Ile
65
t t c ct c acc ggt ct g t cc at g t ac at g gcc acc gag t gc gcg cgc cag      288
Phe Leu Thr Gy Leu Ser Met Tyr Met Ala Thr Gu Cys Ala Arg Gn
85
gca t ac ct c ggc ggc t ac aag ct c t t t ggc aac ccg at g gag aag ggc      336
Ala Tyr Leu Gy Gy Tyr Lys Leu Phe Gy Asn Pro Met Gu Lys Gy
100
acc gag t c g cac gcc ccg ggc at g gcc aac at c at c t ac at c t t c t ac      384
Thr Gu Ser His Ala Pro Gy Met Ala Asn Ile Ile Tyr Ile Phe Tyr
115
gt g agc aag t t c ct c gaa t t c ct c gac acc gt c t t c at g at c ct c ggc      432
Val Ser Lys Phe Leu Gu Phe Leu Asp Thr Val Phe Met Ile Leu Gy
130
aag aag t gg aag cag ct c agc t t t ct c cac gt c t ac cac cac gcg agc      480
Lys Lys Trp Lys Gn Leu Ser Phe Leu His Val Tyr His His Ala Ser
145
at c agc t t c at c t gg ggc at c at c gcc cgc t t c gcg ccc ggt ggc gac      528
Ile Ser Phe Ile Trp Gy Ile Ile Ala Arg Phe Ala Pro Gy Gy Asp
165
gcc t ac t t c t ct acc at c ct c aac agc agc gt g cat gt c gt g ct c t ac      576
Ala Tyr Phe Ser Thr Ile Leu Asn Ser Ser Val His Val Val Leu Tyr
180
ggc t ac t ac gcc t c g acc acc ct c ggc t ac acc t t c at g cgc ccg ct g      624
Gy Tyr Tyr Ala Ser Thr Thr Leu Gy Tyr Thr Phe Met Arg Pro Leu
195
cgc ccg t ac att acc acc att cag ct c acg cag t t c at g gcc at g gt c      672
Arg Pro Tyr Ile Thr Thr Ile Gn Leu Thr Gn Phe Met Ala Met Val
210
gt c cag t cc gt c t at gac t ac t ac aac ccc t gc gac t ac ccg cag ccc      720
Val Gn Ser Val Tyr Asp Tyr Tyr Asn Pro Cys Asp Tyr Pro Gn Pro
225
ct c gt c aag ct g ct c t t c t gg t ac at g ct c acc at g ct c ggc ct c t t c      768
Leu Val Lys Leu Leu Phe Trp Tyr Met Leu Thr Met Leu Gy Leu Phe
245
ggc aac t t c t t c gt g cag cag t ac ct c aag ccc aag gcg ccc aag aag      816
Gy Asn Phe Phe Val Gn Gn Tyr Leu Lys Pro Lys Ala Pro Lys Lys
260
cag aag acc at c t aa
Gn Lys Thr Ile
275

```

<210> 172  
 <211> 276  
 <212> PRT  
 <213> Traust ochyt r i u m sp.

```

<400> 172
Met Asp Val Val Gu Gn Gn Trp Arg Arg Phe Val Asp Ala Val Asp
1          5          10          15
Asn Gly Ile Val Gu Phe Met Gu His Gu Lys Pro Asn Lys Leu Asn
20
Gu Gy Lys Leu Phe Thr Ser Thr Gu Gu Met Met Ala Leu Ile Val
35
Gy Tyr Leu Ala Phe Val Val Leu Gy Ser Ala Phe Met Lys Ala Phe
50
Val Asp Lys Pro Phe Gu Leu Lys Phe Leu Lys Leu Val His Asn Ile
Sei te 211

```

PF58307. txt

65 Phe Leu Thr Gly Leu Ser Met Tyr Met Ala Thr Glu Cys Ala Arg Gln  
 70 75 80  
 85 Tyr Lys Leu Phe Gly Asn Pro Met Glu Lys Gly  
 90 95  
 100 Tyr Lys Leu Phe Gly Asn Ile Ile Tyr Ile Phe Tyr  
 105 110  
 115 Phe Leu Glu Phe Leu Asp Thr Val Phe Met Ile Leu Gly  
 120 125  
 130 Lys Lys Trp Lys Gln Leu Ser Phe Leu His Val Tyr His His Ala Ser  
 135 140 145  
 145 Ile Ser Phe Ile Trp Gly Ile Ile Ala Arg Phe Ala Pro Gly Gly Asp  
 150 155 160  
 165 Thr Ile Leu Asn Ser Val His Val Val Leu Tyr  
 170 175  
 180 Tyr Tyr Ala Ser Thr Thr Leu Gly Tyr Thr Phe Met Arg Pro Leu  
 185 190  
 195 Thr Thr Ile Gln Leu Thr Gln Phe Met Ala Met Val  
 200 205  
 210 Val Gln Ser Val Tyr Asp Tyr Tyr Asn Pro Cys Asp Tyr Pro Gln Pro  
 215 220 225  
 225 Leu Val Lys Leu Leu Phe Trp Tyr Met Leu Thr Met Leu Gly Leu Phe  
 230 235 240  
 245 Val Gln Gln Tyr Leu Lys Pro Lys Ala Pro Lys Lys  
 250 255 260  
 265 Gln Lys Thr Ile  
 270 275

<210> 173  
 <211> 1047  
 <212> DNA  
 <213> Marchantia polymorpha

<220>  
 <221> CDS  
 <222> (1)..(1047)

<400> 173  
 atg gcg acg aag agc ggg agc gga ttg ctg gag tgg ata gca gta gcg 48  
 Met Ala Thr Lys Ser Gly Ser Gly Leu Leu Glu Trp Ile Ala Val Ala  
 1 5 10 15  
 gcg aag atg aag caa gct cgg agc agc ccc gag ggt gag atc gtg ggt 96  
 Ala Lys Met Lys Gln Ala Arg Ser Ser Pro Glu Gly Glu Ile Val Gly  
 20 25 30  
 ggg aat agg atg ggc tct gga aac gga gct gag tgg acc acg agt ctg 144  
 Gly Asn Arg Met Gly Ser Gly Asn Gly Ala Glu Trp Thr Ser Leu  
 35 40 45  
 att cat gca ttt ttg aat gcc acg aat ggg aag agc ggc ggt gct tcg 192  
 Ile His Ala Phe Leu Asn Ala Thr Asn Gly Lys Ser Gly Gly Ala Ser  
 50 55 60  
 aaa gtg agg cct ctg gag gag aga atc ggg gag gcg gtg ttc aga gtt 240  
 Lys Val Arg Pro Leu Glu Glu Arg Ile Gly Glu Ala Val Phe Arg Val  
 65 70 75 80  
 ctt gaa gat gtc gtg ggc gtg gat att agg aag ccg aat cct gtc acg 288  
 Leu Glu Asp Val Val Gly Val Asp Ile Arg Lys Pro Asn Pro Val Thr  
 85 90  
 aag gac ctt ccg atg gtc gag agt ccc gtg ccc gtg ttg gcc tgc att 336  
 Lys Asp Leu Pro Met Val Glu Ser Pro Val Pro Val Leu Ala Cys Ile  
 100 105 110  
 tct ctg tac ttg ctg gtg gtg tgg ctt tgg tct tct cac att aag gcg 384  
 Ser Leu Tyr Leu Leu Val Val Trp Leu Trp Ser Ser His Ile Lys Ala  
 115 120 125  
 tct ggc caa aag ccc agg aag gag gac ccg ctg gcc ctg cgt tgc ctt 432  
 Ser Gly Gln Lys Pro Arg Lys Glu Asp Pro Leu Ala Leu Arg Cys Leu  
 130 135 140  
 gtg att gcc cac aat ctg ttc ctg tgt tgc ttg agc ttg ttc atg tgc 480  
 Val Ile Ala His Asn Leu Phe Leu Cys Cys Leu Ser Leu Phe Met Cys  
 145 150 155 160

PF58307. txt

gt c	ggt	ct c	att	gcc	gca	gct	cga	cat	t ac	ggg	t at	agt	gt a	t gg	ggg	528
Val	G y	Leu	I le	Al a	Al a	Al a	Arg	Hi s	Tyr	G y	Tyr	Ser	Val	Tr p	G y	
				165					170					175		
aac	t ac	t ac	aga	gaa	aga	gaa	ccc	gca	at g	aat	t t g	ct c	at t	t ac	gt g	576
Asn	Tyr	Tyr	Arg	G u	Arg	G u	Pro	Al a	Met	Asn	Leu	Leu	I le	Tyr	Val	
			180					185					190			
t t c	t ac	at g	t cg	aag	ct g	t ac	gaa	t t t	at g	gac	acg	gcc	at t	at g	t t a	624
Phe	Tyr	Met	Ser	Lys	Leu	Tyr	G u	Phe	Met	Asp	Thr	Al a	I le	Met	Leu	
			195				200					205				
t t c	aga	aga	aat	ct g	cga	caa	gt c	acg	t ac	t t g	cat	gt a	t at	cac	cac	672
Phe	Arg	Arg	Asn	Leu	Arg	G n	Val	Thr	Tyr	Leu	Hi s	Val	Tyr	Hi s	Hi s	
	210					215					220					
gca	agc	at c	gca	at g	at t	t gg	t gg	at a	at t	t gc	t at	cgg	t t t	cca	gga	720
Al a	Ser	I le	Al a	Met	I le	Tr p	Tr p	I le	I le	Cys	Tyr	Arg	Phe	Pro	G y	
	225				230					235					240	
gct	gat	t cg	t at	t t c	t cc	gca	gca	t t c	aat	t cc	t gt	at c	cat	gt a	gcg	768
Al a	Asp	Ser	Tyr	Phe	Ser	Al a	Al a	Phe	Asn	Ser	Cys	I le	Hi s	Val	Al a	
				245					250					255		
at g	t ac	ct g	t at	t at	ct a	ct c	gcg	gca	acc	gt c	gcc	aga	gac	gaa	aag	816
Met	Tyr	Leu	Tyr	Tyr	Leu	Leu	Al a	Al a	Thr	Val	Al a	Arg	Asp	G u	Lys	
			260						265				270			
cgg	aga	cgc	aaa	t at	ct c	t t c	t gg	gga	aag	t at	ct g	acc	at c	at a	caa	864
Arg	Arg	Arg	Lys	Tyr	Leu	Phe	Tr p	G y	Lys	Tyr	Leu	Thr	I le	I le	G n	
			275				280					285				
at g	ct t	cag	t t t	t t g	t cc	t t c	at t	ggg	cag	gcg	at t	t at	gca	at g	t gg	912
Met	Leu	G n	Phe	Leu	Ser	Phe	I le	G y	G n	Al a	I le	Tyr	Al a	Met	Tr p	
	290					295					300					
aag	t t t	gaa	t ac	t at	ccc	aag	ggc	t t t	ggc	agg	at g	t t g	t t c	t t t	t ac	960
Lys	Phe	G u	Tyr	Tyr	Pro	Lys	G y	Phe	G y	Arg	Met	Leu	Phe	Phe	Tyr	
	305				310					315					320	
t ct	gt a	t ca	t t g	t t g	gca	t t t	t t c	ggc	aac	t t c	t t t	gt c	aaa	aag	t at	1008
Ser	Val	Ser	Leu	Leu	Al a	Phe	Phe	G y	Asn	Phe	Phe	Val	Lys	Lys	Tyr	
				325					330					335		
t cg	aac	gct	t ca	cag	cct	aag	aca	gt t	aaa	gt g	gag	t ga				1047
Ser	Asn	Al a	Ser	G n	Pro	Lys	Thr	Val	Lys	Val	G u					
			340					345								

<210> 174  
 <211> 348  
 <212> PRT  
 <213> Marchantia polymorpha

<400> 174  
 Met Al a Thr Lys Ser G y Ser G y Leu Leu G u Tr p I le Al a Val Al a  
 1 5 10  
 Al a Lys Met Lys G n Al a Arg Ser Ser Pro G u G y G u I le Val G y  
 20 25 30  
 G y Asn Arg Met G y Ser G y Asn G y Al a G u Tr p Thr Thr Ser Leu  
 35 40 45  
 I le Hi s Al a Phe Leu Asn Al a Thr Asn G y Lys Ser G y G y Al a Ser  
 50 55 60  
 Lys Val Arg Pro Leu G u G u Arg I le G y G u Al a Val Phe Arg Val  
 65 70 75 80  
 Leu G u Asp Val Val G y Val Asp I le Arg Lys Pro Asn Pro Val Thr  
 85 90 95  
 Lys Asp Leu Pro Met Val G u Ser Pro Val Pro Val Leu Al a Cys I le  
 100 105 110  
 Ser Leu Tyr Leu Leu Val Val Tr p Leu Tr p Ser Ser Hi s I le Lys Al a  
 115 120 125  
 Ser G y G n Lys Pro Arg Lys G u Asp Pro Leu Al a Leu Arg Cys Leu  
 130 135 140  
 Val I le Al a Hi s Asn Leu Phe Leu Cys Cys Leu Ser Leu Phe Met Cys  
 145 150 155 160  
 Val G y Leu I le Al a Al a Arg Hi s Tyr G y Tyr Ser Val Tr p G y  
 165 170 175  
 Asn Tyr Tyr Arg G u Arg G u Pro Al a Met Asn Leu Leu I le Tyr Val  
 180 185 190  
 Phe Tyr Met Ser Lys Leu Tyr G u Phe Met Asp Thr Al a I le Met Leu

PF58307. txt

