



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e Comércio Exterior
Instituto Nacional de Propriedade Industrial

(21) **PI0707944-3 A2**

(22) Data de Depósito: 13/02/2007
(43) Data da Publicação: 17/05/2011
(RPI 2106)



(51) *Int.Cl.:*
C12N 15/54
C12N 15/53
C12P 7/64
A01H 5/00

(54) Título: **ÁCIDO NUCLEÍCO**

(30) Prioridade Unionista: 16/02/2006 GB 06 03160.3

(73) Titular(es): Basf Plant Science GMBH

(72) Inventor(es): Johnathan Napier, Monica Venegas Caleron,
Olga Sayanova

(74) Procurador(es): Dannemann, Siemsen, Bigler &
Ipanema Moreira

(86) Pedido Internacional: PCT GB2007000491 de 13/02/2007

(87) Publicação Internacional: WO 2007/093776 de 23/08/2007

(57) Resumo: ÁCIDO NUCLEÍCO. A presente invenção refere-se ao ácido nucleico derivado de *Perkinsus marinus* que codifica uma enzima de 9-elongase, uma $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase. Todas as seqüências de codificação podem ser transcritas como uma única transcrição.

Relatório Descritivo da Patente de Invenção para "ÁCIDO NUCLÉICO".

A presente invenção refere-se ao ácido nucléico derivado de *Perkinsus marinus* que codifica um 9-elongase, $\Delta 8$ -dessaturase e uma enzima de $\Delta 5$ -dessaturase. Todas as seqüências de codificação podem ser transcritas como uma única transcrição, que simplifica o processo de transformação de células requerido para expressar todas as três proteínas. A invenção também refere-se às seqüências de codificação individuais e às proteínas codificadas por estas seqüências como também a um processo para converter o ácido linoléico em ácido araquidônico.

Os ácidos graxos e triacilglicerídeos têm uma multiplicidade de aplicações na indústria alimentícia, em nutrição animal, em cosméticos e no setor farmacológico. Dependendo se eles são ácidos graxos saturados ou não saturados ou então triacilglicerídeos com um conteúdo elevado de ácidos graxos saturados ou não saturados, eles são adequados para muitas aplicações diferentes. Os ácidos graxos poliinsaturados, tal como ácido linoléico e ácido linolênico são essenciais para mamíferos, uma vez que eles não podem ser produzidos pelos mesmos. Os ácidos graxos $\omega 3$ e ácidos graxos $\omega 6$ poliinsaturados são então um componente importante na nutrição animal e humana.

A seguir, os ácidos graxos poliinsaturados são referidos como PUFA, PUFAs, LCPUFA ou LCPUFAs (ácidos graxos poliinsaturados, PUFA, ácidos graxos poliinsaturados de cadeia longa, LCPUFA).

Os vários ácidos graxos e triglicerídeos são principalmente obtidos de microorganismos tal como *Mortierella* e *Schizochytrium* ou de plantas produtoras de óleo, tal como soja, óleo de semente de colza, algas tal como *Cryptothecodinium* ou *Phaeodactylum* e outros, onde eles são obtidos, como uma regra, na forma de seus triacilglicerídeos (= triglicerídeos = trigliceróis). Entretanto, eles também podem ser obtidos de animais, tal como, por exemplo, peixe. Os ácidos graxos livres são preparados vantajosamente através de hidrólise. Os ácidos graxos poliinsaturados de cadeia longa, tal como, ácido docosaexanóico (= DHA, $C_{22}:6^{\Delta 4,7,10,13,16,19}$), ácido eicosapentanóico

(= EPA, C20:5^{Δ5,8, 11, 14, 17}), ácido araquidônico (= ARA, C20:4^{Δ5,8, 11, 14}), ácido di-homo-γ-linolênico (C20:3^{Δ8, 11, 14}) ou ácido de docosapentanóico (DPA, C22:5^{Δ7, 10, 13, 16, 19}), não são sintetizados em colheitas de óleo, tal como óleo de semente de colza, soja, girassol ou açafrão. As fontes naturais convencionais destes ácidos graxos são peixes, tal como arenque, salmão, sardinha, peixe vermelho, enguia, carpa, truta, hipoglosso, cavala, zander ou atum, ou algas.

Dependendo do uso pretendido, os óleos com ácidos graxos saturados ou não saturados são preferidos. Em nutrição humana, por exemplo, os lipídios com ácidos graxos não saturados, especificamente ácidos graxos poliinsaturados, são preferidos. Os ácidos ω3-graxos poliinsaturados são ditos ter um efeito positivo no nível de colesterol no sangue e desse modo na possibilidade de prevenir doença cardíaca. O risco de doença cardíaca, acidente vascular cerebral ou hipertensão pode ser reduzido notadamente adicionando-se estes ácidos ω3-graxos à comida. Além disso, os ácidos ω3-graxos têm um efeito positivo em processos inflamatórios, especificamente em processos cronicamente inflamatórios, em associação com doenças imunológicas, tal como artrite reumática. Eles são então adicionados aos comestíveis, especificamente aos comestíveis dietéticos, ou são empregados em medicamentos. Os ácidos graxos ω6 tal como ácido araquidônico tendem a ter um efeito negativo nestes distúrbios com relação a estas doenças reumáticas por causa do influxo dietético habitual.

Ácidos graxos ω3 e ω6 são precursores de hormônios de tecido, conhecidos como eicosanóides, tal como o prostaglandinas que são derivadas de ácido di-homo-linolênico, ácido araquidônico e ácido eicosapentanóico, e dos tromboxanos e leucotrienos que são derivados de ácido araquidônico e ácido de eicosapentanóico. Os eicosanóides (conhecidos como as séries PG₂) que são geralmente formados de ácidos ω6-graxos promovem as reações inflamatórias, ao mesmo tempo em que os eicosanóides (conhecidos como as séries PG₃) de ácidos graxos ω3 têm pouco ou nenhum efeito pró-inflamatório.

Devido às características positivas dos ácidos graxos poliinsatu-

rados, não houve nenhuma falta de tentativas no passado para fabricar genes disponíveis que estão envolvidos na síntese destes ácidos graxos ou triglicerídeos para a produção de óleos em vários organismos com um conteúdo modificado de ácidos graxos não saturados. Desse modo, WO 91/13972 e sua U.S. equivalente descreve uma $\Delta 9$ -dessaturase. WO 93/1245 reivindica uma $\Delta 15$ -dessaturase e WO 94/11516 uma $\Delta 12$ -dessaturase. Por exemplo, dessaturases adicionais são descritas nos EP-A-O 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP-A-O 794 250, Stukey e outros, J. Biol. Chem., 265, 1990: 20144-20149, Wada e outros, Nature 347, 1990: 200-203 ou Huang e outros, Lipids 34, 1999: 649-659. Entretanto, a caracterização bioquímica das várias dessaturases tem sido insuficiente até hoje uma vez que as enzimas, sendo proteínas ligadas à membrana, apresentam grande dificuldade no seu isolamento e caracterização (McKeon e outros, Methods in Enzymol. 71, 1981: 12141-12147, Wang e outros, Plante Physiol. Biochem., 26, 1988: 777-792). Como uma regra, as dessaturases ligadas à membrana são caracterizadas por serem introduzidas em um organismo adequado que é subsequenteiramente analisado quanto à atividade de enzima analisando-se os materiais de partida e os produtos. $\Delta 6$ -Dessaturases são descritas nos WO 93/06712, U.S. 5.614.393, US5614393, WO 96/21022, WO 00/21557 e WO 99/27111 e a aplicação para a produção de ácidos graxos em organismos transgênicos é descrita nos WO 98/46763, WO 98/46764 e WO 98/46765. Neste contexto, a expressão de várias dessaturases e a formação de ácidos graxos poliinsaturados, é também descrita e reivindicada no WO 99/64616 ou WO 98/46776. Com respeito à expressão e eficácia de dessaturases e seu efeito na formação de ácidos graxos poliinsaturados, deve ser notado que a expressão de uma única dessaturase como descrito até hoje somente resultou em baixos conteúdos de ácidos graxos não saturado /lipídeos tal como, por exemplo, ácido γ -linolênico e ácido estearidônico. Além disso, uma mistura de ácidos graxos $\omega 3$ e $\omega 6$ foi obtida, como uma regra.

Os microorganismos especialmente adequados para a produção de PUFA's são microalgas tal como espécies de *Phaeodactylum tricornutum*,

Porphiridium, espécies de Thraustochytrium, espécies de Schizochytrium ou espécies de Crypthecodinium, ciliados tais como Stylonychia ou Colpidium, fungos tais como Mortierella, Entomophthora ou Mucor e/ou musgos tais como Physcomitrella, Ceratodon e Marchantia (R. Vazhappilly & F. Chen (1998) Botanica Marina 41: 553-558; K. Totani & K. Oba (1987) Lipids 22: 1060-1062; M. Akimoto e outros (1998) Appl. Biochemistry e Biotechnology 73: 269-278). A seleção de cepa resultou no desenvolvimento de várias cepas mutantes dos microorganismos em questão que produzem uma série de compostos desejáveis incluindo PUFA's. Entretanto, a mutação e seleção de cepas com uma produção melhorada de uma molécula particular tal como os ácidos graxos poliinsaturados, é um processo demorado e difícil. Isto é por que os métodos recombinantes como descrito acima são preferidos sempre que possível.

Porém, somente quantidades limitadas dos ácidos graxos poliinsaturados desejados tais como DPA, EPA ou ARA podem ser produzidas com a ajuda dos microorganismos acima mencionados, e, dependendo do microorganismo usado, estes geralmente são obtidos como misturas de ácido graxo de, por exemplo, EPA, DPA e ARA.

Uma variedade de séries de reação sintéticas está sendo descrita pela síntese de ácido araquidônico, ácido de eicosapentanoico (EPA) e ácido docosaexanoico (DHA) (figura 1). Desse modo, são produzidos EPA ou DHA em bactérias marinhas tal como Vibrio sp. ou Shewanella sp. pelas séries de reação de policetideo (Yu, R., e outros, Lipids 35:1061-1064, 2000; Takeyama, H., e outros, Microbiology 143:2725-2731, 1997).

Uma estratégia alternativa é a atividade alternada de dessaturas e elongases (Zank, T.K. e outros, Plant Journal 31:255-268, 2002; Sakuradani, E., e outros, Gene 238:445-453, 1999). Uma modificação da série de reação descrita acima por $\Delta 6$ -dessaturase, $\Delta 6$ -elongase, $\Delta 5$ -dessaturase, $\Delta 5$ -elongase e $\Delta 4$ -dessaturase é a série de reação de Sprecher (Sprecher 2000, Biochim. Biophys. Acta 1486:219-231) em mamíferos. Em vez da $\Delta 4$ -dessaturação, uma outra etapa de alongamento é efetuado aqui para produzir C2A, seguido por uma outra $\Delta 6$ -dessaturação e finalmente β -oxidação

para produzir o comprimento de cadeia C22. Desse modo, o que é conhecido como série de reação de Sprecher (veja figura 1) é, entretanto, não adequado para a produção em plantas e microorganismos uma vez que os mecanismos reguladores não são conhecidos.

5 Dependendo do seu padrão de dessaturação, os ácidos graxos poliinsaturados podem ser divididos em duas classes grandes, viz. Ácidos graxos $\omega 6$ ou $\omega 3$, que diferem com respeito às suas atividades metabólicas e funcionais (figura 1).

O material de partida para a série de reação metabólica $\omega 6$ é o
10 ácido linoléico de ácido graxo ($18:2^{\Delta 9, 12}$) ao mesmo tempo em que a série de reação $\omega 3$ prossegue através de ácido linolênico ($18:3^{\Delta 9, 12, 15}$). O ácido linolênico é formado pela atividade de uma dessaturase $\omega 3$ (Tocher e outros 1998, Prog. Lipid Res. 37, 73-117; Domergue e outros 2002, Eur. J. Biochem. 269, 4105-4113).

15 Os mamíferos, e desse modo também os humanos, não têm nenhuma atividade de dessaturase correspondente ($\Delta 12$ - e $\Delta \omega 3$ -dessaturase) e têm que absorver estes ácidos graxos (ácidos graxo essenciais) pela comida. Começando com estes precursores, os ácidos graxos poliinsaturados fisiologicamente importantes ácido araquidônico (= ARA, $20:4^{\Delta 5, 8, 11, 14}$), um
20 ácido graxo $\omega 6$ e os dois ácidos eicosapentanóicos de ácidos graxos $\omega 3$ (= EPA, $20:5^{\Delta 5, 8, 11, 14, 17}$) e ácido docosaexanóico (DHA, $22:6^{\Delta 4, 7, 10, 13, 17, 19}$) são sintetizados pela seqüência de reações de dessaturase e elongase. A aplicação de ácidos graxos $\omega 3$ mostra a atividade terapêutica descrita acima no tratamento de doenças cardiovasculares (Shimikawa 2001, World Rev. Nutr.
25 Diet. 88, 100-108), Entzündungen (Calder 2002, Proc. Nutr. Soc. 61, 345-358) e Arthritis (Cleland e James 2000, J., Rheumatol. 27, 2305-2307).

O alongamento de ácidos graxo, através de elongases, através de 2 ou 4 átomos de C é de importância crucial para a produção de C_{20} - e C_{22} -PUFAs, respectivamente. Este processo prossegue por 4 etapas. A primeira etapa é a condensação de malonil-CoA com o ácido graxo-acil-CoA
30 através de cetoacil-CoA sintase (KCS, a seguir chamado elongase). Isto é seguido por uma etapa de redução (cetoacil-CoA reductase, KCR), uma eta-

pa de desidratação (desidratase) e uma etapa de redução final (enoi-CoA reductase). Foi postulado que a atividade de elongase afeta a especificidade e taxa do processo inteiro (Millar e Kunst, 1997 Plant Journal 12:121-131).

Houve um número grande de tentativas no passado para obter genes de elongase. Millar e Kunst, 1997 (Plant Journal 12:121-131) e Millar e outros, 1999, (Plant Cell 11:825-838) descreve a caracterização de elongases de planta para a síntese de ácidos graxo de cadeia longa monoinsaturados (C22:1) e para a síntese de ácidos graxos de cadeia muito longa para a formação de ceras em plantas (C₂₈-C₃₂). As descrições com respeito à síntese de ácido araquidônico e EPA, são encontradas, por exemplo, nos WO0159128, WO0012720, WO02077213 e WO0208401. A síntese de ácidos C₂₄-graxos poliinsaturados é descrita, por exemplo, em Tvrdik e outros 2000, JCB 149:707-717 ou WO0244320.

As plantas mais altas compreendem ácidos graxos poliinsaturados tal como ácido linoléico (18:2^{Δ9, 12}) e ácido linolênico (18:3^{Δ9,12, 15}). ARA, EPA e DHA não são encontrados no óleo de semente de plantas mais altas, ou somente em quantidades pequenas (E. Ucciani: Nouveau Dictionnaire des Huiles Vegetales [New Dictionary of Vegetable Oils]. Technique & Documentation - Lavoisier, 1995. ISBN: 2-7430-0009-0). Porém, a produção de LCPUFAs em plantas mais altas, preferivelmente em colheitas de óleo tal como óleo de semente de colza, linhaça, girassol e soja, seria vantajosa uma vez que quantidades grandes de LCPUFAs de alta qualidade para a indústria alimentícia, nutrição animal e propósitos farmacêuticos poderiam ser obtidos economicamente. Para este fim, é vantajoso introduzir, em colheitas de óleo, genes que codificam enzimas da biosíntese de LCPUFA por métodos recombinantes e para expressá-los aqui. Estes genes podem codificar, por exemplo, Δ9-elongases, Δ8-dessaturases e/ou Δ5-dessaturases. Estes genes podem ser isolados vantajosamente de microorganismos e plantas mais baixas que produzem LCPUFAs e os incorporam nas membranas ou triacilglicerídeos. Desse modo, já foi possível isolar os genes de Δ6-dessaturase do musgo *Physcomitrella patens* e genes de Δ6-elongase de *P. patens* e do nematódeo *C. elegans*.

As primeiras plantas transgênicas que compreendem e expressam os genes que codificam enzimas de biossíntese de LCPUFA e que, como uma consequência, produzem LCPUFAs foram descritas pela primeira vez, por exemplo, na Patente DE-A-102 19 203 (processo para a produção de ácidos graxos poliinsaturados em plantas). Porém, estas plantas produzem LCPUFAs em quantidades que requerem otimização adicional para processar os óleos que estão presente nas plantas.

Como pode ser visto da Figura 1, os produtos da $\omega\Delta 6$ -série de reação podem ser modificados usando dessaturases apropriadas e, se necessário, elongases para produzir ácidos graxos $\omega 3$. Portanto, seria sumamente valioso desenvolver um produto que torne possível a produção de ARA em um organismo geneticamente modificado.

O parasita de protozoário de ostra *Perkinsus marinus* é capaz de sintetizar ácidos graxos saturados e não saturados, incluindo o ácido graxo essencial, ácido araquidônico [20:4(n - 6)], pela série de reação de $\Delta -8$ dessaturase. Surpreendentemente, os inventores presentes descobriram que *P. marinus* contém nucléico que codifica um $\Delta 9$ -elongase, um $\Delta 8$ -dessaturase e um $\Delta 5$ - dessaturase que podem todos ser transcritos como uma única transcrição. A seqüência de tamanho natural é mostrada como

SEQ ID NO: 1.

Então, em um primeiro aspecto da invenção, é fornecida uma seqüência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ - dessaturase e que seja selecionado do grupo que consiste em:

a) Seqüência de ácido nucléico que compreende resíduos de ácido nucléico 7668 a 12077 de SEQ ID NO: 1 ou um homólogo deste;

b) uma seqüência de ácido nucléico que hibridiza sob condições rigorosas com uma seqüência de ácido nucléico que compreende resíduos de ácido nucléico 7668 a 12077 de SEQ ID NO: 1;

c) uma seqüência de ácido nucléico, isolada que codifica os polipeptídeos com atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase, em que os polipeptídeos são selecionados do grupo que con-

siste em SEQ ID NOS 2, 3 e 4;

d) Um derivado de uma seqüência de ácido nucléico de SEQ ID NO: 1 que codifica os polipeptídeos com pelo menos 40% de identidade ao nível de aminoácido com SEQ ID NO: 2, SEQ ID NO: 3 e SEQ ID NO: 4; em
5 que os referidos polipeptídeos têm atividade de $\Delta 9$ -elongase, $\Delta 8$ - dessaturase e $\Delta 5$ -dessaturase.

A vantagem da seqüência de ácido nucléico da presente invenção é que, embora codifique três enzimas separadas, ela pode ser transcrita como uma única seqüência que torna muito mais simples preparar os vetores de clonagem e expressão que expressam todas as três enzimas.
10

Preferivelmente, a seqüência de ácido nucléico isolada de acordo com a invenção não é idêntica a SEQ ID NO 1 (seqüência 1047306867) propriamente dita.

No contexto da presente invenção "hibridizar sob condições rigorosas" é pretendido descrever as condições de hibridização e lavagem sob as quais as seqüências de nucleotídeo com pelo menos 60% de homologia para um outro normalmente permanecem hibridizadas com uma outra. As condições são preferivelmente tais que as seqüências com pelo menos cerca de 65%, preferivelmente pelo menos cerca de 70% e especialmente preferivelmente pelo menos 75% ou mais de homologia para uma outra normalmente permaneçam hibridizadas para uma outra. Estas condições rigorosas são conhecidas pelo trabalhador versado e são descritas, por exemplo, em Currents Protocols em Molecular Biology, John Wiley & Sons, N., Y. (1989), 6.3.1-6.3.6. Um exemplo não limitante preferido de condições de hibridização rigorosas, é a hibridização em 6 x de citrato de sódio/cloreto de sódio (= SSC) a aproximadamente 45°C, seguido por uma ou mais etapas de lavagem em 0,2 x SSC, 0,1% de SDS de 50 a 65°C. O trabalhador versado sabe que estas condições de hibridização diferem dependendo do tipo de ácido nucléico e, por exemplo, quando solventes orgânicos estão presentes, relativo à temperatura e concentração de tampão. Sob "condições de hibridização padrão", por exemplo, a temperatura de hibridização está, dependendo do tipo de ácido nucléico, entre 42°C e 58°C em tampão aquoso com
15
20
25
30

uma concentração de 0,1 a 5 x SSC (pH 7,2). Se solventes orgânicos, por exemplo, 50% de formamida, estiverem presentes no tampão acima mencionado, a temperatura sob condições padrão é aproximadamente 42°C. As condições de hibridização para híbridos de DNA:DNA, por exemplo, são 0,1 x SSC e 20°C a 45°C, preferivelmente 30°C a 45°C. As condições de hibridização para híbridos de DNA:RNA são, por exemplo, 0,1 x SSC e 30°C a 55°C, preferivelmente 45°C a 55°C. As condições de hibridização acima mencionadas são determinadas por via de exemplo para um ácido nucléico com aproximadamente 100 pb (= pares de base) em comprimento e com um conteúdo de G + C de 50% na ausência de formamida. O trabalhador versado sabe como determinar as condições de hibridização exigidas em base dos livros de ensino acima mencionados ou livros de ensino tal como Sambrook e outros, "Molecular Cloning", Cold Spring Harbor Laboratory, 1989; Hames e Higgins (Ed.) 1985, "Nucleic Acids Hybridization: A Practical Approach", IRL Press at Oxford University Press, Oxford; Brown (Ed.) 1991, "Essential Molecular Biology: A Practical Approach", IRL Press at Oxford University Press, Oxford.

Além disso, quando o relatório presente refere-se às moléculas de ácido nucléico isoladas de uma seqüência de nucleotídeo que hibridiza com uma das seqüências de nucleotídeo mostrada em SEQ ID NO: 1, SEQ ID NO: 3 ou com uma parte desta sob condições rigorosas, "uma parte desta" é compreendida como significando, de acordo com a invenção, que pelo menos 25 pares de base (= pb), 50 pb, 75 pb, 100 pb, 125 pb ou 150 pb, preferivelmente pelo menos 175 pb, 200 pb, 225 pb, 250 pb, 275 pb ou 300 pb, especialmente preferivelmente 350 pb, 400 pb, 450 pb, 500 pb ou mais pares de base são usados para a hibridização.

No contexto da presente invenção "Homólogos" da seqüência de ácido nucléico com a seqüência SEQ ID NO: 1 significa, por exemplo, variantes alélicas com pelo menos aproximadamente 50 ou 60%, preferivelmente pelo menos aproximadamente 60 ou 70%, mais preferivelmente pelo menos aproximadamente 70 ou 80%, 90% ou 95% e até mesmo mais preferivelmente pelo menos aproximadamente 85%, 86%, 87%, 88%, 89%, 90%,

91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade ou homologia com uma seqüência de nucleotídeo mostrada na SEQ ID NO: 1.

"Variantes alélicas" compreendem, em particular, variantes funcionais que podem ser obtidas por deleção, inserção ou substituição de nucleotídeos de/na seqüência, sendo pretendido, porém, que a atividade de enzima das proteínas resultantes que são sintetizadas seja retida vantajosamente para a inserção de um ou mais genes.

"Homólogos" também quer dizer homólogos bacterianos, fúngicos e de planta, seqüências truncadas, DNA ou RNA de filamento único da seqüência de DNA de codificação e não codificação e derivados tal como, por exemplo, variantes de promotor. Os promotores a montante das seqüências de nucleotídeo detalhadas podem ser modificados por uma ou mais trocas de nucleotídeo, através de inserção(s) deleção(s) sem a funcionalidade ou atividade dos promotores que são adversamente afetados, porém. É, além disso, possível que a modificação da seqüência promotora aumente a sua atividade ou que ela seja substituída completamente por promotores mais ativos, incluindo aqueles de organismos heterólogos.

Para determinar a porcentagem de homologia (= identidade) de duas seqüências de aminoácido, as seqüências são escritas uma sob a outra para uma comparação ideal (por exemplo, podem ser introduzidas aberturas na seqüência de uma proteína ou de um ácido nucléico para gerar um alinhamento ideal com a outra proteína ou o outro ácido nucléico). Então, o resíduo de aminoácido ou nucleotídeos nas posições de aminoácido ou posições de nucleotídeo correspondentes, é comparado. Se uma posição em uma seqüência estiver ocupada pelo mesmo resíduo de aminoácido ou pelo mesmo nucleotídeo como a posição correspondente na outra seqüência, então as moléculas são homólogas a esta posição (isto é, "homologia" de aminoácido ou ácido nucléico como usado no presente contexto corresponde a "identidade" do aminoácido ou ácido nucléico). A porcentagem de homologia entre as duas seqüências é uma função do número de posições que as seqüências compartilham (isto é, % de homologia = número de posições idênticas/número total de posições x 100). Os termos homologia e identidade

serão considerados então como sinônimos.

A homologia foi calculada sobre a região de seqüência de ácido nucléico ou aminoácido total. O trabalhador versado tem disponível uma série de programas que são com base em vários algoritmos para a comparação de várias seqüências. Aqui, os algoritmos de Needleman e Wunsch ou Smith e Waterman dão resultados particularmente seguros. O programa PileUp (o J. Mol. Evolution, 25, 351-360, 1987, Higgins e outros, CABIOS, 5 1989: 151-153) ou os programas Gap e BestFit [Needleman e Wunsch (J. Mol. Biol. 48; 443-453 (1970) e Smith e Waterman (Adv. Appl. Math. 2; 482-10 489 (1981))], que fazem parte do pacote de software GCG [Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, U.S.A. 53711 (1991)], foi usado para o alinhamento de seqüência. Os valores Ogy de seqüência que são indicados acima como uma porcentagem, foram determinado sobre a região de seqüência total que usa o programa GAP as seguintes colocações: 15 Peso GAP: 50, Peso de Comprimento: 3, Equilíbrio Médio: 10.000 e Desequilíbrio Médio: 0,000. A menos que de outro modo especificado, estes ajustes foram sempre usados como ajustes padrões para os alinhamentos de seqüência.

No contexto da presente invenção "atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase" é entendida como significando que uma 20 proteína codificada por um derivado de SEQ ID NO:1 ou resíduos de ácido nucléico 7668 a 12077 de SEQ ID NO: 1 retém uma atividade enzimática de pelo menos 10%, preferivelmente 20%, especialmente preferivelmente 30% e muito especialmente 40% comparada com as proteínas/enzimas codificadas pela seqüência SEQ ID NO: 1 ou resíduos de ácido nucléico 7668 a 25 12077 de SEQ ID NO: 1 e tttt pode desse modo catalisar a conversão de ácido linoléico em ácido araquidônico.

Embora seja freqüentemente extremamente útil transcrever ácido nucléico que codifica polipeptídeos com atividade de $\Delta 9$ -elongase, $\Delta 8$ - 30 dessaturase e $\Delta 5$ -dessaturase como uma única seqüência, pode haver algumas circunstâncias nas quais seja preferível para fazer uso de ácido nucléico que codifica uma única enzima, isto é, uma $\Delta 9$ -elongase, uma $\Delta 8$ -

dessaturase ou uma $\Delta 5$ -dessaturase.

Então, em um segundo aspecto da invenção é fornecida uma sequência de ácido nucléico isolada que codifica um polipeptídeo com atividade de $\Delta 9$ -elongase e que é selecionada do grupo que consiste em:

5 a) uma sequência que compreende resíduos de ácido nucléico 7668 a 9200 de SEQ ID NO: 1, SEQ ID NO: 9 ou um homólogo de uma destas;

b) seqüências de ácido nucléico que hibridizam sob condições rigorosas com uma sequência de ácido nucléico que compreende resíduos
10 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9;

c) uma sequência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 9$ -elongase, em que o polipeptídeo compreende SEQ ID NO: 2 ou SEQ ID NO: 10;

d) um derivado de um uma sequência que compreende resíduos
15 de ácido nucléico 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9, que codificam um polipeptídeo com pelo menos 40% de identidade no nível de aminoácido com SEQ ID NO: 2 ou SEQ ID NO: 10; em que o referido polipeptídeo tem atividade de $\Delta 9$ -elongase.

Em um terceiro aspecto da invenção, é fornecida uma sequência
20 de ácido nucléico isolada que codifica um polipeptídeo com atividade de $\Delta 8$ -dessaturase e que é selecionado do grupo que consiste em:

a) uma sequência que compreende resíduos de ácido nucléico 9351 a 10724 de SEQ ID NO: 1 ou um homólogo deste;

b) seqüências de ácido nucléico que hibridizam sob condições
25 rigorosas com uma sequência de ácido nucléico que compreende resíduos 9351 a 10724 de SEQ ID NO: 1;

c) uma sequência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 8$ -dessaturase, em que o polipeptídeo compreende SEQ ID NO: 3;

d) um derivado de um uma sequência que compreende resíduos
30 de ácido nucléico 9351 a 10724 de SEQ ID NO: 1 que codifica um polipeptídeo com pelo menos 40% de identidade no nível de aminoácido com SEQ ID

NO: 3; em que o referido polipeptídeo tem atividade de $\Delta 8$ -dessaturase.

