



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 345 324**

51 Int. Cl.:

**C12N 15/82** (2006.01)

**C12N 15/53** (2006.01)

**C12N 15/11** (2006.01)

**C12N 9/02** (2006.01)

**C12P 7/62** (2006.01)

**C12P 7/64** (2006.01)

**C12Q 1/68** (2006.01)

**C07K 16/40** (2006.01)

**G01N 33/573** (2006.01)

**C11B 1/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **02704657 .2**

96 Fecha de presentación : **18.01.2002**

97 Número de publicación de la solicitud: **1356067**

97 Fecha de publicación de la solicitud: **29.10.2003**

54

Título: **Método para la producción de ácidos grasos poliinsaturados, nuevos genes de biosíntesis y nuevos constructos vegetales de expresión.**

30

Prioridad: **19.01.2001 DE 101 02 337**

45

Fecha de publicación de la mención BOPI:  
**21.09.2010**

45

Fecha de la publicación del folleto de la patente:  
**21.09.2010**

73

Titular/es: **BASF Plant Science GmbH  
67056 Ludwigshafen, DE**

72

Inventor/es: **Lerchl, Jens;  
Renz, Andreas;  
Heinz, Ernst;  
Domergue, Frederic y  
Zähringer, Ulrich**

74

Agente: **Carvajal y Urquijo, Isabel**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

# ES 2 345 324 T3

## DESCRIPCIÓN

Método para la producción de ácidos grasos poliinsaturados, nuevos genes de biosíntesis y nuevos constructos vegetales de expresión.

5 La presente invención se refiere a un método para la preparación de ácidos grasos poliinsaturados con al menos dos enlaces dobles y/o un método para la preparación de triglicéridos con contenido elevado de ácidos poliinsaturados con al menos dos enlaces dobles. La invención también se refiere al uso ventajoso de las secuencias de ácidos nucleicos SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 ó 11 en el método y para la preparación de un organismo transgénico, preferible de un vegetal transgénico o de un microorganismo transgénico con contenido elevado de ácidos grasos, aceites o lípidos con ácidos grasos insaturados de C<sub>18</sub>, C<sub>20</sub>, o C<sub>22</sub>.

15 La invención también se refiere a desaturasas novedosas con [falta] que en las secuencias SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 y SEQ ID NO: 11 o sus homólogos, derivados o análogos y constructos génicos que comprenden estos genes o sus homólogos, derivados o análogos, así como a su uso solos o en combinación con genes de biosíntesis de ácidos grasos poliinsaturados, tal como se representa ventajosamente en SEQ ID NO: 7 y SEQ ID NO: 9.

20 Adicionalmente, la invención se refiere a secuencias aisladas de ácidos nucleicos; a casetes de expresión que contienen las secuencias de ácidos nucleicos, vectores y organismos transgénicos que contienen al menos una secuencia de ácidos nucleicos o un casete de expresión. Además, la invención se refiere a ácidos grasos insaturados con al menos dos enlaces dobles y a triglicéridos con un contenido elevado de ácidos grasos insaturados con al menos dos enlaces dobles y a su uso.

25 Además, la invención se refiere a casetes de multiexpresión para la expresión específica de semillas y a vectores y a organismos que comprenden un gen de desaturasa solo o en combinación con otras desaturasas con la secuencia SEQ ID NO: 7 y/o genes de elongasa con la secuencia ID NO: 9 o sus homólogos, derivados o análogos usando dichos casetes de expresión.

30 Una serie de productos y productos secundarios de procesos metabólicos de procedencia natural en microorganismos, células animales y vegetales son aprovechables para muchas ramas de la industria, incluyendo la industria de forrajes y alimentos para animales, industria de alimentos, cosméticos y farmacéuticos. Estas moléculas, denominadas de manera general como "productos químicos finos" incluyen, por ejemplo, lípidos y ácidos grasos, entre los cuales los ácidos grasos poliinsaturados son una clase ejemplar. Los ácidos grasos y triglicéridos tienen una gran cantidad de aplicaciones en la industria de productos alimenticios, de la alimentación para animales, de la cosmética y en el campo de la farmacia. Según si se trata de ácidos grasos saturados o insaturados o si se trata de triglicéridos con un contenido elevado de ácidos grasos saturados o insaturados, éstos son adecuados para las más diversas aplicaciones; así, por ejemplo, los ácidos grasos poliinsaturados (polyunsaturated fatty acids, PUFAs) se adicionan a la alimentación para bebés para elevar el valor nutritivo. Los PUFAs tienen además una influencia positiva en el nivel de colesterol en la sangre de las personas y, por lo tanto, son adecuados para la protección frente a enfermedades coronarias. De esta manera encuentran aplicación en diversos productos alimenticios dietéticos o en medicamentos.

45 Microorganismos particularmente adecuados para la producción de PUFAs son microorganismos tales como cepas de thraustochytridos o schizochytridos, algas como del tipo *Phaeodactylum tricorutum* o *Cryptocodinium*, ciliados como *Stylonychia* o *Colpidium*, hongos como *Mortierella*, *Entomophthora* o *Mucor*. Mediante selección de cepas se ha desarrollado una cantidad de cepas mutantes de los microorganismos correspondientes que producen una serie de compuestos deseables, incluyendo PUFAs. La selección de cepas con producción mejorada de una molécula determinada es, no obstante, un proceso difícil que quita tiempo.

50 De manera alterna, la producción de productos químicos finos puede llevarse a cabo a gran escala de una manera adecuada mediante la producción de vegetales que se hayan desarrollado de tal modo que produzcan los PUFAs previamente mencionados. Vegetales adecuados, particularmente buenos para este propósito, son plantas oleaginosas que contienen grandes cantidades de compuestos lípidos, tales como colza, canola, linaza, soya, girasol, borraja y onagra. Pero también son bien adecuadas otras plantas útiles que contienen aceites o lípidos y ácidos grasos, tal como se menciona en la presente descripción de esta invención. Por medio de cultivo convencional se ha desarrollado una serie de plantas mutantes que producen un espectro de lípidos y ácidos grasos deseables, co-factores y enzimas. Sin embargo, la selección de nuevas variedades vegetales con producción mejorada de una molécula determinada es un proceso dispendioso y difícil o incluso imposible si el compuesto no tiene lugar de manera natural en la planta correspondiente, como en el caso de ácidos grasos poliinsaturados de C<sub>18</sub>, C<sub>20</sub> y C<sub>22</sub> y aquellos con cadenas de carbono más largas.

60 Debido a las propiedades positivas de los ácidos insaturados no han faltado intentos en el pasado de poner a disposición genes que participen en la síntesis de ácidos grasos o triglicéridos para la producción de aceites en diversos organismos con contenido modificado de ácidos insaturados. Así, en WO 91/13972 y su equivalente estadounidense se describe una  $\Delta$ -9-desaturasa. En WO93/11245 se reivindica una  $\Delta$ -15-desaturasa, en WO 94/11516 se reivindica una  $\Delta$ -12-desaturasa. Se describen  $\Delta$ -6-desaturasas en WO 93/06712, US 5,614,393, WO 96/21022 y WO 99/27111. Otras desaturasas se describen, por ejemplo, en EP-A-0 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP-A-0 794 250, Stucky y colaboradores, J. Biol. Chem., 265, 1990: 20144-20149, Wada y colaboradores, Nature 347, 1990: 200-203 o Huang y colaboradores, Lipids 34, 1999: 649-659. En WO 96/13591 se describe y se

## ES 2 345 324 T3

reivindica una  $\Delta$ -6-Palmitoil-ACP-desaturasa. No obstante, la caracterización bioquímica de las diferentes desaturasas se ha efectuado hasta ahora de manera insuficiente solo porque las enzimas son muy difíciles de aislar y de caracterizar (McKeon y colaboradores, *Methods in Enzymol.* 71, 1981: 12141-12147, Wang y colaboradores, *Plant Physiol. Biochem.*, 26, 1988: 777-792).

5

En levaduras pudo detectarse tanto un desplazamiento del espectro de ácido graso hacia ácidos grasos insaturados como también un incremento de la productividad (véase Huang y colaboradores, *Lipids* 34, 1999: 649-659, Napier y colaboradores, *Biochem. J.*, Vol. 330, 1998: 611-614). Sin embargo, la expresión de las diversas desaturasas en vegetales transgénicos no mostró el éxito deseado. Pudo mostrarse un desplazamiento del espectro de ácido graso hacia

10

ácidos grasos insaturados, pero simultáneamente se mostró que el desempeño de síntesis de las plantas transgénicas decae fuertemente, lo que significa que frente a las plantas de partida pudieron aislarse solo cantidades bajas de aceites.

Ni en levaduras ni en plantas se producen de manera natural ácidos grasos poliinsaturados de  $C_{20}$  y/o  $C_{22}$  con al menos dos enlaces dobles en la molécula de ácido graso, como ácido araquidónico (ARA) y/o ácido eicosapentaenoico (EPA) y/o ácido docosahexaenoico (DHA).

15

Por lo tanto aún existe una gran demanda de genes novedosos, que codifiquen para enzimas, que participen en la biosíntesis de ácidos grasos insaturados y que hagan posible producirlos a una escala industrial. Ninguna de los procesos biotecnológicos conocidos hasta ahora para la producción de ácidos grasos poliinsaturados proporciona los

20

ácidos grasos ya mencionados en cantidades utilizables económicamente.

Por lo tanto, existía el problema de suministrar otras enzimas adicionales para la síntesis de ácidos grasos poliinsaturados. Y usar opcionalmente estas enzimas con otras enzimas para producir ácidos grasos poliinsaturados. Este problema fue resuelto mediante el proceso según la invención para la producción de ésteres de ácidos grasos con un contenido elevado de ácidos grasos poliinsaturados que tienen al menos dos enlaces dobles, el cual se caracteriza

25

porque al menos una secuencia de ácidos nucleicos seleccionados del grupo de

a) un ácido nucleico con la secuencia representada en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11,

30

b) ácidos nucleicos que se obtienen debido al código genético degenerado mediante re-traducción de las secuencias de aminoácidos representados en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 o SEQ ID NO: 12,

c) Derivados de los ácidos nucleicos representados en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11, que codifican para partes biológicamente activas de una desaturasa y tienen al menos 50% de identidad con la secuencia de aminoácido de las SEQ ID NO: 2, 4, 6 ó 12 se introduce en un organismo no humano que produce ésteres de ácido graso, se cultiva y se aíslan los ésteres de ácido graso contenidos en el organismo.

35

Las secuencias de ácidos nucleicos usadas en el proceso de acuerdo con la invención son secuencias de ácidos nucleicos aisladas que codifican para polipéptidos que tienen actividad  $\Delta$ -5-,  $\Delta$ -6- o  $\Delta$ -12-desaturasa.

40

En el proceso de acuerdo con la invención se producen ventajosamente ésteres de ácido graso con moléculas de ácidos grasos poliinsaturados de  $C_{18}$ ,  $C_{20}$  y/o  $C_{22}$ , que tienen al menos dos enlaces dobles en el éster de ácido graso. Estas moléculas de ácido graso contienen preferiblemente tres, cuatro o cinco enlaces dobles y conducen ventajosamente a la síntesis de ácido araquidónico (ARA), ácido eicosapentaenoico (EPA) o ácido docosahexaenoico (DHA).

45

Los ésteres de ácido graso con moléculas de ácido graso poliinsaturados de  $C_{18}$ ,  $C_{20}$  y/o  $C_{22}$  pueden aislarse de los organismos que se han usado para producir los ésteres de ácidos grasos en forma de un aceite o lípido, por ejemplo en forma de compuestos como esfingolípidos, fosfoglicéridos, lípidos, glicolípidos, fosfolípidos, monoacilglicéridos, diacilglicéridos, triacilglicéridos u otros ésteres de ácido graso que contienen ácidos grasos insaturados que tienen al menos dos enlaces dobles.

50

Como organismo para la producción en el proceso según la invención se toman en consideración teóricamente todos los organismos procariotas o eucariotas, como microorganismos procariotas o eucariotas tales como bacterias gran-positivas o bacterias gran-negativas, hongos, levaduras, algas, ciliados, células animales o vegetales, animales o vegetales como musgos, plantas dicotiledóneas o monocotiledóneas. En el proceso de la invención se usan ventajosamente organismos que pertenecen a los organismos productores de aceite, es decir los que se usan para la producción de aceites, como microorganismos tales como *cryptocodium*, *thraustochytrium*, *phaeodactylum* y *mortierella*, *entomophthora*, *mucor*, *cryptocodium* y otras algas u hongos, así como animales o vegetales, en particular vegetales, preferiblemente plantas oleaginosas que contienen grandes cantidades de compuestos lípidos, como frijol de soya, cacahuete, colza, canola, girasol, cártamo o alazor, onagra, linaza, soya, borraja, árboles (palma de aceite, coco) o frutos del campo como maíz, trigo, centeno, avena, triticale, arroz, cebada, algodón, yuca, pimienta, tagetes, plantas solanáceas como la patata, el tabaco, la berenjena y el tomate, variedades de vicia (algarroba), guisantes (arvejas), alfalfa o arbustos (café, cacao, té), variedades de salix, árboles (palma de aceite, coco) así como gramíneas perennes y frutos del campo para forraje. Plantas particularmente preferidas de acuerdo con la invención son plantas de frutos oleaginosos, como soya, cacahuete (maní), colza, canola, linaza, onagra, girasol, cártamo (alazor) o árboles (palma de aceite, coco).

65

## ES 2 345 324 T3

El proceso según la invención implica, o bien el cultivo de un organismo transgénico adecuado o de un microorganismo transgénico, o bien el cultivo de células vegetales transgénicas, de sus tejidos, órganos o de las plantas completas, los cuales comprenden las secuencias de nucleótidos según la invención de las SEQ ID NO: 1, 3, 5 ó 11 opcionalmente en compuestos con las secuencias representadas en SEQ ID NO: 7 y/o SEQ ID NO: 9, solas o en combinación con secuencias de constructos de expresión de SEQ ID NO: 13-17 o sus homólogos, derivados o análogos o un constructo génico que comprende las SEQ ID NO: 1, 3, 5 u 11 opcionalmente en compuesto con SEQ ID NO: 7 y/o 9 o sus homólogos, derivados o análogos, o un vector que comprende esta secuencia o el constructo génico, el cual ocasiona la expresión de moléculas de ácidos nucleicos de acuerdo con la invención, de modo que se produce un producto químico fino. En una forma preferida de realización el proceso comprende además el paso de obtener una célula que contiene las secuencias de ácidos nucleicos de la invención; se transforma una célula que tiene una secuencia de ácidos nucleicos, un constructo génico o un vector que provocan la expresión de un ácido nucleico de desaturasa, de acuerdo con la invención, solo o en combinación. En otra forma preferida de realización este proceso comprende además el paso de obtener los productos químicos fino a partir del cultivo. En una forma particularmente preferida de realización la célula pertenece al orden de los ciliados, a microorganismos como hongos o al reino vegetal, en particular a plantas oleaginosas; particularmente se prefiere microorganismos o plantas oleaginosas, como por ejemplo cacahuate o maní, colza, canola, linaza, soya, safflower (cártamo o cardo), girasol o borraja.

En el contexto de la invención debe entenderse por transgénico que los ácidos nucleicos de la invención no están en su sitio natural en el genoma de un organismo y en tal caso los ácidos nucleicos pueden expresarse de manera homóloga o heteróloga. Pero transgénico también significa que los ácidos nucleicos de la invención están en su lugar natural en el genoma de un organismo que, no obstante, ha modificado la secuencia en comparación con la secuencia natural y/o que las secuencias de regulación de las secuencias naturales se han modificado. Preferiblemente por transgénico debe entenderse la expresión de los ácidos nucleicos de la invención en sitios no naturales en el genoma, lo cual significa que se presenta una expresión homóloga o heteróloga de los ácidos nucleicos. Organismos transgénicos preferidos son las plantas transgénicas arriba mencionadas, preferible plantas oleaginosas.

De los ésteres de ácido graso producidos en el método de la invención pueden liberarse los ácidos grasos poliinsaturados contenidos, por ejemplo mediante tratamiento con álcali, como KOH o NaOH acuosos, preferible en presencia de un alcohol como metanol o etanol y aislarse mediante separación de fases, por ejemplo, y después acidificación con H<sub>2</sub>SO<sub>4</sub>, por ejemplo.

Otro objeto de la invención son aceites, lípidos y/o ácidos grasos que contienen al menos dos enlaces dobles en las moléculas de ácido graso, preferible tres, cuatro, cinco o seis enlaces dobles, los cuales se han producido según el proceso de la invención descrito arriba. Las composiciones que contienen los aceites, lípidos y/o ácidos grasos mencionados, así como el uso de aceites, lípidos y/o ácidos grasos o de las composiciones en forrajes o alimentos para animales, productos alimenticios, cosméticos o fármacos también son otro objeto de la invención.

Otro aspecto de la invención se refiere a procesos para la modulación de la producción de una molécula mediante un microorganismo. Estos métodos comprenden poner en contacto la célula con una sustancia que modula la actividad de desaturasa de acuerdo con la invención sola o en combinación o la expresión de ácidos nucleicos de desaturasa de modo que se modifica una actividad asociada con la célula en relación con la misma actividad en ausencia de la sustancia. En una forma preferida de realización se modula(n) una o dos rutas de metabolismo de la célula para lípidos y ácidos grasos, cofactores y enzimas o se modula el transporte de compuestos por estas membranas, de modo que se mejore el rendimiento o la rata de la producción de un producto químico fino deseado por este microorganismo. La sustancia que modula la actividad de desaturasa puede ser una sustancia que estimula la actividad de desaturasa o la expresión de ácidos nucleicos de desaturasa o que puede usarse como producto intercambio en la biosíntesis de ácidos grasos. Ejemplos de sustancias que estimulan la actividad de desaturasa o la expresión de ácido nucleico de desaturasa son, entre otras, pequeñas moléculas, desaturasas activas así como ácidos nucleicos que codifican desaturasa que se han introducido a la célula. Ejemplos de sustancias que inhiben la actividad o la expresión de desaturasa son, entre otras, pequeñas moléculas y moléculas de ácidos nucleicos de desaturasa antisense.