```

195          200          205
Phe Arg Arg Asn Leu Arg G n Val Thr Tyr Leu Hi s Val Tyr Hi s Hi s
  210          215          220
Al a Ser Ile Al a Met Ile Trp Trp Ile Ile Cys Tyr Arg Phe Pro Gly
  225          230          235
Al a Asp Ser Tyr Phe Ser Al a Al a Phe Asn Ser Cys Ile Hi s Val Al a
  245          250          255
Met Tyr Leu Tyr Tyr Leu Leu Al a Al a Thr Val Al a Arg Asp Gl u Lys
  260          265          270
Arg Arg Arg Lys Tyr Leu Phe Trp Gly Lys Tyr Leu Thr Ile Ile G n
  275          280          285
Met Leu G n Phe Leu Ser Phe Ile Gly G n Al a Ile Tyr Al a Met Trp
  290          295          300
Lys Phe Gl u Tyr Tyr Pro Lys Gly Phe Gly Arg Met Leu Phe Phe Tyr
  305          310          315
Ser Val Ser Leu Leu Al a Phe Phe Gly Asn Phe Phe Val Lys Lys Tyr
  325          330          335
Ser Asn Al a Ser G n Pro Lys Thr Val Lys Val Gl u
  340          345

```

<210> 175  
 <211> 831  
 <212> DNA  
 <213> Thr aust ochyt ri um sp

<220>  
 <221> CDS  
 <222> (1).. (831)

```

<400> 175
at g gat gt c gt c gag cag caa t gg cgc cgc t t c gt g gac gcc gt g gac      48
Met Asp Val Val 5 Gu G n G n Trp Arg Ar g Phe Val Asp Al a Val Asp
1          5          10
aac gga at c gt g gag t t c at g gag cat gag gag ccc aac aag ct g aac      96
Asn Gly Ile Val Gu Phe Met Gu Hi s Gu Gl u Pro Asn Lys Leu Asn
20          25
gag ggc aag ct c t cc acc t cg acc gag gag at g at g gcg ct t at c gt c      144
Gu Gy Lys Leu Ser Thr Ser Thr Gu Gu Met Met Al a Leu Ile Val
35          40          45
ggc t ac ct g gcg t t c gt g gt c ct c ggg t cc gcc t t c at g aag gcc t t t      192
Gy Tyr Leu Al a Phe Val 55 Leu Gy Ser Al a Phe Met Lys Al a Phe
50          55          60
gt c gat aag cct t t c gag ct c aag t t c ct c aag ct c gt g cac aac at c      240
Val Asp Lys Pro Phe Gu Leu Lys Phe Leu Lys Leu Val Hi s Asn Ile
65          70          75
t t c ct c acc ggt ct g t cc at g t ac at g gcc acc gag t gc gcg cgc cag      288
Phe Leu Thr Gy Leu Ser Met Tyr Met Al a Thr Gu Cys Al a Arg G n
85          90          95
gca t ac ct c ggc ggc t ac aag ct c t t t ggc aac ccg at g gag aag ggc      336
Al a Tyr Leu Gy Gy Tyr Lys Leu Phe Gy Asn Pro Met Gu Lys Gy
100          105          110
acc gag t cg cac gcc ccg ggc at g gcc aac at c at c t ac at c t t c t ac      384
Thr Gu Ser Hi s Al a Pro Gy Met Al a Asn Ile Ile Tyr Ile Phe Tyr
115          120          125
gt g agc aag t t c ct c gaa t t c ct c gac acc gt c t t c at g at c ct c ggc      432
Val Ser Lys Phe Leu Gu Phe Leu Asp Thr Val Phe Met Ile Leu Gy
130          135          140
aag aag t gg aag cag ct c agc t t t ct c cac gt c t ac cac cac gcg agc      480
Lys Lys Trp Lys G n Leu Ser Phe Leu Hi s Val Tyr Hi s Hi s Al a Ser
145          150          155
at c agc t t c at c t gg ggc at c at c gcc cgc t t c gcg ccc ggt ggc gac      528
Ile Ser Phe Ile Trp Gy Ile Ile Al a Arg Phe Al a Pro Gy Gy Asp
165          170          175
gcc t ac t t c t ct acc at c ct c aac agc agc gt g cat gt c gt g ct c t ac      576
Al a Tyr Phe Ser Thr Ile Leu Asn Ser Ser Val Hi s Val Val Val Leu Tyr
180          185          190
ggc t ac t ac gcc t cg acc acc ct c ggc t ac acc t t c at g cgc ccg ct g      624
Gy Tyr Tyr Al a Ser Thr Thr Leu Gy Tyr Thr Phe Met Arg Pro Leu
190          195

```

Seite 214

PF58307. txt

```

195          200          205
cgc ccg tac att acc acc att cag ct c acg cag t t c at g gcc at g gt c      672
Arg Pro Tyr lle Thr Thr lle G n Leu Thr G n Phe Met Ala Met Val
210
gt c cag tcc gt c tat gac tac tac aac ccc tgc gac tac ccg cag ccc      720
Val G n Ser Val Tyr Asp Tyr Tyr Asn Pro Cys Asp Tyr Pro G n Pro
225          230          235          240
ct c gt c aag ct g ct c t t c t gg tac at g ct c acc at g ct c ggc ct c t t c      768
Leu Val Lys Leu Leu Phe Trp Tyr Met Leu Thr Met Leu G y Leu Phe
245          250          255
ggc aac t t c t t c gt g cag cag tac ct c aag ccc aag gcg ccc aag aag      816
G y Asn Phe Phe Val G n G n Tyr Leu Lys Pro Lys Ala Pro Lys Lys
260          265          270
cag aag acc at c t aa
G n Lys Thr lle
275

```

<210> 176  
 <211> 276  
 <212> PRT  
 <213> Thr aust ochyt ri um sp