Em um quarto aspecto da invenção, é fornecida uma seqüência de ácido nucléico isolada que codifica um polipeptídeo com atividade de $\Delta 5$ -dessaturase e que é selecionado do grupo que consiste em:

5 a) uma seqüência que compreende resíduos de ácido nucléico 10842 a 12077 de SEQ ID NO: 1 ou um homólogo desta;

b) seqüências de ácido nucléico que hibridizam sob condições rigorosas com uma seqüência de ácido nucléico que compreende resíduos 10842 a 12077 de SEQ ID NO: 1;

10 c) uma seqüência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 5$ -dessaturase, em que o polipeptídeo compreende SEQ ID NO: 4;

d) um derivado de um uma seqüência que compreende resíduos de ácido nucléico 10842 a 12077 de SEQ ID NO: 1 que codifica um polipeptídeo com pelo menos 40% de identidade no nível de aminoácido com SEQ ID NO: 4; em que o referido polipeptídeo tem atividade de $\Delta 5$ -dessaturase.

Em ainda outro aspecto da invenção, é fornecido um polipeptídeo que é codificado por uma seqüência de ácido nucléico de quaisquer dos primeiro a quarto aspectos da invenção.

20 Vantajosamente, o polipeptídeo codificado por estas moléculas de ácido nucléico têm, pelo menos aproximadamente 50%, preferivelmente pelo menos aproximadamente 60% e mais preferivelmente pelo menos aproximadamente 70%, 80% ou 90% e preferivelmente pelo menos aproximadamente 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
25 96%, 97%, 98%, 99% ou mais de identidade com as seqüências de aminoácido mostradas nas SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 ou SEQ ID NO: 9.

30 As seqüências de ácido nucléico usadas no processo são introduzidas vantajosamente em um cassete de expressão que torna possível a expressão dos ácidos nucléicos em organismos tal como microorganismos ou plantas.

Então, em outro aspecto da invenção é fornecida um construto

de gene compreendendo uma seqüência de ácido nucléico que codifica um ou mais polipeptídeos com atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase como apresentado acima, operavelmente ligado com uma ou mais seqüências reguladoras.

5 No cassete de expressão, a seqüência de ácido nucléico que codifica $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase, é unida operavelmente com uma ou mais seqüências reguladoras, vantajosamente para realçar a expressão de gene. É pretendido que estas seqüências reguladoras tornem possível a expressão específica dos genes e proteínas. Depen-

10 dendo do organismo de hospedeiro, isto pode significar, por exemplo, que o gene é expressado e/ou superexpressado somente após a indução ter ocorrido, ou então que é expressado e/ou superexpressado imediatamente. Por exemplo, estas seqüências reguladoras tomam a forma de seqüências às

15 quais indutores ou repressores se ligam, desse modo controlando a expressão do ácido nucléico. Além destas novas seqüências reguladoras, ou em vez destas seqüências, os elementos reguladores naturais destas seqüên-

20 cias podem ainda estar presentes antes dos genes estruturais atuais e, se apropriado, podem ter sido geneticamente modificados de um tal modo que seu regulamento natural seja eliminado e a expressão dos genes seja inten-

25 sificada. Entretanto, o cassete de expressão (= construto de expressão = construto de gene) também pode ser mais simples na construto, isto é, nenhum sinal regulador adicional foi inserido antes da seqüência de ácido nucléico ou seus derivados, e o promotor natural junto com seu regulamento não foi removido. Ao invés, a seqüência reguladora natural foi mutada de um

30 tal modo que o regulamento já não mais ocorra e/ou expressão de gene, é intensificada. Estes promotores modificados também podem estar posicionados neles mesmos antes do gene natural na forma de partes de seqüên-

 cias (= promotor com partes das seqüências de ácido nucléico usadas de acordo com a invenção) para realçar a atividade. Além disso, o construto de

 gene pode vantajosamente também compreender uma ou mais das que são conhecidas como seqüências realçadoras em ligação operável com o pro-

 motor, o que torna possível uma expressão intensificada da seqüência de

ácido nucléico.

Sequências vantajosas adicionais, tal como outros elementos reguladores ou sequências terminadoras, também podem ser inseridas na extremidade 3' das sequências de DNA. Uma ou mais sequências que codificam enzimas que catalisam a conversão de ARA em um ácido graxo ω 3-insaturado tal como EPA ou DHA também podem estar presentes. Desse modo, por exemplo, as sequências que codificam um Δ 5-elongase, Δ 3-dessaturase e/ou Δ 4-dessaturase, podem estar presentes em uma ou mais cópias do cassete de expressão (= construto de gene). Preferivelmente, somente uma cópia dos genes está presente em cada cassete de expressão. Este construto de gene ou os construtos de gene podem ser expressados juntos no organismo de hospedeiro. Neste contexto, o(s) construto(s) de gene pode(m) ser inserido(s) em um ou mais vetores e está(ão) presente(s) na célula na forma livre, ou então é inserido(s) no genoma. É vantajoso para a inserção de outros genes no genoma quando os genes a serem expressados estão presentes juntos em um construto de gene.

Neste contexto, os fatores ou sequências reguladoras podem, como descrito acima, preferivelmente ter um efeito positivo na expressão de gene dos genes introduzidos, desse modo intensificando. Desse modo, um realce dos elementos reguladores, vantajosamente no nível transcricional, pode ocorrer usando-se sinais de transcrição fortes como intensificadores e/ou promotores. Além disso, porém, a translação intensificada também é possível, por exemplo, melhorando-se a estabilidade do mRNA.

As sequências reguladoras incluem, em particular, sequências de planta tal como as sequências promotoras e terminadoras. Os construtos podem ser vantajosamente estavelmente propagados em microorganismos, em particular em *E. coli* e *Agrobacterium tumefaciens*, sob condições seletivas, e torna possível a transferência de DNA heterólogo em plantas ou microorganismos.

As sequências reguladoras úteis estão presentes, por exemplo, em promotores tal como o promotor *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIq*, 17, T5, T3, *gal*, *trc*, *ara*, SP6, λ -PR ou λ -PL e são empregadas vantagio-

samente em bactérias Gram-negativas. Outras, seqüências reguladoras vantajosas estão, por exemplo, presentes nos promotores Gram-positivos amy e SPO2, nos promotores de levedura ou fúngicos ADC1, MFa, CA, P-60, CYC1, GAPDH, TEF, rp28, ADH ou nos promotores de planta CaMV/35S [Franck e outros, Cell 21 (1980) 285-294], PRP1 [Ward e outros, Plant. Mol. Biol. 22 (1993)], SSU, OCS, iib4, usp, STLS1, B33, nos ou no promotor de ubiquitina ou faseolina. Vantajoso neste contexto também são os promotores induzíveis, tal como os promotores descritos nos EP-A-O 388 186 (induzível por benzenossulfonamida), Plant J. 2, 1992:397-404 (Gatz e outros, induzível por tetraciclina), EP-A-0 335 528 (induzível por ácido abscísico) ou promotores de WO 93/21334 (induzível por etanol ou cicloexenol). Outros promotores de planta adequados são o promotor de FBPase citosólico ou o promotor ST-LSI de batata (Stockhaus e outros, EMBO J. 8, 1989, 2445), o promotor de glicina max fosforibosilpirofosfato amidotransferase (Nº. de Acesso de Genbank U87999) ou o promotor nodo-específico descrito na Patente EP-A-0 249 676.

Os promotores especialmente vantajosos são promotores que tornam possível a expressão em tecidos que estão envolvidos na biossíntese de ácidos graxos. Muito especialmente vantajosos são os promotores específicos de semente, tal como o promotor de USP como descrito, porém também outros promotores tal como LeB4, DC3, promotor de faseolina ou napin. Outros promotores especialmente vantajosos são promotores específicos de semente que podem ser usados para plantas monocotilédneas ou dicotilédneas e que são descritos nos US 5.608.152 (promotor de napin de óleo de semente de colza), WO 98/45461 (promotor de Arabidopsis oleosin), US 5.504.200 (promotor de Phaseolus vulgaris phaseolin), WO 91/13980 (promotor de Brassica Bce4), por Baeumlein e outros, Plant J., 2, 2, 1992:233-239 (promotor LeB4 de um legume), estes promotores que são adequados para dicotilédneas. Os exemplos de promotores que são adequados para monocotilédneas são o lpt-2 de cevada ou promotor de lpt-1 (WO 95/15389 e WO 95/23230), o promotor de proteína da cevada e outros promotores adequados descritos nos WO 99/16890.

Em princípio, é possível usar todos os promotores naturais junto com as suas seqüências reguladoras, tal como aquelas mencionadas acima. Também é possível e vantajoso usar os promotores sintéticos, ou em adição ou sozinho, em particular quando eles medeiam expressão específica de semente, tal como aquelas descritas no WO 99/16890.

Para alcançar um conteúdo de ARA particularmente elevado, especialmente em plantas transgênicas os genes deveriam ser expressados vantajosamente em colheitas de óleo de uma maneira específica de semente. Para este fim, podem ser usados os promotores específicos de semente, ou aqueles promotores que são ativos no embrião e/ou no endosperma. Em princípio, os promotores específicos semente podem ser isolados tanto de plantas dicotiledôneas quanto monocotiledôneas. Os promotores preferidos são listados a seguir: USP (= proteína de semente desconhecida) e vicilina (*Vicia faba*) [Baumlein e outros, Mol. Gen Genet, 1991, 225(3)], napin (óleo de semente de colza) [US 5.608.152], proteína portadora de acila (óleo de semente de colza) [US 5.315.001 e WO 92/18634], oleosina (*Arabidopsis thaliana*) [WO 98/45461 e WO 93/20216], faseolina (*Phaseolus vulgaris*) [US 5.504.200], Bce4 [WO 91/13980], leguminas B4 (promotor de LegB4) [Baumlein e outros, Plant J., 2,2, 1992], Lpt2 e Lpt1 (cevada) [WO 95/15389 e WO95/23230], promotores específicos de semente de arroz, milho e trigo [WO 99/16890], Amy32b, Amy 6-6 e aleurína [US 5.677.474], Bce4 (óleo de semente de colza) [US 5.530.149], glicinina (soja) [EP 571 741], fosfoenol piruvato carboxilase (soja) [JP 06/62870], ADR12-2 (soja) [WO 98/08962], isocitrato liase (óleo de semente de colza) [US 5.689.040] ou α -amilase (cevada) [EP 781 849].

A expressão de gene planta também pode ser facilitada por um promotor quimicamente induzível (veja , revisão em Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108). Os promotores quimicamente induzíveis são particularmente adequados quando é desejado que expressão de gene ocorra de uma maneira específica de tempo. Os exemplos de tais promotores são um promotor induzível por ácido salicílico (WO 95/19443), um promotor induzível por tetraciclina (Gatz e outros (1992) Plant J. 2, 397-

404) e um promotor induzível por etanol.

Para assegurar a integração estável dos genes de biossíntese na planta transgênica em uma pluralidade de gerações, é normalmente necessário para cada dos ácidos nucléico que codifique uma proteína de interesse a ser expressada sob o controle de um promotor separado, preferivelmente um promotor que difere dos outros promotores, uma vez que os motivos de seqüência de repetição podem levar a instabilidade do T-DNA, ou aos eventos de recombinação. Esta é uma razão por que o ácido nucléico da presente invenção é particularmente vantajoso uma vez que seqüências que codificam $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase podem ser transcritas como uma única unidade que precisa de somente um promotor. Certamente, será necessário para outros genes codificar, por exemplo, $\Delta 5$ -elongase, $\omega 3$ -dessaturase e/ou de $\Delta 4$ -dessaturase para estar sob o controle de promotores separados.

Neste contexto, o cassete de expressão é construído vantajosamente de um tal modo que um promotor seja seguido por um sítio de clivagem, vantajosamente em um poliligante, para inserção do ácido nucléico a ser expressado e, se apropriado, uma seqüência terminadora é posicionada atrás do poliligante. Esta seqüência é várias vezes repetida, preferivelmente três, quatro ou cinco vezes, de forma que até cinco genes possam ser combinados em um construto e introduzidos na planta transgênica a ser expressada. Vantajosamente, a seqüência é repetida até três vezes. Para expressar as seqüências de ácido nucléico, a última é inserida atrás do promotor por um sítio de clivagem adequado, por exemplo, no poliligante. Vantajosamente, cada seqüência de ácido nucléico tem seu próprio promotor e, se apropriado, sua própria seqüência terminadora. Tais construtos vantajosos são descritos, por exemplo, em DE 101 02 337 ou DE 101 02 338. Porém, também é possível inserir uma pluralidade de seqüências de ácido nucléico atrás de um promotor e, se apropriado, antes de uma seqüência terminadora. Aqui, o sítio de inserção, ou a seqüência, dos ácidos nucléicos inseridos no cassete de expressão não é de importância crítica, quer dizer que uma seqüência de ácido nucléico pode ser inserida na primeira ou última posição

no cassete sem sua expressão sendo influenciada substancialmente desse modo. Vantajosamente, os promotores diferentes tal como, por exemplo, o promotor USP, LegB4 ou DC3, e seqüências terminadoras diferentes podem ser usadas no cassete de expressão. Porém, também é possível usar somente um tipo de promotor no cassete. Porém, isto pode levar aos eventos de recombinação indesejados.

Como descrito acima, a transcrição dos genes que foram introduzidos, deveria ser vantajosamente terminada por seqüências terminadoras adequadas na extremidade 3' dos genes de biossíntese que foram introduzidos (atrás do codon de interrupção). Um exemplo de uma seqüência que pode ser usada neste contexto é a seqüência terminadora OCS 1. Como é o caso com os promotores, seqüências terminadoras diferentes deveriam ser usadas para cada gene.

O construto de gene da presente invenção também pode compreender genes de biossíntese do ácido graxo ou metabolismo de lipídio selecionados do grupo de desidrogenase(s) de acil-CoA, acil-ACP [= proteína portadora de acila] dessaturase(s), tioesterase(s), acil-ACP aciltransferase(s) de ácido graxo, acil-CoA:lisofosfolípídeo aciltransferase(s), ácido graxo sintases, ácido graxo hidroxilases, acetil-coenzima A carboxilase(s), acil-coenzima A oxidase(s), ácido graxo dessaturase(s), ácido graxo acetilenases, lipoxigenases, triacilglycerol lipases, alenóxido sintases, hidroperóxido liases ou ácido graxo elongase(s) e dessaturase(s) tal como $\Delta 4$ -dessaturase, $\Delta 5$ -dessaturase, $\Delta 6$ – dessaturase, $\Delta 8$ -dessaturase, $\Delta 9$ -dessaturase, $\Delta 12$ -dessaturase ou $\Delta 6$ -elongase.

Estes genes ou ácidos nucléicos adicionais podem ser clonados nos cassetes de expressão, que são então usados para transformar plantas com a ajuda de vetores como *Agrobacterium*.

Aqui, os fatores ou seqüências reguladoras podem, como descrito acima, preferivelmente ter um efeito positivo, e desse modo realçar, nos genes de expressão que foram introduzidos. Desse modo, o realce dos elementos reguladores pode vantajosamente ocorrer no nível transcricional usando sinais de transcrição fortes sinaliza como intensificadores e/ou promo-

tores. Porém, uma translação intensificada também é possível, por exemplo, melhorando a estabilidade do mRNA. Em princípio, os cassetes de expressão podem ser usados diretamente para introdução nas plantas ou então podem ser introduzidos em um vetor.

5 Então, em ainda outro aspecto da invenção, é fornecido um vetor que compreende um ácido nucléico ou um construto de gene em quaisquer dos aspectos da invenção descritos acima.

Em uma modalidade, o vetor pode ser um vetor de clonagem.

As seqüências de ácido nucléico da invenção podem ser introduzidas sozinhas, ou preferivelmente, em combinação com um cassete de expressão (construto de ácido nucléico) em um organismo. Para introduzir os ácidos nucléicos, os ácidos nucléicos são vantajosamente ampliados e ligados de maneira conhecida. Preferivelmente, um procedimento que segue o protocolo para Pfu DNA polimerase ou uma mistura de Pfu/Taq DNA polimerase é seguida. Os iniciadores são selecionados levando em consideração a seqüência a ser ampliada. Os iniciadores deveriam ser escolhidos vantajosamente de tal um modo que o amplificado compreenda a seqüência codogênica total desde o códon de partida ao códon de interrupção. Após a amplificação, o amplificado é convenientemente analisado. Por exemplo, uma separação de gelelectroforético pode ser realizada, a qual é seguida por uma análise quantitativa e qualitativa. Após isso, o amplificado pode ser purificado seguindo um protocolo padrão (por exemplo, Qiagen). Uma alíquota do amplificado purificado é então disponível para a etapa de clonagem subsequente.

10
15
20

25 Os vetores de clonagem adequados são geralmente conhecidos pelo trabalhador versado. Estes incluem, em particular, os vetores que são capazes de replicação em sistemas microbianos, quer dizer principalmente vetores que garantem a clonagem eficiente em leveduras ou fungos e que tornam possível a transformação estável de plantas. Aqueles que devem ser mencionados em particular são vários sistemas de vetor binário e co-integrado que são adequados para a transformação mediada por T-DNA. Tais sistemas de vetor são, como uma regra, caracterizados pelo fato de que

30

eles compreendem os genes vir requeridos para a transformação mediada por *Agrobacterium* e as seqüências de delimitação de T-DNA (borda de T-DNA). Estes sistemas de vetor vantajosamente também compreendem outras regiões cis-reguladoras tal como os promotores e seqüências terminadoras e/ou marcadores de seleção, por meio dos quais os organismos adequadamente transformados podem ser identificados. Ao mesmo tempo em que no caso de sistemas de vetor co-integrado, os genes vir e seqüências de T-DNA estão dispostos no mesmo vetor, os sistemas binários são com base em pelo menos dois vetores, um dos quais transporta os genes vir, porém nenhum T-DNA, ao mesmo tempo em que um segundo transporta o T-DNA, porém nenhum gene vir. Devido a este fato, os últimos vetores mencionados são relativamente pequenos, fáceis de manipular e replicar tanto em *E. coli* quanto em *Agrobacterium*. Estes vetores binários incluem vetores das séries pBIB-HYG, pPZP, pBecks, pGreen. De acordo com a invenção, Bin19, pBI101, pBinAR, pGPTV e pCAMBIA são usados por preferência. Uma avaliação dos vetores binários e de seu uso é encontrada em Hellens e outros, Trends in Plant Science (2000) 5, 446 - 451. Para preparar os vetores, os vetores podem ser primeiro linearizados com endonuclease(s) de restrição e então modificados enzimaticamente de uma maneira adequada. Depois disso, o vetor é purificado, e uma alíquota é empregada para a etapa de clonagem. Na etapa de clonagem, o amplificado enzimaticamente clivado e, se apropriado, purificado é clonado com fragmentos de vetor foram preparados de uma maneira semelhante, usando ligase. Neste contexto, um construto de ácido nucléico particular, ou construto de vetor ou plasmídeo, pode ter um ou então mais de um segmento de gene codogênico. Os segmentos de gene codogênico nestas construtos são preferivelmente ligados operavelmente com seqüências reguladoras. As seqüências reguladoras incluem, em particular, seqüências de planta como os promotores e seqüências terminadoras descritas acima. O construto pode ser vantajosamente estavelmente propagado em microorganismos, em particular em *E. coli* e *Agrobacterium tumefaciens*, sob condições seletivas e torna possível a transferência de DNA heterólogo em plantas ou microorganismos.

Os ácidos nucléicos da invenção podem ser introduzidos em organismos, tais como microorganismos ou vantajosamente plantas, vantajosamente usando vetores de clonagem, e desse modo ser usado na transformação de plantas tal como aquelas que são publicadas e citadas em:

5 Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Capítulo 6/7, p. 71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Vol. 1, Engineering and Utilization, Ed.: Kung e R. Wu, Academic Press, 1993, 15-38; B. Jenes e outros, Techniques for Gene Transfer, em: Transgenic Plants, Vol. 1, Engineering and Utilization, Ed.: Kung e R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225. Desse modo, os ácidos nucléicos, os ácidos nucléicos inventivos e construtos de ácido nucléico e/ou vetores usados no processo podem ser de usados para a modificação de recombinante de um espectro amplo de organismos, vantajosamente plantas, de forma que o último se torne o produtor melhor e/ou mais eficiente de ARA.

Uma série de mecanismos existe pelo qual a modificação das proteínas de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase é possível, de forma que o rendimento, produção e/ou eficiência de produção de ARA em

20 uma planta, preferivelmente em uma planta de colheita de óleo ou um microorganismo, possa ser influenciado diretamente devido a estas proteínas modificadas. O número ou atividade das proteínas ou genes pode ser aumentado, de forma que maiores quantidades dos produtos de gene e, no final das contas, maiores quantidades dos compostos da fórmula geral I, sejam produzidas. Uma síntese *de novo* em um organismo que precisou da atividade e capacidade de biossintetizar os compostos antes de introdução do gene(s) correspondente também é possível. Isto se aplica analogamente à combinação com outras dessaturases ou elongases ou outras enzimas do metabolismo de lipídio e ácido graxo. O uso de várias seqüências divergentes, isto é, seqüências que diferem no nível de seqüência de DNA, também

30 pode ser vantajoso neste contexto, ou então o uso de promotores para expressão de gene que torna possível uma expressão de gene diferente com o

passar do tempo, por exemplo, como uma função do grau de maturidade de uma semente ou um tecido de armazenamento de óleo.

Devido à introdução de um gene que codifica $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase em um organismo, sozinho ou em combinação com outros genes em uma célula, não é somente possível aumentar o fluxo de biossíntese para o produto final, porém também aumentar, ou criar a composição de triacilglicerol correspondente de novo. Igualmente, o número ou atividade de outros genes que estão envolvidos na importação de nutrientes que são requeridos para a biossíntese de um ou mais ácidos graxos, óleos, lipídios neutros e/ou polares pode ser aumentado, de forma que a concentração destes precursores, co-fatores ou intermediários dentro das células ou dentro do compartimento de armazenamento seja aumentada, desse modo, a capacidade das células de produzir ARA como descrito abaixo é intensificada também. Otimizando-se a atividade ou aumentando-se o número de um ou mais genes que codifica $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase que estão envolvidos na ARA de biossíntese, ou destruindo-se a atividade de um ou mais genes que estão envolvidos na degradação de ARA, um rendimento, produção e/ou eficiência aumentados de produção de ácido graxo e moléculas de lipídio em organismos, vantajosamente em plantas, é tornado possível.

Os ácidos nucléicos, que vantajosamente podem ser usados no processo são derivados de bactérias, fungos, diatomáceas, animais tal como, *Caenorhabditis* ou *Oncorhynchus* ou plantas tal como algas ou musgos, tal como os gêneros *Shewanella*, *Physcomitrella*, *Thraustochytrium*, *Fusarium*, *Phytophthora*, *Ceratodon*, *Mantoniella*, *Ostreococcus*, *Isochrysis*, *Aleurita*, *Muscarioides*, *Mortierella*, *Borago*, *Phaeodactylum*, *Cryptocodinium*, especificamente dos gêneros e espécies *Oncorhynchus mykiss*, *Xenopus laevis*, *Ciona intestinalis*, *Thalassiosira pseudonana*, *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 officinalis*, *Phaeodactylum*

tricornutum, *Caenorhabditis elegans* ou especialmente vantajosamente de *Oncorhynchus mykiss*, *Euglena gracilis*, *Thalassiosira pseudonana* ou *Cryptocodinium cohnii*.

5 Em uma modalidade alternativa, o vetor pode ser um vetor de expressão projetado para transformar um organismo no qual o ácido nucléico deve ser expressado e o ácido linoléico convertido em ARA.

Estes vetores vantajosos, preferivelmente vetores de expressão, compreendem o(s) ácido(s) nucléico(s) que codifica(m) o $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase e que são descritos nos primeiro a quarto
10 aspectos da invenção.

Como usado no presente contexto, o termo "vetor" refere-se a uma molécula de ácido nucléico que é capaz de transportar outro ácido nucléico ao qual está ligado. Um tipo de vetor é um "plasmídeo", uma alça de DNA de filamento duplo circular na qual os segmentos de DNA adicionais
15 podem ser ligados. Um tipo adicional de vetor é um vetor viral, sendo possível para os segmentos de DNA adicionais serem ligados no genoma viral. Certos vetores são capazes de replicação autônoma em uma célula hospedeira na qual eles foram introduzidos (por exemplo, vetores bacterianos com origem de replicação bacteriana). Outros vetores são vantajosamente integrados no genoma de uma célula hospedeira quando eles são introduzidos
20 na célula hospedeira, e desse modo replicam junto com o genoma hospedeiro. Além disso, certos vetores podem administrar a expressão de genes com a qual eles estão em ligação operável. Estes vetores são referidos no presente contexto como "vetores de expressão". Normalmente, os vetores de
25 expressão são adequados para técnicas de recombinação de DNA que tomam a forma de plasmídeos.

Na presente descrição onde o termo "plasmídeo" é usado, deveria ser entendido que os plasmídeos podem ser substituídos por outros tipos de vetor de expressão, tal como vetores virais, que exercem funções semelhantes.
30 Além disso, o termo "vetor" também é pretendido compreender outros vetores com os quais o trabalhador versado está familiarizado, tal como fagos, vírus tais como SV40, CMV, TMV, transposons, elementos IS, fasmí-

deos, fagomídeos, cosmídeos, DNA linear ou circular.

Os vetores de expressão recombinantes vantajosamente usados no processo compreendem os ácidos nucléicos descritos abaixo ou o construto de gene descrito acima em uma forma que é adequada para expressar os ácidos nucléicos usados em uma célula hospedeira, o que significa que os vetores de expressão recombinantes compreendem uma ou mais seqüências reguladoras, selecionadas em base das células hospedeiras usadas para a expressão, cuja seqüência(s) reguladora é ligada operavelmente com a seqüência de ácido nucléico a ser expressada. Em um vetor de expressão recombinante, "operavelmente ligado" significa que a seqüência de nucleotídeo de interesse é ligada a seqüência(s) reguladora(s) de um tal modo que a expressão da seqüência de nucleotídeo seja possível, e elas são ligadas uma à outra de um tal modo que ambas as seqüências realizam a função predita que é designada para seqüência (por exemplo, em um sistema de transcrição/translação *in vitro*, ou em uma célula hospedeira se o vetor é introduzido na célula hospedeira). O termo "seqüência reguladora" é pretendido compreender os promotores, intensificadores e outros elementos de controle de expressão (por exemplo, sinais de poliadenilação). Por exemplo, estas seqüências reguladoras são descritas em Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), ou veja: Gruber e Crosby, em: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, Ed.: Glick and Thompson, Capítulo 7, 89-108, incluindo as referências citadas aqui. As seqüências reguladoras compreendem aquelas que administram a expressão constitutiva de uma seqüência de nucleotídeo em muitos tipos de célula hospedeira e aquelas que administram a expressão direta da seqüência de nucleotídeo somente em células hospedeiras específicas sob condições específicas. O trabalhador versado sabe que o desígnio do vetor de expressão pode depender de fatores tal como a escolha da célula hospedeira ser transformada, o nível de expressão desejado da proteína e similares.

Os vetores de expressão recombinantes usados podem ser projetados para a expressão de $\Delta 9$ - elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -

dessaturase em células procarióticas ou eucarióticas. Isto é vantajoso uma vez que as etapas intermediárias da construção de vetor frequentemente são realizadas em microorganismos devido à simplicidade. Por exemplo, o gene de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase pode ser expressado

5 em células bacterianas, células de inseto (usando vetores de expressão de Baculovírus), levedura e outras células fúngicas (veja Romanos, M., A., e outros (1992) "Foreign gene expression in yeast: a review", *Yeast* 8:423-488; van den Hondel, C.A.M. J. J., e outros (1991) "Heterologous gene expression in filamentous fungi", em: *More Gene Manipulations in Fungi*, J. W. Bennet &

10 L. L. Lasure, Ed., pp. 396-428: Academic Press: San Diego; e van den Hondel, C. A.M. J. J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, em: *Applied Molecular Genetics of Fungi*, Peberdy, J. F., e outros, Ed., pp. 1-28, Cambridge University Press: Cambridge), algas (Falcatore e outros, 1999, *Marine Biotechnology* 1, 3:239-

15 251), os ciliados dos tipos: *Holotrichia*, *Peritrichia*, *Spirotrichia*, *Suctorina*, *Tetrahymena*, *Paramecium*, *Colpidium*, *Glaucoma*, *Platyophrya*, *Potomacus*, *Dessaturaseudocohnilembus*, *Euplotes*, *Engelmanniella* e *Stylonychia*, em particular do gênero *Stylonychia lemnae*, usando vetores em um método de transformação como descrito no WO 98/01572 e, preferivelmente, em células

20 de plantas multicelulares (veja Schmidt, R. e 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, Capítulo 6/7, pp.71-119 (1993); F.F. White, B. Jené e outros, *Techniques for Gene Transfer*, em: *Transgenic Plants*, Vol. 1, Engineering and Utilization, Ed.: Kung e R. Wu, Academic Press (1993), 128-43; Potrykus, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991), 205-225 (e referências citadas aqui)).