Otro aspecto de la invención se refiere a procesos para la modulación de los rendimientos de un compuesto deseado desde una célula, que comprenden la introducción a una célula de un gen de desaturasa tipo silvestre o mutante que, o bien se mantiene sobre un plásmido separado o se integra al genoma de la célula huésped. Al integrarse al genoma, la integración puede ser aleatoria o puede efectuarse mediante recombinación de tal modo que el gen nativo se reemplace por la copia introducida, por lo cual se modula la producción del deseado compuesto por la célula, o se usa un gen en trans de modo que el gen con una unidad de expresión funcional que contiene al menos una secuencia que garantiza la expresión de un gen y al menos una secuencia que garantiza la poliadenilación de un gen transcrito funcionalmente, está enlazado funcionalmente.

En una forma preferida de realización se modifican los rendimientos. En otra forma preferida de realización se aumenta el producto químico deseado y pueden disminuirse los compuestos indeseados que interfieren. En una forma preferida de realización el producto químico deseado es un lípido o un ácido graso, un co-factor o una enzima. En una forma particularmente preferida de realización este producto químico es un ácido graso poli-insaturado. Más preferiblemente éste se selecciona de ácido araquidónico (ARA), ácido eicosapentaenoico (EPA) o ácido docosahexaenoico (DHA).

## ES 2 345 324 T3

La presente invención proporciona novedosas moléculas de ácidos nucleicos que son adecuadas para identificar y aislar desaturasas de biosíntesis de PUFAs y que pueden usarse para la modificación de aceites, ácidos grasos, lípidos, compuestos derivados de lípidos y, lo más preferible, para la producción de ácidos grasos poliinsaturados.

5 La invención proporciona además casetes de multi-expresión y constructos para la expresión multiparalela específica para semillas de combinaciones de genes en vegetales.

Microorganismos tales como *Cryptocodinium*, *Thraustochytrium*, *Phaeodactylum* y *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodinium*, así como otras algas y hongos; particularmente las plantas oleaginosas son los organismos preferidos para el proceso de acuerdo con la invención.

En WO 98/01572, o en Falciatore y colaboradores, 1999, *Marine Biotechnology* 1(3): 239-251, y Dunahay y colaboradores, 1995, *Genetic transformation of diatoms*, *J. Phycol.* 31: 10004-1012 y en las referencias allí citadas se describen vectores de clonación y técnicas para la manipulación genética de los microorganismos y ciliados arriba mencionados, algas u organismos relacionados tales como *Phaeodactylum tricorutum*. De esta manera pueden usarse las moléculas de ácidos nucleicos en el proceso de acuerdo con la invención modificando los organismos mediante ingeniería genética de modo que se vuelven productores mejores o más eficientes de uno o varios productos químicos finos. Esta producción mejorada o eficiencia de la producción de un producto químico fino pueden provocarse mediante una acción directa de la manipulación de un gen de acuerdo con la invención o por una acción indirecta de esta manipulación. Por productos químicos finos en el contexto de la invención se entienden ésteres de ácidos grasos, que contienen ácidos grasos poliinsaturados con al menos dos enlaces dobles, tales como esfingolípidos, fosfoglicéridos, lípidos, glicolípidos, fosfolípidos, monoacilglicéridos, diacilglicéridos, triacilglicéridos u otros ésteres de ácidos grasos que contienen ácidos grasos poliinsaturados con al menos dos enlaces dobles. También por estos deben entenderse compuestos tales como vitaminas, por ejemplo vitamina E, vitamina C, vitamina B2, vitamina B6, pantolactona, carotinoides como astaxantina,  $\beta$ -caroteno, zeaxantina y otros.

Musgos y algas son los únicos sistemas vegetales conocidos que producen cantidades sustanciales de ácidos grasos poliinsaturados tales como ácido araquidónico (ARA) y/o ácido eicosapentaenoico (EPA) y/o ácido docosahexaenoico (DHA). Los musgos contienen PUFAs en lípidos de membrana, mientras que las algas, los organismos relacionados con las algas y algunos hongos también acumulan cantidades significativas de PUFAs en la fracción de triacilglicerol. Por lo tanto son adecuadas las moléculas de ácido nucleico que se aíslan de tales cepas que también acumulan PUFAs en la fracción de triacilglicerol, particularmente de manera ventajosa para la modificación del sistema de producción de lípidos y de PUFAs en un huésped, en particular en microorganismos, como los microorganismos previamente mencionados, y en plantas como las plantas oleaginosas, por ejemplo colza, canola, linaza, soya, girasol, borraja. Por eso pueden usarse ventajosamente en el proceso de acuerdo con la invención.

Por esto también son adecuados los ácidos nucleicos de la invención de manera particularmente ventajosa para aislar ácidos nucleicos de microorganismos que acumulan triacilglicerol y para la identificación de aquellas secuencias de ADN y de las enzimas codificadas por ellas en otras especies que son adecuadas para modificar la biosíntesis de moléculas precursoras de PUFAs en los organismos correspondientes.

Microorganismos tales como *Cryptocodinium cohnii*, *Thraustochytrium* y especies de *Phaeodactylum* son microorganismos que son capaces de acumular PUFAs como ARA, EPA o DHA en triacilgliceroles. *Thraustochytrium* también están relacionados estrechamente de manera filogenética con cepas de *Schizochytrium*. La capacidad de identificar desaturasas por medio de ácidos nucleicos de la invención, por ejemplo la predicción de la especificidad de sustrato de las enzimas, puede ser, por lo tanto, de importancia significativa. Estas moléculas de ácido nucleico también pueden servir como secuencias de referencia para el mapeo de genomas relacionados o para la derivación de cebadores (primers) PCR.

Las moléculas de ácidos nucleicos de la invención codifican para proteínas que se denominan desaturasas. Estas desaturasas pueden, por ejemplo, ejercer una función involucrada en el metabolismo (por ejemplo en la biosíntesis o la degradación) de compuestos requeridos para la síntesis de lípidos o ácidos grasos, tales como PUFAs, o pueden participar en el transporte a través de la membrana de uno o varios compuestos de lípido/ácido graso, ya sea adentro de la célula o afuera de la célula.

Las secuencias de ácidos nucleicos de acuerdo con la invención codifican para desaturasas que son adecuadas para la producción de ácidos grasos poliinsaturados de cadena larga, preferiblemente con más de dieciséis, dieciocho o veinte átomos de carbono en el esqueleto de carbonos del ácido graso y/o al menos dos enlaces dobles en la cadena de carbono; un ácido nucleico de acuerdo con la invención codifica para una enzima capaz de introducir enlaces dobles en la posición  $\Delta$ -5, en otro caso en la posición  $\Delta$ -6 y en un caso adicional en la posición  $\Delta$ -12. Con la ayuda de estos ácidos nucleicos pueden obtenerse grandes cantidades de PUFAs en la fracción de triacilglicerol. Adicionalmente se han aislado otras desaturasas que, solas o junto con una desaturasa  $\Delta$ -4, pueden aprovecharse para un proceso para producir ácidos grasos poliinsaturados. En tal caso en la solicitud por singular, es decir por un gen o proteína de desaturasa, también debe entenderse el plural, es decir los genes o las proteínas de desaturasas.

Hasta ahora, con ayuda de desaturasas ha podido mostrarse la producción de ácido trienoico con cadena de carbonos de  $C_{18}$ . En estos procesos conocidos en la literatura se ha reivindicado la producción de ácido  $\gamma$ -linolénico. Sin embargo, hasta ahora nadie había podido mostrar la producción de ácidos grasos poliinsaturados de cadenas más

## ES 2 345 324 T3

largas (con cadena de carbono de C<sub>20</sub> y más larga, así como de ácidos trienoicos y de tipos con mayor insaturación) solo mediante organismos modificados.

Para la producción de PUFAs de cadena larga de acuerdo con la invención, los ácidos grasos poliinsaturados de C<sub>18</sub> primero deben alargarse en al menos dos átomos de carbono mediante actividad enzimática de una elongasa. Después de un ciclo de elongación, esta actividad enzimática conduce a ácidos grasos de C<sub>20</sub>, y después de dos, tres y cuatro ciclos de elongación hasta ácidos grasos de C<sub>22</sub>, C<sub>24</sub> o C<sub>26</sub>. Las secuencias de ácidos nucleicos divulgadas en esta invención, las cuales codifican para diversas desaturasas, en conjunto con elongasas pueden conducir a poliinsaturados de cadenas muy largas. LA actividad de las desaturasas de acuerdo con la invención conduce preferiblemente a ácidos grasos de C<sub>18</sub>, C<sub>20</sub> y/o C<sub>22</sub> con al menos dos enlaces dobles en la molécula de ácido graso, preferentemente con tres, cuatro, cinco o seis enlaces dobles, particularmente preferible hasta ácidos grasos de C<sub>18</sub> y/o C<sub>20</sub> con al menos dos enlaces dobles en la molécula de ácido graso, preferentemente con tres, cuatro o cinco enlaces dobles en la molécula. La elongación de ácido graso puede efectuarse por combinación de las desaturasas de acuerdo con la invención con una actividad de elongasa, y puede usarse ventajosamente la elongasa codificada por SEQ ID NO: 9. Después de que ha tenido lugar la elongación con la(s) enzimas de la invención, pueden efectuarse otros pasos de desaturación tales como, por ejemplo, una desaturación en posición Δ-5. También puede usarse la combinación con otras elongasas como aquellas que conducen a una elongación de cadenas de C<sub>18</sub> a C<sub>20</sub> o de C<sub>20</sub> a C<sub>22-24</sub>, y/o puede emplearse ventajosamente con una desaturasa con actividad para posición Δ-4 para obtener los ácidos grasos altamente de-saturados. Por lo tanto, los productos de las actividades de desaturasa, y de la posible desaturación adicional, conducen a PUFAs preferidos con un alto grado de desaturación, como ácido dihomo-gamma-linolénico, ácido docosadienoico, ácido araquidónico, ácido ω6-eicosatriendihomo-γ-linolénico, ácido eicosapentaenoico, ácido ω3-eicosatrienoico, ácido ω3-eicosatetraenoico, ácido docosapentaenoico o ácido docosahexaenoico. Son sustratos de la actividad enzimática de acuerdo con la invención, por ejemplo, ácido taxoleico; ácido 6,9-octadecadienoico, ácido linoleico, ácido pinolénicosäure, ácido α- o β-linolénico o ácido estearidónico así como ácido araquidónico, ácido eicosatetraenoico, ácido docosapentaenoico, ácido eicosapentaenoico. Sustratos preferidos son ácido linoleico, ácido γ-ácido linolénico y/o ácido α-linolénico así como ácido araquidónico, ácido eicosatetraenoico, Ácido docosapentaenoico, ácido eicosapentaenoico. Particularmente preferidos como productos del proceso son: ácido araquidónico, ácido docosapentaenoico, ácido eicosapentaenoico. Los ácidos grasos de C<sub>18</sub> con al menos dos enlaces dobles en el ácido graso pueden elongarse por la actividad enzimática de acuerdo con la invención en forma del ácido graso libre o en forma de los ésteres tales como fosfolípidos, glicolípidos, esfingolípidos, fosfoglicéridos, monoacilglicéridos, diacilglicéridos o triacilglicéridos.

Para la alimentación humana el ácido linoleico conjugado "CLA" es de significado particular. Por CLA se entienden en particular ácidos grasos como C18:2<sup>9 cis, 11 trans</sup> o el isómero C18:2<sup>10 trans, 12 cis</sup>, que debido al sistema enzimático humano pueden desnaturalizarse o alongarse en el cuerpo después de ingerirlos y pueden contribuir a efectos que incentivan la salud. Con las desaturasas de la invención (desaturasas Δ-12) aquellos ácidos grasos conjugados con al menos dos enlaces dobles en la molécula se desnaturalizan y de esta manera se introducen a la alimentación humana aquellos ácidos grasos que incentivan la salud. Otros ejemplos de ácidos grasos conjugados son ácido alfa-parinárico, ácido punícico, ácido eleosteárico y ácido calendúlico.

Usando vectores de clonación en vegetales y en la transformación de vegetales como aquellas que se han publicado, y allí se citan: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), capítulo 6/7, páginas 71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; en: Transgenic Plants, volumen 1, Engineering and Utilization, Editado por: Kung y R. Wu, Academic Press, 1993, 15-38; B. Jenes y colaboradores, Techniques for Gene Transfer, en: Transgenic Plants, volumen 1, Engineering and Utilization, editores: Kung y R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225), los ácidos nucleicos de acuerdo con la invención pueden usarse para la modificación por ingeniería genética de un amplio espectro de vegetales, de modo que estos se vuelven un productor mejor o más eficiente de uno o varios productos derivados de lípidos, tales como PUFAs. Esta producción o eficiencia mejorada de la producción de un producto derivado de lípidos, como PUFA, puede ocasionarse mediante acción directa de la manipulación o una acción indirecta de esta manipulación.

Existe una serie de mecanismos mediante los cuales la modificación de una proteína de desaturasa de acuerdo con la invención puede afectar directamente el rendimiento, producción y/o eficiencia de la producción de un producto químico fino de una planta oleaginosa o de un microorganismo debido a una proteína modificada. La cantidad o la actividad de la proteína o del gen de desaturasa, así como de las combinaciones génicas de desaturasas y elongasas, pueden elevarse de modo que se produzcan de novo cantidades más grandes de estos compuestos porque a los organismos les faltaba esta actividad y habilidad para la biosíntesis antes de introducir el gen correspondiente. Lo correspondiente es válido para la combinación con otras desaturasas o elongasas u otras enzimas del metabolismo de lípidos. En este caso, también puede ser ventajoso el uso de diferentes secuencias divergentes, es decir secuencias que difieren al nivel de la secuencia de ADN, o bien el uso de promotores para la expresión génica que hace posible otra expresión génica temporal, es decir dependiendo por ejemplo del grado de madurez de un tejido o una semilla que almacena aceite.

Introduciendo un gen de desaturasa o varios genes de desaturasa de la invención a un organismo, solo o en combinación con otros genes a una célula, no solo puede elevarse el flujo de la biosíntesis hasta el producto final, sino también elevarse o crearse de novo la composición-triacilglicerina correspondiente. Así mismo, pueden elevarse el número o la actividad de otros genes que participan en la importación de nutrientes requeridos para la biosíntesis de uno o más productos químicos finos (por ejemplo, ácidos grasos, lípidos polares y neutrales), de modo que la concen-

## ES 2 345 324 T3

tracción de los precursores, co-factores o compuestos intermedios dentro de las células o dentro del compartimiento de almacenamiento, por lo cual se aumenta la capacidad de las células para la producción de PUFAs se sigue aumentando, tal como se describe a continuación. Los ácidos grasos y los lípidos son ellos mismos productos químicos finos deseables; optimizando la actividad o el incremento del número de una o más desaturases que participan en la biosíntesis de estos compuestos, o destruyendo la actividad de una o más desaturases que participan en la degradación de estos compuestos, puede ser posible incrementar el rendimiento, la producción y/o la eficiencia de producción de moléculas de ácido graso y de lípidos a partir de vegetales o microorganismos.

La mutagénesis de el/los gen(es) de desaturasa de la invención puede conducir adicionalmente a una proteína de saturasa con actividades modificadas, las cuales afectan la producción de uno o varios productos químicos finos de manera directa o indirecta. Por ejemplo, el número o la actividad del(de) los gene(s) de desaturasa de la invención pueden incrementarse de modo que los residuos normales o los productos secundarios del metabolismo de la célula (cuya cantidad se eleva posiblemente debido a la sobreproducción de los productos químicos finos deseados) se exporten de manera eficiente antes de que destruyan otras moléculas o procesos dentro de la célula (que disminuiría la viabilidad de la célula) o interfieran las rutas de biosíntesis de los productos químicos finos (por lo cual se disminuyen el rendimiento, la producción o la eficiencia de la producción de los productos químicos finos deseados). Adicionalmente, las cantidades intracelulares relativamente grandes de los productos químicos finos deseados pueden ser, por sí mismas, tóxicas para la célula o interferir con mecanismos enzimáticos de re-acoplamiento, como la regulación aloestérica; aumentando, por ejemplo, la actividad o el número de otras enzimas siguientes en sentido contrario o de enzimas de desintoxicación de la ruta de PUFA, puede aumentar la cuota de los PUFAs en la fracción de triacilglicerina, podría elevarse la viabilidad de células de semilla, lo cual a su vez conduce a un mejor desarrollo de células en cultivo o a semillas que producen los químicos finos deseados. El gen de desaturasa de la invención también puede manipularse de tal modo que se produzcan las cantidades correspondientes de las diversas moléculas de lípido y de ácido graso. Esto puede tener un efecto profundo en la composición de lípido de la membrana de la célula y genera nuevos aceites adicionalmente a la aparición de PUFAs recién sintetizados. Puesto que cada tipo de lípido tiene diversas propiedades físicas, una modificación de la composición del lípido de una membrana puede cambiar significativamente la fluidez de la membrana. Los cambios de la fluidez de la membrana pueden ejercerse sobre el transporte de moléculas por la membrana así como sobre la integridad de la célula, las cuales, ambas, poseen un efecto decisivo en la producción de químicos finos. En vegetales, estos cambios pueden influir además en otras características como tolerancia frente a situaciones de estrés abióticas o bióticas.