```

<400> 176
Met Asp Val Val G u G n G n Trp Arg Arg Phe Val Asp Ala Val Asp
1          5          10          15
Asn G y lle Val G u Phe Met G u His G u G u Pro Asn Lys Leu Asn
20          25          30
G u G y Lys Leu Ser Thr Ser Thr G u G u Met Met Ala Leu lle Val
35          40          45
G y Tyr Leu Ala Phe Val Val Leu G y Ser Ala Phe Met Lys Ala Phe
50          55          60
Val Asp Lys Pro Phe G u Leu Lys Phe Leu Lys Leu Val His Asn lle
65          70          75          80
Phe Leu Thr G y Leu Ser Met Tyr Met Ala Thr G u Cys Ala Arg G n
85          90          95
Ala Tyr Leu G y G y Tyr Lys Leu Phe G y Asn Pro Met G u Lys G y
100          105          110
Thr G u Ser His Ala Pro G y Met Ala Asn lle lle Tyr lle Phe Tyr
115          120          125
Val Ser Lys Phe Leu G u Phe Leu Asp Thr Val Phe Met lle Leu G y
130          135          140
Lys Lys Trp Lys G n Leu Ser Phe Leu His Val Tyr His His Ala Ser
145          150          155          160
lle Ser Phe lle Trp G y lle lle Ala Arg Phe Ala Pro G y G y Asp
165          170          175
Ala Tyr Phe Ser Thr lle Leu Asn Ser Ser Val His Val Val Leu Tyr
180          185          190
G y Tyr Tyr Ala Ser Thr Thr Leu G y Tyr Thr Phe Met Arg Pro Leu
195          200          205
Arg Pro Tyr lle Thr Thr lle G n Leu Thr G n Phe Met Ala Met Val
210          215          220
Val G n Ser Val Tyr Asp Tyr Tyr Asn Pro Cys Asp Tyr Pro G n Pro
225          230          235          240
Leu Val Lys Leu Leu Phe Trp Tyr Met Leu Thr Met Leu G y Leu Phe
245          250          255
G y Asn Phe Phe Val G n G n Tyr Leu Lys Pro Lys Ala Pro Lys Lys
260          265          270
G n Lys Thr lle
275

```

<210> 177  
 <211> 879  
 <212> DNA  
 <213> Ostreococcus tauri

<220>  
 <221> CDS  
 <222> (1)..(879)

PF58307.txt

<400> 177  
atg agt ggc tta cgt gca ccc aac ttt tta cac aga ttc tgg aca aag 48  
Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys  
1 5 10 15  
tgg gac tac gcg att tcc aaa gtc gtc ttc acg tgt gcc gac agt ttt 96  
Trp Asp Tyr Ala Ile Ser Lys Val Val Phe Thr Cys Ala Asp Ser Phe  
20 25 30  
cag tgg gac atc ggg cca gtg agt tcg agt acg gcg cat tta ccc gcc 144  
Gln Trp Asp Ile Gly Pro Val Ser Ser Thr Ala His Leu Pro Ala  
35 40 45  
att gaa tcc cct acc cca ctg gtg act agc ctg ttt ttc tac tta gtc 192  
Ile Glu Ser Pro Thr Pro Leu Val Thr Ser Leu Leu Phe Tyr Leu Val  
50 55 60  
aca gtt ttc ttg tgg tat ggt cgt tta acc agg agt tca gac aag aaa 240  
Thr Val Phe Leu Trp Tyr Gly Arg Leu Thr Arg Ser Ser Asp Lys Lys  
65 70 75  
att aga gag cct acg tgg tta aga aga ttc atg atg tgt cat aat gcg 288  
Ile Arg Glu Pro Thr Trp Leu Arg Arg Phe Ile Ile Cys His Asn Ala  
85 90 95  
ttc ttg ata gtc ctg agt ctt tac atg tgc ctt ggt tgt gtg gcc caa 336  
Phe Leu Ile Val Leu Ser Leu Tyr Met Cys Leu Gly Cys Val Ala Gln  
100 105 110  
gcg tat cag aat gga tat act tta tgg ggt aat gaa ttc aag gcc acg 384  
Ala Tyr Gln Asn Gly Tyr Thr Leu Trp Gly Asn Glu Phe Lys Ala Thr  
115 120 125  
gaa act cag ctt gct ctg tac att tac att ttt tac gta agt aaa ata 432  
Glu Thr Gln Leu Ala Leu Tyr Ile Tyr Ile Phe Tyr Val Ser Lys Ile  
130 135 140  
tac gag ttt gta gat act tac att atg ctt ctg aag aat aac ttg cgg 480  
Tyr Glu Phe Val Asp Thr Tyr Ile Met Leu Leu Lys Asn Asn Leu Arg  
145 150 155 160  
caa gta agt ttc ctg cac att tat cac cac agc acg att tcc ttt att 528  
Gln Val Ser Phe Leu His Ile Tyr His His Ser Thr Ile Ser Phe Ile  
165 170 175  
tgg tgg atc att gct cgg agg gct cca ggt ggt gat gct tac ttc agc 576  
Trp Trp Ile Ile Ala Arg Arg Ala Pro Gly Gly Asp Ala Tyr Phe Ser  
180 185 190  
gcg gcc ttg aac tca tgg gta cac gtg tgc atg tac acc tat tat ctg 624  
Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr Tyr Leu  
195 200 205  
tta tca acc ctt att gga aaa gaa gat cct aag cgt tcc aac tac ctt 672  
Leu Ser Thr Leu Ile Gly Lys Glu Asp Pro Lys Arg Ser Asn Tyr Leu  
210 215 220  
tgg tgg ggt cgc cac ctg acg caa atg cag atg ctt cag ttt ttc ttc 720  
Trp Trp Gly Arg His Leu Thr Gln Met Gln Leu Gln Phe Phe Phe  
225 230 235 240  
aac gta ctt caa gcg ttg tac tgc gct tcg ttc tct acg tat ccc aag 768  
Asn Val Leu Gln Ala Leu Tyr Cys Ala Ser Phe Ser Thr Tyr Pro Lys  
245 250 255  
ttt ttg tcc aaa att ctg ctg gtc tat atg atg agc ctt ctg ggc ttg 816  
Phe Leu Ser Lys Ile Leu Leu Val Tyr Met Met Ser Leu Leu Gly Leu  
260 265 270  
ttt ggg cat ttc tac tat tcc aag cac atg gca gca gct aag ctg cag 864  
Phe Gly His Phe Tyr Tyr Ser Lys His Ile Ala Ala Lys Leu Gln  
275 280 285  
aaa aaa cag cag tga 879  
Lys Lys Gln Gln  
290

<210> 178  
<211> 292  
<212> PRT  
<213> *Ostreococcus tauri*

<400> 178  
Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys  
1 5 10 15

PF58307. txt

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

<210> 179  
 <211> 873  
 <212> DNA  
 <213> Marchantia polymorpha

<220>  
 <221> CDS  
 <222> (1)..(873)

<400> 179  
 atg gag gcg tac gag atg gtg gat agt ttt gtg tcg aag acg gtt ttc 48  
 Met Gu Ala Tyr Gu Met Val Asp Ser Phe Val Ser Lys Thr Val Phe  
 1 5 10  
 gaa acg ctg cag aga ctg agg ggc gga gtc gtg ttg acg gaa tct gcg 96  
 Gu Thr Leu Gn Arg Leu Arg Gly Gy Val Val Leu Thr Gu Ser Ala  
 20 25 30  
 atc acc aaa ggt ttg cca tgc gtc gat agc ccg acg ccg atc gtt ctt 144  
 Ile Thr Lys Gy Leu Pro Cys Val Asp Ser Pro Thr Ile Val Leu  
 35 40 45  
 ggg ttg tcg tcc tac ttg aca ttc gtg ttt ctg ggg ctg att gtc atc 192  
 Gy Leu Ser Ser Tyr Leu Thr Phe Val Phe Leu Gy Leu Ile Val Ile  
 50 55 60  
 aag agc ctg gat ctt aag ccc cgc tcc aag gag ccc gcc att ttg aac 240  
 Lys Ser Leu Asp Leu Lys Pro Arg Ser Lys Gu Pro Ala Ile Leu Asn  
 65 70 75 80  
 ctg ttt gtg atc ttc cac aac ttc gtc tgc ttc gca ctg agt ctg tac 288  
 Leu Phe Val Ile Phe His Asn Phe Val Cys Phe Ala Leu Ser Leu Tyr  
 85 90 95  
 atg tgc gtg gga att gtc cgt caa gct atc ctg aac agg tac tct ctg 336  
 Met Cys Val Gy Ile Val Arg Gn Ala Ile Leu Asn Arg Tyr Ser Leu  
 100 105 110



PF58307. txt

t gg ggc aat gcg t ac aat ccc aaa gaa gt t caa at g ggc cac ct g ct c 384  
 Trp Gly Asn Ala Tyr Asn Pro Lys Gu Val Gn Met Gly His Leu Leu  
 115 120 125  
 t ac att ttc t ac at g t ca aag t ac at c gag ttt at g gac acg gt c at t 432  
 Tyr Ile Phe Tyr Met Ser Lys Tyr Ile Gu Phe Met Asp Thr Val Ile  
 130 135 140  
 at g att ttg aag cgc aac acg cgc cag at c act gt g tt g cat gt g t ac 480  
 Met Ile Leu Lys Arg Asn Thr Arg Gn Ile Thr Val Leu His Val Tyr  
 145 150 155 160  
 cac cac gca tcc at c tcc ttc at c tgg tgg at c at c gcc t ac cat gct 528  
 His His Ala Ser Ile Ser Phe Ile Trp Trp Ile Ile Ala Tyr His Ala  
 165 170 175  
 cct ggc ggt gaa gct t at ttc tct gcc gca ttg aac tcc gga gt a cat 576  
 Pro Gly Gly Glu Ala Tyr Phe Ser Ala Ala Leu Asn Ser Gly Val His  
 180 185 190  
 gt g ct c at g t ac ct c t ac t ac ct t ttg gca gca act ct g gga aag aac 624  
 Val Leu Met Tyr Leu Tyr Tyr Leu Leu Ala Ala Thr Leu Gly Lys Asn  
 195 200 205  
 gag aaa gct cgc cgc aag t ac ct a tgg tgg gga aaa t ac ttg aca cag 672  
 Gu Lys Ala Arg Arg Lys Tyr Leu Trp Trp Gly Lys Tyr Leu Thr Gn  
 210 215 220  
 ct g cag at g ttc cag ttt gt c ct t aac at g at t cag gct t ac t ac gat 720  
 Leu Gn Met Phe Gn Phe Val Leu Asn Met Ile Gn Ala Tyr Tyr Asp  
 225 230 235 240  
 at t aag aac aac t cg cct t ac cca caa ttt ttg at c cag at t ttg ttc 768  
 Ile Lys Asn Asn Ser Pro Tyr Pro Gn Phe Leu Ile Gn Ile Leu Phe  
 245 250 255  
 t ac t ac at g at c t cg ct t tta gcg ct a ttt gga aac ttt t ac gt t cac 816  
 Tyr Tyr Met Ile Ser Leu Leu Ala Leu Phe Gly Asn Phe Tyr Val His  
 260 265 270  
 aaa t ac gt a t ca gcg ccc gca aaa cct gcg aag at c aag agc aaa aag 864  
 Lys Tyr Val Ser Ala Pro Ala Lys Pro Ala Lys Ile Lys Ser Lys Lys  
 275 280 285  
 gca gaa t aa 873  
 Ala Gu  
 290

<210> 180  
 <211> 290  
 <212> PRT  
 <213> Marchantia polymorpha

<400> 180  
 Met Gu Ala Tyr Gu Met Val Asp Ser Phe Val Ser Lys Thr Val Phe  
 1 5 10  
 Gu Thr Leu Gn Arg Leu Arg Gly Gly Val Val Leu Thr Gu Ser Ala  
 20 25 30  
 Ile Thr Lys Gly Leu Pro Cys Val Asp Ser Pro Thr Pro Ile Val Leu  
 35 40 45  
 Gly Leu Ser Ser Tyr Leu Thr Phe Val Phe Leu Gly Leu Ile Val Ile  
 50 55 60  
 Lys Ser Leu Asp Leu Lys Pro Arg Ser Lys Gu Pro Ala Ile Leu Asn  
 65 70 75 80  
 Leu Phe Val Ile Phe His Asn Phe Val Cys Phe Ala Leu Ser Leu Tyr  
 85 90 95  
 Met Cys Val Gly Ile Val Arg Gn Ala Ile Leu Asn Arg Tyr Ser Leu  
 100 105 110  
 Trp Gly Asn Ala Tyr Asn Pro Lys Gu Val Gn Met Gly His Leu Leu  
 115 120 125  
 Tyr Ile Phe Tyr Met Ser Lys Tyr Ile Gu Phe Met Asp Thr Val Ile  
 130 135 140  
 Met Ile Leu Lys Arg Asn Thr Arg Gn Ile Thr Val Leu His Val Tyr  
 145 150 155 160  
 His His Ala Ser Ile Ser Phe Ile Trp Trp Ile Ile Ala Tyr His Ala  
 165 170 175  
 Pro Gly Gly Gu Ala Tyr Phe Ser Ala Ala Leu Asn Ser Gly Val His  
 180 185 190  
 Val Leu Met Tyr Leu Tyr Tyr Leu Leu Ala Ala Thr Leu Gly Lys Asn

Seite 218



PF58307. txt

195  
 G u Lys Al a Arg Arg Lys Tyr 200  
 210 Leu Trp Trp Gly Lys Tyr 205  
 225 Leu G n Met Phe G n Phe Val Leu Asn Met Ile G n Al a Tyr Tyr Asp  
 230 235 240  
 Ile Lys Asn Asn Ser Pro Tyr Pro G n Phe Leu Ile G n Ile Leu Phe  
 245 250 255  
 Tyr Tyr Met Ile Ser Leu Leu Al a Leu Phe Gly Asn Phe Tyr Val Hi s  
 260 265  
 Lys Tyr Val Ser Al a Pro Al a Lys Pro Al a Lys Ile Lys Ser Lys Lys  
 275 280 285  
 Al a G u  
 290

<210> 181  
 <211> 873  
 <212> DNA  
 <213> Physcomitrella patens

<220>  
 <221> CDS  
 <222> (1)..(873)

<400> 181  
 at g gag gt c gt g gag aga t t c t ac ggt gag t t g gat ggg aag gt c t c g 48  
 Met G u Val Val G u Arg Phe Tyr G y G u Leu Asp G y Lys Val Ser  
 1 5 10  
 cag ggc gt g aat gca t t g ct g ggt agt t t t ggg gt g gag t t g acg gat 96  
 G n G y Val Asn Al a Leu Leu G y Ser Phe G y Val G u Leu Thr Asp  
 20 25 30  
 acg ccc act acc aaa ggc t t g ccc ct c gt t gac agt ccc aca ccc at c 144  
 Thr Pro Thr Lys G y Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
 35 40 45  
 gt c ct c ggt gt t t ct gt a t ac t t g act at t gt c at t gga ggg ct t t t g 192  
 Val Leu G y Val Ser Val Tyr Leu Thr Ile Val Ile G y G y Leu Leu  
 50 55 60  
 t gg at a aag gcc agg gat ct g aaa ccg cgc gcc t c g gag cca t t t t t g 240  
 Trp Ile Lys Al a Arg Asp Leu Lys Pro Arg Al a Ser G u Pro Phe Leu  
 65 70 75 80  
 ct c caa gct t t g gt g ct t gt g cac aac ct g t t c t gt t t t gcg ct c agt 288  
 Leu G n Al a Leu Val Leu Val Hi s Asn Leu Phe Cys Phe Al a Leu Ser  
 85 90 95  
 ct g t at at g t gc gt g ggc at c gct t at cag gct at t acc t gg cgg t ac 336  
 Leu Tyr Met Cys Val G y Ile Al a Tyr G n Al a Ile Thr Trp Arg Tyr  
 100 105 110  
 t ct ct c t gg ggc aat gca t ac aat cct aaa cat aaa gag at g gcg at t 384  
 Ser Leu Trp G y Asn Al a Tyr Asn Pro Lys Hi s Lys G u Met Al a Ile  
 115 120 125  
 ct g gt a t ac t t g t t c t ac at g t ct aag t ac gt g gaa t t c at g gat acc 432  
 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val G u Phe Met Asp Thr  
 130 135 140  
 gt t at c at g at a ct g aag cgc agc acc agg caa at a agc t t c ct c cac 480  
 Val Ile Met Ile Leu Lys Arg Ser Thr Arg G n Ile Ser Phe Leu Hi s  
 145 150 155 160  
 gt t t at cat cat t ct t ca at t t cc ct c at t t gg t gg gct at t gct cat 528  
 Val Tyr Hi s Hi s Ser Ser Ile Ser Leu Ile Trp Trp Al a Ile Al a Hi s  