25 As células hospedeiras adequadas são também descritas em Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Como uma alternativa, o vetor de expressão recombinante pode ser transcrito e traduzido *in vitro*, por exemplo, usando as sequências reguladoras T7-promotoras e T7-polimerase.

30

Na maioria dos casos, a expressão de proteínas em procariotas envolve o uso de vetores que compreendem promotores constitutivos ou induzíveis que administram a expressão de proteínas de fusão ou não fusão. Os vetores de expressão de fusão típicos são, *inter alia*, pGEX (Pharmacia Biotech Inc; Smith, D.B., e Johnson, K.S. (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRIT5 (Pharmacia, Piscataway, NJ), onde glutathione S-transferase (GST), proteína de ligação de maltose-E e proteína A, respectivamente, são fundidas com a proteína alvo recombinante.

Os exemplos de vetores de expressão de *E. coli* de não fusão induzíveis adequados são, *inter alia*, pTrc (Amann e outros (1988) gene 69:301-315) e pET 11 d (Studier e outros, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Califórnia (1990) 60-89). A expressão de gene alvo do vetor pTrc é com baseado na transcrição de um promotor de fusão trp-lac híbrido pelo RNA polimerase hospedeiro. A expressão de gene alvo do vetor pET 11 d é com base na transcrição de um promotor de fusão T7-gn10-lac que é mediado por um RNA polimerase viral (T7gn1) que é co-expressado. Esta polimerase viral é fornecida pelas cepas hospedeiras BL21 (DE3) ou HMS174 (DE3) de um λ -profago residente que aloja um gene T7 gn1 sob controle transcricional do promotor lacUV 5.

Outros vetores que são adequados para organismos procarióticos são conhecidos pelo trabalhador versado, estes vetores estão, por exemplo, em pLG338 de *E. coli*, pACYC184, as séries pBR, tal como pBR322, as séries pUC tal como pUC18 ou pUC19, as séries M113mp, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1, λ gt11 ou pBdCl, em *Streptomyces* plJ101, plJ364, plJ702 ou plJ361, em *Bacilo* pUB1 10, pC194 ou pBD214, em *Corynebacterium* pSA77 ou pAJ667.

Em uma outra modalidade, o vetor de expressão é um vetor de expressão de levedura. Os exemplos para vetores para expressão na levedura *S.cerevisiae* compreendem pYeDessaturated (Baldari e outros (1987) Embo J. 6:229-234), pMFa (Kurjan e Herskowitz (1982) Cell 30:933-943), pJRY88 (Schultz e outros (1987) Gene 54:113-123) e pYES2 (Invitrogen

Corporation, San Diego, CA). Os vetores e processos para a construto de vetores que são adequados para uso em outros fungos, tal como os fungos filamentosos, compreendem aqueles que são descritos nos detalhes em: 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 e outros, Ed., pp. 1-28, Cambridge University Press: Cambridge, ou em: More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, Ed., pp. 396-428: Academic Press: San Diego]. Outros vetores de levedura adequados são, por exemplo, pAG-1, YEpl6, YEpl 3 ou pEMBLYe23.

10 Como uma alternativa, $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase podem ser expressados em células de inseto que usam vetores de *Baculovirus*. Os vetores de expressão de *Baculovirus* que estão disponíveis para a expressão de proteínas em células de inseto cultivadas (por exemplo, células de Sf9) compreendem as séries pAc (Smith e outros (1983) 15 Mol. Cell Biol. 3:2156-2165) e as séries pVL (Lucklow e Summers (1989) Virology 170:31-39).

Os vetores de acima mencionados são somente uma pequena avaliação sobre os vetores adequados que são possíveis. Outros plasmídeos são conhecidos por trabalhadores versados e são descritos, por exemplo, em: Cloning Vectors (Ed. Pouwels, P. H., e outros, Elsevier, Amsterdam- 20 Nova Iorque-Oxford, 1985, ISBN 0 444 904018). Para outros sistemas de expressão adequados para células procarióticas e eucarióticas, veja os Capítulos 16 e 17 em Sambrook, J., Fritsch, E. F., e Maniatis, T., Molecular Cloning: A Laboratory Manual, 2ª edição, Cold Spring Harbor Laboratory, 25 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Em uma modalidade adicional do processo, as $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase podem ser expressadas em células de planta unicelulares (tal como, algas), veja Falciatore e outros, 1999, Marine Biotechnology 1 (3):239-251 e referências citadas aqui, e em células de 30 planta de plantas mais altas (por exemplo, espermatófitos, tal como colheitas cultiváveis). Os exemplos de vetores de expressão de planta compreendem aqueles que são descritos nos detalhes: Becker, D., Kemper, E., Schell, J., e

- Masterson, R., (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20:1195-1197; e Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acids Res.* 12:8711-8721 ; Vectors for Gene Transfer in Higher Plants; em: Transgenic Plants, Vol. 1 , Engineering and Utilization, Ed.: Kung e R. Wu, Academic Press, 1993, pp. 15-38.

Um cassete de expressão de planta preferivelmente compreende seqüências reguladoras que são capazes de administrar a expressão de genes em células de planta e são ligadas operavelmente de forma que cada seqüência possa cumprir sua função, tal como terminação transcricional, por exemplo, sinais de poliadenilação. Os sinais de poliadenilação preferidos são aqueles que são derivados de T-DNA de *Agrobacterium tumefaciens*, tal como gene 3 do plasmídeo Ti pTiACH5 (Gielen e outros, *EMBO J.* 3 (1984) 835 et. seq), que é conhecido como octopina sintase, ou equivalentes funcionais destes, porém todas as outras seqüências terminadoras que são funcionalmente ativas em plantas também são adequadas.

Uma vez que a expressão de gene de planta é muito freqüentemente não limitada ao nível transcricional, um cassete de expressão de planta preferivelmente compreende outras seqüências que são operavelmente ligadas, tal como intensificadores de translação, por exemplo, a seqüência "overdrive", que realça a seqüência líder 5'-não translada de vírus em mosaico do tabaco, que aumenta a relação de proteína/RNA (Gallie e outros, 1987, *Nucl. Acids Research* 15:8693-8711).

Como descrito acima, a expressão de gene de planta deve estar ligada operavelmente com um promotor adequado que ativa a expressão de gene com a temporização correta ou de uma maneira específica de célula ou tecido. Os promotores utilizáveis são os promotores constitutivos (Benfey e outros, *EMBO J.* 8 (1989) 2195-2202), tal como aqueles esses que são derivados de vírus de planta, tal como 35S CaMV (Franck e outros, *Cell* 21 (1980) 285-294), 19S CaMV (também veja EUA 5352605 e WO 84/02913), ou promotores de planta, tal como o promotor do subunidade de Rubisco, que é descrito na Patente US 4.962.028.

Outras seqüências preferidas para uso em ligação operável em cassetes de expressão de gene de planta são seqüências de alvejamento, que são requeridas para guiar o produto de gene em seu compartimento de célula correspondente (veja uma revisão em Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423 e referências citadas aqui), por exemplo, no vacúolo, no núcleo, todos os tipos de plastídeos, tais como amiloplastos, cloroplastos, cromoplastos, o espaço extracelular, as mitocôndrias, o retículo de endoplasmídeo, elaioplastos, peroxissomas e outros compartimentos de células de planta.

Como descrito acima, a expressão de gene de planta também pode ser obtida por um promotor quimicamente induzível (veja revisão em Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108). Os promotores quimicamente induzíveis são particularmente adequados quando é desejado que a expressão de gene ocorra de uma maneira específica de tempo. Os exemplos de tais promotores são um promotor induzível por ácido salicílico (WO 95/19443), um promotor induzível por tetraciclina (Gatz e outros (1992) Plant J. 2, 397-404) e um promotor induzível por etanol.

Os promotores que respondem às condições de tensão bióticas ou abióticas também são adequados, por exemplo, o promotor de gene PRP1 induzido por patógeno (Ward e outros, Plant. Mol. Biol. 22 (1993) 361-366), o promotor de tomate hsp80 de tomate induzível por calor (US 5.187.267), o promotor de alfa-amilase de batata induzível por (WO 96/12814) ou o promotor pinII induzível por ferida (EP-A-0 375 091).

Especialmente preferidos são aqueles promotores que provocam a expressão de gene em tecidos e órgãos nos quais a biossíntese de ácidos graxos, lipídios e óleos, ocorre, em células de semente, tal como células do endosperma e do embrião em desenvolvimento. Os promotores adequados são o promotor de napin de óleo de semente de colza (US 5.608.152), o promotor de Vicia faba USP (Baeumlein e outros, Mol Gen Genet, 1991, 225 (3):459-67), o promotor de oleosina Arabidopsis (WO 98/45461), o promotor de faseolina *Phaseolus vulgaris* (US 5.504.200), o promotor de Bce4 *Brassica* (WO 91/13980) ou o promotor de legumina B4 (LeB4; Baeumlein e ou-

tros, 1992, Plant Journal, 2 (2):233-9), e promotores que provocam a expressão específica semente em plantas monocotiledôneas, tal como milho, cevada, trigo, centeio, arroz e similares. Promotores notáveis adequados são o promotor de gene de cevada lpt2 ou lpt1 (WO 95/15389 e WO 95/23230) ou os promotores do gene de proteína de cevada, o gene de glutelina de arroz, o gene de orizina de arroz, o gene de prolamina de arroz, o gene de gliadina de trigo, o gene de glutelina de trigo, o gene de zeína de milho, o gene de glutelina de aveia, o gene de casirina de Sorghum ou o gene de secalina de centeio que são descritos no WO 99/16890.

10 Como descrito acima, pode ser vantajoso incluir em um cassete de expressão, enzimas de codificação de ácidos nucleicos capazes de converter ARA em ácidos graxos $\Delta 3$ -não saturados tal como EPA ou DHA. Desse modo, por exemplo, o cassete de expressão pode também incluir o ácido nucleico que codifica uma $\Delta 5$ -elongase, $\omega 3$ -dessaturase e/ou $\Delta 4$ -dessaturase. Tais cassetes de expressão podem ser introduzidos pela transformação simultânea de uma pluralidade de construtos de expressão individuais ou, preferivelmente, combinando-se uma pluralidade de cassetes de expressão em um construto. Além disso, uma pluralidade de vetores pode ser transformada com, em cada caso, uma pluralidade de cassetes de expressão e então transferida na célula hospedeira.

25 Outros promotores que são igualmente especialmente adequados são aqueles que realizam uma expressão específica de plastídeo, uma vez que os plastídeos constituem o compartimento no qual os precursores e alguns produtos finais são sintetizados de biossíntese de lipídio. Os promotores adequados, tal como o promotor de RNA polimerase viral, são descritos nos WO 95/16783 e WO 97/06250, e o promotor clpP de *Arabidopsis*, descrito no WO 99/46394.

30 O DNA de vetor pode ser introduzido em células procarióticas e eucarióticas por técnicas de transformação ou transfecção convencionais. Os termos "transformação" e "transfecção", conjugação e transdução, como usado no presente contexto, são pretendidos compreender uma multiplicidade de métodos conhecidos na técnica anterior pela introdução de ácido nu-

cléico estrangeiro (por exemplo, o DNA) em uma célula hospedeira, incluindo a co-precipitação de fosfato de cálcio ou cloreto de cálcio, transfecção mediada por DEAE-dextrana, lipofecção, competência natural, transferência quimicamente mediada, eletroporação ou bombardeio de partícula. Os métodos adequados para a transformação ou transfecção de células hospedeiras, incluindo células de planta, podem ser encontrados em Sambrook e outros (Molecular Cloning: A Laboratory Manual., 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) e outros livros de laboratório tal como Methods in Molecular Biology, 1995, Vol. 44, Agrobacterium protocols, Ed.: Gartland and Davey, Humana Press, Totowa, Nova Jérsei.

Em um outro aspecto da invenção é fornecido um organismo não humano transgênico que compreende pelo menos um ácido nucléico, construído de gene ou vetor de acordo com um aspecto anterior da invenção.

O organismo não humano transgênico pode ser um microorganismo, um animal não humano ou uma planta.

As células hospedeiras que são, em princípio, adequadas para absorver o ácido nucléico de acordo com a invenção, o produto de gene de acordo com a invenção ou o vetor de acordo com a invenção são todos os organismos procarióticos ou eucarióticos. Os organismos hospedeiros que são vantajosamente usados são os microorganismos tais como fungos ou leveduras, ou células de planta, preferivelmente plantas ou partes desta. Os fungos, leveduras ou plantas são preferivelmente usados, especialmente plantas, por exemplo, plantas tal como colheitas de óleo, que têm alto teor em compostos de lipídio tal como óleo de semente de colza, primula de noite, linho, cardo, amendoim, óleo de colza, linhaça, soja, açafoa, girassol, borragem, ou plantas tal como milho, trigo, centeio, aveias, tritcale, arroz, cevada, algodão, mandioca, pimenta, Tagetes, plantas Solanacea tal como batata, tabaco, berinjela e tomate, espécies de Vicia, ervilha, alfafa, plantas fechadas (café, cacau, chá), espécies de Salix, árvores (óleo de palma, coco), e gramas perenes e colheitas de forragem. As plantas especialmente preferidas de acordo com a invenção são colheitas de óleo tal como soja,

amendoim, óleo de semente de colza, óleo de colza, linhaça, linho, prímula de noite, girassol, açafroa, árvores (óleo de palma, coco).

Em uma modalidade vantajosa, o termo "ácido nucléico (molécula)" como usado no presente contexto adicionalmente compreende a sequência não transladada nas extremidades 3' e 5' da região de gene de codificação: pelo menos 500, preferivelmente 200, especialmente preferivelmente 100 nucleotídeos da sequência a jusante da extremidade 5' da região de codificação e pelo menos 100, preferivelmente 50, especialmente preferivelmente 20 nucleotídeos da sequência a jusante da extremidade 3' da região de gene de codificação. Uma molécula de ácido nucléico "isolada" é separada de outras moléculas de ácido nucléico que estão presentes na fonte natural do ácido nucléico. Um ácido nucléico "isolado" preferivelmente não tem nenhuma sequência que naturalmente flanqueie o ácido nucléico no DNA genômico do organismo do qual o ácido nucléico é derivado (por exemplo, seqüências que ficam situadas nas extremidades 5' e 3' do ácido nucléico). Em várias modalidades, a molécula de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase ou $\Delta 5$ -dessaturase isolada pode compreender, por exemplo, menos do que aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ou 0,1 kb de seqüências de nucleotídeo que naturalmente flanqueiam a molécula de ácido nucléico no DNA genômico da célula da qual o ácido nucléico é derivado. O mesmo se aplica a outras seqüências de ácido nucléico que podem ser incluídas em um cassete de expressão, por exemplo seqüências que codificam uma $\Delta 5$ -elongase, $\omega 3$ -dessaturase e/ou $\Delta 4$ -dessaturase

As moléculas de ácido nucléico da presente invenção, por exemplo, uma molécula de ácido nucléico com uma sequência de nucleotídeo de SEQ ID NO: 1 ou resíduos 7668 a 12077 desta, ou as partes da SEQ ID NO: 1 especificada pelos segundo a quarto aspectos da invenção, podem ser isoladas usando técnicas padrão molecular-biológicas e as informações de sequência fornecidas nelas. Além disso, por exemplo, uma sequência homóloga ou regiões de sequência homóloga, conservadas podem ser identificadas no nível de aminoácido ou DNA com a ajuda de algoritmos comparativos. Eles podem ser usados como sonda de hibridização e técnicas de

hibridização padrão (tal como, por exemplo, aquelas descritas em Sambrook e outros, *Molecular Cloning: A Laboratory Manual*. 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) para isolar seqüências de ácido nucléico adicionais que podem

5 ser usadas no processo.

Além disso, uma molécula de ácido nucléico de *Perkinsus marinus* que compreende uma seqüência completa de SEQ ID NO: 1 ou uma parte desta pode ser isolada através de reação em cadeia de polimerase, onde os iniciadores de oligonucleotídeo são usados sobre a base desta seqüência ou partes desta (por exemplo, uma molécula de ácido nucléico que
10 compreende a seqüência completa ou parte desta pode ser isolada através de reação em cadeia de polimerase que usa iniciadores de oligonucleotídeo que foram gerados com base nesta mesma seqüência). Por exemplo, mRNA pode ser isolado de células (por exemplo, por meio do método de extração
15 de tiocianato de guanidíneo de Chirgwin e outros (1979) *Biochemistry* 18:5294-5299) e cDNA por meio de transcriptase reversa (por exemplo, transcriptase reversa Moloney MLV, disponibilizado por Gibco/BRL, Bethesda, MD, ou transcriptase reversa AMV, disponibilizado por Seikagaku America, Inc., St.Petersburg, FL).

20 Os iniciadores de oligonucleotídeo sintéticos para a amplificação por meio de reação em cadeia de polimerase podem ser gerados com base em uma das seqüências mostradas na SEQ ID NO: 1 ou com a ajuda das seqüências de aminoácido detalhadas em SEQ ID NO: 2, SEQ ID NO: 3 e SEQ ID NO: 4. São mostrados iniciadores particularmente adequados nos
25 Exemplos como SEQ ID NO: 5 e SEQ ID NO: 6.

Um ácido nucléico de acordo com a invenção pode ser amplificado por técnicas padrão de amplificação de PCR que usam cDNA ou, alternativamente, DNA genômico como modelo (SEQ ID NO 9) e iniciadores de oligonucleotídeo adequados (SEQ ID NO: 5 e SEQ ID NO: 6). O ácido nucléico amplificado desse modo pode ser clonado de em um vetor adequado
30 e pode ser caracterizado por meio de análise de seqüência de DNA. Os oligonucleotídeos que correspondem a uma seqüência de nucleotídeo de des-

saturase ou elongase podem ser gerados através de métodos sintéticos padrão, por exemplo, usando um sintetizador de DNA automático.

Os ácidos nucleicos e moléculas de proteína acima mencionados com atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase, podem ser usados em um processo para a produção de ARA de ácido linoléico em organismos transgênicos.

Então, em um outro aspecto da invenção, é fornecido um processo para a conversão de ácido linoléico ou um derivado deste em ácido araquidônico ou um derivado deste em um organismo, o processo compreendendo introduz em um organismo que compreende ácido linoléico, pelo menos uma seqüência de ácido nucleico que compreende:

a) SEQ ID NO: 1 (seqüência completa 1047306867), seqüência que compreende resíduos de ácido nucleico 7668 a 12077 de SEQ ID NO: 1 ou um homólogo de um destes;

b) seqüências de ácido nucleico que hibridizam sob condições rigorosas com uma seqüência de ácido nucleico de SEQ ID NO: 1 ou uma seqüência que compreende resíduos de ácido nucleico 7668 a 12077 de SEQ ID NO: 1;

c) uma seqüência de ácido nucleico isolada que codifica polipeptídeos com atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase, em que o polipeptídeos é selecionado do grupo que consiste em SEQ ID NOS 2, 3 e 4;

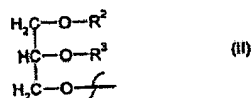
d) um derivado de uma seqüência de ácido nucleico de SEQ ID NO: 1 que codifica os polipeptídeos com pelo menos 40% de identidade no nível de aminoácido com SEQ ID NO: 2, SEQ ID NO: 3 e SEQ ID NO: 4; em que os referidos polipeptídeos têm atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase.

e expressando a referida seqüência de ácido nucleico.

No contexto da presente invenção, um "derivado" de ácido linoléico ou araquidônico é um composto no qual o OH da porção de ácido carboxílico é substituído por uma porção de R^1 , em que:

R^1 é coenzima A (tioéster), lisofosfatidilcolina, lisofosfatidiletano-

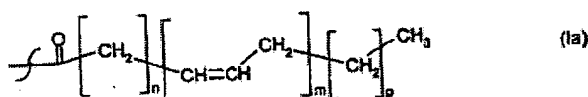
lamina, lisofosfatidilglicerol, liso - difosfatidilglicerol, lisofosfatidilserina, lisofosfatidilinositol, base de esfingo ou um radical da fórmula II



em qual

R^2 = hidrogênio, colina de lisofosfatidila, lisofosfatidiletanolamina, lisofosfatidilglicerol, lisodifosfatidilglicerol, lisofosfatidilserina, lisofosfatidilinositol ou C_2 - C_{24} -alquilcarbonila saturada ou não saturada,

R^3 = hidrogênio, C_2 - C_{24} -alquilcarbonila saturada ou não saturada, ou R^2 e R^3 são independentemente um do outro um radical da fórmula Ia:



em que

$n = 2, 3, 4, 5, 6, 7$ ou 9 , $m = 2, 3, 4, 5$ ou 6 e $p = 0$ ou 3 ;

e em que um oxigênio no radical R^1 pode ser substituído enxofre tal que R^1 seja ligado ao restante da molécula por uma ligação de tioéster.

Os processos de acordo com a invenção preferivelmente produzem ARA total em um conteúdo de pelo menos 1% em peso, vantajosamente pelo menos 3% em peso, com base nos ácidos graxos totais nos organismos transgênicos, preferivelmente em uma planta transgênica.

Uma vez que uma pluralidade de etapas de reação é realizada pelos compostos de partida, o ácido linoléico de ($18:2^{\Delta 9, 12}$) no processo de acordo com a invenção, ARA ($20:4^{\Delta 5, 8, 11, 14}$) não é obtido como um produto puro; traços menores dos precursores sempre estão presentes no produto final.

ARA quimicamente puro também pode ser sintetizado pelo processo descrito acima. Para este fim, ARA ou um derivado deste é isolado dos organismos, tal como os microorganismos ou as plantas ou o meio de cultura dentro ou sobre os quais os organismos foram crescidos, ou do organismo e do meio de cultura, da maneira conhecida, por exemplo, por ex-

tração, destilação, cristalização, cromatografia ou uma combinação destes métodos. Estes derivados de ARA ou ARA quimicamente são vantajosos para aplicações no setor de indústria alimentícia, no setor de cosmético e especialmente no setor de indústria farmacológica.

- 5 O processo pode incluir etapas adicionais de converter o ARA em um ácido graxo ω -3 introduzindo-se no organismo o ácido nucléico que codifica um ω -3 dessaturase e opcionalmente um Δ 5-elongase e/ou um Δ 4-elongase e/ou um Δ 4-dessaturase.

10 Em um outro aspecto a invenção compreende um processo para a conversão de $18:2^{\Delta 9, 12}$ (ácido linoléico) em $20:2^{\Delta 11, 14}$, o processo compreendendo introduzir em um organismo que compreende ácido linoléico pelo menos uma seqüência de ácido nucléico que codifica um polipeptídeo que tem atividade de Δ 9-elongase e que compreende:

a) uma seqüência que compreende resíduos de ácido nucléico
15 7668 a 9200 de SEQ ID NO: 1, SEQ ID NO: 9 ou um homólogo de um destes;

b) uma seqüência de ácido nucléico que hibridiza sob condições rigorosas com uma seqüência de ácido nucléico que compreende resíduos 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9;

20 c) uma seqüência de ácido nucléico isolada que codifica um polipeptídeo com atividade de Δ 9-elongase, em que o polipeptídeo compreende SEQ ID NO: 2 ou SEQ ID NO: 10;

d) um derivado de uma seqüência que compreende resíduos de ácido nucléico 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9 que codifica um polipeptídeo com pelo menos 40% de identidade no nível de amino-
25 ácido com SEQ ID NO: 2 ou SEQ ID NO: 10;

em que o referido polipeptídeo tem atividade de Δ 9-elongase;
e expressando a referida seqüência de ácido nucléico.

30 Em um outro aspecto da invenção é fornecido um processo para a conversão de $20:2^{\Delta 11, 14}$ em $20:3^{\Delta 8, 11, 14}$, o processo compreendendo introduzir em um organismo que compreende $20:2^{\Delta 11, 14}$, ou que compreende ácido linoléico e um Δ 9 elongase, uma seqüência de ácido nucléico isolada

que codifica um polipeptídeo com atividade de $\Delta 8$ -dessaturase e que é selecionado do grupo que consiste em:

- a) uma seqüência que compreende resíduos de ácido nucléico 9351 a 10724 de SEQ ID NO: 1 ou um homólogo desta;
- 5 b) seqüências de ácido nucléico que hibridiza sob condições rigorosas com uma seqüência de ácido nucléico que compreende resíduos 9351 a 10724 de SEQ ID NO: 1;
- c) uma seqüência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 8$ -dessaturase, em que o polipeptídeo compreende
- 10 SEQ ID NO: 3;
- d) um derivado de um uma seqüência que compreende resíduos de ácido nucléico 9351 a 10724 de SEQ ID NO: 1 que codifica um polipeptídeo com pelo menos 40% de identidade no nível de aminoácido com SEQ ID NO: 3;

15 em que o referido polipeptídeo tem atividade de $\Delta 8$ -dessaturase; e expressando a referida seqüência de ácido nucléico.

Em um outro aspecto da invenção, é fornecido um processo para a conversão de $20:3^{\Delta 8,11,14}$ em $20:4^{\Delta 5,8-11,14}$ (ARA), o processo compreendendo introduzir em um organismo que compreende $20:3^{\Delta 8,11,14}$ ou que com-

20 preende $20:2^{\Delta 11,14}$ e um $\Delta 8$ -dessaturase, ou que compreende ácido linoléico, um $\Delta 9$ -elongase e um $\Delta 8$ -dessaturase, uma seqüência de ácido nucléico isolada que codifica um polipeptídeo com atividade de $\Delta 5$ -dessaturase e que é selecionado do grupo que consiste em:

- a) uma seqüência que compreende resíduos de ácido nucléico
- 25 10842 a 12077 de SEQ ID NO: 1 ou um homólogo desta;
- b) seqüências de ácido nucléico que hibridizam sob condições rigorosas com uma seqüência de ácido nucléico que compreende resíduos 10842 a 12077 de SEQ ID NO: 1;
- c) uma seqüência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 5$ -dessaturase, em que o polipeptídeo compreende
- 30 SEQ ID NO: 4;
- d) um derivado de um uma seqüência que compreende resíduos

de ácido nucléico 10842 a 12077 de SEQ ID NO: 1 que codifica um polipeptídeo com pelo menos 40% de identidade no nível de aminoácido com SEQ ID NO: 4;

5 em que o referido polipeptídeo tem atividade de $\Delta 5$ -dessaturase e expressando a referida sequência de ácido nucléico.

O processo pode incluir etapas adicionais para converter o ARA em um ácido graxo ω -3 introduzindo no organismo, o ácido nucléico que codifica um ω -3 dessaturase e opcionalmente um $\Delta 5$ -elongase e/ou um $\Delta 4$ -elongase e/ou um $\Delta 4$ -dessaturase.

10 Para os processos apresentados acima, descobriu-se que a expressão foi mais efetivamente obtida usando indução com galactose.

Os organismos adequados para a produção no processo de acordo com a invenção são, em princípio, qualquer organismo tais como microorganismos, animais não humanos ou plantas.

15 As plantas que são adequadas são, em princípio, todas aquelas plantas que são capazes de sintetizar ácidos graxos, tal como todas as plantas dicotiledôneas ou monocotiledôneas, algas ou musgos. Plantas vantajosas são selecionadas do grupo das famílias de planta *Adelotheciaceae*, *Anacardiaceae*, *Asteraceae*, *Apiaceae*, *Betulaceae*, *Boraginaceae*, *Brassicaceae*, *Bromeliaceae*, *Caricaceae*, *Cannabaceae*, *Convolvulaceae*, *Chenopodiaceae*, *Cryptocodiniaceae*, *Cucurbitaceae*, *Ditrichaceae*, *Elaeagnaceae*, *Ericaceae*, *Euphorbiaceae*, *Fabaceae*, *Geraniaceae*, *Gramineae*, *Juglandaceae*, *Lauraceae*, *Leguminosae*, *Linaceae*, *Euglenaceae*, *Prasinophyceae* ou plantas vegetais ou ornamentais, tal como *Tagetes*.