Tolerancia al estrés biótica y abiótica es una característica general que se desea transmitir a un amplio espectro de plantas tales como maíz, trigo, centeno, avena, triticale, arroz, cebada, frijol de soya, maní o cacahuete, algodón, lino de aceite y de fibra, colza y canola, linaza, yuca, pimienta, girasol y tagetes, plantas oleáceas como patata, tabaco, berenjena y tomate, variedades de vicia (algarroba), guisantes, alfalfa, arbustos (café, cacao, té), variedades de salix, árboles (palmas de aceite, coco) y gramíneas perennes y frutos del campo para forraje. Como otra forma de realización de acuerdo con la invención, estos frutos del campo también son plantas preferidas de destino para la ingeniería genética. Plantas particularmente preferidas de acuerdo con la invención son plantas de frutos oleaginosos tales como frijoles de soya, cacahuete, colza, canola, girasol, linaza, cártamo o cardo, árboles (palmas de aceite, coco) o frutos del campo como maíz, trigo, centeno, avena, triticale, arroz, cebada, alfalfa, o arbustos (café, cacao, té).

Por consiguiente, un aspecto de la invención se refiere a moléculas aisladas de ácidos nucleicos (por ejemplo, cADNs), que comprenden secuencias de nucleótidos que codifican una desaturasa o varias desaturases o partes biológicamente activas de las mismas, o fragmentos de ácidos nucleicos que son adecuados como cebadores (primer) o sondas de hibridación para la detección o para la amplificación de ácidos nucleicos que codifican desaturasa (por ejemplo, ADN o mRNA). En formas particularmente preferidas de realización, la molécula de ácido nucleico comprende una de las secuencias de nucleótidos representadas en la secuencia ID NO: 1 ó 3 y 5 o la región codificante o un complemento de una de estas secuencias de nucleótidos. En otras formas particularmente preferidas de realización, la molécula aislada de ácido nucleico según la invención comprende una secuencia de nucleótidos que se hibrida a una secuencia de nucleótidos, como se representa en la secuencia SEQ ID NO: 1, 3, 5 u 11, o a una parte de la misma o que tiene al menos aproximadamente 50%, preferible al menos aproximadamente 60%, más preferible al menos aproximadamente 70%, 80% o 90% y aún más preferible al menos aproximadamente 95%, 96%, 97%, 98%, 99% o más de homología con la misma. En otras formas preferidas de realización, la molécula de ácido nucleico aislada codifica una de las secuencias de aminoácidos representadas en la secuencia SEQ ID NO: 2, 4, 6 ó 12. El gen de desaturasa preferido según la invención también posee preferentemente al menos una de las actividades de desaturasa aquí descritas.

En otra forma de realización, la molécula aislada de ácido nucleico codifica una proteína o una parte de la misma; la proteína o la parte de la misma contienen una secuencia de aminoácidos que es suficientemente homóloga con una secuencia de aminoácido de la secuencia SEQ ID NO: 2, 4, 6 ó 12 para que la proteína o la parte de la misma mantenga una actividad de desaturasa. La proteína, o la parte de la misma, que se codifica por parte de la molécula de ácido nucleico, preferentemente mantienen la habilidad de participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares de plantas o en el transporte de moléculas a través de estas membranas. En una forma de realización, la proteína codificada por la molécula de ácido nucleico tiene al menos aproximadamente 50%, preferentemente al menos aproximadamente 60% y más preferiblemente al menos alrededor 70%, 80% ó 90% y lo más preferible al menos aproximadamente 95%, 96%, 97%, 98%, 99% o más de homología con una secuencia de aminoácidos de secuencia SEQ ID NO: 2, 4, 6 ó 12. En otra forma preferida de realización, la proteína es una proteína de longitud completa que es esencialmente en partes homóloga a una secuencia total de aminoácidos de las SEQ ID NO: 2, 4, 6 ó 12 (que se deriva del marco de lectura mostrada en SEQ ID NO: 1, 3, 5 u 11).

## ES 2 345 324 T3

En otras formas de realización la desaturasa aislada comprende una secuencia de aminoácidos que tiene al menos aproximadamente 50% de homología con una de las secuencias de las SEQ ID NO: 2, 4, 6 ó 12 y puede participar en el metabolismo de compuestos requeridos para la síntesis de ácidos grasos en un microorganismo o una célula vegetal o en el transporte de moléculas a través de estas membranas; se entienden cadenas de carbono desaturadas de C<sub>18</sub> o C<sub>20-22</sub> con dobles enlaces en al menos dos posiciones.

En otra forma preferida de realización, la molécula aislada de ácido nucleico proviene de *Phaeodactylum tricor- nutum* UTEX646 y codifica una proteína (por ejemplo una proteína de fusión desaturasa) que contiene un dominio biológicamente activo que tiene una homología de al menos aproximadamente 50% o mayor con una secuencia de aminoácidos de la secuencia SEQ ID NR 2, 4, 6 ó 12 y mantiene la habilidad de participar en el metabolismo de com- puestos requeridos en la síntesis de membranas celulares de plantas o en el transporte de moléculas a través de estas membranas, o tiene al menos una de las actividades de desaturación que resultan en PUFAs tales como GLA, ALA, ácido dihomo-gamma linolénico, ARA, EPA o DHA o sus moléculas precursoras, y también comprende secuencias heterólogas de ácidos nucleicos que codifican un polipéptido heterólogo o proteínas regulatorias.

De manera alternativa, la desaturasa aislada puede comprender una secuencia de aminoácidos que se codifica por una secuencia de nucleótidos que hibrida a una secuencia de nucleótidos de las SEQ ID NO: 1, 3, 5 u 11, por ejemplo en condiciones severas, o que tiene una homología de al menos aproximadamente 50%, preferentemente al menos alrededor de 60%, más preferible de al menos aproximadamente 70%, 80% ó 90% y aún más preferible al menos aproximadamente 95%, 96%, 97%, 98%, 99% ó mayor homología con la misma. Así mismo se prefiere que las formas de desaturasa preferidas también posean una de las actividades de desaturasa aquí descritas.

En otra forma de realización, la molécula aislada de ácido nucleico tiene una longitud de al menos 15, 25, 50, 100, 250 ó más nucleótidos e hibrida en condiciones severas a una molécula de ácido nucleico que comprende una secuencia de SEQ ID NO: 1, 3, 5 ó 17. Preferentemente, la molécula aislada de ácidos nucleicos corresponde a una molécula de ácido nucleico de procedencia natural. Con más preferencia la molécula aislada de ácido nucleico codifica desaturasa *Phaeodactylum* de procedencia natural o una parte biológicamente activa de la misma.

Otra forma de realización de la invención son casetes de expresión que hacen posible la expresión de los ácidos nucleicos de la invención con las secuencias SEQ ID NO: 1, 3, 5 u 11 en los diversos organismos tales como micro- organismos, por ejemplo bacterias, hongos, levaduras, ciliados, algas o células animales o vegetales o en animales o plantas.

Por casete de expresión de la invención (= constructo o fragmento de ácido nucleico) deben entenderse las secuen- cias nombradas en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11 que son el resultado del código genéticos y/o sus derivados funcionales o no funcionales que han sido enlazadas funcionalmente con una o varias señales regulatorias para incrementar ventajosamente la expresión génica y que controlas la expresión de la secuencia codificante en la célula huésped. Estas secuencias regulatorias deben hacer posible la expresión objetivo de los genes y de la expresión de proteína. Dependiendo del organismo huésped, esto puede significar, por ejemplo, que el gen primero se induce y solo luego se expresa y/o se sobre-expresa, o que se expresa y/o se sobre-expresa inmediatamente. A manera de ejemplo, estas secuencias regulatorias son secuencias a las que se enlazan inductores o represores y de esta manera regulan la expresión del ácido nucleico. Adicionalmente a estas secuencias regulatorias novedosas, o en lugar de estas secuencias, antes de los propios genes estructurales la regulación natural de estas secuencias puede estar aún presente, y opcionalmente pudo haberse modificado genéticamente de modo que se haya eliminado la regulación natural y se haya incrementado la expresión de los genes. Sin embargo, el constructo génico también pueden tener una construcción más simple, es decir que no se han insertado señales de regulación adicionales antes de la secuencia de ácidos nucleicos o sus derivados. En lugar de eso se mutó la secuencia natural de regulación de tal modo que ya no se efectúa una regulación y/o se incrementa la expresión génica. Para incrementar la actividad, estos promotores modificados también pueden insertarse antes del gen natural en forma de secuencias parciales (= promotor con partes de las secuencias de ácidos nucleicos de la invención) o también solos. El constructo génico también puede contener ventajosamente además una o varias, así llamadas, "secuencias enhancer" que están ligadas funcionalmente con el promotor que hacen posible una expresión incrementada de la secuencia de ácidos nucleicos. Secuencias ventajosas adicionales, tales como otros elementos regulatorios o terminadores, también pueden insertarse en el extremo 3' de las secuencias de ADN. Los genes de desaturasa  $\Delta$ -5 /desaturasa  $\Delta$ -6 y/o desaturasa  $\Delta$ -12 pueden estar contenidos en una o más copias en el casete de expresión (= constructo génico).

Tal como se describe arriba, las secuencias o factores regulatorios pueden tener preferiblemente un efecto posi- tivo sobre la expresión génica de los genes introducidos, incrementándola. Así puede efectuarse ventajosamente un refuerzo de los elementos regulatorios a nivel de transcripción usando fuertes señales de transcripción, tales como promotores y/o "enhancers". Pero además, también es posible un refuerzo de la translación, por ejemplo mejorando la estabilidad del mRNA.

Otro aspecto de la invención se refiere a vectores, por ejemplo vectores de expresión recombinantes que contienen al menos una molécula de ácido nucleico según la invención y a células huésped a las que se han introducido estos vectores, particularmente microorganismos, células vegetales, tejidos vegetales, órganos de vegetales o las plantas enteras. En una forma de realización, una célula huésped así puede almacenar compuestos químicos finos, en particular PUFAs; para aislar el compuesto deseado se cosechan las células. El compuesto (aceites, lípidos, triacilglicéridos, ácidos grasos) o la desaturasa pueden aislarse del medio o de la célula huésped, que en el caso de los vegetales son

células que contienen los productos químicos finos, de mayor preferencia células de tejidos de almacenamiento como las cáscaras de semillas, bulbos, epidermis y células de demillas, endosperma o tejidos embrionales.

Un aspecto más de la invención se refiere a plantas transgénicas, genéticamente modificada, preferible una planta oleaginosa, como se mencionó previamente, particularmente preferible una planta de colza o de linaza a la que se ha introducido un gen de desaturasa. En una forma de realización, el genoma de colza o linaza ha sido modificado introduciendo como transgen una molécula de ácido nucleico de la invención que codifica un tipo silvestre o una secuencia de desaturasa mutada. En otra forma de realización, un gen endógeno de desaturasa en el genoma del organismo donante *Phaeodactylum* se ha destruido funcionalmente mediante mutagénesis y detección por medio de secuencias de ADN o se ha reprimido por medio de tecnología antisense. En una realización preferida también se ha usado colza o linaza para la producción de un compuesto deseado tal como lípidos y ácidos grasos y se prefieren especialmente PUFAs.

En aún otra forma preferida el musgo *Physcomitrella patens* puede usarse para demostrar la función de un gen de desaturasa usando recombinación homóloga sobre la base de los ácidos nucleicos descritos en la presente invención.

Otro aspecto más de la invención se refiere a un gen aislado de desaturasa o a una parte, por ejemplo una parte biológicamente activa, del mismo. En una forma preferida de realización, la desaturasa aislada o una parte de la misma puede participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en un microorganismo o una célula vegetal o en el transporte de moléculas a través de sus membranas. En otra forma de realización preferida, la desaturasa aislada o la parte de la misma tiene suficiente homología con una secuencia de aminoácidos de la SEQ ID NO: 2, 4, 6 ó 12 para que esta proteína o la parte de la misma mantengan la habilidad de participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en Microorganismos o células vegetales o en el transporte de moléculas por estas membranas.

La invención también suministra una preparación aislada de una desaturasa en forma de un extracto crudo o como proteína pura.

El polipéptido de desaturasa, o una parte biológicamente activa del mismo, de manera ventajosa puede estar enlazado funcionalmente a otro polipéptido que tiene otra actividad enzimática distinta de aquella de las desaturasas, por ejemplo una elongasa, aciltransferasa u otra actividad, de modo que se forma una proteína de fusión. Esta proteína de fusión tiene ventajosamente una actividad que difiere de aquella de la desaturasa sola. En otra forma de realización, esta proteína de fusión participa en el metabolismo de compuestos que se requieren para la síntesis de lípidos y ácidos grasos, co-factores y enzimas en microorganismos o plantas, o en el transporte de moléculas por estas membranas. De manera particularmente ventajosa, la introducción de esta proteína de fusión a una célula huésped modula la producción de un compuesto deseado dentro de una célula y por parte de la célula. En una forma de realización preferida estas proteína de fusión también contienen actividades de desaturasa  $\Delta$ -4,  $\Delta$ -5 ó  $\Delta$ -6,  $\Delta$ -8,  $\Delta$ -15,  $\Delta$ -17 ó  $\Delta$ -19 solas o en combinación. En particular, las formas de realización preferidas son aquellas combinación de genes que se seleccionan de SEQ ID NO: 7 ó 9 o partes de las mismas o sus homólogos. En particular, se prefieren aquellas combinaciones que contienen la actividad de proteína completa como en SEQ ID NO: 1, 3, 5 u 11 y se insertan en casetes de expresión múltiple definidos por SEQ ID NO: 13, 14, 15, 16 y 17 y son adecuadas para la transformación vegetales y la expresión en vegetales.

### Descripción detallada de la invención

Un objeto de acuerdo con la invención es/son ácidos nucleicos aislados que codifican para un polipéptido con actividad de desaturasa, seleccionados del grupo:

a) un ácido nucleico con la secuencia representada en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 ó SEQ ID NO: 11,

b) ácidos nucleicos que, debido al código genético degenerado, se obtienen mediante re-traducción de las secuencias de aminoácidos representadas en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 o SEQ ID NO: 12,

c) Derivados de los ácidos nucleicos representados en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11 que codifican para partes biológicamente activas de una desaturasa y presentan al menos 50% de identidad con la secuencias de aminoácidos de las SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 ó SEQ ID NO: 12.

Adicionalmente, la invención se refiere a una secuencia de aminoácidos que se codifican por la(s) arriba mencionada(s) secuencia(s) de ácidos nucleicos (en el contexto de la invención, el singular debe abarcar el plural y viceversa). La invención se refiere especialmente a secuencias de aminoácidos que se codifican mediante la secuencia representada en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 ó SEQ ID NO: 11.

La presente invención proporciona ácidos nucleicos y moléculas de proteína con actividad de desaturasa que participan en el metabolismo de lípidos y ácidos grasos, co-factores de PUFA y enzimas en el musgo *Physcomitrella patens* o en el transporte de compuestos lipofílicos a través de las membranas. Los compuestos de acuerdo con la invención pueden usarse para modular la producción de productos químicos finos desde organismos como, por ejemplo, microorganismos tales como los ciliados, hongos, levaduras, bacterias, algas y/o plantas tales como maíz, trigo,

centeno, avena, triticale, arroz, cebada, frijol de soya, cacahuete o maní, algodón, especies de *Linum* tales como lino de aceite o de fibra, especies de brassica como colza, canola y nabina, pimienta, girasol, borraja, onagra y tagetes, plantas solanáceas como patata, berenjena y tomate, especies de vicia, guisantes, yuca, alfalfa, arbustos (café, cacao, té), especies de salix, árboles (palma de aceite, coco) y gramíneas perennes y frutos del campo para forraje o alimento para animales, ya sea directamente (por ejemplo, si la sobreexpresión u optimización de una proteína para la síntesis de ácido graso tiene una influencia directa en el rendimiento, producción y/o eficiencia de la producción del ácido graso desde organismos modificados) o pueden tener un efecto indirecto que, no obstante, conduzcan a un incremento del rendimiento, producción y/o eficiencia de la producción de un compuesto deseado o a una disminución de compuestos indeseados (por ejemplo, si la modulación del metabolismo de lípidos y ácidos grasos, cofactores y enzimas conduce a modificaciones del rendimiento, producción y/o eficiencia de la producción o a la composición de los compuestos deseados dentro de las células, lo cual puede influir a su vez en la producción de uno o más productos químicos finos) Los aspectos de la invención se ilustran con más detalle a continuación.

#### 15 I. Productos químicos finos y PUFAs

El concepto “producto químico fino” es conocido en el campo de la técnica y abarca moléculas que han sido producidas por un organismo y que se usan en una variedad de industrias tales como, por ejemplo pero sin limitarse a las mismas, la farmacéutica, la agrícola, la alimentaria, la cosmética. Estos compuestos abarcan lípidos, ácidos grasos, co-factores y enzimas, etc. (como se describe, por ejemplo, en Kuninaka, A. (1996) *Nucleotides and related compounds*, páginas 561-612, en *Biotechnology* volumen 6, Rehm y colaboradores, Editorial, VCH: Weinheim y referencias bibliográficas allí contenidas), lípidos, ácidos grasos saturados e insaturados (por ejemplo ácido araquidónico), vitaminas y co-factores (como se describe en Ullmann’s *Encyclopedia of Industrial Chemistry*, vol. A27, *Vitamins*, páginas 443-613 (1996) VCH: Weinheim y citas bibliográficas allí contenidas; y Ong, A.S., Niki, E., & Packer, L. (1995) *Nutrition, Lipids, Health and Disease Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asia*, que tuvo lugar del 1 al 3 de septiembre de 1994 en Penang, Malasia, AOCS Press (1995)), enzimas y todos los otros productos químicos descritos por Gutcho (1983) en *Chemicals by Fermentation*, Noyes Data Corporation, ISBN: 0818805086, y las referencias bibliográficas allí indicadas. El metabolismo y el uso de determinados productos químicos finos se ilustran con mayor detalle a continuación.

La combinación de diversas moléculas precursoras y enzimas de biosíntesis conduce a la producción de diversas moléculas de ácidos grasos, lo cual tiene un efecto decisivo en la composición de la membrana. Puede suponerse que los PUFAs no solo se incorporan a triacilglicerol sino también a los lípidos de membrana.