 165 170 175  
 cac gct cct ggc ggt gaa gca t at t gg t ct gcg gct ct g aac t ca gga 576  
 Hi s Al a Pro G y G y G u Al a Tyr Trp Ser Al a Al a Leu Asn Ser G y  
 180 185 190  
 gt g cat gt t ct c at g t at gcg t at t ac t t c t t g gct gcc t gc ct t cga 624  
 Val Hi s Val Leu Met Tyr Al a Tyr Tyr Phe Leu Al a Al a Cys Leu Arg  
 195 200 205  
 agt agc cca aag t ta aaa aat aag t ac ct t t t t gg ggc agg t ac t t g 672  
 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp G y Arg Tyr Leu  
 210 215 220  
 aca caa t t c caa at g t t c cag t t t at g ct g aac t ta gt g cag gct t ac 720  
 Thr G n Phe G n Met Phe G n Phe Met Leu Asn Leu Val G n Al a Tyr

PF58307.txt

225 230 235 240  
t ac gac at g aaa acg aat gcg cca t at cca caa t gg ct g at c aag at t 768  
Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gl n Trp Leu Ile Lys Ile  
245 250 255  
t t g t t c t ac t ac at g at c t cg t t g ct g t t t ct t t t c ggc aat t t t t ac 816  
Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gl y Asn Phe Tyr  
260 265 270  
gt a caa aaa t ac at c aaa ccc t ct gac gga aag caa aag gga gct aaa 864  
Val Gl n Lys Tyr Ile Lys Pro Ser Asp Gl y Lys Gl n Lys Gl y Ala Lys  
275 280 285  
act gag t ga 873  
Thr Gl u  
290

<210> 182

<211> 290

<212> PRT

<213> Physcomitrella patens

<400> 182

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

<210> 183

<211> 957

<212> DNA

<213> Mbrtierella alpina

<220>

<221> CDS

<222> (1)..(957)

PF58307.txt

<400> 183

at g	gag	t cg	att	gcg	cca	t t c	ct c	cca	t ca	aag	at g	ccg	caa	gat	ct g	48
Met	Gl u	Ser	l l e	Al a	Pro	Phe	Leu	Pro	Ser	Lys	Met	Pro	Gl n	Asp	Leu	
1				5					10					15		
t t t	at g	gac	ct t	gcc	acc	gct	at c	ggt	gt c	cgg	gcc	gcg	ccc	t at	gt c	96
Phe	Met	Asp	Leu	Al a	Thr	Al a	l l e	Gl y	Val	Arg	Al a	Al a	Pro	Tyr	Val	
			20					25					30			
gat	cct	ct c	gag	gcc	gcg	ct g	gt g	gcc	cag	gcc	gag	aag	t ac	at c	ccc	144
Asp	Pro	Leu	Gl u	Al a	Al a	Leu	Val	Al a	Gl n	Al a	Gl u	Lys	Tyr	l l e	Pro	
			35				40					45				
acg	att	gt c	cat	cac	acg	cgt	ggg	t t c	ct g	gt c	gcg	gt g	gag	t cg	cct	192
Thr	l l e	Val	Hi s	Hi s	Thr	Arg	Gl y	Phe	Leu	Val	Al a	Val	Gl u	Ser	Pro	
			50			55					60					
t t g	gcc	cgt	gag	ct g	ccg	t t g	at g	aac	ccg	t t c	cac	gt g	ct g	t t g	at c	240
Leu	Al a	Arg	Gl u	Leu	Pro	Leu	Met	Asn	Pro	Phe	Hi s	Val	Leu	Leu	l l e	
65					70				75						80	
gt g	ct c	gct	t at	t t g	gt c	acg	gt c	t t t	gt g	ggc	at g	cag	at c	at g	aag	288
Val	Leu	Al a	Tyr	Leu	Val	Thr	Val	Phe	Val	Gl y	Met	Gl n	l l e	Met	Lys	
				85					90					95		
aac	t t t	gag	cgg	t t c	gag	gt c	aag	acg	t t t	t cg	ct c	ct g	cac	aac	t t t	336
Asn	Phe	Gl u	Arg	Phe	Gl u	Val	Lys	Thr	Phe	Ser	Leu	Leu	Hi s	Asn	Phe	
			100					105					110			
t gt	ct g	gt c	t cg	at c	agc	gcc	t ac	at g	t gc	ggt	ggg	at c	ct g	t ac	gag	384
Cys	Leu	Val	Ser	l l e	Ser	Al a	Tyr	Met	Cys	Gl y	Gl y	l l e	Leu	Tyr	Gl u	
			115				120					125				
gct	t at	cag	gcc	aac	t at	gga	ct g	t t t	gag	aac	gct	gct	gat	cat	acc	432
Al a	Tyr	Gl n	Al a	Asn	Tyr	Gl y	Leu	Phe	Gl u	Asn	Al a	Al a	Asp	Hi s	Thr	
			130			135					140					
t t c	aag	ggt	ct t	cct	at g	gcc	aag	at g	at c	t gg	ct c	t t c	t ac	t t c	t cc	480
Phe	Lys	Gl y	Leu	Pro	Met	Al a	Lys	Met	l l e	Trp	Leu	Phe	Tyr	Phe	Ser	
145					150				155						160	
aag	at c	at g	gag	t t t	gt c	gac	acc	at g	at c	at g	gt c	ct c	aag	aag	aac	528
Lys	l l e	Met	Gl u	Phe	Val	Asp	Thr	Met	l l e	Met	Val	Leu	Lys	Lys	Asn	
				165					170					175		
aac	cg c	cag	at c	t cc	t t c	t t g	cac	gt t	t ac	cac	cac	agc	t cc	at c	t t c	576
Asn	Arg	Gl n	l l e	Ser	Phe	Leu	Hi s	Val	Tyr	Hi s	Hi s	Ser	Ser	l l e	Phe	
				180				185					190			
acc	at c	t gg	t gg	t t g	gt c	acc	t t t	gt t	gca	ccc	aac	ggt	gaa	gcc	t ac	624
Thr	l l e	Trp	Trp	Leu	Val	Thr	Phe	Val	Al a	Pro	Asn	Gl y	Gl u	Al a	Tyr	
				195			200					205				
t t c	t ct	gct	gcg	t t g	aac	t cg	t t c	at c	cat	gt g	at c	at g	t ac	ggc	t ac	672
Phe	Ser	Al a	Al a	Leu	Asn	Ser	Phe	l l e	Hi s	Val	l l e	Met	Tyr	Gl y	Tyr	
						215					220					
t ac	t t c	t t g	t cg	gcc	t t g	ggc	t t c	aag	cag	gt g	t cg	t t c	at c	aag	t t c	720
Tyr	Phe	Leu	Ser	Al a	Leu	Gl y	Phe	Lys	Gl n	Val	Ser	Phe	l l e	Lys	Phe	
225					230				235						240	
t ac	at c	acg	cg c	t cg	cag	at g	aca	cag	t t c	t gc	at g	at g	t cg	gt c	cag	768
Tyr	l l e	Thr	Arg	Ser	Gl n	Met	Thr	Gl n	Phe	Cys	Met	Met	Ser	Val	Gl n	
				245				250					255			
t ct	t cc	t gg	gac	at g	t ac	gcc	at g	aag	gt c	ct t	ggc	cg c	ccc	gga	t ac	816
Ser	Ser	Trp	Asp	Met	Tyr	Al a	Met	Lys	Val	Leu	Gl y	Arg	Pro	Gl y	Tyr	
			260					265					270			
ccc	t t c	t t c	at c	acg	gct	ct g	ct t	t gg	t t c	t ac	at g	t gg	acc	at g	ct c	864
Pro	Phe	Phe	l l e	Thr	Al a	Leu	Leu	Trp	Phe	Tyr	Met	Trp	Thr	Met	Leu	
				275			280					285				
ggt	ct c	t t c	t ac	aac	t t t	t ac	aga	aag	aac	gcc	aag	t t g	gcc	aag	cag	912
Gl y	Leu	Phe	Tyr	Asn	Phe	Tyr	Arg	Lys	Asn	Al a	Lys	Leu	Al a	Lys	Gl n	
				290		295				300						
gcc	aag	gcc	gac	gct	gcc	aag	gag	aag	gca	agg	aag	t t g	cag	t aa		957
Al a	Lys	Al a	Asp	Al a	Al a	Lys	Gl u	Lys	Al a	Arg	Lys	Leu	Gl n			
305					310					315						

<210> 184

<211> 318

<212> PRT

<213> Mbr t i e r e l l a a l p i n a

<400> 184

PF58307.txt

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

<210> 185  
<211> 251  
<212> PRT  
<213> Artificial sequence

<220>  
<221> Variant  
<222> (2)..(11)  
<223> xaa in position 2 to 11 is any amino acid  
<220>  
<221> Variant  
<222> (14)..(15)  
<223> xaa in position 14 to 15 is any amino acid  
<220>  
<221> Variant  
<222> (17)..(23)  
<223> xaa in position 17 to 23 is any amino acid  
<220>  
<221> Variant  
<222> (25)..(25)  
<223> xaa in position 25 is any amino acid  
<220>  
<221> Variant  
<222> (26)..(28)  
<223> xaa in position 26 to 28 is any or no amino acid  
<220>

<221> Variant  
<222> (30)..(31)  
<223> xaa in position 30 to 31 is any amino acid  
<220>  
<221> Variant  
<222> (32)..(34)  
<223> xaa in position 32 to 34 is any or no amino acid  
<220>  
<221> Variant  
<222> (36)..(37)  
<223> xaa in position 36 to 37 is any amino acid  
<220>  
<221> Variant  
<222> (39)..(44)  
<223> xaa in position 39 to 44 is any amino acid  
<220>  
<221> Variant  
<222> (47)..(51)  
<223> xaa in position 47 to 51 is any amino acid  
<220>  
<221> Variant  
<222> (54)..(54)  
<223> xaa in position 54 is any amino acid  
<220>  
<221> Variant  
<222> (57)..(63)  
<223> xaa in position 57 to 63 is any amino acid  
<220>  
<221> Variant  
<222> (65)..(68)  
<223> xaa in position 65 to 68 is any amino acid  
<220>  
<221> Variant  
<222> (70)..(70)  
<223> xaa in position 70 is any amino acid  
<220>  
<221> Variant  
<222> (72)..(72)  
<223> xaa in position 72 is any amino acid  
<220>  
<221> Variant  
<222> (75)..(82)  
<223> xaa in position 75 to 82 is any amino acid  
<220>  
<221> Variant  
<222> (83)..(86)  
<223> xaa in position 83 to 86 is any or no amino acid  
<220>  
<221> Variant  
<222> (88)..(91)  
<223> xaa in position 88 to 91 is any amino acid  
<220>  
<221> Variant  
<222> (93)..(93)  
<223> xaa in position 93 is any amino acid  
<220>  
<221> Variant  
<222> (96)..(96)  
<223> xaa in position 96 is any amino acid  
<220>  
<221> Variant  
<222> (99)..(100)  
<223> xaa in position 99 to 100 is any amino acid  
<220>  
<221> Variant  
<222> (103)..(103)  
<223> xaa in position 103 is any amino acid  
<220>

<221> Vari ant  
<222> ( 106).. ( 107)  
<223> xaa in position 106 to 107 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 109).. ( 109)  
<223> xaa in position 109 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 111).. ( 115)  
<223> xaa in position 111 to 115 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 117).. ( 119)  
<223> xaa in position 117 to 119 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 126).. ( 126)  
<223> xaa in position 126 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 129).. ( 130)  
<223> xaa in position 129 to 130 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 133).. ( 134)  
<223> xaa in position 133 to 134 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 136).. ( 138)  
<223> xaa in position 136 to 138 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 143).. ( 143)  
<223> xaa in position 143 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 148).. ( 149)  
<223> xaa in position 148 to 149 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 153).. ( 154)  
<223> xaa in position 153 to 154 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 157).. ( 158)  
<223> xaa in position 157 to 158 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 160).. ( 160)  
<223> xaa in position 160 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 163).. ( 179)  
<223> xaa in position 163 to 179 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 180).. ( 185)  
<223> xaa in position 180 to 185 is any or no ami no aci d  
<220>  
<221> Vari ant  
<222> ( 187).. ( 188)  
<223> xaa in position 187 to 188 is any ami no aci d  
<220>  
<221> Vari ant  
<222> ( 190).. ( 191)  
<223> xaa in position 190 to 191 is any ami no aci d  
<220>

<221> Variant  
 <222> (194)..(198)  
 <223> xaa in position 194 to 198 is any amino acid  
 <220>  
 <221> Variant  
 <222> (200)..(201)  
 <223> xaa in position 200 to 201 is any amino acid  
 <220>  
 <221> Variant  
 <222> (203)..(208)  
 <223> xaa in position 203 to 208 is any amino acid  
 <220>  
 <221> Variant  
 <222> (209)..(209)  
 <223> xaa in position 209 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (212)..(217)  
 <223> xaa in position 212 to 217 is any amino acid  
 <220>  
 <221> Variant  
 <222> (219)..(220)  
 <223> xaa in position 219 to 220 is any amino acid  
 <220>  
 <221> Variant  
 <222> (223)..(225)  
 <223> xaa in position 223 to 225 is any amino acid  
 <220>  
 <221> Variant  
 <222> (227)..(227)  
 <223> xaa in position 227 is any amino acid  
 <220>  
 <221> Variant  
 <222> (233)..(246)  
 <223> xaa in position 233 to 246 is any amino acid  
 <220>  
 <221> Variant  
 <222> (247)..(250)  
 <223> xaa in position 247 to 250 is any or no amino acid

<400> 185  
 Ser xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Tyr Leu xaa xaa Val  
 1 5 10 15  
 xaa xaa xaa xaa xaa xaa xaa Lys xaa xaa xaa xaa Lys xaa xaa xaa  
 20 25 30  
 xaa xaa Pro xaa xaa Leu xaa xaa xaa xaa xaa xaa His Asn xaa xaa  
 35 40 45  
 xaa xaa xaa Leu Ser xaa Tyr Met xaa xaa xaa xaa xaa xaa xaa Ala  
 50 55 60  
 xaa xaa xaa xaa Tyr xaa Leu xaa Gly Asn xaa xaa xaa xaa xaa  
 65 70 75 80  
 xaa xaa xaa xaa xaa xaa Met xaa xaa xaa xaa Tyr xaa Phe Tyr xaa  
 85 90 95  
 Ser Lys xaa xaa Glu Phe xaa Asp Thr xaa xaa Met xaa Leu xaa xaa  
 100 105 110  
 xaa xaa xaa Glu xaa xaa xaa Leu His Val Tyr His His xaa Ser Ile  
 115 120 125  
 xaa xaa Ile Trp xaa xaa Ile xaa xaa xaa Ala Pro Gly Gly xaa Ala  
 130 135 140  
 Tyr Phe Ser xaa xaa Leu Asn Ser xaa xaa His Val xaa xaa Tyr xaa  
 145 150 155 160  
 Tyr Tyr xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa  
 165 170 175  
 xaa xaa xaa xaa xaa xaa xaa xaa Thr xaa xaa Glu xaa xaa Glu  
 180 185 190  
 Phe xaa xaa xaa xaa xaa xaa Glu xaa xaa Tyr xaa xaa xaa xaa xaa xaa  
 195 200 205  
 xaa Tyr Pro xaa xaa xaa xaa xaa xaa Leu xaa xaa Tyr Met xaa xaa