25 Os exemplos que podem ser mencionados são as seguintes plantas selecionadas do grupo que consiste em: *Adelotheciaceae* tal como os gêneros *Physcomitrella*, por exemplo, o gênero e espécies *Physcomitrella patens*, *Anacardiaceae* tal como os gêneros *Pistacia*, *Mangifera*, *Anacardium*, por exemplo, o gênero e espécies *Pistacia vera* [pistache], *Mangifer indica* [manga] ou *Anacardium occidentale* [cajueiro], *Asteraceae*, tal como os gêneros *Calendula*, *Carthamus*, *Centaurea*, *Cichorium*, *Cynara*, *Helianthus*, *Lactuca*, *Locusta*, *Tagetes*, *Valeriana*, por exemplo, o gênero e espécies *Ca-*

30

- lendula officinalis* [calêndula comum], *Carthamus tinctorius* [açafroa], *Centaurea cyanus* [centáurea], *Cichorium intybus* [chicória], *Cynara scolymus* [alcachofra], *Helianthus annuus* [girassol], *Lactuca sativa*, *Lactuca crispa*, *Lactuca esculenta*, *Lactuca scariola*, *L. ssp. sativa*, *Lactuca scariola*, *L. var. integrata*, *Lactuca scariola*, *L. var. integrifolia*, *Lactuca sativa* subsp. *romana*,
5 *Locusta communis*, *Valeriana locusta* [legumes de salada], *Tagetes lucida*, *Tagetes erecta* ou *Tagetes tenuifolia* [calêndula africana ou de trincheira],
Apiaceae, tal como o gênero *Daucus*, por exemplo o gênero e espécies *Daucus carota* [cenoura], *Betulaceae*, tal como o gênero *Corylus*, por exem-
10 plo, os gêneros e espécies *Corylus avellana* ou *Corylus columna* [avelã], *Boraginaceae*, tal como o gênero *Borago*, por exemplo, o gênero e espécies *Borago officinalis* [borragem], *Brassicaceae*, tal como os gêneros *Brassica*,
Camelina, *Melanosinapis*, *Sinapis*, *Arabadopsis*, por exemplo, os gêneros e espécies *Brassica napus*, *Brassica rapa* ssp. [óleo de semente de colza],
15 *Sinapis arvensis*, *Brassica juncea*, *Brassica juncea* var. *juncea*, *Brassica juncea* var. *crispifolia*, *Brassica juncea* var. *foliosa*, *Brassica nigra*, *Brassica sinapioides*, *Camelina sativa*, *Melanosinapis communis* [mostarda], *Brassica oleracea* [beterraba de forragem] ou *Arabidopsis thaliana*, *Bromeliaceae*, tal
como os gêneros *Anana*, *Bromelia* (abacaxi), por exemplo, os gêneros e es-
20 pécies *Anana comosus*, *Ananas ananas* ou *Bromelia comosa* [abacaxi], *Caricaceae*, tal como o gênero *Carica*, tal como o gênero e espécies de *Carica*
papaya [mamão], *Cannabaceae*, tal como a *Cannabis* de gênero, tal como o gênero e espécies de *Canabbis sativa* [linho], *Convolvulaceae*, tal como os
gêneros *Ipomea*, *Convolvulus*, por exemplo, os gêneros e espécies *Ipomoea*
25 *batatus*, *Ipomoea pandurata*, *Convolvulus batatas*, *Convolvulus tiliaceus*,
Ipomoea fastigiata, *Ipomoea tiliacea*, *Ipomoea triloba* ou *Convolvulus panduratus* [batata-doce, batata], *Chenopodiaceae*, tal como o gênero *Beta*, tal
como os gêneros e espécies *Beta vulgaris*, *Beta vulgaris* var. *altissima*, *Beta vulgaris* var. *vulgaris*, *Beta maritima*, *Beta vulgaris* var. *perennis*, *Beta vulga-*
30 *ris* var. *conditiva* ou *Beta vulgaris* var. *esculenta* [beterraba doce], *Crypthecodiniaceae*, tal como o gênero *Crypthecodinium*, por exemplo, o gênero e
espécies *Cryptecodinium cohnii*, *Cucurbitaceae*, tal como o gênero *Cucurbi-*

ta, por exemplo, os gêneros e espécies *Cucurbita maxima*, *Cucurbita mixta*, *Cucurbita pepo* ou *Cucurbita moschata* [abóbora], *Cymbellaceae*, tal como os gêneros *Ânfora*, *Cymbella*, *Okedenia*, *Phaeodactylum*, *Reimeria*, por exemplo, o gênero e espécies *Phaeodactylum tricornutum*, *Ditrichaceae*, tal

5 como os gêneros *Ditrichaceae*, *Astomiopsis*, *Ceratodon*, *Chrysoblastella*, *Ditrichum*, *Distichium*, *Eccremidium*, *Lophidion*, *Philibertiella*, *Pleuridium*, *Saelania*, *Trichodon*, *Skottsbergia*, por exemplo os gêneros e espécies *Ceratodon antarcticus*, *Ceratodon columbiae*, *Ceratodon heterophyllus*, *Ceratodon purpurascens*, *Ceratodon purpureus*, *purpureus Ceratodon ssp.*, *convolutus*,

10 *Ceratodon purpureus ssp. stenocarpus*, *Ceratodon purpureus var. rotundifolius*, *Ceratodon ratodon*, *Ceratodon stenocarpus*, *Chrysoblastella Chilensis*, *Ditrichum ambiguum*, *Ditrichum brevisetum*, *Ditrichum crispatisimum*, *Ditrichum difficile*, *Ditrichum falcifolium*, *Ditrichum flexicaule*, *Ditrichum giganteum*, *Ditrichum heteromallum*, *Ditrichum lineare*, *Ditrichum lineare*, *Ditrichum montanum*, *Ditrichum montanum*, *Ditrichum pallidum*, *Ditrichum punctulatum*,

15 *Ditrichum pusillum*, *Ditrichum pusillum var. tortile*, *Ditrichum rhynchostegium*, *Ditrichum schimperi*, *Ditrichum tortile*, *Distichium capillaceum*, *Distichium hagenii*, *Distichium inclinatum*, *Distichium macounii*, *Eccremidium floridanum*, *Eccremidium whiteleggei*, *Lophidion strictus*, *Pleuridium acuminatum*, *Pleuridium altemifolium*, *Pleuridium holdridgei*, *Pleuridium mexicanum*, *Pleuridium ravenelii*, *Pleuridium subulatum*, *Saelania glaucescens*, *Trichodon borealis*, *Trichodon cylindricus* ou *Trichodon cylindricus var. oblongus*, *Elaeagnaceae*, tal como o gênero *Elaeagnus*, por exemplo, o gênero e espécie *Olea europaea* [azeitona], *Ericaceae*, tal como o gênero *Kalmia*, por exemplo, o gênero

25 e espécies *Kalmia latifolia*, *Kalmia angustifolia*, *Kalmia microphylla*, *Kalmia polifolia*, *Kalmia occidentalis*, *Cistus chamaerhodendros* ou *Kalmia lucida* [*loureiro da montanha*], *Euglenaceae*, tal como o gênero *Ascoglena*, *Astasia*, *Colacium*, *Cyclidiopsis*, *Euglena*, *Euglenopsis*, *Hyalaphacus*, *Khawkinea*, *Lepocinclis*, *Phacus*, *Strombomonas*, *Trachelomonas*, por exemplo, o gênero

30 e espécie *Euglena gracilis*; *Euphorbiaceae*, tal como o gênero *Manihot*, *Jani-pha*, *Jatropha*, *Ricinus*, por exemplo, o gênero e espécies *Manihot utilissima*, *Janipha manihot*, *Jatropha manihot*, *Manihot aipil*, *Manihot dulcis*, *Manihot*

- manihot, *Manihot melanobasis*, *Manihot esculenta* [mandioca] ou *Ricinus communis* [planta de óleo de rícino], Fabaceae, tal como o gênero *Pisum*, *Albizia*, *Cathormion*, *Feuillea*, *Inga*, *Pithecolobium*, *Acacia*, *Mimosa*, *Medicago*, *Glycine*, *Dolichos*, *Phaseolus*, soybean, por exemplo, o gênero e espécies
- 5 *Pisum sativum*, *Pisum arvense*, *Pisum humile* [pea], *Albizia berteriana*, *Albizia julibrissin*, *Albizia lebbeck*, *Acacia berteriana*, *Acacia littoralis*, *Albizia berteriana*, *Albizia 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*, *Serican-
rda julibrissin*, *Acacia lebbeck*, *Acacia macrophylla*, *Albizia lebbeck*, *Feuillea lebbeck*, *Mimosa lebbeck*, *Mimosa speciosa*, *Medicago sativa*, *Medicago fal-
cata*, *Medicago varia* [alfafa] *Glycine max* *Dolichos soja*, *Glycine gracilis*, *Glycine hispida*, *Phaseolus max*, *Soja hispida* ou *Soja max* [soja], Funariace-
15 ae, tal como o gênero *Aphanorrhagma*, *Entosthodon*, *Funaria*, *Physcomitrel-
la*, *Physcomitrium*, por exemplo, o gênero e espécies *Aphanorrhagma serra-
tum*, *Entosthodon attenuatus*, *Entosthodon bolanderi*, *Entosthodon bonplandii*, *Entosthodon californicus*, *Entosthodon drummondii*, *Entosthodon jame-
sonii*, *Entosthodon leibergii*, *Entosthodon neoscoticus*, *Entosthodon rubrise-
tus*, *Entosthodon spathulifolius*, *Entosthodon tucsoni*, *Funaria americana*,
20 *Funaria bolanderi*, *Funaria calcarea*, *Funaria californica*, *Funaria calvenscens*, *Funaria convoluta*, *Funaria flavicans*, *Funaria groutiana*, *Funaria hygrometri-
ca*, *Funaria hygrometrica* var. *arctica*, *Funaria hygrometrica* var. *calvenscens*, *Funaria hygrometrica* var. *convoluta*, *Funaria hygrometrica* var. *muralis*, *Fu-
naria hygrometrica* var. *utahensis*, *Funaria microstoma*, *Funaria microstoma*
25 var. *obtusifolia*, *Funaria muhlenbergii*, *Funaria orcuttii*, *Funaria plano-
convexe*, *Funaria polaris*, *Funaria ravenelii*, *Funaria rubriseta*, *Funaria serra-
ta*, *Funaria sonora*, *Funaria sublimbatus*, *Funaria tucsoni*, *Physcomitrella californica*, *Physcomitrella patens*, *Physcomitrella readeri*, *Physcomitrium*
30 *australe*, *Physcomitrium californicum*, *Physcomitrium collenchymatum*, *Phys-
comitrium coloradense*, *Physcomitrium cupuliferum*, *Physcomitrium drum-
mondii*, *Physcomitrium eurystomum*, *Physcomitrium flexifolium*, *Physcomitri-*

um hookeri, *Physcomitrium hookeri* var. *serratum*, *Physcomitrium immersum*, *Physcomitrium kellermanii*, *Physcomitrium megalocarpum*, *Physcomitrium pyriforme*, *pyriforme Physcomitrium* var. *serratum*, *Physcomitrium rufipes*, *Physcomitrium sandbergii*, *Physcomitrium subsphaericum*, *Physcomitrium washingtoniense*, Geraniaceae, tal como os gêneros *Pelargonium*, *Cocos*, *Oleum*, por exemplo, os gêneros e espécies *Cocos nucifera*, *Pelargonium grossularioides* ou *Oleum cocois* [coco], Gramineae, tal como o gênero *Saccharum*, por exemplo, o gênero e espécies *Saccharum officinarum*, Juglandaceae, tal como os gêneros *Juglans*, *Wallia*, por exemplo, os gêneros e espécies *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* ou *Wallia nigra* [noz], Lauraceae, tal como os gêneros *Persea*, *Laurus*, por exemplo, os gêneros e espécies *Laurus nobilis* [baía], *Persea americana*, *Persea gratissima* ou *Persea persea* [abacate], Leguminosae, tal como o gênero *Arachis*, por exemplo, o gênero e espécies *Arachis hypogaea* [amendoim], Linaceae, tal como os gêneros *Adenolinum*, por exemplo, os gêneros e espécies *Linum usitatissimum*, *Linum humile*, *Linum austriacum*, *Linum bienne*, *Linum angustifolium*, *Linum catharticum*, *Linum flavum*, *Linum grandiflorum*, *Adenolinum grandiflorum*, *Linum lewisii*, *Linum narbonense*, *Linum perenne*, *Linum perenne* var. *lewisii*, *Linum pratense* ou *Linum trigynum* [linhaça], Lythrarieae, tal como o gênero *Punica*, por exemplo, o gênero e espécies *Punica granatum* [romã], Malvaceae, tal como o gênero *Gossypium*, por exemplo, os gêneros e espécies *Gossypium hirsutum*, *Gossypium arboreum*, *Gossypium barbadense*, *Gossypium herbaceum* ou *Gossypium thurberi* [algodão], Marchantiaceae, tal como o gênero *Marchantia*, por exemplo, os gêneros e espécies *Marchantia berteroana*, *Marchantia foliacea*, *Marchantia macropora*, Musaceae, tal como o gênero *Musa*, por exemplo os gêneros e espécies *Musa nana*, *Musa acuminata*, *Musa paradisiaca*, *Musa* spp. [banana], Onagraceae, tal como os gêneros *Camissonia*, *Oenothera*, por exemplo, os gêneros e espécies *Oenothera biennis* ou *Camissonia brevipes* [prímula de noite], Palmae, tal como o gênero *Elaeis*, por exemplo, o

- gênero e espécies *Elaeis guineensis* [óleo de palma], Papaveraceae, tal como, por exemplo, o gênero *Papaver*, por exemplo, os gêneros e espécies *Papaver orientale*, *Papaver rhoeas*, *Papaver dubium* [papoula], Pedaliaceae, tal como o gênero *Sesamum*, por exemplo, o gênero e espécies *Sesamum indicum* [gergelim], Piperaceae, tal como o gênero *Piper*, *Artanthe*, *Peperomia*, *Steffensia*, por exemplo, os gêneros e espécies *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* [pimenta], Poaceae, tal como os gêneros *Hordeum*, *Secale*, *Avena*, *Sorghum*, *Andropogon*, *Holcus*, *Panicum*, *Oryza*, *Zea* [milho], *Triticum*, por exemplo, os gêneros e espécies *Hordeum vulgare*, *Hordeum jubatum*, *Hordeum murinum*, *Hordeum secalinum*, *Hordeum distichon*, *Hordeum aegiceras*, *Hordeum hexastichon*, *Hordeum hexastichum*, *Hordeum irregulare*, *Hordeum sativum*, *Hordeum secalinum* [cevada], *Secale cereale* [centeio], *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida* [aveias], *Sorghum bicolor*, *Sorghum halepense*, *Sorghum saccharatum*, *Sorghum vulgare*, *Andropogon drummondii*, *Holcus bicolor*, *Sorghum Holcus*, *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* [milho miúdo], *Oryza sativa*, *Oryza latifolia* de [arroz], *Zea mays* [milho] *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum* ou *Triticum vulgare* [trigo], Porphyridiaceae, tal como os gêneros *Chrootheca*, *Flintiella*, *Petrovanella*, *Porphyridium*, *Rhodella*, *Rhodosorus*, *Vanhoeffenia*, por exemplo, o gênero e espécies *Porphyridium cruentum*, Proteaceae, como o gênero *Macadamia*, por exemplo, o gênero e espécies *Macadamia intergrifolia* [macadâmia], Prasinophyceae, tal como os gêneros *Nephroselmis*, *Prasinococcus*, *Scherffelia*, *Tetraselmis*, *Mantoniella*, *Ostreococcus*, por exemplo, os gêneros e espécies, Olivacea de *Nephroselmis*, Prasi-

- nococcus capsulatus, Scherffelia dubia, Tetraselmis chui, Tetraselmis sueci-
 ca, Mantoniella squamata, Ostreococcus tauri, Rubiaceae, tal como o gênero
 Coffea, por exemplo, os gêneros e espécies Coffea spp., Coffea Arabica,
 Coffea canephora ou Coffea liberica, Scrophulariaceae, tal como o gênero
 5 Verbascum, por exemplo, os gêneros e espécies Verbascum blattaria, Ver-
 bascum chaixii, Verbascum densiflorum, Verbascum lagurus, Verbascum
 longifolium, Verbascum lychnitis, Verbascum nigrum, Verbascum olympicum,
 Verbascum phlomoides, Verbascum phoenicum, Verbascum pulverulentum
 ou Verbascum thapsus [verbascum], Solanaceae, tal como os gêneros Cap-
 10 sicum, Nicotiana, Solanum, Lycopersicon, por exemplo, os gêneros e espé-
 cies Capsicum annuum, Capsicum annuum var. glabriusculum, Capsicum
 frutescens [pimenta], Capsicum annuum [páprica], Nicotiana tabacum, Nico-
 tiana alata, Nicotiana attenuata, Nicotiana glauca, Nicotiana langsdorffii, Ni-
 cotiana obtusifolia, Nicotiana quadrivalvis, Nicotiana repanda, Nicotiana rus-
 15 tica, Nicotiana sylvestris [tabaco], Solanum tuberosum [batata], Solanum me-
 longena [berinjela], Lycopersicon esculentum, Lycopersicon lycopersicum,
 Lycopersicon Pyriforme, Solanum integrifolium ou Solanum lycopersicum
 [tomate], Sterculiaceae, tal como o gênero Theobroma, por exemplo, o gêne-
 ro e espécies Theobroma cacao [cacao] ou Theaceae, tal como o gênero
 20 Camellia, por exemplo, o gênero e espécies de Camellia sinensis [chá].

- Os microorganismos vantajosos são, por exemplo, fungos sele-
 cionados do grupo das famílias Chaetomiaceae, Choanephoraceae, Crypto-
 coccaceae, Cunninghamiaceae, Dematiaceae, Moniliaceae, Mortierellace-
 ae, Mucoraceae, Pythiaceae, Saccharomycetaceae, Saprolegniaceae, Schi-
 25 zosacharomycetaceae, Sodariaceae ou Tuberculariaceae.

- Os exemplos de microorganismos que podem ser mencionados
 são aqueles dos grupos: Choanephoraceae, tal como os gêneros Blakeslea,
 Choanephora, por exemplo, os gêneros e espécies Blakeslea trispora, Cho-
 anephora cucurbitarum, Choanephora infundibulifera var. cucurbitarum, Mor-
 30 tierellaceae, tal como o gênero Mortierella, por exemplo, os gêneros e espé-
 cies Mortierella isabellina, Mortierella polycephala, Mortierella ramanniana,
 Mortierella vinacea, Mortierella zonata, Pythiaceae, tal como os gêneros,

Phytium, Phytophthora, por exemplo, os gêneros e espécies *Pythium debaryanum*, *Pythium intermedium*, *Pythium irregulare*, *Pythium megalacanthum*, *Pythium paroecandrum*, *Pythium sylvaticum*, *Pythium ultimum*, *Phytophthora cactorum*, *Phytophthora cinnamomi*, *Phytophthora citricola*, *Phytophthora citrophthora*, *Phytophthora cryptogea*, *Phytophthora drechsleri*, *Phytophthora erythroseptica*, *Phytophthora lateralis*, *Phytophthora megasperma*, *Phytophthora nicotianae*, *Phytophthora nicotianae* var. *parasitica*, *Phytophthora palmivora*, *Phytophthora parasitica*, *Phytophthora syringae*, *Saccharomycetaceae*, tal como os gêneros *Hansenula*, *Pichia*, *Saccharomyces*, *Saccharomyces*, *Yarrowia*, por exemplo, os gêneros e espécies *Hansenula anomala*, *Hansenula californica*, *Hansenula canadensis*, *Hansenula capsulata*, *Hansenula ciferrii*, *Hansenula glucozyma*, *Hansenula henricii*, *Hansenula holstii*, *Hansenula minuta*, *Hansenula nonfermentans*, *Hansenula philodendri*, *Hansenula polymorpha*, *Hansenula saturnus*, *Hansenula subpelliculosa*, *Hansenula wickerhamii*, *Hansenula wingei*, *Pichia alcoholophila*, *Pichia angusta*, *Pichia anomala*, *Pichia bisporea*, *Pichia burtonii*, *Pichia canadensis*, *Pichia capsulata*, *Pichia carsonii*, *Pichia cellobiosa*, *Pichia ciferrii*, *Pichia farinosa*, *Pichia fermentans*, *Pichia finlandica*, *Pichia glucozyma*, *Pichia guilliermondii*, *Pichia haplophila*, *Pichia henricii*, *Pichia holstii*, *Pichia jadinii*, *Pichia lindnerii*, *Pichia membranaefaciens*, *Pichia methanolica*, *Pichia minuta* var. *minuta*, *Pichia minuta* var. *nonfermentans*, *Pichia norvegensis*, *Pichia ohmeri*, *Pichia pastoris*, *Pichia philodendri*, *Pichia pini*, *Pichia polymorpha*, *Pichia quercuum*, *Pichia rhodanensis*, *Pichia sargentensis*, *Pichia stipitis*, *Pichia strasburgensis*, *Pichia subpelliculosa*, *Pichia toletana*, *Pichia trehalophila*, *Pichia vini*, *Pichia xylosa*, *Saccharomyces acetii*, *Saccharomyces bailii*, *Saccharomyces bayanus*, *Saccharomyces bisporus*, *Saccharomyces capensis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces cerevisiae* var. *ellipsoideus*, *Saccharomyces chevalieri*, *Saccharomyces delbrueckii*, *Saccharomyces diastaticus*, *Saccharomyces drosophilae*, *Saccharomyces elegans*, *Saccharomyces ellipsoideus*, *Saccharomyces fermentati*, *Saccharomyces florentinus*, *Saccharomyces fragilis*, *Saccharomyces heterogenicus*, *Saccharomyces hienipiensis*, *Saccharomyces inusitatus*, *Saccharomyces*

italicus, *Saccharomyces kluyveri*, *Saccharomyces krusei*, *Saccharomyces lactis*, *Saccharomyces marxianus*, *Saccharomyces microellipsoides*, *Saccharomyces montanus*, *Saccharomyces norbensis*, *Saccharomyces oleaceus*, *Paradoxus*, *Saccharomyces*, *Saccharomyces pastorianus*, *Saccharomyces pretoriensis*, *Saccharomyces rosei*, *Saccharomyces rouxii*, *Saccharomyces uvarum*, *Saccharomycodes ludwigii*, *Yarrowia lipolytica*, Schizosaccharomycetaceae tal como os gêneros *Schizosaccharomyces*, por exemplo, as espécies *Schizosaccharomyces japonicus* var. *japonicus*; *Schizosaccharomyces japonicus* var. *versatilis*, *Schizosaccharomyces malidevorans*, *Schizosaccharomyces octosporus*, *Schizosaccharomyces pombe* var. *malidevorans*, *Schizosaccharomyces pombe* var. *pombe*, Thraustochytriaceae tal como os gêneros *Althornia*, *Aplanochytrium*, *Japonochytrium*, *Schizochytrium*, *Thraustochytrium*, por exemplo as espécies *Schizochytrium aggregatum*, *Schizochytrium limacinum*, *Schizochytrium mangrove*, *Schizochytrium minutum*, *Schizochytrium octosporum*, *Thraustochytrium aggregatum*, *Thraustochytrium amoeboides*, *Thraustochytrium antacticum*, *Thraustochytrium arudimentale*, *Thraustochytrium aureum*, *Thraustochytrium benthiicola*, *Thraustochytrium globosum*, *Thraustochytrium indicum*, *Thraustochytrium kerguelense*, *Thraustochytrium kinnei*, *Thraustochytrium motivum*, *Thraustochytrium multirudimentale*, *Thraustochytrium pachydermum*, *Thraustochytrium proliferum*, *Thraustochytrium roseum*, *Thraustochytrium rossii*, *Thraustochytrium striatum* ou *Thraustochytrium visurgense*.

Outros microorganismos vantajosos são, por exemplo, bactérias selecionadas do grupo das famílias Bacillaceae, Enterobacteriaceae ou Rhizobiaceae.

Os exemplos que podem ser mencionados são os seguintes microorganismos selecionados do grupo que consiste em: Bacillaceae, tal como o gênero de *Bacillus*, por exemplo, os gêneros e espécies de *Bacillus acidocaldarius*, *Bacillus acidoterrestris*, *Bacillus alcalophilus*, *Bacillus amyloliquefaciens*, *Bacillus amylolyticus*, *Bacillus brevis*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus sphaericus* subsp. *fusiformis*, *Bacillus galactophilus*, *Bacillus globisporus*, *Bacillus globisporus* subsp. *marinus*, Ba-

cillus halophilus, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus polymyxa*, *Bacillus psychrosaccharolyticus*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus subtilis* subsp. *spizizenii*, *Bacillus* subsp. *subtilis* ou *Bacillus thuringiensis*; Enterobacteriaceae tal

5 como os gêneros *Citrobacter*, *Edwardsiella*, *Enterobacter*, *Erwinia*, *Escherichia*, *Klebsiella*, *Salmonella* ou *Serratia*, por exemplo, os gêneros e espécies *Citrobacter amalonaticus*, *Citrobacter diversus*, *Citrobacter freundii*, *Citrobacter* genomospecies, *Citrobacter gillenbergii*, *Citrobacter intermedium*, *Citrobacter koseri*, *Citrobacter murlinae*, *Citrobacter* sp., *Edwardsiella hoshinae*, *Edwardsiella*

10 *ictaluri*, *Edwardsiella tarda*, *Erwinia alni*, *Erwinia amylovora*, *Erwinia ananatis*, *Erwinia aphidicola*, *Erwinia billingiae*, *Erwinia cacticida*, *Erwinia cancerogena*, *Erwinia carnegieana*, *Erwinia carotovora* subsp. *atroseptica*, *Erwinia carotovora* subsp. *betavascularum*, *Erwinia carotovora* subsp. *odorifera*, *Erwinia carotovora* subsp. *wasabiae*, *Erwinia chrysanthemi*, *Erwinia cypripedii*,

15 *Erwinia dissolvens*, *Erwinia herbicola*, *Erwinia mallotivora*, *Erwinia milletiae*, *Erwinia nigrifluens*, *Erwinia nimipressuralis*, *Erwinia persicina*, *Erwinia psidii*, *Erwinia pyrifoliae*, *Erwinia quercina*, *Erwinia rhapontici*, *Erwinia rubrifaciens*, *Erwinia salicis*, *Erwinia stewartii*, *Erwinia tracheiphila*, *Erwinia uredovora*, *Escherichia adecarboxylata*, *Escherichia anindolica*, *Escherichia*

20 *aureoscens*, *Escherichia blattae*, *Escherichia coli*, *Escherichia coli* var. *communior*, *Escherichia coli*-mutabile, *Escherichia fergusonii*, *Escherichia hermannii*, *Escherichia* sp., *Escherichia Vulneris*, *Klebsiella aerogenes*, *Klebsiella edwardsii* subsp. *atlantae*, *Klebsiella ornithinolytica*, *Klebsiella oxytoca*, *Klebsiella planticola*, *Klebsiella pneumoniae*, *Klebsiella pneumoniae* subsp.

25 *pneumoniae*, *Klebsiella* sp., *Klebsiella terrigena*, *Klebsiella trevisanii*, *Salmonella abony*, *Salmonella arizonae*, *Salmonella bongori*, *Salmonella choleraesuis* subsp. *arizonae*, *Salmonella choleraesuis* subsp. *bongori*, *Salmonella choleraesuis* subsp. *choleraesuis*, *Salmonella choleraesuis* subsp. *diarizonae*, *Salmonella choleraesuis* subsp. *houtenae*, *Salmonella choleraesuis* subsp.

30 *indica*, *Salmonella choleraesuis* subsp. *salamae*, *Salmonella daressalaam*, *Salmonella enterica* subsp. *houtenae*, *Salmonella enterica* subsp. *salamae*, *Salmonella enteritidis*, *Salmonella gallinarum*, *Salmonella heidelberg*, *Salmo-*

nela panama, *Salmonella senftenberg*, *Salmonella typhimurium*, *Serratia entomophila*, *Serratia ficaria*, *Serratia fonticola*, *Serratia grimesii*, *Serratia liquefaciens*, *Serratia marcescens*, *Serratia marcescens* subsp. *marcescens*, *Serratia marinorubra*, *Serratia odorifera*, *Serratia plymouthis*, *Serratia piymuthica*, *Serratia proteamaculans*, *Serratia proteamaculans* subsp. *quinovora*, *Serratia quinivorans* ou *Serratia rubidaea*; Rhizobiaceae, tal como os gêneros *Agrobacterium*, *Carbophilus*, *Cheimatobacter*, *Ensifer*, *Rhizobium*, *Sinorhizobium*, por exemplo, os gêneros e espécies *Agrobacterium atlanticum*, *Agrobacterium ferrugineum*, *Agrobacterium gelatinovorum*, *Agrobacterium larymoorei*, *Agrobacterium meteoris*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium stellulatum*, *Agrobacterium tumefaciens*, *Agrobacterium vitis*, *Carbophilus carboxidus*, *Cheimatobacter heintzii*, *Ensifer adhaerens*, *Ensifer arboris*, *Ensifer fredii*, *Ensifer kostiensis*, *Ensifer kummerowiae*, *Ensifer medicae*, *Ensifer meliloti*, *Ensifer sahelii*, *Ensifer teranga*, *Ensifer xinjiangensis*, *Rhizobium ciceris*, *Rhizobium etli*, *Rhizobium fredii*, *Rhizobium galegae*, *Rhizobium gallicum*, *Rhizobium giardinii*, *Rhizobium hainanense*, *Rhizobium huakuii*, *Rhizobium huautlense*, *Rhizobium indigoferae*, *Rhizobium japonicum*, *Rhizobium leguminosarum*, *Rhizobium loessense*, *Rhizobium loti*, *Rhizobium lupini*, *Rhizobium mediterraneum*, *Rhizobium meliloti*, *Rhizobium mongolense*, *Rhizobium phaseoli*, *Rhizobium radiobacter*, *Rhizobium rhizogenes*, *Rhizobium rubi*, *Rhizobium sulae*, *Rhizobium tianshanense*, *Rhizobium trifolii*, *Rhizobium tropici*, *Rhizobium undicola*, *Rhizobium vitis*, *Sinorhizobium adhaerens*, *Sinorhizobium arboris*, *Sinorhizobium fredii*, *Sinorhizobium kostiensis*, *Sinorhizobium kummerowiae*, *Sinorhizobium medicae*, *Sinorhizobium meliloti*, *Sinorhizobium morelense*, *Sinorhizobium sahelii* ou *Sinorhizobium xinjiangense*.