La síntesis de membranas es un proceso bien caracterizado en el que está involucrada una cantidad de componentes, incluyendo lípidos como parte de la membrana de doble capa. La producción de ácidos grasos novedosos tales como PUFAs puede generar, por lo tanto, propiedades novedosas de funciones de membrana dentro de una célula o un organismo.

Las membranas celulares sirven para un gran número de funciones en una célula. Primero y en primera línea, una membrana delimita los contenidos de una célula del medio ambiente e imparte de esta manera integridad a la célula. Las membranas también pueden actuar como barreras contra la afluencia de compuestos peligrosos o indeseados o sino contra la salida de compuestos deseados.

Para descripciones y participaciones más detalladas de las membranas y los mecanismos involucrados, véase en: Bamberg, E., y colaboradores (1993) *Charge transport of ion pumps on lipid bilayer membranes*, *Q. Rev. Biophys.* 26: 1-25; Gennis, R.B. (1989) *Pores, Channels and Transporters*, en: *Biomembranes, Molecular Structure and Function*, Springer: Heidelberg, páginas 270-322; y Nikaido, H., y Saier, H. (1992) *Transport proteins in bacteria: common themes in their design*, *Science* 258: 936-942, y las citas contenidas en cada una de estas referencias bibliográficas.

La síntesis de lípidos puede dividirse en dos partes: la síntesis de ácidos grasos y su enlazamiento a sn-glicerina-3-fosfato así como la adición o modificación de un grupo polar de cabeza. Lípidos que se usan en membranas comprenden fosfolípidos, glicolípidos, esfingolípidos y fosfoglicéridos. La síntesis de ácidos grasos inicia con la transformación de acetil-CoA en malonil-CoA por la acetil-CoA-carboxilase o en acetil-ACP por la acetiltransacilasa. Después de una reacción de condensación estas dos moléculas de producto forman juntas acetoacetil-ACP, que se convierte mediante una serie de reacciones de condensación, reducción y deshidratación para producir una molécula de ácido graso saturado con la longitud de cadena deseada. La producción de los ácidos grasos insaturados a partir de estas moléculas se cataliza por desaturasas específicas, o bien aerobicamente por medio de oxígeno molecular o anaeróbicamente (respecto de la síntesis de ácidos grasos en microorganismos véase F.C. Neidhardt y colaboradores (1996) *E. coli* y *Salmonella*. ASM Press: Washington, D.C., S. 612-636 y las referencias bibliográficas allí contenidas; Lengeler y colaboradores (editores) (1999) *Biology of Prokaryotes*. Thieme: Stuttgart, New York, y las referencias bibliográficas contenidas, así como Magnuson, K., y colaboradores (1993) *Microbiological Reviews* 57: 522-542 y las referencias bibliográficas contenidas).

Precusores para la biosíntesis de PUFA son, por ejemplo, ácido oleico, ácido linoleico y ácido linolénico. Estos ácidos grasos de C<sub>18</sub> carbonos deben alongarse a C<sub>20</sub> y C<sub>22</sub>, para que se obtengan ácidos grasos del tipo de cadena eicosa y docosa. Con ayuda de diversas desaturasas, como enzimas que presentan actividad de desaturasa Δ-12, desaturasa Δ-

## ES 2 345 324 T3

15, desaturasa  $\Delta$ -6, desaturasa  $\Delta$ -5 y  $\Delta$ -4 desaturasa, pueden obtenerse ácido araquidónico, ácido eicosapentaenoico y ácido docosahexaenoico, así como otros PUFAs diferentes de cadena larga, extraerse y usarse para diversos propósitos en aplicaciones de productos alimenticios, forrajes, cosméticos y productos farmacéuticos.

5 Para producir PUFAs de cadena larga, los ácidos grasos poliinsaturados de  $C_{18}$  o  $C_{20}$  deben polidesaturarse como se mencionó arriba. Las secuencias de ácidos nucleicos según la invención codifican primero desaturasas funcionalmente activas de *Phyeodactylum tricorutum*, un microorganismo que contiene PUFAs en la fracción de triacilglicerol. Pueden introducirse enlaces dobles en la posición  $\Delta$ -5,  $\Delta$ -6 ó  $\Delta$ 12 con las desaturasas según la invención. Las actividades de las desaturasas según la invención conducen preferentemente a ácidos grasos de  $C_{18} + C_{20}$  con al menos  
10 dos, tres, cuatro o cinco enlaces dobles en la molécula de ácido graso, preferentemente a ácidos grasos de  $C_{20}$  preferentemente con tres, cuatro o cinco enlaces dobles en la molécula de ácido graso. La de-saturación puede efectuarse antes o después de la elongación de los ácidos grasos correspondientes. Por lo tanto, los productos de las actividades de desaturasa y de la posible de-saturación y elongación conducen a PUFAs preferidos con un más alto grado de desaturación, incluyendo una elongación adicional de ácidos grasos de  $C_{20}$  a  $C_{22}$ , a ácidos grasos tales como ácido linoleico, ácido docosadienoico, ácido dihomo- $\gamma$ -linolénico, ácido araquidónico, ácido  $\omega$ 6-eicosatriendihomo- $\gamma$ -linolénico, ácido eicosapentaenoico, ácido  $\omega$ 3-eicosatrienoico, ácido  $\omega$ 3-eicosatetraenoico, ácido docosapentaenoico o ácido docosahexaenoico. Sustratos de esta actividad enzimática de acuerdo con la invención son, por ejemplo, ácido taxólico, ácido 6,9-octadecadienoico, ácido oleico, ácido linoleico, ácido  $\gamma$ -linolénico, ácido pinolénico, ácido  $\alpha$ -linolénico, ácido araquidónico, ácido eicosapentaenoico o ácido estearidónico. Sustratos preferidos son ácido linoleico, ácido  $\gamma$ -linolénico y/o ácido  $\alpha$ -linolénico, ácido dihomo- $\gamma$ -linolénico o ácido araquidónico, ácido eicosatetraenoico o ácido eicosapentaenoico. Los ácidos grasos de  $C_{18}$  o  $C_{20}$  con al menos dos enlaces dobles en el ácido graso pueden alongarse por la actividad enzimática de acuerdo con la invención en forma de ácido graso libre o en forma de los ésteres tales como fosfolípidos, glicolípidos, esfingolípidos, fosfoglicéridos, monoacilglicéridos, diacilglicéridos, triacilglicéridos u otros ésteres.

25 Además, los ácidos grasos deben transportarse a continuación a diferentes lugares de modificación e incorporarse al lípido de almacenamiento de triacilglicerina. Otro paso importante en la síntesis de lípidos es la transferencia de ácidos grasos a los grupos polares de cabeza, por ejemplo por la glicerina-ácido graso-aciltransferasa (véase Frentzen, 1998, Lipid, 100(4-5): 161-166).

30 Véanse publicaciones sobre la biosíntesis vegetal de ácido graso, de-saturación, el metabolismo lipídico y el transporte de membrana de compuestos con contenido de grasa, la betaoxidación, modificación de ácido graso y co-factores, almacenamiento y ensamblaje de triacilglicerina, en los siguientes artículos, incluyendo las referencias bibliográficas de allí: Kinney, 1997, Genetic Engineering, editor: JK Setlow, 19: 149-166; Ohlrogge y Browse, 1995, Plant Cell 7: 957-970; Shanklin y Cahoon, 1998, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 611-641; Voelker, 1996, Genetic Engineering, editor: JK Setlow, 18: 111-13; Gerhardt, 1992, Prog. Lipid R. 31: 397-417; Gühnemann-Schäfer & Kindl, 1995, Biochim. Biophys Acta 1256: 181-186; Kunau y colaboradores, 1995, Prog. Lipid Res. 34: 267-342; Szymne y colaboradores, 1993, en: Biochemistry and Molecular Biology of Membrane and Storage Lipids of Plants, editor: Murata y Somerville, Rockville, American Society of Plant Physiologists, 150-158, Murphy & Ross 1998, Plant Journal. 13 (1): 1-16.

45 Vitaminas, co-factores y productos nutracéuticos como PUFAs, comprenden un grupo de moléculas que ya pueden sintetizar los animales superiores y deben por lo tanto ingerirse o los animales superiores ya no pueden producirlos por sí mismos en cantidad suficiente y deben ingerirse por lo tanto adicionalmente, aunque se sintetizan fácilmente por otros organismos, tales como las bacterias. La biosíntesis de estas moléculas en organismos que pueden producirlas, como en las bacterias, ha sido caracterizada en gran medida (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", vol. A27, páginas 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E., & Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research Asia, que tuvo lugar del 1 al 3 de septiembre de 1994 en Penang, Malasia, AOCs Press, Champaign, IL X, 374 pág).

55 Las moléculas mencionadas arriba son, o bien ellas mismas biológicamente activas, o precursores de sustancias biológicamente activas que sirven ya sea como vehículo de electrones o como productos intermedios en una cantidad de rutas metabólicas. Estos compuestos tienen, además de su valor nutricional, también un valor industrial significativo como colorantes, antioxidantes y catalizadores u otros auxiliares de procesamiento (véase una vista general sobre estructura, actividad y aplicaciones industriales de estos compuestos, por ejemplo, en Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", vol. A27, pág. 443-613, VCH: Weinheim, 1996). Los ácidos grasos poliinsaturados tienen diversas funciones y efectos que incentivan la salud, por ejemplo en el caso de enfermedad coronaria, mecanismos de inflamación, nutrición infantil, etc. Véanse publicaciones y referencias bibliográficas, incluyendo citas bibliográficas allí citadas, en: Simopoulos, 1999, Am. J. Clin. Nutr. 70 (3. Suppl.): 560-569, Takahata y colaboradores, Biosci. Biotechnol. Biochem. 1998, 62(11): 2079-2085, Willich y Winther, 1995, Deutsche Medizinische Wochenschrift (Semanaario médico alemán) 120(7):229 y siguientes.

65

## II. Elementos y métodos de la invención

La presente invención se basa entre otras en el descubrimiento de moléculas nuevas que aquí se denominan moléculas de ácidos nucleicos y proteínas de desaturasa, las cuales ejercen un efecto en la producción de membranas celulares y lípidos en *Phaeodactylum tricornutum* y, por ejemplo, tienen un efecto en el movimiento de moléculas a través de estas membranas. En una forma de realización, las moléculas de desaturasa participan en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en organismos, tales como microorganismos y vegetales, o indirectamente afectan el transporte de moléculas por estas membranas. En una forma preferida de realización, la actividad de las moléculas de desaturasa de acuerdo con la invención para regular la producción de componentes de membrana y transporte de membrana tiene un efecto en la producción del producto químico fino deseado mediante este organismo. En una forma particularmente preferida de realización, la actividad de las moléculas de desaturasa de acuerdo con la invención se modula para que el rendimiento, la producción y/o la eficiencia de la producción de las rutas metabólicas de microorganismos o vegetales que regulan las desaturasas de la invención se modulen y se modifique la eficiencia de transporte de compuestos a través de estas membranas, lo cual directa o indirectamente modula el rendimiento, producción y/o eficiencia de la producción de un producto químico fino deseado mediante microorganismos y vegetales.

El término “desaturasa” o “polipéptido de desaturasa” comprende proteínas que participan en la desaturación de ácidos grasos. Ejemplos de desaturasas se divulgan en las SEQ ID NO: 1, 3, 5, 11 o sus homólogos, derivados o análogos. Los términos desaturasa o secuencia(s) de ácidos nucleicos de desaturasa comprenden secuencias de ácidos nucleicos que codifican una desaturasa y parte de las cuales puede ser una región codificante y también regiones correspondientes de secuencia no traducidas 5' y 3'. Ejemplos de genes de desaturasa son los que se muestran en las SEQ ID NO: 1, 3, 5 u 11. Los términos producción o productividad son conocidos en el campo técnico y abarcan la concentración del producto de fermentación (por ejemplo del producto químico fino deseado) que se forma dentro de un período específico y en un volumen específico de fermentación (por ejemplo kg de producto por hora por litro). El término eficiencia de producción abarca el tiempo requerido para lograr una cantidad de producción particular (por ejemplo, el tiempo requerido por la célula para establecer una tasa de rendimiento determinada de un producto químico fino). El término rendimiento o rendimiento de producto/carbono es conocido en el campo técnico y comprende la eficiencia de la transformación de la fuente de carbono en el producto (es decir, en el producto químico fino). Esto se expresa usualmente, por ejemplo, en kg de producto por kg de fuente de carbono. Incrementando el rendimiento o la producción del compuesto se incrementa la cantidad de las moléculas obtenidas o de las moléculas adecuadas de este compuesto obtenido en una cantidad determinada de cultivo por un período fijado. Los términos biosíntesis o ruta de biosíntesis son conocidos en el campo técnico y comprenden la síntesis de un compuesto, preferiblemente de un compuesto orgánico, por parte de una célula de compuestos intermedios, por ejemplo en un proceso de pasos múltiples fuertemente regulado. Los términos catabolismo o ruta catabólica son conocidos en el campo técnico y comprenden la descomposición de un compuesto, preferentemente de un compuesto orgánico, por parte de una célula en productos de catabolismo (dicho de manera general, moléculas más pequeñas o de menor complejidad) por ejemplo en un proceso de múltiples pasos, fuertemente regulado. El término metabolismo es conocido en el campo técnico y comprende la totalidad de las reacciones bioquímicas que tienen lugar en un organismo. El metabolismo de un compuesto determinado (por ejemplo, el metabolismo de un ácido graso) comprende entonces la totalidad de las rutas de biosíntesis, modificación y catabolismo de este compuesto en la célula, los cuales se refieren a este compuesto.

En otra forma de realización, las secuencias de ácido nucleico de acuerdo con la invención, que codifican moléculas de desaturasa, pueden modular la producción de una molécula deseada, como la de un producto químico fino, en un microorganismo o en vegetales. Existe una serie de mecanismos por los cuales la modificación de una secuencia de la invención puede afectar directamente el rendimiento, la producción y/o la eficiencia de producción de un producto químico fino desde un microorganismo o una cepa vegetal que comprende esta proteína modificada. La cantidad o actividad de desaturasas que participan en el transporte de moléculas de productos químicos finos dentro o hacia afuera de la célula puede elevarse para que se transporten cantidades mayores de estos compuestos por las membranas, a partir de las cuales pueden obtenerse y convertirse una en otra de manera más fácil. Adicionalmente, los ácidos grasos, la triacilglicerina y/o los lípidos son, por sí mismos, productos químicos finos deseables; optimizando la actividad o aumentando el número de una o más desaturasas según la invención, que participan en la biosíntesis de estos compuestos, o interfiriendo con la actividad de una o más desaturasas que participan en el catabolismo de estos compuestos, puede ser posible incrementar el rendimiento, la producción y/o la eficiencia de la producción de moléculas de ácido graso y de lípidos de organismos tales como microorganismos o vegetales.

La mutagénesis de las secuencias de ácidos nucleicos de acuerdo con la invención puede generar desaturasas con actividades modificadas que influyen indirectamente la producción de uno o más productos químicos finos deseados desde microorganismos o vegetales. Por ejemplo, las desaturasas de acuerdo con la invención que participan en la exportación de productos de desecho pueden exhibir un número mayor o actividad superior, de modo que los productos de desecho catabólicos de la célula (cuya cantidad puede incrementarse debido a la sobreproducción del producto químico fino deseado) se exportan de manera eficiente antes de que puedan dañar las moléculas en la célula (lo que reduciría la viabilidad de la célula) o interfieran con las rutas de biosíntesis de los productos químicos finos (lo que reduciría el rendimiento, la producción o la eficiencia de la producción de un producto químico fino deseado). Las cantidades intracelulares relativamente grandes de los productos químicos finos pueden, por sí mismas, ser tóxicas para la célula de modo que incrementar la actividad o número de los transportadores capaces de exportar estos compuestos desde las células da lugar a una viabilidad incrementada de la célula en el cultivo, lo que a su vez conduce a un número mayor de células en el cultivo que producen el producto químico fino deseado. Las desaturasas de acuerdo

## ES 2 345 324 T3

con la invención también pueden manipularse de tal manera que se producen las cantidades correspondientes de las diferentes moléculas de lípidos y las moléculas de ácido graso. Esto puede tener un efecto considerable en la concentración de lípido de la membrana celular. Puesto que cada tipo de lípido tiene propiedades físicas diferentes, una modificación de la composición de lípido de una membrana puede modificar significativamente la fluidez de la membrana. Las modificaciones de la fluidez de la membrana pueden afectar el transporte de moléculas a través de la membrana, así como a la integridad de la célula, lo cual tiene respectivamente un efecto considerable en la producción de productos químicos finos desde microorganismos y vegetales en un cultivo de fermentación a gran escala. Las membranas vegetales confieren propiedades específicas tales como tolerancia al calor, al frío, a la sal, a la sequía, así como tolerancia frente a patógenos, como bacterias y hongos. Por lo tanto, la modulación de los componentes de membrana puede tener un efecto crítico en la capacidad de las plantas para sobrevivir bajo los parámetros de estrés arriba mencionados. Esto puede efectuarse mediante modificaciones en señales de cascadas o directamente a través de la composición modificada de membrana (véase, por ejemplo: Chapman, 1998, Trends in Plant Science, 3(11): 419-426) y cascadas de señales (véase Wang 1999, Plant Physiology, 120: 645-651) o la tolerancia al frío, como se divulga en WO 95/18222.

Las secuencias aisladas de ácidos nucleicos de acuerdo con la invención se encuentran, por ejemplo, en el genoma de una cepa de *Phaeodactylum tricorutum* UTEX646 que se encuentra disponible por la colección de algas de la University of Texas, Austin.