```

      210                215                220
xaa Leu xaa Leu Phe Gly Asn Phe xaa xaa xaa xaa xaa xaa xaa xaa
225                230                235                240
xaa xaa xaa xaa xaa xaa xaa xaa xaa xaa Lys
                245                250

```

```

<210> 186
<211> 60
<212> PRT
<213> Artificial sequence

<220>
<221> Variant
<222> (3)..(4)
<223> xaa in position 3 to 4 is any amino acid
<220>
<221> Variant
<222> (7)..(7)
<223> xaa in position 7 is Leu, Met or Val
<220>
<221> Variant
<222> (10)..(10)
<223> xaa in position 10 is any amino acid
<220>
<221> Variant
<222> (11)..(11)
<223> xaa in position 11 is Phe or Ile
<220>
<221> Variant
<222> (13)..(13)
<223> xaa in position 13 is Ile, Leu or Val
<220>
<221> Variant
<222> (15)..(18)
<223> xaa in position 15 to 18 is any amino acid
<220>
<221> Variant
<222> (19)..(19)
<223> xaa in position 19 is Lys or Arg
<220>
<221> Variant
<222> (21)..(21)
<223> xaa in position 21 is Ile, Leu or Val
<220>
<221> Variant
<222> (22)..(22)
<223> xaa in position 22 is Ser or Thr
<220>
<221> Variant
<222> (23)..(23)
<223> xaa in position 23 is Phe or Val
<220>
<221> Variant
<222> (26)..(26)
<223> xaa in position 26 is Ile or Val
<220>
<221> Variant
<222> (30)..(30)
<223> xaa in position 30 is Ala or Ser
<220>
<221> Variant
<222> (31)..(31)
<223> xaa in position 31 is Ser or Thr
<220>
<221> Variant
<222> (33)..(34)
<223> xaa in position 33 to 34 is any amino acid

```



<220>  
 <221> Variant  
 <222> (37)..(37)  
 <223> xaa in position 37 is any amino acid  
 <220>  
 <221> Variant  
 <222> (38)..(38)  
 <223> xaa in position 38 is Ala, Ile or Leu  
 <220>  
 <221> Variant  
 <222> (39)..(39)  
 <223> xaa in position 39 is Ile or Val  
 <220>  
 <221> Variant  
 <222> (40)..(40)  
 <223> xaa in position 40 is Ala or Thr  
 <220>  
 <221> Variant  
 <222> (41)..(42)  
 <223> xaa in position 41 to 42 is any amino acid  
 <220>  
 <221> Variant  
 <222> (45)..(45)  
 <223> xaa in position 45 is Gly or Asn  
 <220>  
 <221> Variant  
 <222> (47)..(47)  
 <223> xaa in position 47 is Asp or Glu  
 <220>  
 <221> Variant  
 <222> (50)..(50)  
 <223> xaa in position 50 is Phe or Trp  
 <220>  
 <221> Variant  
 <222> (52)..(52)  
 <223> xaa in position 52 is Ala or Thr  
 <220>  
 <221> Variant  
 <222> (53)..(53)  
 <223> xaa in position 53 is Ala or Ile  
 <220>  
 <221> Variant  
 <222> (57)..(57)  
 <223> xaa in position 57 is any amino acid  
 <220>  
 <221> Variant  
 <222> (58)..(58)  
 <223> xaa in position 58 is Ile or Val

<400> 186  
 Ser Lys xaa xaa Glu Phe xaa Asp Thr xaa xaa Met xaa Leu xaa xaa  
 1 5 10 15  
 xaa xaa xaa Gln xaa xaa xaa Leu His xaa Tyr His His xaa xaa Ile  
 20 25 30  
 xaa xaa Ile Trp xaa xaa xaa xaa xaa Ala Pro xaa Gly xaa Ala  
 35 40 45  
 Tyr xaa Ser xaa xaa Leu Asn Ser xaa xaa His Val  
 50 55 60

<210> 187  
 <211> 23  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (3)..(3)

<223> xaa in position 3 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is Phe or Tyr  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Ala or Cys  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is any amino acid  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Glu or Gly  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is any amino acid  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is Ala, Ile or Val  
 <220>  
 <221> Variant  
 <222> (11)..(11)  
 <223> xaa in position 11 is any amino acid  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is Ala or Gln  
 <220>  
 <221> Variant  
 <222> (14)..(17)  
 <223> xaa in position 14 to 17 is any amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is any amino acid  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Leu or Val  
 <220>  
 <221> Variant  
 <222> (21)..(21)  
 <223> xaa in position 21 is Phe or Trp  
  
 <400> 187  
 Leu Ser xaa xaa Met xaa xaa xaa xaa xaa xaa xaa Ala xaa xaa xaa  
   1                  5                                  10                                  15  
 xaa Tyr xaa xaa xaa Gly Asn  
                   20

<210> 188  
 <211> 24  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (3)..(4)  
 <223> xaa in position 3 to 4 is any amino acid  
 <220>  
 <221> Variant

<222> (5)..(5)  
 <223> xaa in position 5 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (6)..(7)  
 <223> xaa in position 6 to 7 is any amino acid  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (10)..(11)  
 <223> xaa in position 10 to 11 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Ser or Thr  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (21)..(22)  
 <223> xaa in position 21 to 22 is any amino acid  
 <220>  
 <221> Variant  
 <222> (24)..(24)  
 <223> xaa in position 24 is Phe or Tyr  
  
 <400> 188  
 Tyr Pro xaa xaa xaa xaa xaa Leu xaa xaa Tyr Met xaa xaa xaa  
   1                  5                  10                  15  
 Leu xaa Leu Phe xaa xaa Phe xaa  
                   20

<210> 189  
 <211> 20  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (4)..(5)  
 <223> xaa in position 4 to 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Leu or Met  
 <220>  
 <221> Variant  
 <222> (8)..(8)  
 <223> xaa in position 8 is any amino acid  
 <220>

<221> Variant  
 <222> (11)..(14)  
 <223> xaa in position 11 to 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Gly, Ile or Val  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is Ala or Ser  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is any amino acid  
 <220>  
 <221> Variant  
 <222> (19)..(19)  
 <223> xaa in position 19 is Trp or Tyr  
 <220>  
 <221> Variant  
 <222> (20)..(20)  
 <223> xaa in position 20 is Ala or Asp  
  
 <400> 189  
 Tyr xaa Thr xaa xaa Gln xaa xaa Gln Phe xaa xaa xaa xaa xaa Gln  
   1                  5                  10                  15  
 xaa xaa xaa xaa  
                   20

<210> 190  
 <211> 18  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(5)  
 <223> xaa in position 2 to 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (8)..(9)  
 <223> xaa in position 8 to 9 is any amino acid  
 <220>  
 <221> Variant  
 <222> (11)..(12)  
 <223> xaa in position 11 to 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (13)..(13)  
 <223> xaa in position 13 is Phe or Leu  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (15)..(15)  
 <223> xaa in position 15 is Ile or Leu  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is any amino acid

<400> 190  
 Lys xaa xaa xaa xaa xaa Pro xaa xaa Leu xaa xaa xaa xaa xaa xaa  
 1 5 10 15  
 His Asn

<210> 191  
 <211> 18  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Pro or Thr  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is any amino acid  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is Glu or Pro  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Ile, Leu, Met or Val  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Leu, Met or Val  
 <220>  
 <221> Variant  
 <222> (7)..(8)  
 <223> xaa in position 7 to 8 is any amino acid  
 <220>  
 <221> Variant  
 <222> (9)..(9)  
 <223> xaa in position 9 is Ile, Leu or Val  
 <220>  
 <221> Variant  
 <222> (10)..(11)  
 <223> xaa in position 10 to 11 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(15)  
 <223> xaa in position 14 to 15 is any amino acid  
 <220>  
 <221> Variant  
 <222> (17)..(17)  
 <223> xaa in position 17 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(18)  
 <223> xaa in position 18 is Gly or Leu

<400> 191  
 Ser xaa xaa xaa xaa xaa xaa xaa xaa xaa Tyr Leu xaa xaa Val  
 1 5 10 15  
 xaa xaa

<210> 192  
 <211> 10  
 <212> PRT

<213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(2)  
 <223> xaa in position 2 is Ala, Gly or Asn  
 <220>  
 <221> Variant  
 <222> (3)..(3)  
 <223> xaa in position 3 is any amino acid  
 <220>  
 <221> Variant  
 <222> (4)..(4)  
 <223> xaa in position 4 is Ile, Leu or Met  
 <220>  
 <221> Variant  
 <222> (5)..(5)  
 <223> xaa in position 5 is Ile, Leu or Val  
 <220>  
 <221> Variant  
 <222> (6)..(6)  
 <223> xaa in position 6 is Trp or Tyr  
 <220>  
 <221> Variant  
 <222> (7)..(7)  
 <223> xaa in position 7 is Ile, Leu or Val  
 <220>  
 <221> Variant  
 <222> (10)..(10)  
 <223> xaa in position 10 is Phe, Met or Val

<400> 192  
 Met xaa xaa xaa xaa xaa xaa Phe Tyr xaa  
 1 5 10

<210> 193  
 <211> 1086  
 <212> DNA  
 <213> Phyt opt hor a i n f e s t a n s

<220>  
 <221> CDS  
 <222> (1)..(1086)

<400> 193  
 at g g c g a c g a a g g a g g c g t a t g t g t t c c c c a c t c t g a c g g a g a t c a a g 48  
 Met Ala Thr Lys Gu Ala Tyr Val Phe Pro Thr Leu Thr Glu Ile Lys  
 1 5 10 15  
 c g g t c g c t a c c t a a a g a c t g t t t c g a g g c t c g g t g c c t c t g t c g c t c 96  
 Arg Ser Leu Pro Lys Asp Cys Phe Glu Ala Ser Val Pro Leu Ser Leu  
 20 25 30  
 t a c t a c a c c g t g c g t t g t c t g g t g a t c g c g g t g g c t a a c c t t c g g t 144  
 Tyr Tyr Thr Val Arg Cys Leu Val Ile Ala Val Ala Leu Thr Phe Gly  
 35 40 45  
 c t c a a c t a c g c t c g c g c t c t g c c c g a g g t c g a g a g c t t c t g g g c t c t g 192  
 Leu Asn Tyr Ala Arg Ala Leu Pro Gu Val Gu Ser Phe Trp Ala Leu  
 50 55 60  
 g a c g c c g c a c t c t g c a c g g g c t a c a t c t t g c t g c a g g g c a t c g t g t t c 240  
 Asp Ala Ala Leu Cys Thr Gy Tyr Ile Leu Leu Gn Gy Ile Val Phe  
 65 70 75 80  
 t g g g g c t t c t t c a c g g t g g g c c a c g a t g c c g g c c a c g g c g c c t t c t c g 288  
 Trp Gy Phe Phe Thr Val Gy His Asp Ala Gy His Gy Ala Phe Ser  
 85 90 95  
 c g c t a c c a c c t g c t t a a c t t c g t g g t g g g c a c t t t c a t g c a c t c g c t c 336  
 Arg Tyr His Leu Leu Asn Phe Val Val Gy Thr Phe Met His Ser Leu  
 100 105 110  
 a t c c t c a c g c c c t t c g a g t c g t g g a a g c t c a c g c a c c g t c a c c a c c a c 384  
 Ile Leu Thr Pro Phe Gu Ser Trp Lys Leu Thr His Arg His His His

PF58307. txt