Outros exemplos de microorganismos vantajosos para o processo de acordo com a invenção são protistas ou diatomáceas selecionadas do grupo das famílias Dinophyceae, Turaniellidae ou Oxytrichidae, tal como os gêneros e espécies: *Cryptocodinium cohnii*, *Phaeodactylum tricoratum*, *Stylonychia mytilus*, *Stylonychia pustulata*, *Stylonychia putrina*, *Stylonychia notophora*, *Stylonychia* sp., *Colpidium campylum* ou *Colpidium* sp.

Aqueles que são vantajosamente aplicados no processo de acordo com a invenção são organismos transgênicos, tal como fungos, tal como *Mortierella* ou *Thraustochytrium*, leveduras tal como *Saccharomyces* ou *Schizosaccharomyces*, musgos tal como *Physcomitrella* ou *Ceratodon*, animais não humanos tal como *Caenorhabditis*, algas tal como *Nephroselmis*, *Pseudoscourfielda*, *Prasinococcus*, *Scherffelia*, *Tetraselmis*, *Mantoniella*, *Ostreococcus*, *Cryptocodinium* ou *Phaeodactylum* ou plantas tal como plantas dicotiledôneas ou monocotiledôneas. Os organismos que são especialmente vantajosamente usados no processo de acordo com a invenção são organismos que pertencem aos organismos produtores de óleo, quer dizer que são usados para a produção de óleo, tal como fungos, tal como *Mortierella* ou *Thraustochytrium*, algas tal como *Nephroselmis*, *Pseudoscourfielda*, *Prasinococcus*, *Scherffelia*, *Tetraselmis*, *Mantoniella*, *Ostreococcus*, *Cryptocodinium*, *Phaeodactylum*, ou plantas, em particular plantas, preferivelmente plantas de semente de óleo ou de colheita de óleo que compreendem quantidades grandes de compostos de lipídio, tal como amendoim, óleo de semente de colza, óleo de colza, girassol, açafoa (*Carthamus tinctoria*), papoula, mostarda, linho, planta de óleo de rícino, azeitona, gergelim, *Calêndula*, *Punica*, *prímula de noite*, *verbascum*, cardo, rosas selvagens, avelã, amêndoa, macadâmia, abacate, baía, abóbora, linhaça, soja, pistaches, borragem, árvores (óleo de palma, coco ou noz) ou colheitas cultiváveis tal como milho, trigo, centeio, aveias, tritcale, arroz, cevada, algodão, mandioca, pimenta, plantas *Tagetes*, *Solanaceae*, tal como batata, tabaco, berinjela e tomate, espécies de *Vicia*, ervilha, alfafa ou plantas fechadas (café, cacau, chá), espécies de *Salix*, e gramas perenes e colheitas de forragem. As plantas preferidas de acordo com a invenção são plantas de colheita de óleo tal como amendoim, óleo de semente de colza, óleo de colza, girassol, açafoa, papoula, mostarda, linho, planta de óleo de rícino, azeitona, *Calêndula*, *Punica*, *prímula de noite*, abóbora, linhaça, soja, borragem, árvores (óleo de palma, coco). Especialmente preferido para o processo são plantas que têm o teor alto de ácidos graxos C18:2-, tal como girassol, açafoa, tabaco, *verbascum*, gergelim, algodão, abóbora, papoula, *prímula de noite*, noz, linhaça,

linho ou cardo. Muito especialmente preferidas plantas são plantas tal como açafroa, girassol, papoula, prímula de noite, noz, linhaça ou linho.

Também deve ser vantajoso para o método acima descrito de acordo com a invenção adicionalmente introduzir, no organismo, outros ácidos nucleicos que codificam as enzimas do metabolismo de ácido graxo ou lipídio, além dos ácidos nucleicos dos primeiro a quarto aspectos da invenção.

Tais ácidos nucleico são vantajosamente derivados de plantas tal como algas, por exemplo, algas da família do Prasinophyceae tal como os gêneros *Heteromastix*, *Mammella*, *Mantoniella*, *Micromonas*, *Nephroselmis*, *Ostreococcus*, *Prasinocladus*, *Prasinococcus*, *Pseudoscourfielda*, *Pycnococcus*, *Pyramimonas*, *Scherffelia* ou *Tetraselmis*, tal como os gêneros e espécies *Heteromastix longifilllis*, *Mamiella gilva*, *Mantoniella squamata*, *Micromonas pusilla*, *Nephroselmis olivacea*, *Nephroselmis pyriformis*, *Nephroselmis rotunda*, *Ostreococcus tauri*, *Ostreococcus sp.*, *Prasinocladus Ascus*, *Prasinocladus lubricus*, *Pycnococcus provasolii*, *Pyramimonas amylifera*, *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* ou *Tetraselmis sp.* ou de algas da família Euglenaceae tal como o gênero *Ascoglena*, *Astasia*, *Colacium*, *Cyclidiopsis*, *Euglena*, *Euglenopsis*, *Hyalophacus*, *Khawkinea*, *Lepocinclis*, *Phacus*, *Strombomonas* ou *Trachelomonas*, tal como o gênero e espécie *Euglena acus*, *Euglena geniculate*, *Euglena gracilis*, *Euglena mixocylindracea*, *Euglena rostrifera*, *Euglena viridis*, *Colacium stentorium*, *Trachelomonas cylindrica* ou *Trachelomonas volvocina*. São derivados os ácidos nucleicos usados vantajosamente de algas dos gêneros *Euglena*, *Mantoniella* ou *Ostreococcus*.

Outras plantas vantajosas são algas tal como *Isochrysis* ou *Crypthecodinium*, algas/diatomáceas, tal como *Thalassiosira* ou *Phaeodactylum*, musgos tal como *Physcomitrella* ou *Ceratodon*, ou plantas mais altas tal como *Primulaceae*, tal como *Aleuritia*, *Calendula stellata*, *Osteospermum spinescens* ou *Osteospermum hyoseroides*, microorganismos tal como fungos, tal como *Aspergillus*, *Thraustochytrium*, *Phytophthora*, *Entomophthora*, *Mucor* ou *Mortierella*, bactérias tal como *Shewanella*, leveduras ou animais tal como nematódeos como *Caenorhabditis*, insetos, rãs, abalone, ou peixe. As seqüências de ácido nucléico isoladas de acordo com a invenção são

10 vantajosamente derivadas de um animal da ordem dos vertebrados. Preferivelmente, as seqüências de ácido nucléico são derivadas das classes do Vertebrata; Euteleostomi, Actinopterygii; Neopterygii; Teleostei; Euteleostei, Protacanthopterygii, Salmoniformes; Salmonidae ou *Oncorhynchus* ou Vertebrata, Amphibia, Anura, Pipidae, *Xenopus* ou Evertebrata tal como Proto-

15 chordata, Tunicata, Holothuroidea, Cionidae tal como *Amaroucium constellatum*, *Botryllus schlosseri*, *Ciona intestinalis*, *Molgula citrina*, *Molgula manhatensis*, *Perophora viridis* ou *Styela partita*. Os ácidos nucléicos são especialmente vantajosamente derivados de fungos, animais, ou de plantas tal como algas ou musgos, preferivelmente da ordem do Salmoniformes, tal como

20 a família do Salmonidae, tal como o gênero *Salmo*, por exemplo, dos gêneros e espécies *Oncorhynchus mykiss*, *Trutta trutta* ou *Salmo trutta fario*, de algas, tal como os gêneros *Mantoniella* ou *Ostreococcus*, ou das diatomáceas tal como os gêneros *Thalassiosira* ou *Phaeodactylum* ou de algas tal como *Crypthecodinium*.

25 Em uma modalidade preferida, o processo compreende a etapa de obter uma célula ou um organismo intacto que compreende as seqüências de ácido nucléico usadas no processo, onde a célula e/ou organismo é transformado com uma seqüência de ácido nucléico de acordo com a invenção que codifica o $\Delta 9$ -elongase, e/ou $\Delta 8$ -dessaturase e/ou $\Delta 5$ -dessaturase,

30 um construto de gene ou um vetor como descrito acima, sozinho ou em combinação com outras seqüências de ácido nucléico que codificam as proteínas do metabolismo de ácido graxo ou lipídio. Em uma outra modalidade

preferida, este processo também compreende a etapa de obter os óleos, lipídios ou ácidos graxos livres do organismo ou da cultura. A cultura pode, por exemplo, tomar a forma de uma cultura de fermentação, por exemplo, no caso do cultivo de microorganismos, tal como, por exemplo, *Mortierella*, *Thalassiosira*, *Mantoniella*, *Ostreococcus*, *Saccharomyces* ou *Thraustochytrium*,
 5 ou uma cultura de crescimento em estufa ou campo de uma planta. A célula ou o organismo produzido desse modo é vantajosamente uma célula de um organismo produtor de óleo, tal como uma colheita de óleo, tal como, por exemplo, amendoim, óleo de semente de colza, óleo de colza, linhaça, linho,
 10 amendoim, soja, açafrão, linho, girassóis ou borragem.

No caso de células de planta, tecido de planta ou órgãos de planta, "crescer" é entendido como significando, por exemplo, o cultivo dentro ou sobre um meio de nutriente, ou da planta intacta sobre ou dentro de um substrato, por exemplo, em uma cultura hidropônica, composto de plan-
 15 tação ou em terra cultivável.

Para os propósitos da invenção, "transgênico" ou "recombinante" significa com respeito a, por exemplo, uma sequência de ácido nucléico, um cassete de expressão (=construto de gene) ou um vetor que compreende a sequência de ácido nucléico ou um organismo transformado com as se-
 20 quências de ácido nucléico, cassetes de expressão ou vetores de acordo com a invenção, todas essas construtos são realizadas através de métodos recombinantes nos quais ou

a) a sequência de ácido nucléico de acordo com a invenção, ou
 b) uma sequência de controle genética que é operavelmente li-
 25 gada com a sequência de ácido nucléico de acordo com a invenção, por exemplo, um promotor, ou

c) a) e b)

não ficam situadas no seu ambiente genético natural ou foi modificada através de métodos recombinantes, isto sendo possível para a modifi-
 30 cação tomar a forma de, por exemplo, uma substituição, adição, deleção, inversão ou inserção de um ou mais resíduos de nucleotídeo. O ambiente genético natural é compreendido como significando o loco genômico ou

cromossômico natural no organismo original ou a presença em uma biblioteca genômica. No caso de uma biblioteca genômica, o ambiente genético natural da seqüência de ácido nucléico é preferivelmente retido, pelo menos em parte. O ambiente flanqueia a seqüência de ácido nucléico pelo menos em um lado e tem um comprimento de seqüência de pelo menos 50 pb, preferivelmente pelo menos 500 pb, especialmente preferivelmente pelo menos 1000 pb, preferivelmente pelo menos 5000 pb. Um cassete de expressão de ocorrência natural - por exemplo, a combinação de ocorrência natural do promotor natural das seqüências de ácido nucléico com o gene de $\Delta 5$ -
10 dessaturase correspondente - se torna um cassete de expressão transgênico quando este cassete de expressão é modificado por métodos não-naturais, sintéticos ("artificial") tal como, por exemplo, tratamento mutagênico. Os métodos adequados são descritos, por exemplo, na US 5.565.350 ou WO 00/15815.

15 Um organismo transgênico ou plantas transgênicas para os propósitos da invenção é então compreendido como significando, como acima, que os ácidos nucléicos usados no processo não estão no seu lugar natural no genoma de um organismo, sendo possível que os ácidos nucléicos sejam expressados homologamente ou heterologamente. Porém, como menciona-
20 do, transgênico também significa que, ao mesmo tempo em que os ácidos nucléicos de acordo com a invenção estão em sua posição natural no genoma de um organismo, a seqüência foi modificada com respeito à seqüência natural, e/ou que as seqüências reguladoras das seqüências naturais foram modificadas. Transgênico é preferivelmente entendido como significando a
25 expressão dos ácidos nucléicos de acordo com a invenção em um lugar antinatural no genoma, isto é, homólogo ou, preferivelmente, a expressão heteróloga dos ácidos nucléicos ocorre. Os organismos transgênicos preferidos são fungos tal como *Mortierella* ou *Phytophthora*, musgos tal como *Physcomitrella*, algas tal como *Mantoniella*, *Euglena*, *Cryptocodinium* ou *Ostreococcus*, diatomáceas tal como *Thalassiosira* ou *Phaeodactylum*, ou plantas tal
30 como as colheitas de óleo.

Os organismos ou organismos hospedeiros para os ácidos nu-

cléico, o cassete de expressão ou o vetor usado no processo de acordo com a invenção são, em princípio, vantajosamente todos os organismos que são capazes de sintetizar ácidos graxos, especificamente ácidos graxos não saturados, e/ou que são adequados para a expressão de genes recombinantes. Os exemplos que podem ser mencionados são plantas tal como *Arabidopsis*, *Asteraceae* tal como *Calendula* ou plantas de colheita tal como soja, amendoim, planta de óleo de rícino, girassol, milho, algodão, linho, óleo de semente de colza, coco, óleo de palma, açafroa (*Carthamus tinctorius*) ou semente de cacau, microorganismos, tal como fungos, por exemplo, o gênero *Mortierella*, *Thraustochytrium*, *Saprolegnia*, *Phytophthora* ou *Pythium*, bactérias, tal como o gênero *Escherichia* ou *Shewanella*, leveduras, tal como o gênero *Saccharomyces*, cyanobacteria, ciliates, algas tal como *Mantoniella*, *Euglena*, *Thalassiosira* ou *Ostreococcus*, ou protozoários tal como dinoflagellates, tal como *Cryptocodinium*. Os organismos preferidos são aqueles que são naturalmente capazes de sintetizar quantidades significativas de óleo, tal como fungos, tal como *Mortierella alpina*, *Pythium insidiosum*, *Phytophthora infestans*, ou plantas tal como soja, óleo de semente de colza, coco, óleo de palma, açafroa, linho, linho, planta de óleo de rícino, calêndula, amendoim, semente de cacau ou girassol, ou leveduras, tal como *Saccharomyces cerevisiae*, tal com soja, linho, óleo de semente de colza, açafroa, girassol, Calêndula, *Mortierella* ou *Saccharomyces cerevisiae* que são especialmente preferidos. Em princípio, os organismos hospedeiros são, além dos organismos transgênicos acima mencionados, também animais transgênicos, vantajosamente os animais não humanos, por exemplo, *C. elegans*, *Ciona intestinalis* ou *Xenopus laevis*.

Também utilizáveis são as células hospedeiras detalhadas em: Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990).

A expressão cepas que pode ser usada, por exemplo, aquelas com uma atividade de protease mais baixa, é descrita em: Gottesman, Sr, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128.

Estes incluem células de planta e certos tecidos, órgãos e partes de plantas em todas suas formas fenotípicas tal como anteras, fibras, cabelos da raiz, talos, embriões, caule, cotilédneas, pecíolos, material colhido, tecido de planta, tecido reprodutivo e culturas de célula que sejam derivados da planta transgênica atual e/ou possam ser usados para realizar a planta transgênica.

As plantas transgênicas que compreendem os ácidos graxos poliinsaturados sintetizados no processo de acordo com a invenção podem ser comercializadas vantajosamente e diretamente sem haver qualquer necessidade de que os óleos, lipídios ou ácidos graxos sintetizados sejam isolados. As plantas para o processo de acordo com a invenção são listadas como significando plantas intactas e todas as partes da planta, órgãos de planta ou partes de planta tal como folha, talo, sementes, raiz, tubérculos, anteras, fibras, cabelos da raiz, talos, embriões, caule, cotilédneas, pecíolos, material colhido, tecido de planta, tecido reprodutivo e cultura de célula que são derivados da planta transgênica atual e/ou podem ser usados para realizar a planta transgênica. Neste contexto, a semente compreende todas as partes da semente tal como a casca da semente, células epidérmicas, células da semente, endosperma ou tecido embrionário. Porém, os compostos produzidos no processo de acordo com a invenção também podem ser isolados dos organismos, vantajosamente plantas, na forma de seus óleos, gorduras, lipídios e/ou ácidos graxos livres. Os ácidos graxos poliinsaturados produzidos por este processo podem ser obtidos colhendo-se os organismos, ou da colheita na qual eles crescem, ou do campo. Isto pode ser feito por pressão ou extração das partes da planta, preferivelmente as sementes de planta. Neste contexto, os óleos, gorduras, de lipídios e/ou ácidos graxos livres podem ser obtidos pelo que é conhecido como batimento a frio ou pressionamento a frio sem aplicar calor. Para permitir maior facilidade de rompimento das partes da planta, especificamente as sementes, elas são previamente fragmentadas, cozidas em vapor ou assadas. As sementes que foram pré-tratadas desta maneira podem ser subsequente apertadas ou extraídas com solventes como hexano quente. O solvente é subsequen-

temente removido. No caso de microorganismos, os últimos são, após colheita, por exemplo, extraídos diretamente sem outras etapas de processamento ou ainda, após rompimento, extraídos por vários métodos com os quais o operário qualificado está familiarizado. Desta maneira, pode ser isolado mais que 96% dos compostos produzidos no processo. Depois disso, os produtos resultantes são processados mais adiante, isto é, refinados. Neste processo, substâncias como as mucilagens de planta e substâncias suspensas são primeiro removidas. O que é conhecido como desenlameamento pode ser realizado enzimaticamente ou, por exemplo, quimiofisicamente por adição de ácido tal como ácido fosfórico. Depois disso, os ácidos graxos livres são removidos através de tratamento com uma base, por exemplo, solução de hidróxido de sódio. O produto resultante é lavado completamente com água para remover o álcali que permanece no produto e então é secado. Para remover o pigmento que permanece no produto, os produtos são sujeitados a branqueamento, por exemplo, usando a terra da carga ou carvão ativo. No final, o produto é desodorizado, por exemplo, usando vapor.

Os ácidos graxos produzidos pelos processos da presente invenção podem ser isolados do organismo na forma de um óleo, um lipídio ou um ácido graxo livre. Por exemplo, os organismos adequados são aqueles mencionados acima. Os organismos preferidos são plantas transgênicas.

Uma modalidade da invenção é então óleos, lipídios ou ácidos graxos da fórmula I ou frações destes que foram produzidos pelo processo acima descrito, especialmente preferivelmente óleo, lipídio ou uma composição de ácido graxo que compreendem um composto de fórmula I e sendo derivado de plantas transgênicas.

Uma outra modalidade de acordo com a invenção é o uso do óleo, lipídio, dos ácidos graxos e/ou da composição de ácido graxo em produtos alimentícios, comestíveis, cosméticos ou farmacêuticos. Os óleos, lipídios, ácidos graxos ou misturas de ácido graxo de acordo com a invenção podem ser usados da maneira com a qual o trabalhador versado está familiarizado para misturar com outros óleos, lipídios, ácidos graxos ou misturas

de ácido graxo de origem animal, tal como, por exemplo, óleos de peixe. Também podem ser usados estes óleos, lipídios, ácidos graxos ou misturas de ácido graxo que sejam compostos de legume e componentes de animais para a preparação de produtos alimentícios, comestíveis, cosméticos ou farmacológicos.

O termo "óleo", "lipídio" ou "gordura" é entendido como significando uma mistura de ácido graxo que compreende ácido(s) graxo não saturado, saturado, preferivelmente esterificado. O óleo, lipídio ou gordura é preferivelmente elevado em ácido(s) graxo poliinsaturado livre ou, vantajosamente, esterificado, em particular ácido linoléico, ácido γ -linolênico, ácido di-homo- γ -linolênico, ácido araquidônico, ácido α -linolênico, ácido estearidônico, ácido eicosatetraenóico, ácido eicosapentaenóico, ácido docosapentaenóico ou ácido docosaexaenóico.

A quantidade de ácidos graxos esterificados não saturados preferivelmente quantidades de aproximadamente 30%, um conteúdo de 50% é mais preferido, um conteúdo de 60%, 70%, 80% ou mais é ainda mais preferido. Para a análise, o conteúdo de ácido graxo pode, por exemplo, ser determinado por cromatografia de gás depois de converter os ácidos graxos nos ésteres de metila através de transesterificação. O óleo, lipídio ou gordura podem compreender vários outros ácidos graxo saturados ou não saturados, por exemplo ácido calendúlico, ácido palmítico, ácido palmitoléico, ácido esteárico, ácido oléico e outros. O conteúdo dos vários ácidos graxos no óleo ou gordura pode variar, em particular dependendo do organismo de partida.

O ARA produzido no processo pode estar, como descrito acima, na forma de derivado de ácido graxo, por exemplo, esfingolipídeos, fosfoglicerídeos, lipídios, glicolipídeos, fosfolipídeos, monoacilglicerol, diacilglicerol, triacilglicerol ou outros ésteres de ácido graxo.

O ARA e outros ácidos graxos poliinsaturados que estão presentes podem ser liberados, por exemplo, por tratamento com álcali, por exemplo, KOH ou NaOH aquoso, ou hidrólise de ácido, vantajosamente na presença de um álcool tal como metanol ou etanol, ou por clivagem enzimática,

e isolado por, por exemplo, separação de fase e acidificação subsequente por, por exemplo, H_2SO_4 . Os ácidos graxos também podem ser liberados diretamente sem a etapa de processamento acima descrita.

Depois da sua introdução em um organismo, vantajosamente
5 uma célula de planta ou planta, os ácidos nucleicos usados no processo, ou podem estar presentes em um plasmídeo separado ou, vantajosamente, integrado no genoma da célula hospedeira. No caso de integração no genoma, a integração pode ser aleatória ou então pode se efetuar através de recombinação tal que o gene nativo seja substituído pela cópia introduzida, por
10 meio do que a produção do composto desejado pela célula é modulada, ou pelo uso de um gene em trans, de forma que o gene seja operavelmente ligado com uma unidade de expressão funcional que compreende pelo menos uma seqüência que garante a expressão de um gene, e pelo menos uma seqüência que garante a poliadenilação de um gene funcionalmente
15 transcrito. Os ácidos nucleicos são introduzidos vantajosamente nos organismos por cassetes de multiexpressão ou construtos para expressão multiparalela, vantajosamente nas plantas para a expressão específica de semente multiparalela de genes.

Se os microorganismos tal como leveduras, tal como *Saccharomyces* ou *Schizosaccharomyces*, fungos tal como *Mortierella*, *Aspergillus*, *Phytophthora*, *Entomophthora*, *Mucor* ou *Thraustochytrium*, algas tal como *Isochrysis*, *Mantoniella*, *Euglena*, *Ostreococcus*, *Phaeodactylum* ou *Cryptocodium* forem usados como organismos no processo de acordo com a invenção, estes organismos serão vantajosamente crescidos em culturas de
25 fermentação.

Se os microorganismos forem usados como organismos no processo de acordo com a invenção, eles são crescidos ou cultivados da maneira com qual o trabalhador versado está familiarizado, dependendo do organismo hospedeiro. Como uma regra, os microorganismos são crescidos em
30 um meio líquido que compreende uma fonte de carbono, normalmente na forma de açúcares, uma fonte de nitrogênio, normalmente na forma de fontes de nitrogênio orgânico tal como extrato de levedura ou sais tal como sul-

fato de amônio, elementos traço, tal como sais de ferro, manganês e magnésio e, se apropriado, vitaminas, a temperaturas de entre 0°C e 100°C, preferivelmente entre 10°C e 60°C, ao mesmo tempo em que passando em oxigênio. O pH do meio líquido pode ou ser qualquer um mantido constante, quer dizer regulado durante o período de cultivo, ou não. As culturas podem
5 ser cultivadas em bateladas, semibateladas ou continuamente. Os nutrientes podem ser fornecidos no começo da fermentação ou podem ser alimentados semicontinualmente ou continuamente. Os ácidos graxos poliinsaturados produzidos podem ser isolados dos organismos como descrito acima por
10 processos conhecidas pelo trabalhador versado, por exemplo, através de extração, destilação, cristalização, precipitação apropriada com sal, e/ou cromatografia. Para este fim, os organismos podem ser rompidos vantajosamente anteriormente.

Se os organismos hospedeiros forem microorganismos, o processo de acordo com a invenção é realizado vantajosamente a uma temperatura de entre 0°C e 95°C, preferivelmente entre 10°C e 85°C, especialmente preferivelmente entre 15°C e 75°C, muito especialmente preferivelmente entre 15°C e 45°C.

Neste processo, o valor de pH é mantido vantajosamente entre
20 pH 4 e 12, preferivelmente entre pH 6 e 9, especialmente preferivelmente entre pH 7 e 8. O processo de acordo com a invenção pode ser operado em bateladas, semibateladas ou continuamente. Uma avaliação sobre os métodos de cultivo conhecidos pode ser encontrada no livro de Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik [Bioprocess technology 1. Introduction to Bioprocess technology] (Gustav Fischer Verlag, Stuttgart, 1991)) ou no livro de Storhas (Bioreaktoren und periphere Einrichtungen [Bioreactors and peripheral equipment] (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)).
25

O meio de cultura a ser usado tem que satisfazer adequadamente
30 te às exigências das cepas em questão. As descrições dos meios de cultura para vários microorganismos podem ser encontradas no livro "Manual of Methods for General Bacteriology" of the American Society for Bacteriology

(Washington D. C, USA, 1981).

Como descrito acima, estes meios que normalmente podem ser empregados de acordo com a invenção compreendem uma ou mais fontes de carbono, fontes de nitrogênio, sais inorgânicos, vitaminas e/ou elementos traços.

As fontes de carbono preferidas são açúcares, tal como mono-, di- ou polissacarídeos. Os exemplos de fontes de carbono muito boas são glicose, frutose, manose, galactose, ribose, sorbose, ribulose, lactose, maltose, sacarose, rafinose, amido ou celulose. Os açúcares também podem ser adicionados aos meios por compostos complexos tal como melados ou outros subprodutos de refinação de açúcar. A adição de misturas de uma variedade de fontes de carbono também pode ser vantajosa. Outras possíveis fontes de carbono são óleos e gorduras tal como, por exemplo, óleo de soja, óleo de girassol, óleo de amendoim e/ou gordura de coco, ácidos graxos tal como, por exemplo, ácido palmítico, ácido esteárico e/ou ácido linoléico, poliálcoois e/ou álcoois tal como, por exemplo, glicerol, etanol e/ou metanol, e/ou ácidos orgânicos tal como, por exemplo, ácido láctico e/ou ácido acético.

As fontes de nitrogênio são compostos de nitrogênio normalmente orgânicos ou inorgânicos ou materiais compreendendo estes compostos. Os exemplos de fontes de nitrogênio compreendem amônio na forma líquida ou na forma gasosa ou sais de amônio tal como sulfato de amônio, cloreto de amônio, fosfato de amônio, carbonato de amônio ou nitrato de amônio, nitrato, uréia, aminoácidos ou fontes de nitrogênio complexas tal como licor macerado de milho, carne de soja, proteína de soja, extrato de levedura, extrato de carne e similares. As fontes de nitrogênio podem ser usadas individualmente ou como uma mistura.

Os compostos de sal inorgânicos que podem estar presentes nos meios compreendem os sais de cloreto, fósforo e sulfato de cálcio, magnésio, sódio, cobalto, molibdênio, potássio, manganês, zinco, cobre e ferro.

Os compostos contendo enxofre inorgânico tal como, por exemplo, sulfatos, sulfetos, ditionetos, tetracionatos, tiosulfatos, sulfetos, ou então

compostos de enxofre orgânico tal como mercaptanos e tióis, podem ser usados como fontes de enxofre para a produção de substâncias químicas finas contendo enxofre, em particular de metionina.

Ácido fosfórico, fosfato de diidrogênio de potássio ou fosfato de hidrogênio de dipotássio ou sais contendo sódio podem ser usados como fontes de fósforo.

Os agentes de quelação podem ser adicionados ao meio para manter os íons de metal na solução. Os agentes de quelação particularmente adequados incluem dihidroxifenóis, tal como catecol ou protocatecuato e ácidos orgânicos tal como ácido cítrico.

Os meios de fermentação usados de acordo com a invenção para cultivar microorganismos normalmente também compreendem outros fatores de crescimento, por exemplo, tal como vitaminas ou promotores de crescimento que incluem, biotina, riboflavina, tiamina, ácido fólico, ácido nicotínico, pantotenato e piridoxina. Os fatores de crescimento e sais frequentemente são derivados de componentes de meios complexos tal como extrato de levedura, melados, licor macerado de milho e similares. Também é possível adicionar precursores adequados ao meio de cultura. A composição exata dos compostos do meio pesadamente depende da experiência particular e é decidida individualmente para cada caso específico. Informações sobre a otimização do meio podem ser encontradas no livro "Applied Microbiol. Physiology, A Practical Approach" (Editors P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19 963577 3). Os meios de crescimento também podem ser obtidos de fornecedores comerciais, por exemplo Standard 1 (Merck) ou BHI (infusão de cérebro coração, DIFCO) e similares.