La secuencia de nucleótidos del ADNc de *Phaeodactylum tricorutum* y las secuencias de aminoácidos derivadas de las desaturasas se indican en las SEQ ID NO: 1 a 6 así como en las 11 y 12. Se han realizado análisis por ordenador que clasifican y/o identifican estas secuencias de nucleótidos como secuencias que codifican proteínas que participan en el metabolismo de componentes de membrana celular o que participan en el transporte de compuestos por membranas celulares, o en la biosíntesis de PUFA. EST's con la entrada al banco de datos NO: PT001070010R y PT001078032R por los inventores representan las secuencias de acuerdo con la invención en SEQ ID NO: 1 y 3. La secuencia del fragmento de EST PT001070010R fue determinada y es tal como se representa en SEQ ID NO: 5. De manera análoga la secuencia del clon PT001078032R se representa en SEQ ID NO: 1. Se asignaron nombres de genes a los clones. Las abreviaturas significan: Pp = *Phycomitrella patens*, Pt = *Phaeodactylum tricorutum*. PT001070010R de la SEQ ID NO: 5 codifica para un nuevo gen homólogo a la desaturasa  $\Delta$ -12 y PT001078032R codifica para una desaturasa  $\Delta$ -5 novedosa. Pt\_des6 puede aislarse de acuerdo con el ejemplo 5a por medio de una reacción de cadena de polimerasa recurriendo a oligonucleótidos degenerados. Un fragmento obtenido así puede aislarse para clasificar un banco de ADNc de *Phaeodactylum tricorutum* y la región codificante de una desaturasa  $\Delta$ -6 de *Phaeodactylum tricorutum*. Un gen aislado de esta manera se denomina en la tabla 1 como Pt\_des6 y se representa en la SEQ ID NO: 3. Las secuencias correspondientes de aminoácidos se obtienen por traducción del código genético de la secuencia ID NO: 1, 3 y 5 y se definen como SEQ ID NO: 2, 4 y 6 (véase también Tabla 1). De la tabla 1 también puede inferirse otra secuencia de ácidos nucleicos que codifica para una desaturasa de  $\Delta$ -12. Esta lleva el número de clon PT001072031R.

TABLA 1

	Nombre del gen	Nombre del clon	Ácido nucleico SEQ ID NO:	Polipéptido SEQ ID NO:
Desaturasa D5	Pt_des5	PT001078032R	1	2
Desaturasa D6	Pt_des6	Pt_des6	3	4
Desaturasa D12	Pt_des12	PT001070010R	5	6
Desaturasa D6	Pp_des6	Pp_des6	7	8
Elongasa D6	Pp_PSE1	PP001019019F	9	10
Desaturasa $\Delta$ 12	Pt des12.2	PT001072013R	11	12

La presente invención también se refiere a proteínas con una secuencia de aminoácidos que es esencialmente homóloga a una secuencia de aminoácidos de la SEQ ID NO: 2, 4, 6 o 12. Tal como aquí se usa, una proteína con una secuencia de aminoácidos que es esencialmente homóloga a una secuencia de aminoácidos seleccionada tiene una homología de al menos aproximadamente 50% a la secuencia de aminoácidos seleccionada, por ejemplo la secuencias de aminoácidos completa seleccionada. Una proteína con una secuencia de aminoácidos, que es esencialmente homóloga a una secuencia seleccionada de aminoácidos, también puede tener una homología de al menos aproximadamente 50 a 60%, preferentemente de al menos aproximadamente 60 a 70% y más preferentemente de al menos aproximadamente 70 a 80%, 80 a 90% ó 90 a 95% y lo más preferible de al menos aproximadamente 96%, 97%, 98%, 99% ó más homología a una secuencia de aminoácidos.

La desaturasa de acuerdo con la invención o la parte biológicamente activa o el fragmento de la misma puede participar en el metabolismo de lípidos requeridos para la síntesis de membranas o lípidos de almacenamiento en microorganismos y puede, en combinación con otros genes, en particular con aquellos con actividad de elongasa, contribuir a actividades requeridas para la elongación de PUFAs de C<sub>18</sub> o C<sub>20-22</sub> de modo que se obtienen PUFAs de

## ES 2 345 324 T3

C<sub>18</sub>, C<sub>20</sub>, C<sub>22</sub> o C<sub>24</sub>, así como PUFAs relacionados. En este caso, las desaturaras de acuerdo con la invención pueden clonarse en combinación con elongasas y otras desaturaras en casetes de expresión de acuerdo con la invención y se emplean para la transformación de plantas con la ayuda de *Agrobacterium*.

5 Diversos aspectos de la invención se describen con más detalle en las siguientes sub-secciones.

### A. Moléculas de ácidos nucleicos

10 Una forma de realización de la invención son ácidos nucleicos que provienen de microorganismos que producen PUFA y codifican para polipéptidos que desaturan ácidos grasos de C<sub>18</sub> o C<sub>20-22</sub> con al menos uno, dos, tres o cuatro enlaces dobles en el ácido graso.

15 Otra forma de realización de acuerdo con la invención son ácidos nucleicos que comprenden secuencias de nucleótidos que codifican para polipéptidos que desaturan ácidos grasos C<sub>18</sub> o C<sub>20</sub> con al menos uno, dos, tres o cuatro enlaces dobles en el ácido graso y se seleccionan del grupo que se compone

20 a) una secuencia de ácidos nucleicos con la secuencia representada en las SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 ó SEQ ID NO: 11,

b) Secuencias de ácidos nucleicos que debido a la degeneración del código genético se obtienen mediante re-traducción de las secuencias de aminoácidos representadas en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 o SEQ ID NO: 12,

25 c) Derivados de las secuencias de ácidos nucleicos representadas en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11 que codifican para polipéptidos con las secuencias de aminoácidos representadas en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 o SEQ ID NO: 12 y presentan al menos 50% de homología a nivel de aminoácido sin que se reduzca esencialmente la acción enzimática de los polipéptidos.

30 El ácido nucleico de acuerdo con la invención arriba mencionado proviene de organismos como ciliados, hongos, algas o dinoflagelados que son capaces de sintetizar PUFAs, preferentemente de *Phaeodactylum tricorutum* o de organismos estrechamente relacionados.

35 Un aspecto de la invención se refiere a moléculas aisladas de ácido nucleico que codifican polipéptidos de desaturasa o partes biológicamente activas de las mismas, así como a fragmentos de ácido nucleico que son suficientes para usar como muestras de hibridación o cebadores para identificar o amplificar un ácido nucleico que codifica desaturasa (por ejemplo, ADN de desaturasa). El término "molécula de ácido nucleico", tal como aquí se usa, debe comprender moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de ARN (por ejemplo, ARNm) así como análogos de ADN o de ARN que se generan por medio de análogos de nucleótidos. Este término comprende además 40 la secuencia no traducida en los extremos 3' y 5' de la región de gen codificante: al menos 500, preferible 200, particularmente preferible 100 nucleótidos de la secuencia hacia arriba del extremo 5' de la región codificante y al menos 100, preferible 50, particularmente preferible 20 nucleótidos de la secuencia hacia abajo del extremo 3' de la región génica codificante. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria, pero preferentemente es ADN bicatenario. Una molécula "aislada" de ácido nucleico se separan de otras moléculas de ácido nucleicos que están presentes en la fuente natural del ácido nucleico. Un ácido nucleico "aislado" no tiene, preferentemente, secuencias que flanquean naturalmente el ácido nucleico en el ADN genómico del organismo del cual se deriva el ácido nucleico (por ejemplo, secuencias localizadas en los extremos 5' y 3' del ácido nucleico). En diversas formas de realización, la molécula de ácido nucleico aislada de desaturasa puede comprender, por ejemplo, menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ó 0,1 kb de secuencias de nucleótidos, que flanquean naturalmente la molécula de ácido 45 nucleico en el ADN genómico de la célula, a partir de la cual procede el ácido nucleico (por ejemplo, una célula de *Physcomitrella patens*). Una molécula "aislada" de ácido nucleico, como una molécula de ADNc, puede estar además esencialmente libre de otro material celular o medio de cultivo cuando se produce mediante técnicas recombinantes, o libres de precursores químicos u otros productos químicos cuando se sintetiza químicamente.

55 Una molécula de ácido nucleico de acuerdo con la invención, por ejemplo una molécula de ácido nucleico con una secuencia de nucleótidos de la SEQ ID NO: 1 o de una parte de la misma, puede aislarse usando técnicas estándar de biología molecular y la información de secuencia aquí proporcionada. Con ayuda de algoritmos de comparación también pueden identificarse una secuencia homóloga o regiones de secuencia homólogas, conservadas a nivel de ADN o de aminoácido. Por ejemplo, puede aislarse un ADNc de *Phaeodactylum tricorutum* de un banco de *Phaeodactylum tricorutum* usando las SEQ ID NO: 1, 3, 5 u 11 completas o una parte de las mismas como sondas de hibridación (como, por ejemplo, se describe en Sambrook y colaboradores, Molecular Cloning: A Laboratory Manual. 2.Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Además, puede aislarse una molécula de ácido nucleico que comprende una secuencia completa de SEQ ID NO: 1, 3, 5 u 11 o una 60 parte de la misma mediante reacción en cadena de polimerasa, y se usan cebadores de oligonucleótidos que se generan sobre la base de esta secuencia o partes de la misma, en particular regiones alrededor de motivos del Ejemplo 5a o modificaciones de los mismos aminoácidos individuales definidos (por ejemplo, puede aislarse una molécula de ácido nucleico, que comprende la secuencia completa de las SEQ ID NO: 1, 3, 5 u 11 o una parte de la misma, mediante reacción en cadena de polimerasa usando cebadores de oligonucleótidos los cuales se han generado a base de esta misma

## ES 2 345 324 T3

secuencia de la SEQ ID NO: 1, 3, 5 ó 11). Por ejemplo, ARNm puede prepararse de células (por ejemplo mediante el método de extracción de tiocianato de guanidinio de Chirgwin y colaboradores (1979) *Biochemistry* 18: 5294-5299) y ADNc por medio de transcriptasa inversa (por ejemplo transcriptasa inversa MLV de Moloney, disponible de Gibco/BRL, Bethesda, MD, o transcriptasa inversa de AMV, disponible de Seikagaku America, Inc., St. Petersburg, FL).

5 Cebadores sintéticos de oligonucleótidos para la amplificación por medio de reacción en cadena de polimerasa pueden generarse sobre la base de una de las secuencias mostradas en SEQ ID NO: 1, 3, 5 u 11, así como las secuencias mostradas en la Figura 5a o con ayuda de las secuencias de aminoácidos representadas en SEQ ID NO: 2, 4, 6 ó 12. Un ácido nucleico de acuerdo con la invención puede amplificarse usando ADNc o, de manera alterna, ADN genómico como matriz y cebadores adecuados de oligonucleótidos según técnicas de amplificación PCR estándares. El ácido nucleico amplificado de esta manera puede clonarse en un vector adecuado y caracterizarse mediante análisis de secuencia de ADN. Oligonucleótidos que corresponden a una secuencia de nucleótidos de desaturasa pueden producirse mediante procesos de síntesis estándar, por ejemplo con un sintetizador automáticos de ADN.

El ADNc mostrado en SEQ ID NO: 1,3, 5 u 11 comprende secuencias que codifican desaturasas (es decir, la “región codificante”) así como secuencias no traducidas 5’ y secuencias no traducidas 3’. De manera alterna, la molécula de ácido nucleico puede comprender solo la región codificante de una de las secuencias en SEQ ID NO: 1, 3, 5 u 11 o puede contener fragmentos genómicos enteros que se aíslan de ADN genómico.

En otra forma preferida de realización una molécula aislada de ácido nucleico de acuerdo con la invención comprende una molécula de ácido nucleico que es un complemento de una secuencia de nucleótidos mostrada en SEQ ID NO: 1, 3, 5 u 11 o de una parte de la misma. Una molécula de ácido nucleico, que es complementaria a una de las secuencias de nucleótidos mostradas en SEQ ID NO: 1, 3, 5 u 11, es luego suficientemente complementaria si puede hibridizarse con una de las secuencias indicadas en SEQ ID NO: 1, 3, 5 u 11 por lo cual se genera un dúplex estable.

Homólogos de las nuevas secuencias de ácidos nucleicos de desaturasa con la secuencia SEQ ID NO: 1, 3, 5 u 11 significan, por ejemplo, variantes alélicas con al menos aproximadamente 50 a 60%, preferentemente al menos aproximadamente 60 a 70%, más preferible al menos aproximadamente 70 a 80%, 80 a 90% o 90 a 95% y aún más preferible al menos aproximadamente 95%, 96%, 97%, 98%, 99% o más de homología a una de las secuencias de nucleótidos mostradas en SEQ ID NO: 1, 3, 5 u 11 o sus homólogos, derivados o análogos o partes de los mismos. En otra forma preferida de realización una molécula aislada de ácido nucleico de acuerdo con la invención comprende una secuencia de nucleótidos que hibrida en una de las secuencias de nucleótidos mostradas en SEQ ID NO: 1, 3, 5 u 11 o una parte de las mismas, por ejemplo en condiciones estrictas. Variantes alélicas comprenden particularmente variantes funcionales que pueden obtenerse mediante delección, inserción o sustitución de nucleótidos desde/a la secuencia mostrada en SEQ ID NO: 1, 3, 5 u 11, pero la intención es que la actividad enzimática de las proteínas sintetizadas a partir de ellas se mantenga ventajosamente para la inserción de uno o varios genes. Las proteínas que mantienen la actividad enzimática de desaturasa, es decir cuya actividad no se reduce esencialmente, significan proteínas con al menos 10%, preferentemente 20%, particularmente preferible 30%, muy particularmente preferible 40% de la actividad enzimática original, comparada con la proteína codificada por SEQ ID NO: 2, 4, 6 ó 12.

Homólogos de la SEQ ID NO: 1, 3, 5 u 11 también significan, por ejemplo, homólogos bacterianos, de hongos y vegetales, secuencias truncadas, ADN o ARN monocatenarios de la secuencia de ADN codificante y no codificante.

Homólogos de la SEQ ID NO: 1, 3, 5 u 11 también significan derivados tales como, por ejemplo, variantes de promotor. Los promotores hacia arriba de las secuencias indicadas de nucleótidos pueden modificarse mediante una o más sustituciones de nucleótidos, mediante inserción(es) y/o delección(es), sin que por esto, no obstante, se destruya la funcionalidad o actividad de los promotores. Adicionalmente es posible que la actividad de los promotores se incremente modificando su secuencia o reemplazándolos completamente por promotores más activos, incluso de organismos heterólogos.

Además, la molécula de ácido nucleico de acuerdo con la invención puede comprender sólo una parte de la región codificante de una de las secuencias en SEQ ID NO: 1, 3, 5 u 11, por ejemplo un fragmento que puede usarse como sonda o cebador, o un fragmento que codifica un segmento biológicamente activo de una desaturasa. Las secuencias de nucleótidos determinadas de clonar el gen de desaturasa de *Phaeodactylum tricornutum* hacen posible la generación de sondas y cebadores que están configuradas para identificar y/o clonar homólogos de desaturasa en otros tipos de células y organismos, así como homólogos de desaturasa de otras microalgas o especies relacionadas. La sonda/el cebador comprende usualmente un oligonucleótido purificado esencialmente. El oligonucleótido comprende usualmente una región de secuencia de nucleótidos que hibrida en condiciones severas a al menos aproximadamente 12, preferentemente aproximadamente 16, más preferible aproximadamente 25, 40, 50 ó 75 nucleótidos sucesivos de un cordón sense de una de las secuencias indicadas en SEQ ID NO: 1, 3, 5 u 11, de un cordón antisense de una de las secuencias indicadas en SEQ ID NO: 1, 3, 5 u 11 o sus homólogos, derivados o análogos o mutantes de los mismos de procedencia natural. Pueden usarse cebadores a base de una secuencia de nucleótidos de la SEQ ID NO: 1, 3, 5 u 11 en reacciones PCR para la clonación de homólogos de desaturasa. Pueden usarse sondas a base de secuencias de nucleótidos de desaturasa para detectar transcritos o secuencias genómicas que codifican una proteína igual u homóloga. En formas preferidas de realización la sonda comprende además un grupo de marcación enlazada a ella, por ejemplo un radioisótopo, un compuesto fluorescente, una enzima o un co-factor de enzima. Estas sondas pueden usarse como parte de un kit de ensayo para marcadores genómicos para identificar células que malexpresan una desaturasa, por ejemplo midiendo una cantidad de un ácido nucleico que codifica desaturasa en una muestra celular, por ejemplo midiendo el nivel de ARNm de desaturasa, o para determinar si un gen de desaturasa genómica se muta o de deleta.

En una forma de realización la molécula de ácido nucleico de acuerdo con la invención codifica una proteína o una parte de la misma que comprende una secuencia de aminoácidos, la cual es suficientemente homóloga a una secuencia de aminoácidos de la SEQ ID NO: 2, 4, 6 ó 12 como para que la proteína o la parte de la misma mantengan la capacidad de participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en microorganismos o vegetales o en el transporte de moléculas por estas membranas. Tal como aquí se usa, el término “suficientemente homólogo” se refiere a proteínas o a partes de la misma cuyas secuencias de aminoácidos tienen un número mínimo de residuos de aminoácidos idénticos o equivalentes (por ejemplo, un residuo de aminoácido con una cadena lateral similar, como un residuo de aminoácidos en una de las secuencias de la SEQ ID NO: 2) a una secuencia de aminoácidos de la SEQ ID NO: 2, de modo que la proteína o la parte de la misma pueden participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en microorganismos o plantas o en el transporte de moléculas a través de estas membranas. Tal como se describe aquí, los componentes de proteína de estas rutas metabólicas para componentes de membrana o sistemas de transporte de membranas pueden desempeñar un papel en la producción y secreción de uno o más productos químicos finos. Ejemplos de estas actividades también se describen aquí. Así, la “función de una desaturasa” contribuye ya sea directa o indirectamente al rendimiento, producción y/o eficiencia de producción de uno o más productos químicos finos. Se indican ejemplos de especificidad de sustrato de desaturasa de la actividad catalítica en tablas 5 y 6.