```

115          120          125
aag aac acg ggc aac att gac cgt gac gag gtc ttc tac ccg caa cgc      432
Lys Asn Thr Gly Asn Ile Asp Arg Asp Glu Val Phe Tyr Pro Gn Arg
130
aag gcc gac gac cac ccg ctg tct cgc aac ctg att ctg gcg ct c ggg      480
Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Ile Leu Ala Leu Gly
145
gca gcg tgg ct c gcc tat ttg gtc gag gcc ttc cct cct cgt aag gtc      528
Ala Ala Trp Leu Ala Tyr Leu Val Glu Gly Phe Pro Pro Arg Lys Val
165
aac cac ttc aac ccg ttc gag cct ctg ttc gtg cgt cag gtg tca gct      576
Asn His Phe Asn Pro Phe Glu Pro Leu Phe Val Arg Gn Val Ser Ala
180
gtg gta at c tct ct t ct c gcc cac ttc ttc gtg gcc gga ct c tcc at c      624
Val Val Ile Ser Leu Leu Ala His Phe Phe Val Ala Gly Leu Ser Ile
195
tat ctg agc ct c cag ctg gcc ct t aag acg at g gca at c tac tac tat      672
Tyr Leu Ser Leu Gn Leu Gly Leu Lys Thr Met Ala Ile Tyr Tyr Tyr
210
gga cct gtt ttt gtg ttc gcc agc at g ctg gtc at t acc acc ttc ct a      720
Gly Pro Val Phe Val Phe Gly Ser Met Leu Val Ile Thr Thr Phe Leu
225
cac cac aat gat gag gag acc cca tgg tac gcc gac t cg gag tgg acg      768
His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser Glu Trp Thr
245
tac gtc aag gcc aac ct c t cg tcc gtg gac cga t cg tac gcc gcg ct c      816
Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala Leu
260
att gac aac ctg agc cac aac at c gcc acg cac cag at c cac cac ct t      864
Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gn Ile His His Leu
275
ttc cct at c att ccg cac tac aaa ct c aag aaa gcc act gcg gcc ttc      912
Phe Pro Ile Ile Pro His Tyr Lys Leu Lys Lys Ala Thr Ala Ala Phe
290
cac cag gct ttc cct gag ct c gtg cgc aag agc gac gag cca at t at c      960
His Gn Ala Phe Pro Glu Leu Val Arg Lys Ser Asp Glu Pro Ile Ile
305
aag gct ttc ttc cgg gtt gga cgt ct c tac gca aac tac gcc gtt gtg      1008
Lys Ala Phe Phe Arg Val Gly Arg Leu Tyr Ala Asn Tyr Gly Val Val
325
gac cag gag gcg aag ct c ttc acg ct a aag gcc aag gcg gcg acc      1056
Asp Gn Glu Ala Lys Leu Phe Thr Leu Lys Glu Ala Lys Ala Ala Thr
340
gag gcg gcg gcc aag acc aag tcc acg taa      1086
Glu Ala Ala Ala Lys Thr Lys Ser Thr
355
360

```

<210> 194  
 <211> 361  
 <212> PRT  
 <213> Phyt opht hor a i nf est ans

```

<400> 194
Met Ala Thr Lys Glu Ala Tyr Val Phe Pro Thr Leu Thr Glu Ile Lys
1 5 10 15
Arg Ser Leu Pro Lys Asp Cys Phe Glu Ala Ser Val Pro Leu Ser Leu
20 25 30
Tyr Tyr Thr Val Arg Cys Leu Val Ile Ala Val Ala Leu Thr Phe Gly
35 40 45
Leu Asn Tyr Ala Arg Ala Leu Pro Glu Val Glu Ser Phe Trp Ala Leu
50 55 60
Asp Ala Ala Leu Cys Thr Gly Tyr Ile Leu Leu Gn Gly Ile Val Phe
65 70 75 80
Trp Gly Phe Phe Thr Val Gly His Asp Ala Gly His Gly Ala Phe Ser
85 90 95
Arg Tyr His Leu Leu Asn Phe Val Val Gly Thr Phe Met His Ser Leu
100 105 110
Ile Leu Thr Pro Phe Glu Ser Trp Lys Leu Thr His Arg His His His
Sei te 233

```

PF58307. txt

115 120 125  
 Lys Asn Thr Gly Asn Ile Asp Arg Asp Glu Val Phe Tyr Pro Gln Arg  
 130 135 140  
 Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Ile Leu Ala Leu Gly  
 145 150 155  
 Ala Ala Trp Leu Ala Tyr Leu Val Glu Gly Phe Pro Pro Arg Lys Val  
 165 170 175  
 Asn His Phe Asn Pro Phe Glu Pro Leu Phe Val Arg Gln Val Ser Ala  
 180 185 190  
 Val Val Ile Ser Leu Leu Ala His Phe Phe Val Ala Gly Leu Ser Ile  
 195 200 205  
 Tyr Leu Ser Leu Gln Leu Gly Leu Lys Thr Met Ala Ile Tyr Tyr Tyr  
 210 215 220  
 Gly Pro Val Phe Val Phe Gly Ser Met Leu Val Ile Thr Thr Phe Leu  
 225 230 235  
 His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser Glu Trp Thr  
 245 250 255  
 Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala Leu  
 260 265 270  
 Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile His His Leu  
 275 280 285  
 Phe Pro Ile Ile Pro His Tyr Lys Leu Lys Lys Ala Thr Ala Ala Phe  
 290 295 300  
 His Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Asp Glu Pro Ile Ile  
 305 310 315 320  
 Lys Ala Phe Phe Arg Val Gly Arg Leu Tyr Ala Asn Tyr Gly Val Val  
 325 330 335  
 Asp Gln Glu Ala Lys Leu Phe Thr Leu Lys Glu Ala Lys Ala Thr  
 340 345 350  
 Gu Ala Ala Ala Lys Thr Lys Ser Thr  
 355 360

<210> 195  
 <211> 1077  
 <212> DNA  
 <213> Saproligna dieliana

<220>  
 <221> CDS  
 <222> (1)..(1077)

<400> 195  
 atg act gag gat aag acg aag gtc gag ttc ccg acg ct c acg gag ct c 48  
 Met Thr Glu Asp Lys Thr Lys Val Glu Phe Pro Thr Leu Thr Glu Leu  
 1 5 10 15  
 aag cac t cg at c ccg aac gcg tgc ttt gag t cg aac ct c ggc ct c t cg 96  
 Lys His Ser Ile Pro Asn Ala Cys Phe Glu Ser Asn Leu Gly Leu Ser  
 20 25 30  
 ct c t ac t ac acg gcc cgc gcg at c ttc aac gcg t cg gcc t cg gcg gcg 144  
 Leu Tyr Tyr Thr Ala Arg Ala Ile Phe Asn Ala Ser Ala Ser Ala Ala  
 35 40 45  
 ct g ct c t ac gcg gcg cgc t cg acg ccg ttc att gcc gat aac gt t ct g 192  
 Leu Leu Tyr Ala Ala Arg Ser Thr Pro Phe Ile Ala Asp Asn Val Leu  
 50 55 60  
 ct c cac gcg ct c gtt tgc gcc acc t ac at c t ac gt g cag ggc gt c at c 240  
 Leu His Ala Leu Val Cys Ala Thr Tyr Ile Tyr Val Gln Gly Val Ile  
 65 70 75 80  
 ttc tgg ggc ttc ttc acg gtc ggc cac gac tgc ggc cac t cg gcc ttc 288  
 Phe Trp Gly Phe Phe Thr Val Gly His Asp Cys Gly His Ser Ala Phe  
 85 90 95  
 t cg cgc t ac cac agc gtc aac ttt at c at c ggc t gc at c at g cac t ct 336  
 Ser Arg Tyr His Ser Val Asn Phe Ile Ile Gly Cys Ile Met His Ser  
 100 105 110  
 gcg att ttg acg ccg ttc gag agc tgg cg c gt g acg cac cgc cac cac 384  
 Ala Ile Leu Thr Pro Phe Glu Ser Trp Arg Val Thr His Arg His His  
 115 120 125  
 cac aag aac acg ggc aac att gat aag gac gag at c ttt t ac ccg cac 432  
 His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr Pro His



PF58307. txt