Todos os componentes do meio são esterilizados, ou através de calor (20 min a 1,5 bar e 121°C) ou através de esterilização de filtro. Os componentes ou podem ser esterilizados juntos ou, se preciso for, separadamente. Todos os componentes do meio podem estar presentes no começo do cultivo ou adicionados continuamente ou em batelada, como desejado.

A temperatura de cultura normalmente está entre 15°C e 45°C, preferivelmente de 25°C a 40°C, e pode ser mantida constante ou pode ser

alterada durante a experiência. O pH do meio deveria estar na faixa de 5 a 8,5, preferivelmente em cerca de 7,0. O pH para cultivo pode ser controlado durante cultivo adicionando-se compostos básicos tal como hidróxido de sódio, hidróxido de potássio, amônio e amônia aquosa ou compostos ácidos tal como ácido fosfórico ou ácido sulfúrico. A espumação pode ser controlada empregando antiespumantes tal como, por exemplo, ésteres de poliglicol de ácido graxo. Para manter a estabilidade de plasmídeos é possível adicionar ao meio, substâncias adequadas que têm um efeito seletivo, por exemplo, antibiótico. As condições aeróbias são mantidas introduzindo oxigênio ou misturas de gás contendo oxigênio tal como, por exemplo, ar ambiente na cultura. A temperatura da cultura normalmente é 20° a 40°C e preferivelmente 25°C a 40°C. A cultura é continuada até que formação do produto desejado esteja em um máximo. Este objetivo normalmente é alcançado dentro de 10 a 160 horas.

Os caldos de fermentação obtidos deste modo, em particular aqueles contendo ácidos graxos poliinsaturados, normalmente contêm uma massa seca de 7,5 a 25% em peso.

O caldo de fermentação pode então ser processado também. A biomassa pode, de acordo com a exigência, ser removida completamente ou parcialmente do caldo de fermentação através de métodos de separação tal como, por exemplo, centrifugação, filtração, decantação ou uma combinação destes métodos ou ser deixada completamente no referido caldo. É vantajoso processar a biomassa após sua separação.

Porém, o caldo de fermentação também pode ser engrossado ou pode ser concentrado sem separar as células, usando métodos conhecidos tal como, por exemplo, com a ajuda de um evaporador giratório, evaporador de película fina, evaporador de película de queda, por osmose inversa ou através de nanofiltração. Finalmente, este caldo de fermentação concentrado pode ser processado para obter os ácidos graxos presentes nele.

Os ácidos graxos obtidos no processo também são adequados como material de partida para a síntese química de produtos adicionais de interesse. Por exemplo, eles podem ser usados em combinação com um

outro ou sozinhos para a preparação de farmacêuticos, comestíveis, alimentos animais ou cosméticos. Todas as seqüências de ácido nucléico usadas no processo de acordo com a invenção são vantajosamente derivadas de um organismo eucariótico, tal como uma planta, um microorganismo ou um animal. As seqüências de ácido nucléico são derivadas preferivelmente da

5 A invenção será descrita agora em maiores detalhes com referência aos seguintes Exemplos e aos desenhos nos quais:

Figura 1 mostra várias séries de reação sintéticas para a biossíntese de ácidos graxo ω -6 e ω -3.

Figura 2 é um traço de cromatografia de gás que mostra a conversão de Δ 9, 12-18:2 (ácido linoléico) em Δ 11, 14-20:2 por expressão heteróloga da seqüência de Δ 9-elongase de *P. marinus* (SEQ ID NO: 1, resíduos 7668 a 9200) em levedura ou induzido por galactose (Figura 2A) ou glicose (Figura 2B).

Exemplo 1 – Clonagem de um elongase de FAE1 de *Perkinsus marinus*

Perkinsus marinus é um parasita de protozoário de ostra capaz de sintetizar ácidos graxos saturados e não saturados incluindo o ácido graxo essencial, ácido araquidônico [20:4(n-6)]. *P. marinus* emprega a série de reação delta-8 (Δ -8) dessaturase para sintetizar o ácido araquidônico.

Materiais e Métodos.

Crescimento e colheita de *P. marinus*

25 *Perkinsus marinus* meronts foram cultivados a 28°C em um meio preparado como descrito por La Peyre e outros (J Eukaryot Microbiol 1993;40:304-10) e conteve aminoácidos, nucleotídeos, carboidrato, e vitaminas, porém nenhum soro bovino fetal.

Manipulação de ácido nucléico e clonagem com base em PCR

30 O DNA foi extraído de células usando um minikit de DNA DNeasy (Qiagen). O DNA foi amplificado com iniciadores específicos para gene de delta5 dessaturase como segue: as reações foram aquecidas a 95°C durante 2

min seguido por 35 ciclos a 95°C durante 1 minutos, 2 minutos a 52 e 72°C durante 4 minutos, em seguida uma única etapa a 72°C durante 5 minutos. Os produtos de amplificação de PCR foram clonados em vetor de TOPO (Invitrogen) e verificados através de sequenciamento. O gen FAE elongase foi amplificado com iniciadores específicos de gene (Tabela I) designados para as extremidades 5' e 3' da região de codificação, com locais de restrição para facilitar a clonagem no vetor de levedura (Tabela I). Os iniciadores dianteiros para clonagem em vetor de expressão de levedura pYES2 (Invitrogen) foram projetados para conter um G/A na posição -3 e um G na posição +4 para melhorar a iniciação de translação em células eucarióticas.

Iniciadores de oligonucleotídeo usados neste estudo

As transcrições de *Perkinsus marinus* foram analisadas através de transcriptase reversa PCR (RT - PCR). O RNA total foi extraído de células que usam um minikit de planta RNeasy (Qiagen). Primeiro o cDNA do filamento foi sintetizado de RNA total que usa o *kit* de Amplificação de cDNA SMARTE RACE (BD-Clontech, Basingstoke, UK) de acordo com as instruções do fabricante. Os cDNAs de filamento único foram amplificados com os seguintes iniciadores.

FAEoperon dianteiro 5' - GGAATTCGAGGAGTAGGATCTTATCTGAGGATAGTC A-
 CACTAGTCGTACT-3' FAEoperon reverso 5'-CATCTGCGAATACTAACCATACATT

As reações foram aquecidas a 95°C durante 2 minutos seguidos por 30 ciclos a 94°C durante 30 segundos, 30 segundos em temperaturas que variam de 55 a 72 de acordo com o iniciador designado e 72°C durante 2 minutos, em seguida uma única etapa a 72°C durante 10 minutos. Os produtos de amplificação de PCR foram clonados em vetor TOPO (Invitrogen) e verificados através de sequenciamento. Surpreendentemente foi mostrado que as transcrições da $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase são todas encontradas no mesmo mRNA. Este é o primeiro exemplo mostrando os genes de PUFA a serem organizados em uma estrutura tipo operon.

Em uma investigação adicional foi analisada a especificidade da $\Delta 9$ -elongase. Para este propósito a sequência de codificação deste gene foi amplificada por RT-PCR como descrito acima usando os seguintes iniciado-

res.

Elo2For: 5' - ATGCAAGTTCCCGCGGAGCATCACTCC -3'

Elo2Rev: 5' - CGTTACGCATCAATATTATGCATAGCCAACC -3'

O produto de PCR amplificado foi então clonado em um vetor de pCRscript de acordo com as recomendações do fabricante (Stratagen). Em uma segunda etapa de PCR as seqüências modificadas para expressão de levedura foram introduzidas usando os seguintes iniciadores.

Expressão de levedura

Kpn Elo2 dianteira 5' - TTGGTACCATGGGATTTCCTGCGGAG -3'

10 Sac Elo1 Reversa 5' - GGGAGCTCTTACGCATCAATATTATGCATAGC-3'

A seqüência dos iniciadores é determinada na orientação 5' a 3'.

Os sítios de restrição usados para clonagem estão em negrito.

RESULTADOS

Isolamento de elongase de FAE1 de *P.marinus*

15 Usando dados publicamente disponíveis derivados de um projeto de sequenciamento de genoma de *P.marinus* realizado por TIGR (<http://www.tigr.org/tdb/e2k1/pmg/>) foi identificado um contig (1047306867) que mostrou homologia significativa a elongases conhecidas, com a seqüência alvo (designada Elo1 Para, SEQ ID NO: 9) consistindo em uma estrutura de leitura aberta de 511 resíduos e nenhum íntron. O aminoácido putativo derivado da seqüência designada é SEQ ID NO: 10.

Caracterização funcional em *Saccharomyces cerevisiae*

O cDNA de tamanho natural que corresponde a $\Delta 9$ elongase de ácido graxo putativo (SEQ ID NO: 9) foi clonado em vetor expressão de levedura pYES2 para determinar um construto pYPmFAE designada. A cepa de *S. cerevisiae* W303-1A foi transformada com o pYPmFAE ou o vetor vazio como um controle. As células transformadas foram crescidas em um meio mínimo que contém rafinose, e foram induzidas com 2% de galactose. Depois de 48 horas de crescimento os ácidos graxos de levedura totais foram extraídos e os FAMES resultantes analisados por GC.

A análise de GC (Figura 2) revelou que as células de levedura transformadas com pYPmFAE produziram um ácido graxo adicional que foi

identificado como ácido eicosadienóico que indica que o gene que foi clonado codificou um ácido graxo delta 9 elongase. A célula de levedura que expressa o ácido graxo delta 9 elongase de *P. marinus* é capaz de reconhecer o C18:2 (c9,12) substrato com uma porcentagem de 8,2% de taxa de conversão.

5

A tabela 1 mostra o conteúdo de ácido graxo das células de levedura após a transformação com pYPmFAE (+) ou com o vetor vazio pYES2 (-) e indução com 2% de galactose. A porcentagem de conversão em 18:2^{Δ9,12} a 20:2^{Δ11,14}, por exemplo, é calculado pela equação:

10

$$\% \text{ conversão} = \frac{[20:2^{\Delta 11,14}]}{[18:2^{\Delta 8,12}] + [20:2^{\Delta 11,14}]}$$

Tabela 1

	ácidos graxos											
	%	16:0	16:1 ^{Δ5}	18:0	18:1 ^{Δ5}	18:2 ^{Δ9,12}	20:2 ^{Δ11,14}	18:3 ^{Δ6,9,12}	20:3 ^{Δ5,11,14}	20:4 ^{Δ5,5,11,14}	22:4 ^{Δ1,10,12,16}	% con
FAE 18:2+	19,05	19,05	22,79	5,15	12,54	37,14	3,33	0,00	0,00	0,00	0,00	8,2
FAE 18:2-	20,81	20,81	19,50	5,50	12,11	42,07	0,00	0,00	0,00	0,00	0,00	0,00
FAE 18:3+	18,86	18,86	19,77	4,81	11,15	0,00	0,00	44,93	0,48	0,00	0,00	1,1
FAE 18:3-	20,35	20,35	18,15	4,84	10,33	0,00	0,00	46,27	0,07	0,00	0,00	0,1
FAE 20:4+	20,84	20,84	31,09	5,48	15,91	0,00	0,00	0,00	0,00	26,68	0,00	0,0
FAE 0:4-	22,13	22,13	31,00	4,55	14,65	0,00	0,00	0,00	0,00	27,67	0,00	0,0

Os resultados apresentados na Tabela 1 mostram que nenhuma atividade de elongase foi detectada com 20:4^{Δ5,8-11,14}, e uma atividade mínima (1% de conversão) em 18:3^{Δ6/Δ11,12}. Portanto parece que o ácido graxo Δ9 elongase é seletivo para ácido linoléico e não age para prolongar outros PUFAs.

Listagem de Sequência

<110> BASF

<120> Ácido Nucléico

<130> DY-03

<160> 10

<170> Patente em versão 3.1

<210> 1

<211> 18342

<212> DNA

<213> *Perkinsus marinus*

<220>

<221> CDS

<222> (7668) .. (9200)

<223>

<220>

<221> CDS

<222> (9351) .. (10724)

<223>

<220>

<221> CDS

<222> (10842) .. (12077)

<223>

<400> 1

tctaagacat catacctaca gtggcttcat ctactgtgat ggttatgatg atgacgacga
60

attcagagtt ggaaaaggaa tttgaggtca gcgatggtga gtttctcgag ggctccattg
120

aagttaccgc cgatagtgc atcgaagatc aattgcttct tctcctgtaa caggaaatga
180

tcatcactcg agtatccaag gaaaactgta tcgcccgggg acgttgatct ctattagctc
240

aggtttcgtg tagagttgag tcattagata aataaattcc atgaagccaa ggtagtgta
300

cctggagagc cttgattcgc tctcgtatgg tgttctcgga gataagatgg tgactcctca
360

caatctgctg ctgtcctaga cgggtggcaac gctgtatggc ctgcttctct acggccgggt
420

tccaccatgg atctatacag gagtatacca ccctacgtgg gccagcgcag tctcaccgag
480

tatgaagacg tgactcgccc tttggaggtt caacccttca ccaccagcgg aaatcgatac
540

cagcaggagg ctgcagtccg acgctggcga ctgatatagc agaccagggt caagatcaca
600

acacagacat ctcccacctc gatgaaagcc cggatgatcc tggtagcaac ggcgaggggt
660

atcgatccgt ggatgatctg aggacagtag aactcgacg agctactgta cgtattttca
720

tggttaacga taaccgagag atctgcctta cgaaagtgct aatccctctg gctttgaggt
780

ggtattgtga gagctccagc aacttgacaa agcaggagaa taccagcctg aacgaaggag
840

gactcagaga tccaacgc atgtactgctc ccgttgctaa aaatgagact atgctctctc
900

cggcggggcg tacgagaaga cttgttagac attacgaagt aagagactcc gcggaattga
960

aagtaattgc tccccataa tacgggatcg atgacctcac cctttagcat tgggatcctg
1020

tctacgcac tccaagactc tggatgatg ggcacgatc ttagacgatg atggtatacc
1080

agggccatgc atctcttgca gttgactgat gacgttatto cttcgggata agcttgaaag
1140

agacagtaaa gctctgattt cctctgagcg tggcatcgct gatgtagtag cagcagtcctg
1200

agcctaataa ggagaccagt tgatccacga atgccccca aagacctacc caggactggg
1260

gtagaaccga gctgctctca caaacgggac actgcacgca agcgctcgac aatgagctga
1320

tattcgtact agctgtaccc acacagtccc tacagccttt aatacgtgag aggggggaag
1380

cggctagcca acctgtggaa ctcatgaccg cagtccatcc ataccagatc atcttcccca
1440

taccccgatga tgatatcgtc gtcacatacg tctaccaat acagtgtgcc tcgcgttatg
1500

catcaacacc tacgcatttc ctcatcctt catcgtcacc agaattcacc aaatacttat
1560

ggtttgcagc ctgaaagcat cagtactgtc aaaattaggg agaccaccta ccaatctgag
1620

cttcatgata agcgagagca cctcaaccat atgctcgccg atttcatttc gtctcagata
1680

gccgcgtatt ctgtcctggg agagttcgaa aagattgtcg tagacctcac ggtcctttgg
1740

tttcatcgaa agtgtatgca gtgcactcgt taattctggg aggcggactt ccatctgctt
1800

ctccgctttc gtgcgectca gcatgagacg cttgagaacc ctcgctcctga gcagctccaa
1860

cccctcgcgt ccttcaaacy catcgaagcc gtggctgatg cgaactccat cgtctggtaa
1920

agcaccgcgc caaaagggtg cctcttgatg ggattggcaa tgaatcgtgt aaagtaaag
1980

tagtggctaa ttctgggatg cccgcaggcc ttgcagttct ctatggggtg gtccagcaac
2040

acgcagtcac aaccgcacca gtcgcacatc ctagcgccct ccggcagaac ccggaggaat
2100

ctatgggcac aacaatgatg aatttcggtc acgactgtca ccttatcagc gaaaaaaggt
2160

ctccaacttt gttttggaga ggagttcccg acaggcacca tcggaacctc gaccgcaagg
2220

cacacctgag catcgatgtt gacacgaagg ggacggaccg ggaggattct tacgcggcct

2280

tagcggtcga ggacccggct gccttgatct tgtgggcctc gtctaagact actctctccc
2340

agcgcaaacc atgcaatagc cctgcacggc gcaatcgggt ctccaagtcc gacctgtcag
2400

tccctggaag gaaagaacga tcctcgatac ctgaggagac cccaacagag cccacactgt
2460

ctcgtcgcta ctgtcctcga cagctgcttg cgctaggcga accctccggg gtttaaccgc
2520

acgcctcaca atctccgctg ccttacgacg ggccccaccc atatcttgcc cacacgatcg
2580

aggcatcccc atagtagccc ctttccatcc cacttcgcta ctgctcatgt tggcagtaga
2640

gagtttcggg agaggactgc attgcggtga tggtttctgg taagatcgga ccggtccggg
2700

gaactccac gcattatcat cgacgttctc tttgttccct gccatttctc cagccttgat
2760

acgtttcccc aagcccagca gaggactctc ctccaggcgc cttccgggca gtcctccct
2820

ctttaogaga ggactataca tottctact gacaggtaca gtcggaccat cgagcactgc
2880

tgaactcccg gtcacaccag gtctgtgtag accattactg aagacgagac cttcttgagt
2940

gataggccca acactcgacc gccattcatg tcgatcggag tcccacgagg ctccgttctc
3000

caccttgata taacctagat tgctgcactc cctcagaggg ctgctgggct tcaatcggac
3060

caccataggc tttgcctctt cctccacgga aatgctcgca ccactcggtc ggccgtgacc
3120

aaggacatga tatcgggtgt ggtttaaaga cgagctgata ctgccagccg ctgccgagga
3180

gggtccacg agactgtccc ttgtcgtcca ctgttctctc ttaaaacctc ctgtatgatc
3240

gatacctctc gaagcaacat gccaccaat agtgcgttgt aaactagcag aaatagagag
3300

gttgatgcgc aacgggtctg tggaagggcc tatgctcggg ctcaaggcac cactagagcc
3360

acccacggac ggtcttgagt acggtgtagt tctccaagtc ggcttatage agttgtctat
3420

ggagtctctg ccctgctcag ctttcggccg taatctcttg agcatgtcga ctagatccct
3480

cagtgtcttc tggatcatcat cgacatcttc ctgtttgacg agatggctct cccgcactct
3540

cgagccggct tgaccagtga agaagtcctt gatcgtctcc atggatcccg gattacgata
3500

ttgctgccga ttagaattaa ccgagacatt agaaccatac gaaggcgtc tgacctccct
3660

cgaacctccg ccggccaaaa ctgaacgggt ctcttcttgg aagagattct tccgcaatga
3720

ctggcctgtc ggaggagccc catcctgatg atggtaggtg gatgatgatg acgatgggtc
3780

gctcggctga tacactggtg cactgggatg ggctcggccg ctggcgtcat aggtccaagt
3840

ggtacgaggc ctcttcgggc acttctcttt atgcgagaag agcagttctg ggagtatctt
3900

gcgcttgacg tagggacacg cctttttcag ctgatcgagc tcctttctgt actcggcctc
3960

cagagttctg taagtcgtca gaacgatatc atatctcttc aaggacttag cagtgatctt
4020

agattttcca tggtagagga atattttcac agtgccaggc cgtaggtatc tattgatctc
4080

ggcctccac tggagcatcg aggtagtggc gcatacgatc agagtggggt tattcctccc
4140

cacaagtttc ttagtaatca acgagaccgt ttgtatggtc ttcccatgc ccctctcatc
4200

agccaatata ccccccttac agtcctcctt ggtctcctgt tcacacatcc aagcgagacc
4260

ttcacgttgg tatcccagca gcggaattat taactgaaaa taaccaagca atcaatcctc
4320

agagtcggg agccgcacgg tcatccgaag cattccttat atattataag ctgcccagcc
4380

tgctctggag ggggatactt gagcaccaac tcgggaggta tcttctggac catcctcccc
4440

agctcccggt tcacctggc ttcttcagcc ggctgggtat gtgtgggatt gggatcggcg
4500

ttctcgtctc tgatgaacgc gttagggtag attaatgggt aaagttggaa ctccacgaaa
4560

tcaattctga aggatcggag acattagttc gacctgtacc aaatgaccta aaactcaagt
4620

ccattcttcc aacgaaaaac ctccgaagat catggcgccct tcggtagtta acatgtcaac
4680

gtagcgtgtg ttgagaaaaa atacaacagc aatacaagat cgaaagccag cctggtagct
4740

aggatttgaa cctgatgtct tctcctgagt agcaattcag attctagtaa tgtctatcat
4800

actccaacac atgggcctct ctttttaccc tataatgatg agcaaatatg caaaaccaca
4860

aggggcgcta cttgacccta ccgogaccat tcgctctcct gtgggtatcat ggatcaccat
4920

ctccatgatc tgatcaacag ttcgtcgtga acgatctgca gtagggttga agacgtagtg
4980

atgaagctgg cattcttcaa tctatgcaga tacatcctca atagtaggga acaccagagg
5040

tcaattaaat gccacaatt tcaccagaaa gctgagaacc tcagtccac gaagtgttg
5100

ctagtcatca gaccgagagt taaacgccga gcctattcgg aaaaccgact atgatgagac
5160

tctcatcgtg tcatattgta tgccatcatc ataactctgat ccaatcacgt catggtatca
5220

gatggcgagt gactcgccgc gagtcagtac ctgaaagtgc atcctatttc cacggcgccg
5280

gcatgctcat cagcctgtct cttggcagtc gtacgcatgc gtcagcctat tgaatgccac
5340

cacagactca agtcaatgac tcttaacca gagattaacg aaagtccaga gcatatgtgt
5400

ggacctgtga agaggaagca atgatgagag aatactctgt tcgggcctgt ataagaacac
5460

actgaatagc agagtggaaa tgctgagatc actctggcac acacggagtg agtcggttca
5520

ggtcagctac gaattatttt ccatttcacc cagctgtatg cccgagtata acaaatttct
5580

tcactggggc aaatatgcc cgtcaagata tgccaaagtg gggccactat gccacaccaa
5640

gatctgcttg agtatgatat aaatttcttg tctggggcaa atatgccaca ccaagatatg
5700

ccaaagtggg gccactatgc cacaccaaga tctgcttgag tatgatataa atttcttggtg
5760

tagggcaaat atgccacgcc aagatatgcc aaagtggggc cactatgcc caccaagttc

5820

tgcttgagca caatgaattt ctggtgtggg actcatatac cagccaaaa tacgcctaag
5880

cggggtaact atgccacgcc aaaatatgcc agagtagggc aagtatgaca cgccaagata
5940

tgctcaagta aaatgaattt cttcactggg gcaaatatgc catgccaaga tatgcctgag
6000

tataaacaca caatttcttc agtggggcag acatgccaaa ccaagatctg ctcgagtaca
6060

atgaaactggt gtgggtcaaa tataccaagt caaaatatgc caaatcgggg taactatgcc
6120

acgtcaacat aggccagagt ggggcaacta ttatgcgtta agatatgctt aagtacaata
6180

tatttcattt ttctactggg acaaatatgc cagccaaga tgtgtccgag tatgaaaaaa
6240

tcttcagtga gacacctatg ccaggacaag atatatcgga gaacaataaa ttctgtaagt
6300

gggataaata tgccacgccg agatatacct gagtggggca aatatgccaa gccgagatat
6360

gctcgagtac agtgaacttc ttcagttagg ctgatatgcc cgccaaaaca tgcttgggtg
6420

tacgaaatta ttcagtga aaatttctt cagtgggggt acccacttat cttggcgtgc
6480

cataggtagc ccactggggg aaccatgcc cccaagata tgcccgagta caacaaaatt
6540

attcggtgag gcagatatac caccocgata catcactgaa aacaacaaat ttccgccagt
6600

gggcagctat gccatcccaa gaaatgccc agtacacgca atctcccgag agcgcgacct
6660

aggcatcca ccttggtcgg aaaaacaaac gaaacacaat ccggcttga cgaattgcgg
6720

aagataaatg cgcatagtaa agtcgactcc gttcgagagt agacatgcc tttccgagct
6780

taactgcgac atgctatctt ctcaggagag actctgacga gtatttctt ggtaatttcc
6840

agtggagagt ggacactgaa cacagacctc tgaccccgcg cagtactcac aagatgaagg
6900

tccgactgag aaaagattga tcttactaac tgtgaagcag acatacgaaa ggtgtattcc
6960

aatcttgctg ttgttggtat tgttgtagt ctccgtgtgg gactgaaccg catcccattg
7020

tcgagtgaat cgagcggcgc accgaaagat gagcaatact cagactgatg tgcattgagt
7080

caccatagaa atacgaaaca gcgtgatagt atcagcattt tcttgctgt ctgattcggg
7140

atttggtgac atgcaccatc cagcattctc cagaggggag cggcacgcac ggcaaaattg
7200

tcaaattcaa tactattgag agcaccaa atcattcgagt ggtcgcgcac tagcgttaact
7260

ctaattgctg atcaacccca ataaccctc ggcgctcta tacgggccct caaatactgg
7320

tgagatgcaa aatcaagcag tttaggtcgg aaattgtcgt ggttcgtcaa aactcaattc
7380

atcgagtccc tcgggtcatt cgtcgcgcgc gttgaagacg aacggcgcac caatcatttg
7440

cgcacgcgtg caaaacgctc attgcgcagg cctacattcc agaattgagg gccacgggat
7500

gtgctgcact aacatcgcca gactctacac tcgtagcata atatattgaa ttcgaggagt
7560

aggatcttat ctgaggatag tcacactagt cgtactgctg aatcagcgtt gctgattctg
7620

caacggttat actattaaga cgatcggttag gacgtcatta cgtcgtc atg caa gtt
7676

Met Gln Val
1

ccc gcg gag cat cac tcc act cgg gtt ata tct atc tgt gat atc gtc
7724

Pro Ala Glu His His Ser Thr Arg Val Ile Ser Ile Cys Asp Ile Val
5 10 15

atc tca ggc ccc ttt gga atg tgt aac cat gat tac tcc gct tct ata
7772

Ile Ser Gly Pro Phe Gly Met Cys Asn His Asp Tyr Ser Ala Ser Ile
20 25 30 35

cct gcc tct agt agt ggt agc act cga cgt atg cgt ttg gta gcc tac
7820

Pro Ala Ser Ser Ser Gly Ser Thr Arg Arg Met Arg Leu Val Ala Tyr
40 45 50

atc aca ttg gtc tct atc cac tat caa cag cta ctc ttt tac tct tct
7868

Ile Thr Leu Val Ser Ile His Tyr Gln Gln Leu Leu Phe Tyr Ser Ser
55 60 65

atc ata act ata atc act ggc tat cac tac tat gtc gca gct ctg ccc

7916
Ile Ile Thr Ile Ile Thr Gly Tyr His Tyr Tyr Val Ala Ala Leu Pro
70 75 80

ctt tac gac atc tca att gct cta tct gtg ctt tgc gga cta aca cta
7964
Leu Tyr Asp Ile Ser Ile Ala Leu Ser Val Leu Ser Gly Leu Thr Leu
85 90 95

ctg tgg tta tgt aat tgc tat tac aac agc aag ccc aat gta ttc tgc
8012
Leu Trp Leu Cys Asn Cys Tyr Tyr Asn Ser Lys Pro Asn Val Phe Cys
100 105 110 115

atc gat cat gtt gag ttt gac gct cct ccc tct tgg aag gtc agt cat
8060
Ile Asp His Val Glu Phe Asp Ala Pro Pro Ser Trp Lys Val Ser His
120 125 130

gaa gac atc atc aac att gcc aag ata caa ggt tgc tac acg gaa gat
8108
Glu Asp Ile Ile Asn Ile Ala Lys Ile Gln Gly Cys Tyr Thr Glu Asp
135 140 145

tca ctc aac ttc atg cag cgt ctt ctc gag agg tct ggt act tgc cct
8156
Ser Leu Asn Phe Met Gln Arg Leu Leu Glu Arg Ser Gly Thr Cys Pro
150 155 160

gat aag agt gct gct tac cct cca gtg gtt gtt gag tca ctg agg act
8204
Asp Lys Ser Ala Ala Tyr Pro Pro Val Val Val Glu Ser Leu Arg Thr
165 170 175

aac gcc ccc gcc gat gcc tct gct gtc aat act aga gag gaa gcg agg
8252
Asn Ala Pro Ala Asp Ala Ser Ala Val Asn Thr Arg Glu Glu Ala Arg
180 185 190 195

gaa gtg atc ata act acg gtc aaa gat ctg ctc aag aaa act ggt gtg
8300
Glu Val Ile Ile Thr Thr Val Lys Asp Leu Leu Lys Lys Thr Gly Val
200 205 210

cat cct aaa tct atc gac tac atc atc gtc aat tgc gcc atg tac aac
8348
His Pro Lys Ser Ile Asp Tyr Ile Ile Val Asn Cys Ala Met Tyr Asn
215 220 225