En otra forma de realización los derivados de la molécula de ácido nucleicos de acuerdo con la invención codifican proteínas con al menos aproximadamente 50 a 60%, preferentemente al menos aproximadamente 60 hasta 70% y más preferible al menos aproximadamente 70 hasta 80%, 80 hasta 90%, 90 hasta 95% y lo más preferible al menos aproximadamente 96%, 97%, 98%, 99% o más de homología a una secuencia completa de aminoácidos de la SEQ ID NO: 2. La homología de la secuencia de aminoácidos puede determinarse por toda la región de secuencia con el programa PileUp (J. Mol. Evolution., 25, 351-360, 1987, Higgins y colaboradores, CABIÓS, 5, 1989: 151-153) o BESTFIT o GAP (Henikoff, S. and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.).

Las partes de proteínas que se codifican por las moléculas de ácido nucleico de desaturasa de acuerdo con la invención son preferentemente partes activas biológicamente de una de las desaturasas. Tal como aquí se usa, el término “parte biológicamente activa de una desaturasa” debe comprender un segmento, por ejemplo un dominio/un motivo, de una desaturasa que puede participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en microorganismos o vegetales o en el transporte de moléculas a través de estas membranas o tiene una actividad indicada en las Tablas 5 y 6. Para determinar si una desaturasa o una parte biológicamente activa de la misma pueden participar en el metabolismo de compuestos requeridos para la síntesis de membranas celulares en microorganismos o vegetales o en el transporte de moléculas a través de estas membranas puede realizarse un ensayo de la actividad enzimática. Este proceso de ensayo, tal como se escribe detalladamente en el ejemplo 8 de de la sección de ejemplos, es corriente para el técnico en la materia. Fragmentos adicionales de ácidos nucleicos que codifican segmentos biológicamente activos de una desaturasa pueden generarse aislando parte de una de las secuencias en SEQ ID NO: 1, 3, 5 u 11, expresando el segmento codificado de la desaturasa o del péptido (por ejemplo mediante expresión recombinante *in vitro*) y determinando la actividad de la parte codificada de la desaturasa o del péptido.

La invención comprende además moléculas de ácidos nucleicos que se diferencian de una de las secuencias de nucleótidos mostradas en SEQ ID NO: 1, 3, 5 u 11 (y de partes de las mismas) debido a la degeneración del código genético y de este modo, las desaturasas iguales codifican como aquella que se codifica por las secuencias de nucleótidos mostradas en SEQ ID NO: 1, 3, 5 u 11. En otra forma de realización una molécula aislada de ácido nucleico de acuerdo con la invención tiene una secuencia de nucleótidos que codifica para una proteína con una secuencia de aminoácidos mostrada en SEQ ID NO: 2, 4, 6 ó 12. En otra forma de realización la molécula de ácido nucleico de acuerdo con la invención codifica una proteína de desaturasa de longitud completa, la cual es esencialmente homóloga a una secuencia de aminoácido de la SEQ ID NO: 2,4, 6 ó 12 (la cual se codifica por un marco de lectura abierto mostrado en SEQ ID NO: 1, 3, 5 u 11) y la cual puede identificarse y aislarse mediante métodos habituales.

En adición a las secuencias de nucleótidos de desaturasa mostradas en SEQ ID NO: 1, 3, 5 u 11 el técnico en la materia reconoce que pueden existir polimorfismos de secuencia de ADN que conducen a cambios en las secuencias de aminoácidos de las desaturasas dentro de una población (por ejemplo la población de *Phaeodactylum tricorutum*). Estos polimorfismos genéticos en el gen de desaturasa pueden existir entre individuos dentro de una población debido a la variación natural. Tal como se usa aquí, los términos “gen” y “gen recombinante” significan moléculas de ácidos nucleicos con un marco de lectura que codifica una desaturasa, preferentemente una desaturasa de *Phaeodactylum tricorutum*. Estas variantes naturales causan usualmente una varianza de 1 hasta 5% en la secuencia de nucleótidos del gen de desaturasa. Todas estas variaciones de nucleótidos y los polimorfismos de aminoácidos resultantes de las mismas en la desaturasa, los cuales son el resultado de variación natural y no modifican la actividad funcional de desaturasas deben estar contenidos en el alcance de la invención.

Pueden aislarse moléculas de ácidos nucleicos que corresponden a variantes naturales y no homólogos, derivados o análogos de *Phaeodactylum tricorutum*, del ADNc de *Phaeodactylum tricorutum*, debido a su homología con el ácido nucleico de desaturasa de *Phaeodactylum tricorutum* usando el ADNc de *Phaeodactylum tricorutum* o de una parte del mismo como sonda de hibridación según técnicas estándares de hibridación en condiciones estrictas de hibridación. En otra forma de realización, una molécula aislada de ácido nucleico de acuerdo con la invención tiene una longitud mínima de 15 nucleótidos e hibrida en condiciones estrictas a una molécula de ácido nucleico que comprende una secuencia de nucleótidos de SEQ ID NO: 1, 3, 5 u 11. En otras formas de realización, el ácido nucleico

## ES 2 345 324 T3

tiene una longitud de al menos 25, 50, 100, 250 o más nucleótidos. El término “hibrida en condiciones estrictas”, tal como se usa aquí, debe describir condiciones de hibridación y de lavado en las que las secuencias de nucleótidos, que son homólogas unas a otras en al menos 60%, usualmente permanecen hibridadas unas en otras. Las condiciones son preferentemente de tal tipo que las secuencias que tienen una homología de al menos aproximadamente 65%, más preferible de al menos aproximadamente 70% y aún más preferible de al menos aproximadamente 75% o más entre ellas, usualmente se mantienen hibridadas unas a otras. Estas condiciones severas son conocidas para el técnico en la materia y pueden encontrarse en Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6. Un ejemplo preferido, no limitante, de condiciones severas de hibridación son hibridaciones en 6 x cloruro de sodio/citrato de sodio (sodium chloride/sodium citrate = SSC) a aproximadamente 45°C, seguidas de uno o más pasos de lavado en 0,2 x SSC, 0,1% SDS a 50 hasta 65°C. El técnico en la materia conoce que estas condiciones de hibridación se diferencian según el tipo del ácido nucleico y cuando se encuentran presentes solventes orgánicos, por ejemplo, respecto de la temperatura y de la concentración del búfer. La temperatura se diferencia, por ejemplo, en “condiciones estándar de hibridación” según el tipo del ácido nucleico entre 42°C y 58°C en búfer acuoso con una concentración de 0,1 hasta 5 x SSC (pH 7,2). Si está presente solvente orgánico en el búfer arriba mencionado, por ejemplo formamida al 50%, en condiciones estándar la temperatura es de aproximadamente 42°C. Preferentemente, las condiciones de hibridación para híbridos de ADN:ADN son de, por ejemplo, 0,1 x SSC y 20°C hasta 45°C, preferentemente entre 30°C y 45°C. las condiciones de hibridación para híbridos ADN:ARN son preferentemente de, por ejemplo, 0,1 x SSC y 30°C hasta 55°C, preferentemente entre 45°C y 55°C. Las temperaturas de hibridación mencionadas previamente se determinan, por ejemplo, para un ácido nucleico con aproximadamente 100 bp (= pares de bases) de longitud y un contenido G + C de 50% en ausencia de formamida. El técnico en la materia sabe cómo pueden determinarse las condiciones de hibridación requeridas con referencia a libros de enseñanza como el mencionado previamente o de los siguientes libros de enseñanza: Sambrook y colaboradores, “Molecular Cloning”, Cold Spring Harbor Laboratory, 1989; Hames y Higgins (editores) 1985, “Nucleic Acids Hybridization: A Practical Approach”, IRL Press at Oxford University Press, Oxford; Brown (editores) 1991, “Essential Molecular Biology: A Practical Approach”, IRL Press at Oxford University Press, Oxford.

Una molécula aislada de ácido nucleico de acuerdo con la invención, que hibrida en condiciones severas a una secuencia de la SEQ ID NO: 1, 3, 5 u 11, corresponde preferentemente a una molécula de ácido nucleico de procedencia natural. Tal como se usa aquí, una molécula de “procedencia natural” se refiere a una molécula de ARN o ADN con una secuencia de nucleótidos que tiene lugar en la naturaleza (por ejemplo, que codifica una proteína natural). En una forma de realización, el ácido nucleico codifica una desaturasa de *Phaeodactylum tricorutum* de procedencia natural.

Adicionalmente a las variantes de procedencia natural de la secuencia de desaturasa que pueden existir en la población, el técnico en la materia reconoce además que también pueden introducirse cambios por medio de mutación a la secuencia de nucleótidos de la SEQ ID NO: 1, 3, 5 u 11, lo cual conduce a cambios de la secuencia de aminoácidos de la desaturasa codificada, si que se perjudique la capacidad funcional de la proteína de desaturasa. Por ejemplo, pueden producirse sustituciones de nucleótidos que conducen a sustituciones de aminoácidos en residuos de aminoácidos “no esenciales” en una secuencia de las SEQ ID NO: 2, 4, 6 ó 12. Un residuo de aminoácido “no esencial” es un residuo que puede modificarse en una secuencia del tipo silvestre de una de las desaturasas (SEQ ID NO: 2, 4, 6 ó 12) sin que se modifique, es decir sin que se reduzca esencialmente, la actividad de la desaturasa, mientras que para la actividad de desaturasa se requiere un residuo de aminoácido “esencial”. Otros residuos de aminoácidos (por ejemplo, aquellos que no se conservan con actividad de desaturasa en el dominio o solo se semiconservan) pueden, sin embargo, no ser esenciales para la actividad y pueden, por lo tanto, modificarse sin modificar la actividad de desaturasa.

Por consiguiente, otro aspecto de la invención se refiere a moléculas de ácido nucleico que codifican desaturasas que comprenden residuos modificados de aminoácidos que no son esenciales para la actividad de desaturasa. Estas desaturasas se diferencian en la secuencia de aminoácidos de una secuencia en SEQ ID NO: 2, 4, 6 ó 12 y mantienen no obstante al menos una de las actividades de desaturasa aquí descritas. La molécula aislada de ácido nucleico comprende en una forma de realización una secuencia de nucleótidos que codifica una proteína y la proteína comprende una secuencia de aminoácidos con al menos aproximadamente 50% de homología a una secuencia de aminoácidos de la SEQ ID NO: 2, 4, 6 ó 12 y puede participar en el metabolismo de compuestos para la síntesis de membranas celulares en *Phaeodactylum tricorutum* o en el transporte de moléculas a través de estas membranas. La proteína codificada por la molécula de ácido nucleico tiene una homología de preferentemente al menos aproximadamente 50 hasta 60% a una de las secuencias en SEQ ID NO: 2, 4, 6 ó 12, más preferiblemente de al menos aproximadamente 60 hasta 70% de homología a una de las secuencias en SEQ ID NO: 2, 4, 6 ó 12, aún más preferible de al menos aproximadamente 70 hasta 80%, 80 hasta 90%, 90 hasta 95% de homología a una de las secuencias en SEQ ID NO: 2, 4, 6 ó 12 y lo más preferible de al menos aproximadamente 96%, 97%, 98% ó 99% de homología a una de las secuencias en SEQ ID NO: 2, 4, 6 ó 12.

Para la determinación de la homología porcentual de dos secuencias de aminoácidos (por ejemplo de una de las secuencias de la SEQ ID NO: 2, 4, 6 ó 12 y una forma mutada de la misma) o de dos ácidos nucleicos se escriben las secuencias una debajo de la otra con el propósito de comparar óptimamente (por ejemplo, pueden introducirse espacios vacíos en la secuencia de una proteína o de un ácido nucleico para generar una alineación óptima con la otra proteína o el otro ácido nucleico). Los residuos de aminoácidos o nucleótidos se comparan luego en las posiciones correspondientes de aminoácidos o nucleótidos. Si una posición en una secuencia (por ejemplo de una de las secuencias de la SEQ ID NO: 2, 4, 6 ó 12) se ocupa por el mismo residuo de aminoácidos o el mismo nucleótido que la posición correspondiente en la otra secuencia (por ejemplo en una forma mutada de la secuencia seleccionada de SEQ ID NO:

## ES 2 345 324 T3

2, 4, 6 ó 12), entonces las moléculas son homólogas en esta posición (es decir, “homología” de aminoácido o de ácido nucleico, tal como aquí se usa, corresponde a “identidad” de aminoácido o de ácido nucleico). La homología porcentual entre ambas secuencias es una función del número de posiciones idénticas que son comunes a las secuencias (es decir homología % = número de las posiciones idénticas/número total de las posiciones x 100). Los términos homología e identidad deben considerarse como sinónimos.

Una molécula aislada de ácido nucleico que codifica una desaturasa que es homóloga a una secuencia de proteína de las SEQ ID NO: 2, 4, 6 ó 12 puede generarse introduciendo una o más sustituciones, adiciones o deleciones de nucleótidos en una secuencia de nucleótidos de las SEQ ID NO: 1, 3, 5 u 11 de modo que una o varias sustituciones, adiciones o deleciones de aminoácidos se introducen en la proteína codificada. Pueden introducirse mutaciones en una de las secuencias de las SEQ ID NO: 1, 3, 5 u 11 mediante técnicas estándar, como mutagénesis específica de la posición y mutagénesis mediada por PCR. Preferentemente se producen sustituciones conservativas de aminoácidos en uno o más de los residuos de aminoácidos no esenciales previamente dichos. En una “sustitución conservativa de aminoácidos” el residuo de aminoácido se intercambia por un residuo de aminoácido con una cadena lateral similar. En el campo técnico se han definido familias de residuos de aminoácidos con cadenas laterales similares. Estas familias comprenden aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas polares descargadas (por ejemplo glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales apolares, (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano), cadenas laterales beta-ramificadas (por ejemplo treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptofano, histidina). Un residuo de aminoácido no esencial previamente dicho en una desaturasa se intercambia de esta manera, preferentemente, por otro residuo de aminoácido de la misma familia de cadena lateral. De manera alterna, en otra forma de realización, las mutaciones pueden introducirse aleatoriamente por toda la secuencia que codifica desaturasa o por una parte de ella, por ejemplo mediante mutagénesis de saturación, y los mutantes resultantes pueden analizarse en cuanto a la actividad de desaturasa para identificar mutantes que mantienen actividad de desaturasa. Después de la mutagénesis de una de las secuencias de la SEQ ID NO: 1, 3, 5 u 11, proteína codificada puede expresarse de manera recombinante, y puede determinarse la actividad de la proteína, usando por ejemplo los ensayos descritos aquí (véase la sección de ejemplos).

Adicionalmente a las moléculas de ácido nucleico que codifican las desaturasas descritas previamente, otro aspecto de la invención se refiere a moléculas aisladas de ácidos nucleicos que son “antisense” para las secuencias de ácidos nucleicos de acuerdo con la invención. Un ácido nucleico “antisense” comprende una secuencia de nucleótidos que es complementaria a un ácido nucleico “sense” que codifica una proteína, por ejemplo complementaria al cordón codificante de una molécula de ADNc bicatenaria o complementaria a una secuencia de ARNm. Un ácido nucleico antisense puede enlazarse, por lo tanto, mediante enlaces de puente de hidrógeno a un ácido nucleico. El ácido nucleico antisense puede ser complementario a un cordón completo que codifica desaturasa o solo a una parte del mismo. En una forma de realización una molécula de ácido nucleico antisense es “antisense” a una “región codificante” del cordón codificante de una secuencia de nucleótidos que codifica una desaturasa. La expresión “región codificante” se refiere a la región de la Secuencia de nucleótidos que comprende codones que se traducen a residuos de aminoácidos (por ejemplo, la región codificante completa que inicia y finaliza con el stop codón, es decir el último codón antes del stopcodón). En otra forma de realización, la molécula de ácido nucleico antisense es “antisense” a una “región no codificante” del cordón codificante de una secuencia de nucleótidos que codifica la desaturasa. La expresión “región no codificante” se refiere a secuencias 5' y 3' que flanquean la región codificante y no se traducen en aminoácidos (es decir, las que también se denominan como regiones no traducidas 5' y 3').

Condicionando las secuencias aquí divulgadas del cordón codificante que codifican desaturasas (por ejemplo, las secuencias representadas en SEQ ID NO: 1, 3, 5 u 11) pueden configurarse ácidos nucleicos antisense de acuerdo con la invención de conformidad con las reglas de apareamiento de bases de Watson-Crick. La molécula de ácido nucleico antisense puede ser complementaria a toda la región codificante de desaturasa-ARNm, pero se prefiere más un oligonucleótido que es complementario solo a una parte de la región codificante o no codificante de ARNm-desaturasa “antisense”. El oligonucleótido antisense puede ser complementario, por ejemplo, a la región que circunda la posición de inicio de traducción de ARNm-desaturasa. Un oligonucleótido antisense puede tener, por ejemplo, una longitud de aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45 ó 50 y más nucleótidos. Un ácido nucleico antisense de acuerdo con la invención puede estructurarse mediante procesos conocidos en la técnica usando síntesis química y reacciones de enlazamiento enzimático. Un ácido nucleico antisense (por ejemplo un oligonucleótido antisense) puede, por ejemplo, sintetizarse químicamente usando nucleótidos que tienen lugar naturalmente o nucleótidos modificados de manera variada, los cuales están configurados de tal manera que incrementan la estabilidad biológica de las moléculas o incrementan la estabilidad física del dúplex formado entre el ácido nucleico antisense y el sense; por ejemplo, pueden usarse los derivados de fosforotioato y los nucleótidos sustituidos con acridina. Ejemplos de nucleótidos modificados que pueden usarse para la generación de ácido nucleico antisense son, entre otros, 5-fluoruracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroximetil)uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopentiladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metil-N6-isopentiladenina, ácido uracil-5-oxiacético (v), wybutoxosina, desaturasaudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico de ácido uracil-5-oxiacético, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w y 2,6-diaminopurina. El ácido nucleico antisense puede producirse

de manera alterna biológicamente usando un vector de expresión que el que se ha subclonado un ácido nucleico en dirección antisense (es decir que el ARN que se transcribe por parte del ácido nucleico introducido se orienta hacia un ácido nucleico objetivo de interés en dirección antisense, lo cual se describe en la sub-sección que sigue a continuación).