```

130          135          140
cgg t cg gt c aag gac ct c cag gac gt g cgc caa t gg gt c t ac acg ct c      480
Arg Ser Val Lys Asp Leu G n Asp Val Arg G n Tr p Val Tyr Thr Leu
145          150          155
ggc ggt gcg t gg t t t gt c t ac t t g aag gt c ggg t at gcc ccg cgc acg      528
G y G y Al a Tr p Phe Val Tyr Leu Lys Val G y Tyr Al a Pro Arg Thr
165          170          175
at g agc cac t t t gac ccg t gg gac ccg ct c ct c ct t cgc cgc gcg t cg      576
Met Ser Hi s Phe Asp Pro Tr p Asp Pro Leu Leu Leu Arg Arg Al a Ser
180          185          190
gcc gt c at c gt g t cg ct c ggc gt c t gg gcc gcc t t c t t c gcc gcg t ac      624
Al a Val Il e Val Ser Leu G y Val Tr p Al a Al a Phe Phe Al a Al a Tyr
195          200          205
gcg t ac ct c aca t ac t cg ct c ggc gt c t gg gcc gt c at g ggc ct c t ac t ac      672
Al a Tyr Leu Thr Tyr Ser Leu G y Phe Al a Val Met G y Leu Tyr Tyr
210          215          220
t at gcg ccg ct c t t t gt c t t t gct t cg t t c ct c gt c at t acg acc t t c      720
Tyr Al a Pro Leu Phe Val Phe Al a Ser Phe Leu Val Il e Thr Thr Phe
225          230          235
t t g cac cac aac gac gaa gcg acg ccg t gg t ac ggc gac t cg gag t gg      768
Leu Hi s Hi s Asn Asp G u Al a Thr Pro Tr p Tyr G y Asp Ser G u Tr p
240          245          250
acg t ac gt c aag ggc aac ct c t cg agc gt c gac cgc t cg t ac ggc gcg      816
Thr Tyr Val Lys G y Asn Leu Ser Ser Val Asp Arg Ser Tyr G y Al a
255          260          265
t t c gt g gac aac ct g agc cac cac at t ggc acg cac cag gt c cac cac      864
Phe Val Asp Asn Leu Ser Hi s Hi s Il e G y Thr Hi s G n Val Hi s Hi s
270          275          280
t t g t t c ccg at c att ccg cac t ac aag ct c aac gaa gcc acc aag cac      912
Leu Phe Pro Il e Il e Pro Hi s Tyr Lys Leu Asn G u Al a Thr Lys Hi s
285          290          300
t t t gcg gcc gcg t ac ccg cac ct c gt g cgc agg aac gac gag ccc at c      960
Phe Al a Al a Al a Tyr Pro Hi s Leu Val Arg Arg Asn Asp G u Pro Il e
305          310          315
at c acg gcc t t c t t c aag acc gcg cac ct c t t t gt c aac t ac ggc gct      1008
Il e Thr Al a Phe Phe Lys Thr Al a Hi s Leu Phe Val Asn Tyr G y Al a
320          325          330
gt g ccc gag acg gcg cag at c t t c acg ct c aaa gag t cg gcc gcg gcc      1056
Val Pro G u Thr Al a G n Il e Phe Thr Leu Lys G u Ser Al a Al a Al a
335          340          345
gcc aag gcc aag t cg gac t aa
Al a Lys Al a Lys Ser Asp
350          355

```

<210> 196  
 <211> 358  
 <212> PRT  
 <213> Saprol egni a di cl i na