ccg aca ccg tca cat gct gct atg ata gtg aat gaa gtc ggt atg agg
8396
Pro Thr Pro Ser His Ala Ala Met Ile Val Asn Glu Val Gly Met Arg
230 235 240

aat gac gtt atc acc tat aac ctc agt ggt atg ggg tgt agt gcc ggt
8444
Asn Asp Val Ile Thr Tyr Asn Leu Ser Gly Met Gly Cys Ser Ala Gly
245 250 255

gtt atc aca att gat cta gca acg cgt ctg ttg aga gag acc aga ggt
8492

Val Ile Thr Ile Asp Leu Ala Thr Arg Leu Leu Arg Glu Thr Arg Gly
260 265 270 275

agg gca ctg att gtc tca act gag ata cta act cgt tgc ttc tat cgt
8540

Arg Ala Leu Ile Val Ser Thr Glu Ile Leu Thr Arg Cys Phe Tyr Arg
280 285 290

ggg aat gat cgt gaa cca ctg atg ggt aac aca tta ttc aga tgt ggt
8588

Gly Asn Asp Arg Glu Pro Leu Met Gly Asn Thr Leu Phe Arg Cys Gly
295 300 305

ggg gct gct gct ttg cta tct tca ttg cct aaa gac cta tct cgt gcc
8636

Gly Ala Ala Ala Leu Leu Ser Ser Leu Pro Lys Asp Leu Ser Arg Ala
310 315 320

aaa tat aag ttg tta cat acc gta aga acg caa gtt ctc ggt aat gag
8684

Lys Tyr Lys Leu Leu His Thr Val Arg Thr Gln Val Leu Gly Asn Glu
325 330 335

agt ttt gaa acg att atg gag act gat gac agt acc aag ccc aac agt
8732

Ser Phe Glu Thr Ile Met Glu Thr Asp Asp Ser Thr Lys Pro Asn Ser
340 345 350 355

att gtt aca cta agg ctc cag aag agc atc atc aaa gtt gct gct gtt
8780

Ile Val Thr Leu Arg Leu Gln Lys Ser Ile Ile Lys Val Ala Ala Val
360 365 370

gct att aaa caa aat ttt act aag ctt gct tat atg gtt ctc cct ctg
8828

Ala Ile Lys Gln Asn Phe Thr Lys Leu Ala Tyr Met Val Leu Pro Leu
375 380 385

aga gaa ctg ttg aag gtc tta tac tct atg gtg acg atg aaa atg aga
8876

Arg Glu Leu Leu Lys Val Leu Tyr Ser Met Val Thr Met Lys Met Arg
390 395 400

aga aag tct tca aaa gaa ggt cgc gag ttg tac gta cct gat ttt aga
8924

Arg Lys Ser Ser Lys Glu Gly Arg Glu Leu Tyr Val Pro Asp Phe Arg
405 410 415

aag ggt act gat cat tgg tgt att cat gct ggt ggc cgt ggt gta ttg
8972

Lys Gly Thr Asp His Trp Cys Ile His Ala Gly Gly Arg Gly Val Leu
420 425 430 435

gat acc tta caa gat tct ctc cag ctg tca gac tac gat atc caa gca
9020

Asp Thr Leu Gln Asp Ser Leu Gln Leu Ser Asp Tyr Asp Ile Gln Ala
440 445 450

agc cgt agt gtt ctc tat gag aga ggc aac acc agt agc agt agc ata
9068

Ser Arg Ser Val Leu Tyr Glu Arg Gly Asn Thr Ser Ser Ser Ser Ile
455 460 465

tgg tat gag ttg gca tgg ctc gaa cgt gac caa cgt att aag cgt gga
9116

Trp Tyr Glu Leu Ala Trp Leu Glu Arg Asp Gln Arg Ile Lys Arg Gly
470 475 480

gat agg gta tta cag ttg gct ttt ggt agt ggt ttc aaa tgt aac tca
9164

Asp Arg Val Leu Gln Leu Ala Phe Gly Ser Gly Phe Lys Cys Asn Ser
485 490 495

tca gta tgg ttg gct atg cat aat att gat gcg taa cgacaatcag
9210

Ser Val Trp Leu Ala Met His Asn Ile Asp Ala
500 505 510

tttttctcac tatgagttgg ctccaccgta atcaatggcc atcatctcct ttcttagtta
9270

ttatogatga ttatagtcag tgccgatgtg tgctagtgtt ttactcttta tcaacttgtg
9330

agtttcaggc ccccttttccc atg act act tca acc act act gtg caa cta caa
9383

Met Thr Thr Ser Thr Thr Thr Val Gln Leu Gln
515 520

gaa gac ctg tca agt ggt gac cag aac gcc cac ccc agt cca agc cga
9431

Glu Asp Leu Ser Ser Gly Asp Gln Asn Ala His Pro Ser Pro Ser Arg
525 530 535

gct act cct agt gtt ggt gat act aag gag gat gcg agg gtt gtg atc
9479

Ala Thr Pro Ser Val Gly Asp Thr Lys Glu Asp Ala Arg Val Val Ile
540 545 550

aaa cta ttt ggt aca tgg gtt gat gtt aca gct tgg ttg aat gac cat
9527

Lys Leu Phe Gly Thr Trp Val Asp Val Thr Ala Trp Leu Asn Asp His
555 560 565

cct ggt ggt tct aaa gtg ctc aga gca ttc aac aag aag gac gcg act
9575

Pro Gly Gly Ser Lys Val Leu Arg Ala Phe Asn Lys Lys Asp Ala Thr
570 575 580 585

gat gct gtt atg gcc atg cac act gat gaa gct atc aag cgc atc atc
9623

Asp Ala Val Met Ala Met His Thr Asp Glu Ala Ile Lys Arg Ile Ile
590 595 600

aga ttt tca aat gtg gtc tcc tgg gcc ccc atc aac gcc tct att ggt
9671

Arg Phe Ser Asn Val Val Ser Ser Ala Pro Ile Asn Ala Ser Ile Gly
 605 610 615
 gat gtc cag gtt att gag aaa tct cta tcg aga gaa cag ttg atg tat
 9719
 Asp Val Gln Val Ile Glu Lys Ser Leu Ser Arg Glu Gln Leu Met Tyr
 620 625 630
 tac aag ctc cgc act ctt gct aga aac cag ggc tgg ttt caa agc aat
 9767
 Tyr Lys Leu Arg Thr Leu Ala Arg Asn Gln Gly Trp Phe Gln Ser Asn
 635 640 645
 cta tta tat gaa gga gtg aaa gca atg ata gcc ttc ggt ttg ctc atc
 9815
 Leu Leu Tyr Glu Gly Val Lys Ala Met Ile Ala Phe Gly Leu Leu Ile
 650 655 660 665
 atc ggg ttt gct act ctc tac ttt gac tat ggt att tgg tca acc gca
 9863
 Ile Gly Phe Ala Thr Leu Tyr Phe Asp Tyr Gly Ile Trp Ser Thr Ala
 670 675 680
 ctg ata ggt ttc gct tgg ttt cag ctg ggg tgg ttg gga cat gac tgg
 9911
 Leu Ile Gly Phe Ala Trp Phe Gln Leu Gly Trp Leu Gly His Asp Trp
 685 690 695
 tca cat cat aca gct cta cct aag tct act act aac tgt gcg aac tac
 9959
 Ser His His Thr Ala Leu Pro Lys Ser Thr Thr Asn Cys Ala Asn Tyr
 700 705 710
 aac gac tat ctt ggc tgg ctt acc ggt ttg gcc aga ggg aat aca ctt
 10007
 Asn Asp Tyr Leu Gly Trp Leu Thr Gly Leu Ala Arg Gly Asn Thr Leu
 715 720 725
 ctg tgg tgg aaa ctg agg cac aat act cat cac gtg ctg acc aat cag
 10055
 Leu Trp Trp Lys Leu Arg His Asn Thr His His Val Leu Thr Asn Gln
 730 735 740 745
 tac gag aat gat cct gat att cta act caa cca ccg ttg cat ttt ttc
 10103
 Tyr Glu Asn Asp Pro Asp Ile Leu Thr Gln Pro Pro Leu His Phe Phe
 750 755 760
 gag gac ttc gat gtt ggt aat gtg aac aga tat caa gct gtc tac tac
 10151
 Glu Asp Phe Asp Val Gly Asn Val Asn Arg Tyr Gln Ala Val Tyr Tyr
 765 770 775
 cta cca atg tta act cta ctg cat cta ttt tgg ttg tac gag tcg gta
 10199
 Leu Pro Met Leu Thr Leu Leu His Leu Phe Trp Leu Tyr Glu Ser Val
 780 785 790
 ttg gtt tgc ttg aga caa agt agg tct att aat aga tac aac cgt atg

10247
 Leu Val Cys Leu Arg Gln Ser Arg Ser Ile Asn Arg Tyr Asn Arg Met
 795 800 805

 cat gcc agg agg gat acc gta gct ttg gta ctt cac ata ctc att gtt
 10295
 His Ala Arg Arg Asp Thr Val Ala Leu Val Leu His Ile Leu Ile Val
 810 815 820 825

 gcc atc ata tcg tac acc agt ggt aag tat ttg ctc atc ctt ctg gcc
 10343
 Gly Ile Ile Ser Tyr Thr Ser Gly Lys Tyr Leu Leu Ile Leu Leu Ala
 830 835 840

 tac atg ctt agt gcc ttt cta act gct gtt gtt gta ttt gcc agc cac
 10391
 Tyr Met Leu Ser Gly Phe Leu Thr Ala Val Val Val Phe Ala Ser His
 845 850 855

 tac aac gag cct agg gta gct tct ggt gaa tcc tta tca ctc gtt cgt
 10439
 Tyr Asn Glu Pro Arg Val Ala Ser Gly Glu Ser Leu Ser Leu Val Arg
 860 865 870

 cag aca ttg tta acc act atc aat ata gcc tca ttc agt gat act cat
 10487
 Gln Thr Leu Leu Thr Thr Ile Asn Ile Gly Ser Phe Ser Asp Thr His
 875 880 885

 tgg gag aag aag ttg tgg ttc tat cta act ggt ggt ctt aat atg caa
 10535
 Trp Glu Lys Lys Leu Trp Phe Tyr Leu Thr Gly Gly Leu Asn Met Gln
 890 895 900 905

 atc gag cat cat ctc ttc cca aca atg ccc cgc cat aat ctt ccg aag
 10583
 Ile Glu His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Pro Lys
 910 915 920

 aca act ttt ctg gtc aag tca cta gcc cag gag cta gga ctg cca tac
 10631
 Thr Thr Phe Leu Val Lys Ser Leu Ala Gln Glu Leu Gly Leu Pro Tyr
 925 930 935

 aag gaa acc aac att gtc agt tta acc aag gcg gcc gtt act act ttg
 10679
 Lys Glu Thr Asn Ile Val Ser Leu Thr Lys Ala Ala Val Thr Thr Leu
 940 945 950

 cat cat aat gct ctg cgt aac atc gag aga ttg ctt gct agg tag
 10724
 His His Asn Ala Leu Arg Asn Ile Glu Arg Leu Leu Ala Arg
 955 960 965

 ttctcatcat tgcaaccgca acaagaacat ggtcaactcg tagtggtggt agaggattgt
 10784

 cgtcatggtc agtagctttc tgatagaatc tttttcatct tcttctgtt gtgttga
 10841

atg tct tct ctt acc ctc tac aga ggc ccc ttt tcc cga atg gtg ctc
10889

Met Ser Ser Leu Thr Leu Tyr Arg Gly Pro Phe Ser Arg Met Val Leu
970 975 980

cct cgt cag gaa atc tgc atc aat ggt cgc ata tac gat gtc act gag
10937

Pro Arg Gln Glu Ile Cys Ile Asn Gly Arg Ile Tyr Asp Val Thr Glu
985 990 995

ttc atc aat cgt cat cca ggt ggt aag att atc ctc ttc caa gtt
10982

Phe Ile Asn Arg His Pro Gly Gly Lys Ile Ile Leu Phe Gln Val
1000 1005 1010

ggt gct gat gcc act gat gct ttt cgt gag ttt cat gct ggc agt
11027

Gly Ala Asp Ala Thr Asp Ala Phe Arg Glu Phe His Ala Gly Ser
1015 1020 1025

gag aag gca gag aag atc ctc aaa acc cta cca tcc cgt gat gat
11072

Glu Lys Ala Glu Lys Ile Leu Lys Thr Leu Pro Ser Arg Asp Asp
1030 1035 1040

gac ggt act ttc ctt cct tca acc caa cgc tcc atc atg gat gat
11117

Asp Gly Thr Phe Leu Pro Ser Thr Gln Arg Ser Ile Met Asp Asp
1045 1050 1055

ttc aaa cgc cta aga gat gac ctc gtc agc aga ggt gtc ttc aag
11162

Phe Lys Arg Leu Arg Asp Asp Leu Val Ser Arg Gly Val Phe Lys
1060 1065 1070

cca agc gtc atg cat gtt gta tac cgc tgc ttg gaa gtc gtt gct
11207

Pro Ser Val Met His Val Val Tyr Arg Cys Leu Glu Val Val Ala
1075 1080 1085

ctc tat ctc att ggc ttc tat ttg gct ctg tgc acc agt aat gtg
11252

Leu Tyr Leu Ile Gly Phe Tyr Leu Ala Leu Cys Thr Ser Asn Val
1090 1095 1100

tac gtt ggg tgt gct gta ctt ggt gta gct caa ggt cgt gct ggt
11297

Tyr Val Gly Cys Ala Val Leu Gly Val Ala Gln Gly Arg Ala Gly
1105 1110 1115

tgg ttg atg cat gaa gga ggt cat cac tct ctg act ggt aac tgg
11342

Trp Leu Met His Glu Gly Gly His His Ser Leu Thr Gly Asn Trp
1120 1125 1130

aaa gtt gac cag ttc ctc caa gaa cta ttt ttc ggc att ggt tgt
11387

Lys Val Asp Gln Phe Leu Gln Glu Leu Phe Phe Gly Ile Gly Cys

1135		1140		1145
ggt atg tca gct gcg tgg	tgg cgc aat gca cac	aac aag cat cac		
11432				
Gly Met Ser Ala Ala Trp	Trp Arg Asn Ala His	Asn Lys His His		
1150	1155	1160		
gct gct cct cag cat tta	ggg aaa gat gtt gat	ctc gag aca ttg		
11477				
Ala Ala Pro Gln His Leu	Gly Lys Asp Val Asp	Leu Glu Thr Leu		
1165	1170	1175		
cct ctg gtc gcc ttc aat	aag gcc gta ctt cga	ggc cgt cta ccg		
11522				
Pro Leu Val Ala Phe Asn	Lys Ala Val Leu Arg	Gly Arg Leu Pro		
1180	1185	1190		
tct gtc tgg atc aga tca	caa gct gtg tgc ttt	gca ccg ata tca		
11567				
Ser Val Trp Ile Arg Ser	Gln Ala Val Cys Phe	Ala Pro Ile Ser		
1195	1200	1205		
aca cta ctg gta tgc ttc	ttt tgg caa ttc tac	cta cac ccg agg		
11612				
Thr Leu Leu Val Ser Phe	Phe Trp Gln Phe Tyr	Leu His Pro Arg		
1210	1215	1220		
cat att att agg aca ggt	cga cga atg gag tct	ttc tgg cta ctc		
11657				
His Ile Ile Arg Thr Gly	Arg Arg Met Glu Ser	Phe Trp Leu Leu		
1225	1230	1235		
gta cgc tac tta gtt att	gtg tac ctc ggg ttc	agc tat gga ttg		
11702				
Val Arg Tyr Leu Val Ile	Val Tyr Leu Gly Phe	Ser Tyr Gly Leu		
1240	1245	1250		
gta tgc gtc ttg tta tgt	tac atc gca agt gtg	cat gtt ggt ggt		
11747				
Val Ser Val Leu Leu Cys	Tyr Ile Ala Ser Val	His Val Gly Gly		
1255	1260	1265		
atg tac atc ttt gta cac	ttc gct cta tca cat	aca cat tta cct		
11792				
Met Tyr Ile Phe Val His	Phe Ala Leu Ser His	Thr His Leu Pro		
1270	1275	1280		
gtc att aac cag cat ggt	aga gct aac tgg ttg	gaa tac gca tct		
11837				
Val Ile Asn Gln His Gly	Arg Ala Asn Trp Leu	Glu Tyr Ala Ser		
1285	1290	1295		
aag cac aca gtt aat gtg	tca act aac aat tat	ttc gtc aca tgg		
11882				
Lys His Thr Val Asn Val	Ser Thr Asn Asn Tyr	Phe Val Thr Trp		
1300	1305	1310		
ctc atg agt tat ttg aat	tat caa ata gag cat	cat ctc ttc ccg		
11927				

Leu Met Ser Tyr Leu Asn Tyr Gln Ile Glu His His Leu Phe Pro
 1315 1320 1325
 tca tgt ccc cag ttt aga ttc cct ggt tac gtc agt atg agg gtt
 11972
 Ser Cys Pro Gln Phe Arg Phe Pro Gly Tyr Val Ser Met Arg Val
 1330 1335 1340
 cga gaa ttt ttt cat aag cat gga ttg aag tat aac gag gtc ggc
 12017
 Arg Glu Phe Phe His Lys His Gly Leu Lys Tyr Asn Glu Val Gly
 1345 1350 1355
 tat cta cat gca ctc aat ctc aca ttt tca aat ctg gct gct gtt
 12062
 Tyr Leu His Ala Leu Asn Leu Thr Phe Ser Asn Leu Ala Ala Val
 1360 1365 1370
 gcc ata gtg gaa tag ctcaatagtg taggagagca tagtagacgc ttatgattgg
 12117
 Ala Ile Val Glu
 1375
 tatgtaaaat caataatggt tgggtgagag tgatcaatgt tttctaccag cggagcacta
 12177
 ataattctgca gcataatacca taacatatcg aaaactctgc atttcttata gttctcattc
 12237
 ctagtaagga ccacttcgtc ctagtaatta tcgccaaca cgggaggagg ttagttcgat
 12297
 gacaagggtg agtatgcctg ctgcaattac cgtgtctgga ctgaaacgta tcacccctat
 12357
 cgttacaact cattccccgc actctgggtgc ataatatgcg cgattcagca ccctttattt
 12417
 ctccccctca acgacccctt ctccgcaca ccgtccttag cagggtcaaat ctaccctgca
 12477
 attatcgtga taccaccctt gagttgcca actaatata ctcaactaca ataagtacac
 12537
 tgcgaatgca tcattaacca tcaacaatat gtcattaaa taataatcta tcatcttcgg
 12597
 acatctcctg ggtgctctcc tcttcactac cggggccagt ctcaatcgat gacctacta
 12657
 tgggcaactc tgacgatgac caagatgta aaacgtgaag ctgtgaatcc aagcactcac
 12717
 tcggatcgag aggtcagtt cctctggacc gaactgggt gttcaacatt ccaccagacc
 12777
 ggctatctct catccatagc agaatgcca ccctatcga tgctgtacog atgaagaaga
 12837

agaacttctc tacggctacg taggagccag tggatgaacc tgcgccgag accaatgggg
12897

acaggaacag tggagcagta tgtaggcgtg gacctcccat attaatgcgt accatcatat
12957

gagagttttg ctatgcaagg cttacacacg gattcgaata taagtgtaga aaaaccaatg
13017

aaattttccc gattttgctc agaatttcca tattgtcagc aacagatgaa gttgaccatg
13077

tgcattatct tccatacagg gtacagattc tacggttgca ctatacgata acattggggg
13137

gtgatacgac aacactgagc aattgatagg atggaccccc tgagtcgggg gccctgagca
13197

aagtgccttag atctagagca cagtattgaa gttgaattta cactgaggat tgcataccta
13257

tcaagttgcc accatacacc tcgatcagcc tttgtccact tacgtatatg agccactata
13317

acaggcgtaa ccgccaagcc gaagttttgt gcagccgttg taacaccata ggccgcgcct
13377

aacagacaaa ccggagtcac aaccatcgat gcagcgatga gtccctgggt ctatccccgg
13437

caacatcatt ttaatgtoga aaactcaagt tctacgatga ttgccaatca cacataatgc
13497

agcttggtta gagatcaagc cagaaattct ccacagcaac ctaccagtct tccataggagg
13557

tactgcagct gctaccaacg gcatactgc agaggcgaac accgtatagg atagccccaa
13617

aacggtcata gatatacct acacaacaac taaacttaga acgaataaag cctcgtaccg
13677

ggtgcaatat aggcaggatg atatgtgcca atgtcgccat agtaaaacta gccagaagaa
13737

aatagaggcg gtaacccaat ttgtctatgc atataccagc gataggagag cccaatgctg
13797

acacccccaa aagtatctga tcgaatattc tccaactaaa tgatcataat aataacgagc
13857

cattgaaggt cactcgagag cggatgagt tgcaccgcca tacggtaga aacagtatca
13917

ctatgttgtg ctaatgctca ttccctagca gaatcacaac atagcagaaa agaagggcga
13977

tcagttctgc actgcattcg aattacacag gaatgctctc taaaaaccg ttatgcaaat
14037

caccttggtc aatcgagaga gttcatcaat aactcttgag gggcatgacg tcaataatag
14097

tgaaaaacgg ttaacatcat ttactatct actgttgagg gcgacaatat gacaactgta
14157

taggaaaaac tgtagtaat tctatctct tactcccatg atctgtccaa ctttaacatg
14217

ggcagccgct ttatccatgg ttggatacca cacctcgacg aaaaatgccg cggcgatggt
14277

attgaaaggc agtataaccg cataggccag cacgcagctg agacatacca accagaagtc
14337

taccgataat gacaatgacg attggagact gagaaaaag gaagaacttc aacacagtgg
14397

tgacaatatt tcgcataag tagccgttat aactgcgcat ttctcagtag aaaaaccgaa
14457

aaatcgtggt gaaaactgac aaacattctt tcgaataaca aagttaaggt atgtgagagc
14517

tccaagccac cactagcatt acttctctat aaagccctga ctgggagctt cctcactagt
14577

gatatcatca tcgatgtgct tcgtgaaata acacaacact aacgttagaa taaccgtgcc
14637

tgaaaaaagt ccaccgatgg tgacaccata actacgagaa cagcttactt gtcaccgcta
14697

taagccccgc tatgaacgct cccgttacct aggaaacata atgtagaaat atatgtgtgg
14757

tttctcagc tgaggctgta aggttctca ctacattatg gaaatccaca aacgattttg
14817

ggtagaagcg tacaaaaacc aacaataagt cctcaaaatt ttccagaat ttccataatt
14877

ttaccaacac atgtagttga ccatgtgtat ccatagatag cgatagggtc atctcaaaca
14937

cttactccat atcgtgcctc tacaccaggt gacagcaaat cgttgagaac cgagccaaat
14997

cgggcaactg agaggtttag tcccaatgca aaggcgatct cctgccctcg gaaccagcct
15057

gctagaagag tcgtctgtgc cacctgaaga taaagggtata agaggaaggt caccgagagt
15117

tccccttaca ttcatgctct ccccgcttag tccgtagatt acccttcta gcagcatagt
15177

cctcatagaa aaataatag acccgacaca ttgtagcgt tgcccagagca ataacagtac

15237

cgccagtagg aggagcatgg acctgagcag agcttcttga catgctaaaa ggcctacagg
15297

gataaaactgc atttttaatg ggagccttcg gaatcttact actacatgag acattgaaaa
15357

agtttaggtt ttgcatgga ctttgaagaa ctattccaca cagctacttt aatattcatt
15417

gacacactct gaaatgttgt acaagtttca caaaattcac tcaagtgcg aagcacatca
15477

tcacatcata acagcatatc aaacgctcgt tacctcatac caatccactc tatagagact
15537

cccagtatta atggcgttac catattgggc agggcataga cggtatataa gagattgaag
15597

tagtaggtat attcctcgac agtcataccg ctccctccac cttaaagtagc tgatgcggtc
15657

acttctgaca acaacctacc ggtgaatcga gcctgcagtt gatcgtttat cgcggaggga
15717

atatcatagg cgtagtactg gccagagagc gctagacagg ccacgatgag aacagcccat
15777

cttagttttt caggagggtc gggagtagac ataataaaat catagctacg aaacagaaat
15837

tcagccaagt ttgagaaaat accttcccca ggatttgata acaacaggaa aaatgtttat
15897

aatgtgatga tgttaaactg attaaagaat gtgttcctag cagtgtagga aattcaaaat
15957

gagtttcaact cggaacatt ctggaagaat cactgcgacg aaagcgatct ggacacagca
16017

acgagcagcg gttgactttt cacccgaaac cctcagcaca gtcccttcgc gaaaatcata
16077

caggcaaagt ttcataaaaa aaatgcaatt ttccggtaaa aaaccgcga gaaaaaccgt
16137

cagtgatgat tgacgatccg attacaccaa aaatgctgat atttttacgg aaaatttttc
16197

cgtaaaagct caggccagcc caggcgtctc gcactgattc atcccttccc ataataacac
16257

aattcctcta ctatggacca atacgagtct tagctgtaaa gggctaccac cagtagcaga
16317

gccaaacgta acagtaagaa aatgatcaac catagaagag atgtagacta aacgcatgaa
16377

gtgaggccag ggacagctca acagctactc aatgtgagag tcttgttgaa ttttgccgag
16437

gtaataatcg ataaaatcaa aacaatgctg tttcaccogt tttacgcgaa gtgaocgtaaa
16497

actcgggtgta aaacgcgata ttccgcgctat gcagtagtcg tgaaatatgt ttttcatcag
16557

tgtttttcat atcactaggg gaaactccaa cacattcaca agtagaaagg aacaacaact
16617

gcacaaaacg tcataaacca caccaagttc ggttccgtca aattgggtatt acgatcgaac
16677

cctctagaat tttccccac atcatgtcgt cattggctca ctctgcgcca ataatttagc
16737

gcgcgcccga ccgaaaagat ctcaaaaata ctcaagccta ccttgcatct gtogtattaa
16797

tgggcggcac gcttcagacg tgttacgctc cggcagtgag gattatatgc atattttact
16857

tatattttctt atagaatata agaacactga cgtactaaac tatacaaacc atgtgttgtt
16917

ctattttcaa cctcaagcag tctgattgac agatcccgcg ggaaaacaga caaaagaatg
16977

gtaggcggta tctgtacact acacagtgag gggaaactttc agagcccgaa aagtatgctt
17037

aggcgtcttc gaccgcgtgg tgtccagggt tcgcaacgca acaggctggg aactctacac
17097

attttcaatc tcaaattcag cgaatgaaag aacaaattga atagtgaag ttgatttcca
17157

tagagtgaca ggaattttct tcggctgaaa actcatttag cccatgggaa tcaaattgta
17217

ttgaaaaagt gtataaaaca accaatatcg acgatcgagg aattatgtat tctccgatat
17277

ttaaatctta taattttttc actcatgaag tttctcgaga gcggtataac caggtaaagt
17337

atatcatgat tgaacctaat acccgggtga acgaccgtat ttcccaaccg tgaatcatcg
17397

atcacatgga ccacaagctc agtgcgctca actgatatac gcctactgta ttcacataac
17457

tatagaactt catcaacaca actgetgcat atgtgaaaac ttgaaagtgt cttatcatca
17517

atctatgaaa ggcaggaacg ataacatagt cgcattttcta acgactacga acaatgacgt
17577

ttaatatattt gggaagatca cgatctttac atagtacttg cgcattgagcg acaatgagat
17637

cggtctatcg tacgtaacag acaacaattg cticgaatcg gtaacatcac cgattgaaga
17697

gttttaacag gagataaacg atgaagattt acagtatagg tagccatcaa acattaataa
17757

cagaagagct tatacagccg aatagcatta gcttggttatt ggatagcagt agatagatat
17817

ttctttataa tacttaggcg aattcatcac caaacgcctt gatgaacgaa cttgtctaac
17877

tgcttaccag aaaccaagca accaccaata acccagcac accgtgttca taacaatcag
17937

attcacaacc atcacgatga caccctggaa ccaccatagc ctcaacggga cgtatcccaa
17997

gttgaagaaa ggtggattgg tagagctggg gtagtgagtt aaatgccag atagagttgt
18057

atacgcaagc gctctgcaac caagctccat cggaacacca gcggcctgca tgatctgcag
18117

gaacactggg aacatggcag agacatgtgc agtaatactt gcaaacagat actgggacac
18177

atagaaaata gccgtgcaaa tccatacact gccgtaagcc gccacacttg acaagtogat
18237

cttttcggca aaccactcaa agaagccagt ttcttcaac tcagtagcca acatcagcaa
18297

cacaccgaac caaaggaaca aatcccaagc tccactgtta ttcac
18342

<210> 2

<211> 510

<212> PRT

<213> *Perkinsus marinus*

<400> 2

Met Gln Val Pro Ala Glu His His Ser Thr Arg Val Ile Ser Ile Cys
1 5 10 15

Asp Ile Val Ile Ser Gly Pro Phe Gly Met Cys Asn His Asp Tyr Ser
20 25 30

Ala Ser Ile Pro Ala Ser Ser Ser Gly Ser Thr Arg Arg Met Arg Leu
35 40 45

Val Ala Tyr Ile Thr Leu Val Ser Ile His Tyr Gln Gln Leu Leu Phe
50 55 60

Tyr Ser Ser Ile Ile Thr Ile Ile Thr Gly Tyr His Tyr Tyr Val Ala
65 70 75 80

Ala Leu Pro Leu Tyr Asp Ile Ser Ile Ala Leu Ser Val Leu Ser Gly
85 90 95

Leu Thr Leu Leu Trp Leu Cys Asn Cys Tyr Tyr Asn Ser Lys Pro Asn
100 105 110

Val Phe Cys Ile Asp His Val Glu Phe Asp Ala Pro Pro Ser Trp Lys
115 120 125

Val Ser His Glu Asp Ile Ile Asn Ile Ala Lys Ile Gln Gly Cys Tyr
130 135 140

Thr Glu Asp Ser Leu Asn Phe Met Gln Arg Leu Leu Glu Arg Ser Gly
145 150 155 160

Thr Cys Pro Asp Lys Ser Ala Ala Tyr Pro Pro Val Val Val Glu Ser
165 170 175

Leu Arg Thr Asn Ala Pro Ala Asp Ala Ser Ala Val Asn Thr Arg Glu
180 185 190

Glu Ala Arg Glu Val Ile Ile Thr Thr Val Lys Asp Leu Leu Lys Lys
195 200 205

Thr Gly Val His Pro Lys Ser Ile Asp Tyr Ile Ile Val Asn Cys Ala
210 215 220

Met Tyr Asn Pro Thr Pro Ser His Ala Ala Met Ile Val Asn Glu Val
225 230 235 240

Gly Met Arg Asn Asp Val Ile Thr Tyr Asn Leu Ser Gly Met Gly Cys
245 250 255

Ser Ala Gly Val Ile Thr Ile Asp Leu Ala Thr Arg Leu Leu Arg Glu

260	265	270
Thr Arg Gly Arg Ala Leu Ile Val Ser Thr Glu Ile Leu Thr Arg Cys 275 280 285		
Phe Tyr Arg Gly Asn Asp Arg Glu Pro Leu Met Gly Asn Thr Leu Phe 290 295 300		
Arg Cys Gly Gly Ala Ala Ala Leu Leu Ser Ser Leu Pro Lys Asp Leu 305 310 315 320		
Ser Arg Ala Lys Tyr Lys Leu Leu His Thr Val Arg Thr Gln Val Leu 325 330 335		
Gly Asn Glu Ser Phe Glu Thr Ile Met Glu Thr Asp Asp Ser Thr Lys 340 345 350		
Pro Asn Ser Ile Val Thr Leu Arg Leu Gln Lys Ser Ile Ile Lys Val 355 360 365		
Ala Ala Val Ala Ile Lys Gln Asn Phe Thr Lys Leu Ala Tyr Met Val 370 375 380		
Leu Pro Leu Arg Glu Leu Leu Lys Val Leu Tyr Ser Met Val Thr Met 385 390 395 400		
Lys Met Arg Arg Lys Ser Ser Lys Glu Gly Arg Glu Leu Tyr Val Pro 405 410 415		
Asp Phe Arg Lys Gly Thr Asp His Trp Cys Ile His Ala Gly Gly Arg 420 425 430		
Gly Val Leu Asp Thr Leu Gln Asp Ser Leu Gln Leu Ser Asp Tyr Asp 435 440 445		
Ile Gln Ala Ser Arg Ser Val Leu Tyr Glu Arg Gly Asn Thr Ser Ser 450 455 460		
Ser Ser Ile Trp Tyr Glu Leu Ala Trp Leu Glu Arg Asp Gln Arg Ile 465 470 475 480		
Lys Arg Gly Asp Arg Val Leu Gln Leu Ala Phe Gly Ser Gly Phe Lys 485 490 495		