5

Las moléculas de ácido nucleico antisense de acuerdo con la invención usualmente se administran a una célula o se generan *in situ* de modo que hibridan con, o se enlazan con, el ARNm celular y/o el ADN genómico que codifica una desaturasa, inhibiendo así la expresión de la proteína, por ejemplo inhibiendo transcripción y/o traducción. La hibridación puede efectuarse mediante complementariedad de nucleótido convencional con la formación de un dúplex estable o, por ejemplo en el caso de una molécula de ácido nucleico antisense que enlaza duplicados de ADN, mediante interacciones específicas en el surco mayor de la hélice doble. La molécula antisense puede modificarse de tal manera que se enlace específicamente con un receptor o con un antígeno expresado en la superficie celular seleccionada, enlazándose por ejemplo la molécula del ácido nucleico antisense con un péptido o un anticuerpo que se enlaza a un receptor de superficie celular o a un antígeno. La molécula de ácido nucleico antisense también puede introducirse a las células usando los vectores aquí descritos. Para alcanzar concentraciones intracelulares suficientes de la molécula antisense se prefiere constructos de vectores en los que la molécula de ácido nucleico antisense se encuentra bajo el control de un promotor procariótico, viral o eucariótico, incluso vegetal, fuerte.

En otra forma de realización, la molécula antisense de ácido nucleico de acuerdo con la invención es una molécula de ácido nucleico  $\alpha$ -anomérica. Una molécula de ácido nucleico  $\alpha$ -anomérica forma híbridos bicatenarios específicos con ARN complementario y los cordones corren paralelos unos a otros en contraste con las habituales  $\beta$ -unidades. (Gaultier y colaboradores (1987) *Nucleic Acids Res.* 15: 6625-6641). La molécula de ácido nucleico antisense puede abarcar además un 2'-o-metilribonucleótido (Inoue y colaboradores (1987) *Nucleic Acids Res.* 15: 6131-6148) o análogos quiméricos de ARN-ADN (Inoue y colaboradores (1987) *FEBS Lett.* 215: 327-330).

25

En otra forma de realización, un ácido nucleico antisense de acuerdo con la invención es una ribozima. Ribozimas son moléculas catalíticas de ARN con actividad de ribonucleasa que pueden descomponer un ácido nucleico monocatenario, como un ARNm, hacia el cual tienen una región complementaria. De esta manera, los ribozimas (por ejemplo, ribozimas Hammerhead o cabeza de martillo (descritos en Haselhoff y Gerlach (1988) *Nature* 334: 585-591)) pueden usarse para la descomposición catalítica de transcriptos de ARNm-desaturasa, con el fin de inhibir la traducción de ARNm-desaturasa. Una ribozima con especificidad para un ácido nucleico que codifica desaturasa puede configurarse a base de la secuencia de uno de los nucleótidos de un ADNc-desaturasa divulgados en SEQ ID NO: 1, 3, 5 u 11 (es decir, a base de una secuencia heteróloga a aislar de acuerdo con los procesos enseñados en esta invención. Por ejemplo, puede construirse un derivado de un ARN Tetrahimena-L-19-IVS donde la secuencia de nucleótidos de la posición activa es complementaria a la secuencia de nucleótidos que debe descomponerse en un ARNm que codifica desaturasa. Véase, por ejemplo, Cech y colaboradores, patente de Estados Unidos No. 4,987,071 y Cech y colaboradores, patente de Estados Unidos No. 5,116,742. De manera alterna puede usarse ARNm-desaturasa para la selección de un ARN catalítico con una actividad específica de ribonucleasa de un acervo (pool) de moléculas de ARN. Véase, por ejemplo, Bartel, D., y Szostak, J.W. (1993) *Science* 261: 1411-1418.

40

De manera alterna puede inhibirse la expresión de gen-desaturasa dirigiendo las secuencias de nucleótidos que son complementarias hacia la región regulatoria de una secuencia de nucleótidos -desaturasa (por ejemplo, a un promotor-y/o enhancer-desaturasa) de tal manera que se formen estructuras de hélices triples que inhiban la transcripción de un gen-desaturasa. Véase, en general, Helene, C. (1991) *Anticancer Drug Res.* 6(6) 569-84; Helene, C., y colaboradores (1992) *Ann. N. Y. Acad. Sci.* 660: 27-36; y Maher, L.J. (1992) *Bioassays* 14(12): 807-815.

45

#### B. Constructo génico (= constructo o fragmento de ácido nucleico o casete de expresión)

Por casetes de expresión de acuerdo con la invención deben entenderse las secuencias nombradas en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11 que son el resultado del código genético y/o sus derivados funcionales o no funcionales que se han enlazado ventajosamente de manera funcional con una o más señales regulatorias para incrementar la expresión génica y que controlan ventajosamente la expresión de la secuencia codificante en la célula huésped. Estas secuencias regulatorias deben hacer posible la expresión dirigida de los genes y de la expresión de proteína. Dependiendo del organismo huésped esto puede significar, por ejemplo, que el gen se expresa y/o sobre-expresa solo después de la inducción o que se expresa y/o sobre-expresa inmediatamente. Por ejemplo, estas secuencias regulatorias son secuencias a las que se enlazan inductores o represores y regulan así la expresión del ácido nucleico. Adicionalmente a estas nuevas secuencias de regulación o en lugar de estas secuencias, la regulación natural de estas secuencias aún puede estar presente antes de los propios genes estructurales y pueden opcionalmente haberse modificado de modo se haya eliminado la regulación natural y se haya incrementado la expresión de los genes. El constructo génico puede, sin embargo, tener una estructura más simple, es decir que no deben haberse insertado señales regulatorias adicionales antes de la secuencia de ácido nucleico o sus derivados y no se ha removido el promotor natural junto con su regulación. En vez de esto, la secuencia regulatoria natural ha mutado de tal manera que ya no se efectúa la regulación y/o se aumenta la expresión génica. Estos promotores modificados también pueden introducirse en forma de secuencias parciales (= promotor con partes de la secuencia de ácidos nucleicos según la invención) solas antes del gen natural para incrementar la actividad. Además, el constructo génico también puede contener de manera ventajosa una o varias de las llamadas "secuencias de enhancer" vinculadas funcionalmente con el promotor, las cuales hacen posible una expresión elevada del ácido nucleico. También es posible insertar secuencias adicionales

65

## ES 2 345 324 T3

ventajosas en el extremo 3' de las secuencias de ADN, así como elementos regulatorios o terminadores. Los genes de desaturasa  $\Delta 5$ /desaturasa  $\Delta 6$  y/o desaturasa  $\Delta 12$  pueden estar contenidas en una o más copias en el casete de expresión (= constructo génico).

5 En este caso, las secuencias regulatorias o factores pueden preferiblemente influir positivamente y elevar así la expresión génica de los genes introducidos, tal como se ha descrito arriba. De este modo puede efectuarse ventajosamente un refuerzo de los elementos regulatorios al nivel de la transcripción usando señales fuertes de transcripción como promotores y/o "enhancers". Además, también es posible un refuerzo de la traducción, por ejemplo mejorando la estabilidad del ARNm.

10 Otra forma de realización de la invención son uno o más constructos génicos que contienen una o más secuencias que se definen por SEQ ID NO: 1, 3, 5, 7, 9 u 11 y codifican polipéptidos según SEQ ID NO: 2, 4, 6, 8, 10 ó 12. En tal caso, las SEQ ID NO: 1, 3, 5, 7 y 11 se derivan de desaturasas mientras que la SEQ ID NO: 9 codifica para una elongasa. Desaturasas codifican enzimas que introducen un enlace doble en posición  $\Delta 5$ ,  $\Delta 6$  o  $\Delta 12$  y el sustrato tiene uno, dos, tres o cuatro enlaces dobles. La secuencia representada en SEQ ID NO: 9 codifica para una actividad enzimática que alarga un ácido graso en al menos dos átomos de carbono, y sus homólogos, derivados o análogos que están vinculados funcionalmente a una o más señales regulatorias para incrementar ventajosamente la expresión de genes. Ejemplos de estas secuencias de regulación son secuencias a las que se enlazan inductores o represores y regulan así la expresión del ácido nucleico. Adicionalmente a estas nuevas secuencias de regulación, aún puede estar presente la regulación natural de estas secuencias antes de los propios genes de estructura y, si es adecuado, pueden haber sido modificada genéticamente de modo que la regulación natural ha sido eliminada y la expresión de los genes se ha incrementado.

25 El constructo génico también puede tener, sin embargo, una estructura más simple, es decir que no se han insertado señales de regulación adicionales antes de la secuencia SEQ ID NO: 1, 3, 5 u 11 o sus homólogos y no se ha deletado el promotor natural con su regulación. En lugar de esto se ha mutado la secuencia de regulación natural de tal manera que ya no tiene lugar una regulación y se intensifica la expresión génica. El constructo génico puede comprender además ventajosamente una o más de las llamadas secuencias enhancer que están enlazadas funcionalmente con el promotor y hacen posible la expresión incrementada de la secuencia de ácidos nucleicos. También es posible insertar adicionalmente secuencias ventajosas en el extremo 3' de la secuencia de ADN, por ejemplo otros elementos de regulación o terminadores. Los genes de desaturasa y los genes de elongasa pueden estar presentes en el constructo génico en una o más copias. Pueden presentarse en un constructo génico o más constructos génicos. Este constructo génico o los constructos génicos pueden expresarse juntos en el organismo huésped. En tal caso, el constructo génico, o los constructos génicos, puede insertarse en uno o varios vectores y estar presente libre en la célula o, no obstante, insertarse en el genoma. Es ventajoso para la inserción de otros genes en organismos si otros genes están presentes en el constructo génico.

40 Secuencias ventajosas de regulación para el nuevo proceso se encuentran presentes, por ejemplo, en promotores como el promotor *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacI<sup>q-</sup>*, T7, T5, T3, *gal*, *trc*, *ara*, SP6,  $\lambda$ -P<sub>R</sub>- o  $\lambda$ -P<sub>L</sub> y se aplican ventajosamente en bacterias Gram-negativas. Otras secuencias de regulación ventajosas están presentes, por ejemplo, en los promotores Gram-positivos *amy* y SPO2, en los promotores de levaduras u hongos ADC1, MF $\alpha$ , AC, P-60, CYC1, GAPDH, TEF, *rp28*, ADH o en los promotores vegetales CaMV/35S [Franck y colaboradores, Cell 21 (1980) 285-294], PRP1 [Ward y colaboradores, Plant. Mol. Biol. 22 (1993)], SSU, OCS, *lib4*, *usp*, STLS1, B33, nos o promotor en la ubiquitina o faseolina. En este contexto también son ventajosos los promotores inducibles como los promotores descritos en EP-A-0 388 186 (inducible por benzilsulfonamida), Plant J. 2, 1992: 397-404 (Gatz y colaboradores, inducible por tetraciclina), EP-A-0 335 528 (inducible por ácido abscísico) o WO 93/21334 (inducible por etanol o ciclohexanol). Otros promotores vegetales adecuados son el promotor de la fase FB citosólica o el promotor ST-LSI de la patata (Stockhaus y colaboradores, EMBO J. 8, 1989, 2445), el promotor de fosforibosilpirofosfatamidotransferasa de glicina *max* (Número de acceso del banco genético U87999) o el promotor específico de nodos descrito en EP-A-0 249 676. Promotores particularmente ventajosos son promotores que hace posible la expresión en tejidos que participan en la biosíntesis de ácido graso. Son muy particularmente ventajosos los promotores específicos de semillas, como los promotores USO de acuerdo con la descripción pero también otros promotores como el LeB4 (Baeumlein y colaboradores, Plant J., 1992, 2 (2): 233-239), DC3 (Thomas, Plant Cell 1996, 263: 359-368), promotor faseolina o napina. Otros promotores particularmente ventajosos son promotores específicos de semillas que pueden usarse para plantas monocotiledóneas o dicotiledóneas y se describen en US 5,608,152 (promotor-napina de colza), WO 98/45461 (promotor oleosina de *arabidopsis*), US 5,504,200 (promotor faseolina de *Faseolus vulgaris*), WO 91/13980 (promotor Bce4 de brassica), de Baeumlein y colaboradores, Plant J., 1992, 2 (2): 233-239 (promotor LeB4 de una leguminosa); estos promotores son adecuados para dicotiledóneas. Los siguientes promotores son adecuados, por ejemplo, para monocotiledóneas promotor *1pt-2* o *1pt-1* de cebada (WO 95/15389 y WO 95/23230), promotor Hordein de cebada y otros promotores adecuados descritos en WO 99/16890.

En teoría para el nuevo proceso es posible usar todos los promotores naturales con sus secuencias de regulación como las arriba mencionadas. Así mismo es posible y ventajoso usar adicionalmente promotores sintéticos.

65 Tal como se describe arriba, el constructo génico también puede comprender otros genes que deben introducirse en los organismos. Es posible y ventajoso introducir a los organismos huésped, y expresar allí, genes regulatorios tales como genes para inductores, represores o enzimas que, debido a la actividad enzimática, intervienen en la regulación de uno o más genes de una ruta biosintética. Estos genes pueden ser de origen heterólogo u homólogo. Además, en

el constructo de ácido nucleico o el constructo génico pueden estar contenidos ventajosamente otros genes de biosíntesis del metabolismo de ácido graso o de lípido o, no obstante, estos genes pueden encontrarse en otro y otros varios constructos de ácido nucleico. Como genes de biosíntesis del metabolismo de ácido graso o de lípido se usa un gen seleccionado ventajosamente del grupo de acil-CoA-dehidrogenasa(s), acil-ACP[= acyl carrier protein]-desaturasa(s), acil-ACP-tioesterasa(s), ácido graso-aciltransferasa(s), ácido graso-sintasa(s), ácido graso-hidroxilasa(s), acetil-coenzima A-carboxilasa(s), acil-coenzima, A-oxidasa(s), ácido graso-desaturasa(s), ácido graso-acetilenasas, lipoxigenasas, triacilglicerol-lipasas, alenóxido-sintasas, hidroperóxido-lipasas o ácido graso-elongasa(s) o sus combinaciones. Para expresar los otros genes presentes los constructos génicos comprenden ventajosamente otras secuencias de regulación 3'- y/o 5'-terminales para el incremento de la expresión, los cuales se seleccionan dependiendo del organismo huésped elegido y del gen o genes para la expresión óptima. Como ya se mencionó, estas secuencias de regulación deben hacer posible la expresión de los genes y la expresión de proteína. Dependiendo del organismo huésped, esto puede significar que el gen se expresa o sobre-expresa solo después de inducción o que se expresa y/o sobre-expresa inmediatamente.

Las secuencias o factores de regulación pueden tener además un efecto ventajoso en la expresión de los genes introducidos y de esta manera incrementarlos. De este modo es posible que los elementos de regulación se refuercen usando señales más fuertes de transcripción, tales como promotores y/o enhancers, ventajosamente a nivel de transcripción. Sin embargo, también es posible reforzar la traducción, por ejemplo mejorando la estabilidad de ARNm.

### C. Vectores de expresión recombinantes y células huésped

Otro aspecto de la invención se refiere a vectores, preferentemente vectores de expresión que contienen un ácido nucleico que codifica una desaturasa sola (o una parte de la misma) o un constructo de ácido nucleico descrito bajo el punto B en el que el ácido nucleico de acuerdo con la invención está contenido solo o en combinación con otros genes de biosíntesis del metabolismo de ácido graso o de lípido, tales como desaturasas o elongasas. Tal como aquí se usa, el término "vector" se refiere a una molécula de ácido nucleico que puede transportar otro ácido nucleico al cual está enlazado. Un tipo de vector es un "plásmido", el cual representa un bucle de ADN bicatenario circular al cual pueden ligarse segmentos adicionales de ADN. Otro tipo de vector es un vector viral y es posible que segmentos adicionales de ADN se ligan al genoma viral. Determinados vectores pueden replicarse autónomamente en una célula huésped a la que se han introducido (por ejemplo, vectores de bacterias con origen bacteriano de replicación y vectores episomales de mamíferos). Otros vectores (por ejemplo vectores de mamíferos no episomales) se integran al genoma de una célula huésped al introducirse a la célula huésped y de esa manera se replican con el genoma huésped. Además, determinados vectores pueden controlar la expresión de genes con los que están enlazados funcionalmente. Estos vectores se denominan aquí "vectores de expresión". Habitualmente, los vectores de expresión que son adecuados para técnicas de recombinación de ADN tienen la forma de plásmidos. En la presente descripción pueden intercambiarse "plásmido" y "vector" puesto que el plásmido es la forma de vector que se usa con mayor frecuencia. Sin embargo, la invención debe comprender estas otras formas de vector de expresión, como vectores virales (por ejemplo, retrovirus deficientes de replicación, adenovirus y virus adeno-relacionados), que ejercen funciones similares. Además, el término vector también debe comprender otros vectores que son conocidos para la persona técnica en la materia, como fagos, virus tales como SV40, CMV, baculovirus, adenovirus, transposones, IS-elementos, fásmidos, fagémidos, cósmidos, ADN lineales o circulares.