```

<400> 196
Met Thr G u Asp Lys Thr Lys Val G u Phe Pro Thr Leu Thr G u Leu
1          5          10          15
Lys Hi s Ser Il e Pro Asn Al a Cys Phe G u Ser Asn Leu G y Leu Ser
20          25          30
Leu Tyr Tyr Thr Al a Arg Al a Il e Phe Asn Al a Ser Al a Ser Al a Al a
35          40          45
Leu Leu Tyr Al a Al a Arg Ser Thr Pro Phe Il e Al a Asp Asn Val Leu
50          55          60
Leu Hi s Al a Leu Val Cys Al a Thr Tyr Il e Tyr Val G n G y Val Il e
65          70          75          80
Phe Tr p G y Phe Phe Thr Val G y Hi s Asp Cys G y Hi s Ser Al a Phe
85          90          95
Ser Arg Tyr Hi s Ser Val Asn Phe Il e Il e G y Cys Il e Met Hi s Ser
100          105          110
Al a Il e Leu Thr Pro Phe G u Ser Tr p Arg Val Thr Hi s Arg Hi s Hi s
115          120          125
Hi s Lys Asn Thr G y Asn Il e Asp Lys Asp G u Il e Phe Tyr Pro Hi s
130          135          140

```

PF58307.txt

Arg Ser Val Lys Asp Leu G n Asp Val Arg G n Tr p Val Tyr Thr Leu  
 145 Gly Gly Ala Trp Phe Val Tyr Leu Lys Val Gly Tyr Ala Pro Arg Thr  
 165 Met Ser His Phe Asp Pro Trp Asp Pro Leu Leu Leu Arg Arg Ala Ser  
 180 Ala Val Ile Val Ser Leu Gly Val Trp Ala Ala Phe Phe Ala Ala Tyr  
 195 Ala Tyr Leu Thr Tyr Ser Leu Gly Phe Ala Val Met Gly Leu Tyr Tyr  
 210 Tyr Ala Pro Leu Phe Val Phe Ala Ser Phe Leu Val Ile Thr Thr Phe  
 225 Leu His His Asn Asp Gu Ala Thr Pro Trp Tyr Gly Asp Ser Gu Trp  
 245 Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala  
 260 Phe Val Asp Asn Leu Ser His His Ile Gly Thr His G n Val His His  
 275 Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Gu Ala Thr Lys His  
 290 Phe Ala Ala Ala Tyr Pro His Leu Val Arg Arg Asn Asp Gu Pro Ile  
 305 Ile Thr Ala Phe Phe Lys Thr Ala His Leu Phe Val Asn Tyr Gly Ala  
 325 Val Pro Gu Thr Ala G n Ile Phe Thr Leu Lys Gu Ser Ala Ala Ala  
 340 Ala Lys Ala Lys Ser Asp  
 355

<210> 197  
 <211> 359  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <221> Variant  
 <222> (2)..(5)  
 <223> xaa in position 2 to 5 is any amino acid  
 <220>  
 <221> Variant  
 <222> (12)..(12)  
 <223> xaa in position 12 is any amino acid  
 <220>  
 <221> Variant  
 <222> (14)..(14)  
 <223> xaa in position 14 is any amino acid  
 <220>  
 <221> Variant  
 <222> (16)..(16)  
 <223> xaa in position 16 is any amino acid  
 <220>  
 <221> Variant  
 <222> (18)..(19)  
 <223> xaa in position 18 to 19 is any amino acid  
 <220>  
 <221> Variant  
 <222> (23)..(26)  
 <223> xaa in position 23 to 26 is any amino acid  
 <220>  
 <221> Variant  
 <222> (33)..(33)  
 <223> xaa in position 33 is any amino acid  
 <220>  
 <221> Variant  
 <222> (35)..(38)  
 <223> xaa in position 35 to 38 is any amino acid  
 <220>

<221> Variant  
<222> (40)..(40)  
<223> xaa in position 40 is any amino acid  
<220>  
<221> Variant  
<222> (42)..(45)  
<223> xaa in position 42 to 45 is any amino acid  
<220>  
<221> Variant  
<222> (47)..(48)  
<223> xaa in position 47 to 48 is any amino acid  
<220>  
<221> Variant  
<222> (51)..(52)  
<223> xaa in position 51 to 52 is any amino acid  
<220>  
<221> Variant  
<222> (54)..(60)  
<223> xaa in position 54 to 60 is any amino acid  
<220>  
<221> Variant  
<222> (62)..(62)  
<223> xaa in position 62 is any amino acid  
<220>  
<221> Variant  
<222> (64)..(65)  
<223> xaa in position 64 to 65 is any amino acid  
<220>  
<221> Variant  
<222> (67)..(68)  
<223> xaa in position 67 to 68 is any amino acid  
<220>  
<221> Variant  
<222> (71)..(72)  
<223> xaa in position 71 to 72 is any amino acid  
<220>  
<221> Variant  
<222> (75)..(76)  
<223> xaa in position 75 to 76 is any amino acid  
<220>  
<221> Variant  
<222> (87)..(87)  
<223> xaa in position 87 is any amino acid  
<220>  
<221> Variant  
<222> (90)..(90)  
<223> xaa in position 90 is any amino acid  
<220>  
<221> Variant  
<222> (97)..(98)  
<223> xaa in position 97 to 98 is any amino acid  
<220>  
<221> Variant  
<222> (101)..(102)  
<223> xaa in position 101 to 102 is any amino acid  
<220>  
<221> Variant  
<222> (104)..(105)  
<223> xaa in position 104 to 105 is any amino acid  
<220>  
<221> Variant  
<222> (109)..(109)  
<223> xaa in position 109 is any amino acid  
<220>  
<221> Variant  
<222> (118)..(119)  
<223> xaa in position 118 to 119 is any amino acid  
<220>

<221> Variant  
<222> ( 133).. ( 133)  
<223> xaa in position 133 is any amino acid  
<220>  
<221> Variant  
<222> ( 136).. ( 136)  
<223> xaa in position 136 is any amino acid  
<220>  
<221> Variant  
<222> ( 140).. ( 140)  
<223> xaa in position 140 is any amino acid  
<220>  
<221> Variant  
<222> ( 142).. ( 144)  
<223> xaa in position 142 to 144 is any amino acid  
<220>  
<221> Variant  
<222> ( 146).. ( 149)  
<223> xaa in position 146 to 149 is any amino acid  
<220>  
<221> Variant  
<222> ( 151).. ( 155)  
<223> xaa in position 151 to 155 is any amino acid  
<220>  
<221> Variant  
<222> ( 158).. ( 158)  
<223> xaa in position 158 is any amino acid  
<220>  
<221> Variant  
<222> ( 161).. ( 162)  
<223> xaa in position 161 to 162 is any amino acid  
<220>  
<221> Variant  
<222> ( 165).. ( 166)  
<223> xaa in position 165 to 166 is any amino acid  
<220>  
<221> Variant  
<222> ( 168).. ( 169)  
<223> xaa in position 168 to 169 is any amino acid  
<220>  
<221> Variant  
<222> ( 172).. ( 174)  
<223> xaa in position 172 to 174 is any amino acid  
<220>  
<221> Variant  
<222> ( 177).. ( 177)  
<223> xaa in position 177 is any amino acid  
<220>  
<221> Variant  
<222> ( 179).. ( 180)  
<223> xaa in position 179 to 180 is any amino acid  
<220>  
<221> Variant  
<222> ( 183).. ( 184)  
<223> xaa in position 183 to 184 is any amino acid  
<220>  
<221> Variant  
<222> ( 186).. ( 187)  
<223> xaa in position 186 to 187 is any amino acid  
<220>  
<221> Variant  
<222> ( 191).. ( 192)  
<223> xaa in position 191 to 192 is any amino acid  
<220>  
<221> Variant  
<222> ( 195).. ( 195)  
<223> xaa in position 195 is any amino acid  
<220>

<221> Variant  
<222> (196)..(197)  
<223> xaa in position 196 to 197 is any or no amino acid  
<220>  
<221> Variant  
<222> (199)..(199)  
<223> xaa in position 199 is any amino acid  
<220>  
<221> Variant  
<222> (202)..(202)  
<223> xaa in position 202 is any amino acid  
<220>  
<221> Variant  
<222> (204)..(205)  
<223> xaa in position 204 to 205 is any amino acid  
<220>  
<221> Variant  
<222> (206)..(207)  
<223> xaa in position 206 to 207 is any or no amino acid  
<220>  
<221> Variant  
<222> (210)..(212)  
<223> xaa in position 210 to 212 is any amino acid  
<220>  
<221> Variant  
<222> (215)..(217)  
<223> xaa in position 215 to 217 is any amino acid  
<220>  
<221> Variant  
<222> (219)..(220)  
<223> xaa in position 219 to 220 is any amino acid  
<220>  
<221> Variant  
<222> (224)..(224)  
<223> xaa in position 224 is any amino acid  
<220>  
<221> Variant  
<222> (226)..(226)  
<223> xaa in position 226 is any amino acid  
<220>  
<221> Variant  
<222> (230)..(230)  
<223> xaa in position 230 is any amino acid  
<220>  
<221> Variant  
<222> (232)..(232)  
<223> xaa in position 232 is any amino acid  
<220>  
<221> Variant  
<222> (245)..(245)  
<223> xaa in position 245 is any amino acid  
<220>  
<221> Variant  
<222> (250)..(250)  
<223> xaa in position 250 is any amino acid  
<220>  
<221> Variant  
<222> (271)..(272)  
<223> xaa in position 271 to 272 is any amino acid  
<220>  
<221> Variant  
<222> (278)..(278)  
<223> xaa in position 278 is any amino acid  
<220>  
<221> Variant  
<222> (284)..(284)  
<223> xaa in position 284 is any amino acid  
<220>

<221> Variant  
 <222> (297)..(298)  
 <223> xaa in position 297 to 298 is any amino acid  
 <220>  
 <221> Variant  
 <222> (301)..(302)  
 <223> xaa in position 301 to 302 is any amino acid  
 <220>  
 <221> Variant  
 <222> (304)..(305)  
 <223> xaa in position 304 to 305 is any amino acid  
 <220>  
 <221> Variant  
 <222> (307)..(307)  
 <223> xaa in position 307 is any amino acid  
 <220>  
 <221> Variant  
 <222> (309)..(309)  
 <223> xaa in position 309 is any amino acid  
 <220>  
 <221> Variant  
 <222> (313)..(314)  
 <223> xaa in position 313 to 314 is any amino acid  
 <220>  
 <221> Variant  
 <222> (320)..(320)  
 <223> xaa in position 320 is any amino acid  
 <220>  
 <221> Variant  
 <222> (324)..(327)  
 <223> xaa in position 324 to 327 is any amino acid  
 <220>  
 <221> Variant  
 <222> (329)..(330)  
 <223> xaa in position 329 to 330 is any amino acid  
 <220>  
 <221> Variant  
 <222> (334)..(334)  
 <223> xaa in position 334 is any amino acid  
 <220>  
 <221> Variant  
 <222> (336)..(338)  
 <223> xaa in position 336 to 338 is any amino acid  
 <220>  
 <221> Variant  
 <222> (340)..(341)  
 <223> xaa in position 340 to 341 is any amino acid  
 <220>  
 <221> Variant  
 <222> (347)..(348)  
 <223> xaa in position 347 to 348 is any amino acid  
 <220>  
 <221> Variant  
 <222> (349)..(352)  
 <223> xaa in position 349 to 352 is any or no amino acid  
 <220>  
 <221> Variant  
 <222> (357)..(357)  
 <223> xaa in position 357 is any amino acid

<400> 197  
 Lys xaa xaa xaa xaa Phe Pro Thr Leu Thr Glu xaa Lys xaa Ser xaa  
 1 5 10 15  
 Pro xaa xaa Cys Phe Glu xaa xaa xaa Leu Ser Leu Tyr Tyr Thr  
 20 25 30  
 xaa Arg xaa xaa xaa xaa Ala xaa Ala xaa xaa xaa xaa Leu xaa xaa  
 35 40 45  
 Ala Arg xaa xaa Pro xaa xaa xaa xaa xaa xaa xaa Leu xaa Ala xaa  
 Seite 240

PF58307.txt

```

50          55          60
xaa Cys xaa xaa Tyr Ile xaa xaa G n G y xaa xaa Phe Trp G y Phe
65      70      75      80
Phe Thr Val G y H i s Asp xaa G y H i s xaa Al a Phe Ser Arg Tyr H i s
85      90      95
xaa xaa Asn Phe xaa xaa G y xaa xaa M et H i s Ser xaa Ile Leu Thr
100     105     110
Pro Phe G u Ser Tr p xaa xaa Thr H i s Arg H i s H i s H i s Lys Asn Thr
115     120     125
G y Asn Ile Asp xaa Asp G u xaa Phe Tyr Pro xaa Arg xaa xaa xaa
130     135     140
Asp xaa xaa xaa xaa Arg xaa xaa xaa xaa xaa Leu G y xaa Al a Tr p
145     150     155     160
xaa xaa Tyr Leu xaa xaa G y xaa xaa Pro Arg xaa xaa xaa H i s Phe
165     170     175
xaa Pro xaa xaa Pro Leu xaa xaa Arg xaa xaa Ser Al a Val xaa xaa
180     185     190
Ser Leu xaa xaa xaa Al a xaa Phe xaa Al a xaa xaa xaa Tyr
195     200     205
Leu xaa xaa xaa Leu G y xaa xaa xaa M et xaa xaa Tyr Tyr Tyr xaa
210     215     220
Pro xaa Phe Val Phe xaa Ser xaa Leu Val Ile Thr Thr Phe Leu H i s
225     230     235     240
H i s Asn Asp G u xaa Thr Pro Tr p Tyr xaa Asp Ser G u Tr p Thr Tyr
245     250     255
Val Lys G y Asn Leu Ser Ser Val Asp Arg Ser Tyr G y Al a xaa xaa
260     265     270
Asp Asn Leu Ser H i s xaa Ile G y Thr H i s G n xaa H i s H i s Leu Phe
275     280     285
Pro Ile Ile Pro H i s Tyr Lys Leu xaa xaa Al a Thr xaa xaa Phe xaa
290     295     300
xaa Al a xaa Pro xaa Leu Val Arg xaa xaa Asp G u Pro Ile Ile xaa
305     310     315     320
Al a Phe Phe xaa xaa xaa xaa Leu xaa xaa Asn Tyr G y xaa Val xaa
325     330     335
xaa xaa Al a xaa xaa Phe Thr Leu Lys G u xaa xaa xaa xaa xaa
340     345     350
Al a Al a Al a Lys xaa Lys Ser
355

```

- <210> 198
- <211> 60
- <212> PRT
- <213> Artificial sequence
  
- <220>
- <221> Variant
- <222> ( 10) .. ( 10)
- <223> xaa in position 10 is Al a or G u
- <220>
- <221> Variant
- <222> ( 15) .. ( 15)
- <223> xaa in position 15 is Al a or G y
- <220>
- <221> Variant
- <222> ( 36) .. ( 36)
- <223> xaa in position 36 is Phe or Leu
- <220>
- <221> Variant
- <222> ( 37) .. ( 37)
- <223> xaa in position 37 is Ile or Val
- <220>
- <221> Variant
- <222> ( 43) .. ( 43)
- <223> xaa in position 43 is any amino acid
- <220>
- <221> Variant

&lt;222&gt; (49)..(49)

&lt;223&gt; xaa in position 49 is Ile or Val

&lt;400&gt; 198

```

Thr Thr Phe Leu His His Asn Asp Gu xaa Thr Pro Trp Tyr xaa Asp
 1   1 5   5   5   5   5   5   5   5   5   5   5   5   5   5   5
Ser Gu Trp Thr Tyr Val Lys Gy Asn Leu Ser Ser Val Asp Arg Ser
   10 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
Tyr Gy Ala xaa xaa Asp Asn Leu Ser His xaa Ile Gy Thr His Gn
   30 35 35 35 35 35 35 35 35 35 35 35 35 35 35 35
xaa His His Leu Phe Pro Ile Ile Pro His Tyr Lys
   40 45 45 45 45 45 45 45 45 45 45 45 45 45 45 45
   50 55 55 55 55 55 55 55 55 55 55 55 55 55 55 55

```

&lt;210&gt; 199

&lt;211&gt; 60

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (6)..(6)

&lt;223&gt; xaa in position 6 is Ala or Cys

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (9)..(9)

&lt;223&gt; xaa in position 9 is Gly or Ser

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (16)..(16)

&lt;223&gt; xaa in position 16 is any amino acid

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (17)..(17)

&lt;223&gt; xaa in position 17 is Leu or Val

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (20)..(21)

&lt;223&gt; xaa in position 20 to 21 is Ile or Val

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (23)..(23)

&lt;223&gt; xaa in position 23 is Cys or Thr

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (24)..(24)

&lt;223&gt; xaa in position 24 is Phe or Ile

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (28)..(28)

&lt;223&gt; xaa in position 28 is Ala or Leu

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (37)..(37)

&lt;223&gt; xaa in position 37 is Lys or Arg

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (38)..(38)

&lt;223&gt; xaa in position 38 is Leu or Val

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (52)..(52)

&lt;223&gt; xaa in position 52 is Lys or Arg

&lt;220&gt;

&lt;221&gt; Variant

&lt;222&gt; (55)..(55)

&lt;223&gt; xaa in position 55 is Ile or Val

&lt;220&gt;



<221> Variant

<222> (59)..(59)

<223> xaa in position 59 is any amino acid

<400> 199

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