Cys Asn Ser Ser Val Trp Leu Ala Met His Asn Ile Asp Ala
 500 505 510

5 <210> 3
 <211> 457
 <212> PRT
 10 <213> *Perkinsus marinus*

15 <400> 3
 Met Thr Thr Ser Thr Thr Thr Val Gln Leu Gln Glu Asp Leu Ser Ser
 1 5 10 15

20 Gly Asp Gln Asn Ala His Pro Ser Pro Ser Arg Ala Thr Pro Ser Val
 20 25 30

25 Gly Asp Thr Lys Glu Asp Ala Arg Val Val Ile Lys Leu Phe Gly Thr
 35 40 45

30 Trp Val Asp Val Thr Ala Trp Leu Asn Asp His Pro Gly Gly Ser Lys
 50 55 60

35 Val Leu Arg Ala Phe Asn Lys Lys Asp Ala Thr Asp Ala Val Met Ala
 65 70 75 80

Met His Thr Asp Glu Ala Ile Lys Arg Ile Ile Arg Phe Ser Asn Val
 85 90 95

40 Val Ser Ser Ala Pro Ile Asn Ala Ser Ile Gly Asp Val Gln Val Ile
 100 105 110

45 Glu Lys Ser Leu Ser Arg Glu Gln Leu Met Tyr Tyr Lys Leu Arg Thr
 115 120 125

50 Leu Ala Arg Asn Gln Gly Trp Phe Gln Ser Asn Leu Leu Tyr Glu Gly
 130 135 140

Val Lys Ala Met Ile Ala Phe Gly Leu Leu Ile Ile Gly Phe Ala Thr
 145 150 155 160

55 Leu Tyr Phe Asp Tyr Gly Ile Trp Ser Thr Ala Leu Ile Gly Phe Ala
 165 170 175

Trp Phe Gln Leu Gly Trp Leu Gly His Asp Trp Ser His His Thr Ala
180 185 190

Leu Pro Lys Ser Thr Thr Asn Cys Ala Asn Tyr Asn Asp Tyr Leu Gly
195 200 205

Trp Leu Thr Gly Leu Ala Arg Gly Asn Thr Leu Leu Trp Trp Lys Leu
210 215 220

Arg His Asn Thr His His Val Leu Thr Asn Gln Tyr Glu Asn Asp Pro
225 230 235 240

Asp Ile Leu Thr Gln Pro Pro Leu His Phe Phe Glu Asp Phe Asp Val
245 250 255

Gly Asn Val Asn Arg Tyr Gln Ala Val Tyr Tyr Leu Pro Met Leu Thr
260 265 270

Leu Leu His Leu Phe Trp Leu Tyr Glu Ser Val Leu Val Cys Leu Arg
275 280 285

Gln Ser Arg Ser Ile Asn Arg Tyr Asn Arg Met His Ala Arg Arg Asp
290 295 300

Thr Val Ala Leu Val Leu His Ile Leu Ile Val Gly Ile Ile Ser Tyr
305 310 315 320

Thr Ser Gly Lys Tyr Leu Leu Ile Leu Leu Ala Tyr Met Leu Ser Gly
325 330 335

Phe Leu Thr Ala Val Val Val Phe Ala Ser His Tyr Asn Glu Pro Arg
340 345 350

Val Ala Ser Gly Glu Ser Leu Ser Leu Val Arg Gln Thr Leu Leu Thr
355 360 365

Thr Ile Asn Ile Gly Ser Phe Ser Asp Thr His Trp Glu Lys Lys Leu
370 375 380

Trp Phe Tyr Leu Thr Gly Gly Leu Asn Met Gln Ile Glu His His Leu
385 390 395 400

Phe Pro Thr Met Pro Arg His Asn Leu Pro Lys Thr Thr Phe Leu Val
405 410 415

Lys Ser Leu Ala Gln Glu Leu Gly Leu Pro Tyr Lys Glu Thr Asn Ile
 420 425 430

Val Ser Leu Thr Lys Ala Ala Val Thr Thr Leu His His Asn Ala Leu
 435 440 445

Arg Asn Ile Glu Arg Leu Leu Ala Arg
 450 455

<210> 4

<211> 411

<212> PRT

<213> *Perkinsus marinus*

<400> 4

Met Ser Ser Leu Thr Leu Tyr Arg Gly Pro Phe Ser Arg Met Val Leu
 1 5 10 15

Pro Arg Gln Glu Ile Cys Ile Asn Gly Arg Ile Tyr Asp Val Thr Glu
 20 25 30

Phe Ile Asn Arg His Pro Gly Gly Lys Ile Ile Leu Phe Gln Val Gly
 35 40 45

Ala Asp Ala Thr Asp Ala Phe Arg Glu Phe His Ala Gly Ser Glu Lys
 50 55 60

Ala Glu Lys Ile Leu Lys Thr Leu Pro Ser Arg Asp Asp Asp Gly Thr
 65 70 75 80

Phe Leu Pro Ser Thr Gln Arg Ser Ile Met Asp Asp Phe Lys Arg Leu
 85 90 95

Arg Asp Asp Leu Val Ser Arg Gly Val Phe Lys Pro Ser Val Met His
 100 105 110

Val Val Tyr Arg Cys Leu Glu Val Val Ala Leu Tyr Leu Ile Gly Phe
 115 120 125

Tyr Leu Ala Leu Cys Thr Ser Asn Val Tyr Val Gly Cys Ala Val Leu

130		135		140
Gly Val Ala Gln Gly Arg Ala Gly Trp Leu Met His Glu Gly Gly His 145 150 155 160				
His Ser Leu Thr Gly Asn Trp Lys Val Asp Gln Phe Leu Gln Glu Leu 165 170 175				
Phe Phe Gly Ile Gly Cys Gly Met Ser Ala Ala Trp Trp Arg Asn Ala 180 185 190				
His Asn Lys His His Ala Ala Pro Gln His Leu Gly Lys Asp Val Asp 195 200 205				
Leu Glu Thr Leu Pro Leu Val Ala Phe Asn Lys Ala Val Leu Arg Gly 210 215 220				
Arg Leu Pro Ser Val Trp Ile Arg Ser Gln Ala Val Cys Phe Ala Pro 225 230 235 240				
Ile Ser Thr Leu Leu Val Ser Phe Phe Trp Gln Phe Tyr Leu His Pro 245 250 255				
Arg His Ile Ile Arg Thr Gly Arg Arg Met Glu Ser Phe Trp Leu Leu 260 265 270				
Val Arg Tyr Leu Val Ile Val Tyr Leu Gly Phe Ser Tyr Gly Leu Val 275 280 285				
Ser Val Leu Leu Cys Tyr Ile Ala Ser Val His Val Gly Gly Met Tyr 290 295 300				
Ile Phe Val His Phe Ala Leu Ser His Thr His Leu Pro Val Ile Asn 305 310 315 320				
Gln His Gly Arg Ala Asn Trp Leu Glu Tyr Ala Ser Lys His Thr Val 325 330 335				
Asn Val Ser Thr Asn Asn Tyr Phe Val Thr Trp Leu Met Ser Tyr Leu 340 345 350				
Asn Tyr Gln Ile Glu His His Leu Phe Pro Ser Cys Pro Gln Phe Arg 355 360 365				

Phe Pro Gly Tyr Val Ser Met Arg Val Arg Glu Phe Phe His Lys His
 370 375 380

Gly Leu Lys Tyr Asn Glu Val Gly Tyr Leu His Ala Leu Asn Leu Thr
 385 390 395 400

Phe Ser Asn Leu Ala Ala Val Ala Ile Val Glu
 405 410

<210> 5

<211> 27

<212> DNA

<213> Sequência Artificial

<220>

<223> Iniciador de sequência

<400> 5
 atgcaagttc ccgcggagca tcactcc
 27

<210> 6

<211> 31

<212> DNA

<213> Sequência Artificial

<220>

<223> Iniciador de sequência

<400> 6
 cgttacgcac caatattatg catagccaac c
 31

<210> 7

<211> 26

<212> DNA

<213> Sequência Artificial

<220>

<223> Iniciador de sequência com sítios de restrição para clonagem em vetor de expressão de levedura

<400> 7
 ttggtaccat gggatttcct gcggag
 26

<210> 8

<211> 32

<212> DNA

<213> Sequência Artificial

<220>

<223> Iniciador de sequência

<400> 8
 gggagctctt acgcataaat attatgcata gc
 32

<210> 9

<211> 1533

<212> DNA

<213> *Perkinsus marinus*

<220>

<221> CDS

<222> (1)..(1533)

<223>

<400> 9
 atg cga ttt cct gcg gag cgt cac ttc act cgg gtt ata tct atc tgt
 48
 Met Arg Phe Pro Ala Glu Arg His Phe Thr Arg Val Ile Ser Ile Cys
 1 5 10 15
 gat atc atc atc tca ggc ccc ttt gga atg tgt aac cat gat tac tcc
 96

Asp Ile Ile Ile Ser Gly Pro Phe Gly Met Cys Asn His Asp Tyr Ser
 20 25 30
 tct tct ata cct gcc tct tgt agt ggt agc act cga cgc atg cgt ttg
 144
 Ser Ser Ile Pro Ala Ser Cys Ser Gly Ser Thr Arg Arg Met Arg Leu
 35 40 45
 gta gcc tac atc aca ttg gtc tct atc cac tat caa cag cta ctc ttt
 192
 Val Ala Tyr Ile Thr Leu Val Ser Ile His Tyr Gln Gln Leu Leu Phe
 50 55 60
 tac tct tct atc gta act cta atc act ggc tat cac tac tat gtc gca
 240
 Tyr Ser Ser Ile Val Thr Leu Ile Thr Gly Tyr His Tyr Tyr Val Ala
 65 70 75 80
 gct ctg ccc ctt tac gac ata tca ctt gct cta tct gtg ctt tcg gga
 288
 Ala Leu Pro Leu Tyr Asp Ile Ser Leu Ala Leu Ser Val Leu Ser Gly
 85 90 95
 ata acg cta ctt tgg tta tgt aat tgc tat tac aac agc aag ccc aat
 336
 Ile Thr Leu Leu Trp Leu Cys Asn Cys Tyr Tyr Asn Ser Lys Pro Asn
 100 105 110
 gta ttc tgc atc gat cat gct gag ttt gac gct ccc ccc tct tgg aag
 384
 Val Phe Cys Ile Asp His Ala Glu Phe Asp Ala Pro Pro Ser Trp Lys
 115 120 125
 gtc agc cat gaa gac atc atc aac att gcc aag ata caa ggt tgc tac
 432
 Val Ser His Glu Asp Ile Ile Asn Ile Ala Lys Ile Gln Gly Cys Tyr
 130 135 140
 acg gaa gat tca ctc aac ttc atg cag cgt ctt ctc gaa agg tct ggt
 480
 Thr Glu Asp Ser Leu Asn Phe Met Gln Arg Leu Leu Glu Arg Ser Gly
 145 150 155 160
 act tgc cct ggt aag agt gct gct tac cct cca gtg gtt gtt gag tca
 528
 Thr Cys Pro Gly Lys Ser Ala Ala Tyr Pro Pro Val Val Val Glu Ser
 165 170 175
 ttg agg act aac gcc ccc gct gat gcc tct gct gtc aat act aga gag
 576
 Leu Arg Thr Asn Ala Pro Ala Asp Ala Ser Ala Val Asn Thr Arg Glu
 180 185 190
 gaa gcg agg gaa gtg atc ata act acg gtc aaa gat ctg ctt aag aag
 624
 Glu Ala Arg Glu Val Ile Ile Thr Thr Val Lys Asp Leu Leu Lys Lys
 195 200 205
 act ggt gtg cat cct aaa tct att gac tat atc atc gtc aat tgc gcc

672
 Thr Gly Val His Pro Lys Ser Ile Asp Tyr Ile Ile Val Asn Cys Ala
 210 215 220

 atg tac aac ccg aca ccg tca cat gct gct atg ata gtg aat gaa gtc
 720
 Met Tyr Asn Pro Thr Pro Ser His Ala Ala Met Ile Val Asn Glu Val
 225 230 235 240

 ggt atg agg aat gac gtc atc acc tat aac ctc agt ggt atg ggg tgt
 768
 Gly Met Arg Asn Asp Val Ile Thr Tyr Asn Leu Ser Gly Met Gly Cys
 245 250 255

 agt gcc ggt gtt atc aca att gat cta gca acg cgt ctg ttg aga gag
 816
 Ser Ala Gly Val Ile Thr Ile Asp Leu Ala Thr Arg Leu Leu Arg Glu
 260 265 270

 acc aga ggt agg gca ctg att gtg tca act gag ata cta act cgt tgc
 864
 Thr Arg Gly Arg Ala Leu Ile Val Ser Thr Glu Ile Leu Thr Arg Cys
 275 280 285

 ttc tat cgt ggt aat gat cgt gaa cca ctg atg ggt aac aca tta ttc
 912
 Phe Tyr Arg Gly Asn Asp Arg Glu Pro Leu Met Gly Asn Thr Leu Phe
 290 295 300

 aga tgt ggt ggt gct gct gct ttg cta tcg tca ttg cct aaa gac tta
 960
 Arg Cys Gly Gly Ala Ala Ala Leu Leu Ser Ser Leu Pro Lys Asp Leu
 305 310 315 320

 tct cgt ggt aaa tat aag ttg tta cat acc gta aga acg caa gtt ctc
 1008
 Ser Arg Gly Lys Tyr Lys Leu Leu His Thr Val Arg Thr Gln Val Leu
 325 330 335

 ggt aat gag agt ttt gaa acg att atg gag act gat gac agc acc aag
 1056
 Gly Asn Glu Ser Phe Glu Thr Ile Met Glu Thr Asp Asp Ser Thr Lys
 340 345 350

 ccc aac agc att gtt aca ctc agg ctc cag aag agt att atc aaa gtt
 1104
 Pro Asn Ser Ile Val Thr Leu Arg Leu Gln Lys Ser Ile Ile Lys Val
 355 360 365

 gct gct gtt gct att aaa caa aat ttt act aag ctt gct tat gtg gtt
 1152
 Ala Ala Val Ala Ile Lys Gln Asn Phe Thr Lys Leu Ala Tyr Val Val
 370 375 380

 ctg cct ctg aga gaa ctg ttg aag gtc gta tat tcg atg gtg atg atg
 1200
 Leu Pro Leu Arg Glu Leu Leu Lys Val Val Tyr Ser Met Val Met Met
 385 390 395 400

aag atg agg agg aag tgg tca aaa gaa ggt cgt gag ttg tac gta cct
1248

Lys Met Arg Arg Lys Ser Ser Lys Glu Gly Arg Glu Leu Tyr Val Pro
405 410 415

gat ttt aga aag ggc att gat cat tgg tgt att cat gct ggt ggc cgt
1296

Asp Phe Arg Lys Gly Ile Asp His Trp Cys Ile His Ala Gly Gly Arg
420 425 430

ggt gta ttg gat acc tta caa gat tct ctc cag cta tca gac tat gat
1344

Gly Val Leu Asp Thr Leu Gln Asp Ser Leu Gln Leu Ser Asp Tyr Asp
435 440 445

atc caa gca agc cgt agt gtt ctt tat gag aga ggc aac acc agt agc
1392

Ile Gln Ala Ser Arg Ser Val Leu Tyr Glu Arg Gly Asn Thr Ser Ser
450 455 460

agt agc ata tgg tat gag ttg gca tgg ctc gaa cgt gac caa cgt att
1440

Ser Ser Ile Trp Tyr Glu Leu Ala Trp Leu Glu Arg Asp Gln Arg Ile
465 470 475 480

aag cgt gga gat agg gta tta cag gtg gct ttt ggt agt ggt ttc aaa
1488

Lys Arg Gly Asp Arg Val Leu Gln Val Ala Phe Gly Ser Gly Phe Lys
485 490 495

tgt aac tca tca gta tgg ttg gct atg cat aat att gat gcg taa
1533

Cys Asn Ser Ser Val Trp Leu Ala Met His Asn Ile Asp Ala
500 505 510

<210> 10

<211> 510

<212> PRT

<213> Sequência Artificial

<220>

<400> 10

Met Arg Phe Pro Ala Glu Arg His Phe Thr Arg Val Ile Ser Ile Cys
1 5 10 15

Asp Ile Ile Ile Ser Gly Pro Phe Gly Met Cys Asn His Asp Tyr Ser
20 25 30

Ser Ser Ile Pro Ala Ser Cys Ser Gly Ser Thr Arg Arg Met Arg Leu

35

40

45

Val Ala Tyr Ile Thr Leu Val Ser Ile His Tyr Gln Gln Leu Leu Phe
50 55 60

Tyr Ser Ser Ile Val Thr Leu Ile Thr Gly Tyr His Tyr Tyr Val Ala
65 70 75 80

Ala Leu Pro Leu Tyr Asp Ile Ser Leu Ala Leu Ser Val Leu Ser Gly
85 90 95

Ile Thr Leu Leu Trp Leu Cys Asn Cys Tyr Tyr Asn Ser Lys Pro Asn
100 105 110

Val Phe Cys Ile Asp His Ala Glu Phe Asp Ala Pro Pro Ser Trp Lys
115 120 125

Val Ser His Glu Asp Ile Ile Asn Ile Ala Lys Ile Gln Gly Cys Tyr
130 135 140

Thr Glu Asp Ser Leu Asn Phe Met Gln Arg Leu Leu Glu Arg Ser Gly
145 150 155 160

Thr Cys Pro Gly Lys Ser Ala Ala Tyr Pro Pro Val Val Val Glu Ser
165 170 175

Leu Arg Thr Asn Ala Pro Ala Asp Ala Ser Ala Val Asn Thr Arg Glu
180 185 190

Glu Ala Arg Glu Val Ile Ile Thr Thr Val Lys Asp Leu Leu Lys Lys
195 200 205

Thr Gly Val His Pro Lys Ser Ile Asp Tyr Ile Ile Val Asn Cys Ala
210 215 220

Met Tyr Asn Pro Thr Pro Ser His Ala Ala Met Ile Val Asn Glu Val
225 230 235 240

Gly Met Arg Asn Asp Val Ile Thr Tyr Asn Leu Ser Gly Met Gly Cys
245 250 255

Ser Ala Gly Val Ile Thr Ile Asp Leu Ala Thr Arg Leu Leu Arg Glu
260 265 270

Thr Arg Gly Arg Ala Leu Ile Val Ser Thr Glu Ile Leu Thr Arg Cys
 275 280 285

Phe Tyr Arg Gly Asn Asp Arg Glu Pro Leu Met Gly Asn Thr Leu Phe
 290 295 300

Arg Cys Gly Gly Ala Ala Ala Leu Leu Ser Ser Leu Pro Lys Asp Leu
 305 310 315 320

Ser Arg Gly Lys Tyr Lys Leu Leu His Thr Val Arg Thr Gln Val Leu
 325 330 335

Gly Asn Glu Ser Phe Glu Thr Ile Met Glu Thr Asp Asp Ser Thr Lys
 340 345 350

Pro Asn Ser Ile Val Thr Leu Arg Leu Gln Lys Ser Ile Ile Lys Val
 355 360 365

Ala Ala Val Ala Ile Lys Gln Asn Phe Thr Lys Leu Ala Tyr Val Val
 370 375 380

Leu Pro Leu Arg Glu Leu Leu Lys Val Val Tyr Ser Met Val Met Met
 385 390 395 400

Lys Met Arg Arg Lys Ser Ser Lys Glu Gly Arg Glu Leu Tyr Val Pro
 405 410 415

Asp Phe Arg Lys Gly Ile Asp His Trp Cys Ile His Ala Gly Gly Arg
 420 425 430

Gly Val Leu Asp Thr Leu Gln Asp Ser Leu Gln Leu Ser Asp Tyr Asp
 435 440 445

Ile Gln Ala Ser Arg Ser Val Leu Tyr Glu Arg Gly Asn Thr Ser Ser
 450 455 460

Ser Ser Ile Trp Tyr Glu Leu Ala Trp Leu Glu Arg Asp Gln Arg Ile
 465 470 475 480

Lys Arg Gly Asp Arg Val Leu Gln Val Ala Phe Gly Ser Gly Phe Lys
 485 490 495

Cys Asn Ser Ser Val Trp Leu Ala Met His Asn Ile Asp Ala
 500 505 510

REIVINDICAÇÕES

1. Seqüência de ácido nucléico isolada, caracterizada pelo fato de que codifica um polipeptídeo com atividade de $\Delta 9$ - elongase e que é selecionada a partir do grupo consistindo em:

5 a) uma seqüência compreendendo resíduos de ácido nucléico 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9;

 b) uma seqüência de ácido nucléico que hibridiza sob condições rigorosas com uma seqüência de ácido nucléico compreendendo resíduos 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9;

10 c) uma seqüência de ácido nucléico isolada que, codifica um polipeptídeo com atividade de $\Delta 9$ -elongase, em que o polipeptídeo compreende SEQ ID NO: 2 ou SEQ ID NO: 10;

 d) um derivado de uma seqüência compreendendo resíduos de ácido nucléico 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9 que codifica
15 um polipeptídeo com pelo menos 50% de identidade no nível de aminoácido com SEQ ID NO: 2 ou SEQ ID NO: 10; em que o referido polipeptídeo tem atividade de $\Delta 9$ -elongase.

2. Polipeptídeo, caracterizado pelo fato de que é codificado por uma seqüência de ácido nucléico, como definida na reivindicação 1.

20 3. Construto de gene, caracterizado pelo fato de que compreende uma seqüência de ácido nucléico, como definida na reivindicação 1, operavelmente ligados a uma ou mais seqüências reguladoras.

 4. Vetor, caracterizado pelo fato de que compreende ácido nucléico, como definido na reivindicação 1, ou um construto de gene, como
25 definido na reivindicação 3.

5. Organismo não humano transgênico, caracterizado pelo fato de que compreende pelo menos um ácido nucléico, como definido na reivindicação 1, um construto de gene, como definido na reivindicação 3, ou um vetor, como definido na reivindicação 4.

30 6. Organismo não humano transgênico de acordo com a reivindicação 5, caracterizado pelo fato de que o organismo é um microorganismo, um animal não humano ou uma planta.

7. Organismo não humano transgênico de acordo com a reivindicação 6, caracterizado pelo fato de que o organismo é uma planta.

8. Processo para a conversão de ácido linoléico ou um derivado do mesmo em ácido araquidônico ou um derivado do mesmo em um organismo, caracterizado pelo fato de que compreende introduzir em um organismo que compreende ácido linoléico pelo menos uma sequência de ácido nucléico compreendendo:

uma sequência de ácido nucléico isolada que codifica polipeptídeos com atividade de $\Delta 9$ -elongase, $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase, em que os polipeptídeos são selecionados a partir do grupo consistindo em SEQ ID NOS 2, 3 e 4, e expressando a referida sequência de ácido nucléico.

9. Processo de acordo com a reivindicação 8, caracterizado pelo fato de que incluindo ainda a etapa adicional de converter o ácido araquidônico em um ácido graxo $\omega 3$ introduzindo-se no organismo um ácido nucléico codificando uma $\omega 3$ dessaturase e opcionalmente uma $\Delta 5$ -elongase, e/ou uma $\Delta 4$ -elongase e/ou uma $\Delta 4$ -dessaturase.

10. Processo para a conversão de $18:2^{\Delta 9,12}$ (ácido linoléico) em $20:2^{\Delta 11,14}$, caracterizado pelo fato de que compreende introduzir em um organismo que compreende ácido linoléico pelo menos uma sequência de ácido nucléico que codifica um polipeptídeo que tem atividade de $\Delta 9$ -elongase e que compreende:

a) uma sequência compreendendo resíduos de ácido nucléico 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9;

b) uma sequência de ácido nucléico que hibridiza sob condições rigorosas com uma sequência de ácido nucléico compreendendo resíduos 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9;

c) uma sequência de ácido nucléico isolada que, codifica um polipeptídeo com atividade de $\Delta 9$ - elongase, em que o polipeptídeo compreende SEQ ID NO: 2 ou SEQ ID NO: 10;

d) um derivado de uma sequência compreendendo resíduos de ácido nucléico 7668 a 9200 de SEQ ID NO: 1 ou SEQ ID NO: 9 que codifica um polipeptídeo com pelo menos 50% de identidade no nível de aminoácido

com SEQ ID NO: 2 ou SEQ ID NO: 10; em que o referido polipeptídeo tem atividade de $\Delta 9$ -elongase; e expressando a referida seqüência de ácido nucléico.

5 11. Processo de acordo com qualquer uma das reivindicações 8 a 10, caracterizado pelo fato de que inclui ainda a etapa de indução com galactose.

12. Processo de acordo com qualquer uma das reivindicações 8 a 11, caracterizado pelo fato de que o organismo é um microorganismo, um animal não humano ou uma planta.

RESUMO

Patente de Invenção: "ÁCIDO NUCLÉICO".

A presente invenção refere-se ao ácido nucléico derivado de *Perkinsus marinus* que codifica uma enzima de 9-elongase, uma $\Delta 8$ -dessaturase e $\Delta 5$ -dessaturase. Todas as seqüências de codificação podem ser transcritas como uma única transcrição.