Los vectores de expresión recombinantes de acuerdo con la invención comprenden un ácido nucleico de acuerdo con la invención o un constructo génico de acuerdo con la invención en una forma que es adecuada para la expresión del ácido nucleico en una célula huésped, lo cual significa que los vectores de expresión recombinantes comprenden una o más secuencias de regulación seleccionadas sobre la base de las células huésped usadas para la expresión, las cuales se enlazan funcionalmente con la secuencia de ácidos nucleicos a expresar. En un vector de expresión recombinante, "enlazado funcionalmente" significa que la secuencia de nucleótidos de interés está enlazada de tal forma a la(s) secuencia(s) de regulación que es posible la expresión de la secuencia de nucleótidos y se enlazan unas con otras para que ambas secuencias cumplan la función predicha que se atribuye a la secuencia (por ejemplo, en un sistema de transcripción *in vitro*/traducción o en una célula huésped, si el vector se introduce en la célula huésped). El término "secuencia de regulación" debe comprender promotores, enhancers y otros elementos de control de expresión (por ejemplo, señales de poliadenilización). Estas secuencias de regulación se describen, por ejemplo, en Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), o véase: Gruber y Crosby, en: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, editores: Glick y Thompson, capítulo 7, 89-108, incluyendo las referencias bibliográficas de allí. Secuencias de regulación comprenden aquellas que controlan la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células huésped y aquellas que controlan la expresión directa de la secuencia de nucleótidos solo en determinadas células huésped en determinadas condiciones. El técnico en la materia sabe que la configuración del vector de expresión puede depender de factores tales como la selección de la célula huésped a transformar, la extensión a la cual se expresa la proteína deseada, etc. Los vectores de expresión de acuerdo con la invención pueden introducirse en células huésped para que de esta manera produzcan proteínas o péptidos, incluyendo proteínas o péptidos de fusión que se codifican por los ácidos nucleicos, tal como los descritos (por ejemplo, desaturasas, formas mutantes de desaturasas, proteínas de fusión, etc.).

Los vectores de expresión recombinantes de acuerdo con la invención pueden configurarse para la expresión de desaturasas y elongasas en células procarióticas o eucarióticas. Por ejemplo, los genes de desaturasa pueden expresarse en células bacterianas como *C. glutamicum*, células de insectos (usando vectores de expresión de baculovirus),

levaduras y otras células de hongos (véase Romanos, M.A., y colaboradores (1992) "Foreign gene expression in yeast: a review", *Yeast* 8: 423-488; van den Hondel, C.A.M.J.J., y colaboradores (1991) "Heterologous gene expression in filamentous fungi", en: *More Gene Manipulations in Fungi*, J.W. Bennet & L.L. Lasure, editores, páginas 396-428: Academic Press: San Diego; y van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi", en: *Applied Molecular Genetics of Fungi*, Peberdy, J.F., y colaboradores, editores, páginas 1-28, Cambridge University Press: Cambridge), algas (Falcitore y colaboradores, 1999, *Marine Biotechnology*, 1, 3: 239-251), ciliados de los tipos: holotrichia, peritrichia, spirotrichia, suctorina, tetrahymena, paramecium, colpidium, glaucoma, platyophrya, potomacrus, pseudocohnilembus, euplotes, engelmaniella y stilonychia, en especial del género *stilonychia lemnae*, usando vectores según un método de transformación como se describe en WO 98/01572, así como células de vegetales multicelulares (véase Schmidt, R. y 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, páginas 71-119 (1993); F.F. White, B. Jenes y colaboradores, *Techniques for Gene Transfer*, en: *Transgenic Plants*, volumen 1, Engineering and Utilization, editores: Kung y R. Wu, Academic Press (1993), 128-43; Potrikus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225 (y las referencias bibliográficas allí citadas)) o células de mamíferos. Células huésped adecuadas se discuten adicionalmente en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). De manera alterna, el vector de expresión recombinante puede transcribirse *in vitro* y traducirse usando, por ejemplo, secuencias de regulación-promotor-T7 y polimerasa-T7.

La expresión de proteínas en procariotas se efectúa en la mayoría de casos con vectores que contienen promotores constitutivos o inducibles, los cuales controlan la expresión de proteínas de fusión o de no fusión. Los vectores de fusión añaden una serie de aminoácidos a una proteína allí codificada, usualmente en el terminal amino de la proteína recombinante, pero también en el terminal C o fusionada dentro de una región adecuada en las proteínas. Estos vectores de fusión tienen usualmente tres tareas: 1) el refuerzo de la expresión de proteína recombinante; 2) el incremento de la solubilidad de la proteína recombinante y 3) el apoyo a la purificación de la proteína recombinante por efecto como ligando en la purificación de afinidad. En el caso de vectores de expresión de fusión con frecuencia se introduce una posición de escisión proteolítica en la posición del compuesto de la unidad de fusión y de la proteína recombinante, de modo que la separación de la proteína recombinante de la unidad de fusión es posible después de la purificación de la proteína de fusión. Estas enzimas y sus secuencias correspondientes de reconocimiento comprenden factor Xa, trombina y enterocinasas.

Vectores de fusión típicos son, entre otros, pGEX (Pharmacia Biotech Inc; Smith, D.B., y Johnson, K.S. (1988) *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Pharmacia, Piscataway, NJ), en los que la transferasa-S-glutaciona (GST), proteína que enlaza maltosa E o proteína A se fusiona a la proteína recombinante objetivo. En una forma de realización, la secuencia que codifica desaturasa se clona en un vector de expresión pGEX, de modo que se genera un vector que codifica proteína de fusión, la cual comprende proteína -X- posición de escisión GST-trombina, desde el N-terminal hasta el C-terminal. La proteína de fusión puede purificarse mediante cromatografía de afinidad usando resina de agarosa-glutanionina. Desaturasa recombinante que no se fusiona a GST puede obtenerse descomponiendo la proteína de fusión con trombina.

Ejemplos de vectores de expresión de *E. coli* inducibles de no fusión son, entre otros, pTrc (Amann y colaboradores (1988) *Gene* 69: 301-315) y pET 11d (Studier y colaboradores, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). La expresión génica objetivo del vector pTrc se basa en la transcripción por la polimerasa de ARN huésped de un promotor de fusión trp-lac híbrido. La expresión génica objetivo del vector pET 11d se basa en transcripción de un promotor de fusión T7-gn10-lac que se transmite por una polimerasa de ARN viral co-expresada (T7 gn1). Esta polimerasa viral se suministra por las cepas huésped BL21 (DE3) o HMS174 (DE3) de un profago  $\lambda$  residente que alberga un gen T7 gnl bajo el control de transcripción del promotor lacUV 5.

Otros vectores que son adecuados en organismos procarióticos son conocidos al técnico en la materia; estos vectores son, por ejemplo, en *E. coli* pLG338, pACIC184, la serie pBR, como pBR322, la serie pUC, como pUC18 o pUC19, la serie M113mp, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III<sup>113</sup>-B1,  $\lambda$ gt11 o pBdCI, en *Streptomyces* pIJ101, pIJ364, pIJ702 o pIJ361, en bacilos pUB110, pC194 o pBD214, en *Corynebacterium* pSA77 o pAJ667. Una estrategia para maximizar la expresión de proteína recombinante es expresar la proteína en una bacteria huésped cuya habilidad para descomponer proteolíticamente la proteína recombinante es destruida (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Otra estrategia es la modificación de la secuencia de ácidos nucleicos del ácido nucleico a insertar en un vector de expresión, de modo que los codones individuales para cada aminoácido son aquellos que se usan preferentemente en una bacteria seleccionada para expresión, tal como *C. glutamicum* (Wada y colaboradores (1992) *Nucleic Acids Res.* 20: 2111-2118). Esta modificación de la secuencia de ácidos nucleicos de la invención se efectúa mediante técnicas de síntesis de ADN estándar.

En otra forma de realización el vector de expresión de desaturasa es un vector de expresión de levadura. Ejemplos de vectores para expresión en la levadura *S. cerevisiae* comprenden pYeDesaturasecl (Baldari y colaboradores (1987) *Embo J.* 6: 229-234), pMFa (Kurjan y Herskowitz (1982) *Cell* 30: 933-943), pJRY88 (Schultz y colaboradores (1987) *Gene* 54: 113-123) así como pYES2 (Invitrogen Corporation, San Diego, CA). Vectores y métodos para la construcción de vectores que son adecuados para usar en otros hongos, como los hongos filamentosos, comprenden aquellos que se describen detalladamente en: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector

development for filamentous fungi”, en: Applied Molecular Genetics of fungi, J.F. Peberdy y colaboradores, editores, páginas 1-28, Cambridge University Press: Cambridge, o en: More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, Editores, páginas 396-428: Academic Press: San Diego]. Otros vectores adecuados de levadura son, por ejemplo, pAG-1, YEp6, YEp13 o pEMBLEY23.

5

De manera alterna, las desaturadas de acuerdo con la invención pueden expresarse en células de insectos usando vectores de expresión de baculovirus. Vectores de baculovirus que se encuentran disponibles para expresar proteínas en células de insecto cultivadas (por ejemplo, células Sf9), comprenden la serie pAc (Smith y colaboradores (1983) Mol. Cell Biol. 3: 2156-2165) y la serie pVL (Lucklow y Summers (1989) Virology 170: 31-39).

10

Los vectores arriba mencionados son solo una pequeña sinopsis de posibles vectores adecuados. Otros plásmidos son conocidos para el técnico en la materia y se describen, por ejemplo, en: Cloning Vectors (Editores Pouwels, P.H., y colaboradores, Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018).

15

En una modalidad más, se expresa un ácido nucleico de acuerdo con la invención en células de mamífero usando un vector de expresión de mamífero. Se entienden por mamíferos en el sentido de la invención todos los mamíferos no humanos. Ejemplos de vectores de expresión de mamíferos comprenden pCDM8 (Seed, B. (1987) Nature 329:840) y pMT2PC (Kaufman y colaboradores (1987) EMBO J. 6: 187-195). En el caso del uso en células de mamíferos, las funciones de control del vector de expresión se suministran frecuentemente por elementos regulatorios virales. Promotores que se usan usualmente se derivan, por ejemplo, de polioma, adenovirus, citomegalovirus y virus 40 Simian. Otros sistemas de expresión adecuados para células procariótidas y eucariótidas pueden encontrarse en los capítulos 16 y 17 de Sambrook, J., Fritsch, E.F., y Maniatis, T., Molecular Cloning: A Laboratory Manual, 2. Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

20

25

En otra forma de realización, el vector de expresión recombinante de mamífero puede controlar la expresión del ácido nucleico preferible en un tipo de célula específico (por ejemplo, elementos regulatorios específicos de tejido se usan para expresar el ácido nucleico). En la técnica se conocen elementos regulatorios específicos de tejido. Ejemplos no limitantes de promotores específicos de tejido son, entre otros, el promotor de albúmina (específico de hígado; Pinkert y colaboradores (1987) Genes Dev. 1: 268-277), promotores específicos de linfocitos (Calame y Eaton (1988) Adv. Immunol. 43: 235-275), en particular promotores de receptores de células T (Winoto y Baltimore (1989) EMBO J. 8: 729-733) e inmunoglobulinas (Banerji y colaboradores (1983) Cell 33: 729-740; Queen y Baltimore (1983) Cell 33: 741-748), promotores específicos de neuronas (por ejemplo promotor de neurofilamento; Byrne y Ruddle (1989) PNAS 86: 5473-5477), promotores específicos de páncreas (Edly y colaboradores, (1985) Science 230: 912-916) y promotores específicos de glándula mamaria (por ejemplo, promotor de suero láctico; patente de Estados Unidos No. 4,873,316 y publicación de solicitud europea de patente No. 264,166). También están incluidos los promotores regulados por desarrollo, por ejemplo los promotores hox de ratón (Kessel y Gruss (1990) Science 249: 374-379) y el promotor de proteína de feto (Campes y Tilghman (1989) Genes Dev. 3: 537-546).

30

35

40

En otra forma de realización, las desaturadas de acuerdo con la invención pueden expresarse en células vegetales unicelulares (como algas), véase Falciatore y colaboradores, 1999, Marine Biotechnology 1 (3): 239-251 y los datos bibliográficos allí citados, y células vegetales de vegetales superiores (por ejemplo, espermatofitos tales como frutos del campo). Ejemplos de vectores de expresión vegetales comprenden aquellos que se describen detalladamente en: Becker, D., Kemper, E., Schell, J., y Masterson, R. (1992) “New plant binary vectors with selectable markers located proximal to the left border”, Plant Mol. Biol. 20: 1195-1197; y Bevan, M.W. (1984) “Binary Agrobacterium vectors for plant transformation”, Nucl. Acids Res. 12: 8711-8721; Vectors for Gene Transfer in Higher Plants; en: Transgenic Plants, volumen 1, Engineering and Utilization, Editores: Kung y R. Wu, Academic Press, 1993, páginas 15-38.

45

50

Un casete de expresión de vegetales contiene preferentemente secuencias de regulación que pueden controlar la expresión génica en células vegetales y se enlazan de manera funcional de modo que cada secuencia pueda desempeñar su función, como terminación de la transcripción, por ejemplo señales de poliadenilación. Señales preferidas de poliadenilación son aquellas que provienen de agrobacterium tumefaciens-t-ADN, tal como el gen 3, conocido como octopinsintasa, del Ti-plásmido pTiACH5 (Gielen y colaboradores, EMBO J. 3 (1984) 835ff.) o equivalentes funcionales de la misma pero también son adecuadas todos los otros terminadores funcionalmente activos en vegetales.

55

Puesto que la expresión génica vegetal con mucha frecuencia no se limita al nivel de transcripción, un casete de expresión vegetal comprende preferiblemente otras secuencias enlazadas funcionalmente, tal como enhancers de traducción, por ejemplo la secuencia overdrive, que contienen la secuencia líder 5'-no traducida del virus mosaico de tabaco, la cual eleva la proporción proteína/ARN (Gallie y colaboradores, 1987, Nucl. Acids Research 15: 8693-8711).

60

65

La expresión génica vegetal debe enlazarse funcionalmente con un promotor adecuado que realice expresión génica de una manera específica para célula o tejido en el tiempo correcto. Promotores preferidos son aquellos que causan expresión constitutiva (Benfey y colaboradores, EMBO J. 8 (1989) 2195-2202), como aquellos que se derivan de virus vegetales tales como 35S CAMV (Franck y colaboradores, Cell 21 (1980) 285-294), 19S CaMV (véase también US 5352605 y WO 84/02913) o promotores vegetales como el descrito en US 4,962,028 de la pequeña subunidad de RuBisCo.

Otras secuencias preferidas para el uso para el enlazamiento funcional en casetes de expresión génicos son secuencias targeting (direccionadoras) que se requieren para dirigir el producto génico hacia su correspondiente compartimiento celular (véase una sinopsis en Kermodé, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423 y referencias bibliográficas allí citadas), por ejemplo a la vacuola, el núcleo, todos los tipos de plástidos tales como amiloplastos, cloroplastos, cromoplastos, el espacio eextracelular, la mitocondria, el retículo endoplasmático, oleoplastos, peroxisomas y otros compartimientos de las células vegetales.

La expresión génica vegetal puede facilitarse mediante un promotor inducible químicamente (véase una sinopsis en Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48: 89-108). Los promotores químicamente inducibles son adecuados particularmente si se desea que tenga lugar expresión génica de una manera específica en el tiempo. Ejemplos de tales promotores son un promotor inducible por ácido salicílico (WO 95/19443), un promotor inducible por tetraciclina (Gatz y colaboradores (1992) Plant J. 2, 397-404) y un promotor inducible por etanol.

Otros promotores adecuados también son promotores que reaccionan a condiciones de estrés bióticas o abióticas, por ejemplo el promotor de gen PRP1 inducido por patógeno (Ward y colaboradores, Plant. Mol. Biol. 22 (1993) 361-366), el promotor hsp80 del tomate, inducido por calor, (US 5,187,267), el promotor de alfaamilasa de la patata, inducido por frío, (WO 96/12814) o el promotor pinII incible por heridas (EP-A-0 375 091).

En particular se prefieren aquellos promotores que ocasionan la expresión génica en tejidos y órganos en los que tiene lugar la biosíntesis de lípidos y aceite, en células de semillas, como en las células del endosperma y del embrión que se desarrolla. Promotores adecuados son el promotor de gen napina de la colza (US 5,608,152), el promotor USP de Vicia faba o haba (Baumlein y colaboradores, Mol Gen Genet, 1991, 225 (3): 459-67), el promotor oleosina de arabidopsis (WO 98/45461), el promotor faseolina de *Faseolus vulgaris* (US 5,504,200), el promotor Bce4 de Brassica (WO 91/13980) o el promotor B4 de legumina (LeB4; Baumlein y colaboradores, 1992, Plant Journal, 2 (2): 233-9), así como promotores que ocasionan expresión específica de semillas en plantas monocotiledóneas, tales como maíz, cebada, trigo, centeno, arroz, etc. Promotores adecuados notables son promotor - gen lpt2 o lpt1 de cebada (WO 95/15389 y WO 95/23230) o los promotores descritos en WO 99/16890 (promotores del gen hordeina de la cebada, del gen glutelina de arroz, del gen orizina de arroz, del gen prolamina de arroz, del gen gliadina de trigo, del gen glutelina de trigo, del gen zeina de maíz, del gen glutelina de avena, del gen casirina de sorgo, del gen secalina de centeno).

En especial la expresión multiparalela de las desaturasas de acuerdo con la invención, sola o en combinación con otras desaturasas o elongasas, puede ser deseable. La introducción de unos casetes de expresión así puede efectuarse por una transformación simultánea de varios constructos de expresión sobre un constructo. Varios vectores también pueden transformarse con varios casetes de expresión respectivamente y transferirse a la célula huésped.

Así mismo son particularmente adecuados aquellos promotores que conducen a una expresión específica de plástido ya que los plástidos son el compartimiento en el que se sintetizan los precursores y algunos productos finales de biosíntesis lipídica. Promotores adecuados como el promotor de polimerasa ARN viral se describen en WO 95/16783 y WO 97/06250, y el promotor clpP de arabidopsis, descrito en WO 99/46394.

La invención proporciona además un vector de expresión recombinante que comprende una molécula de ADN de acuerdo con la invención, la cual se clona en dirección antisense en el vector de expresión. Es decir que la molécula de ADN se enlaza funcionalmente con una secuencia regulatoria de tal modo que se hace posible la expresión (por transcripción de la molécula de ADN) de una molécula de ARN, que es "antisense" para la desaturasa de ARNm. Pueden seleccionarse secuencias de regulación que se enlazan funcionalmente con un ácido nucleico clonado en dirección antisense y que controlan la expresión continua de la molécula de ARN antisense en un gran número de tipos de células, por ejemplo pueden seleccionarse promotores y/o enhancers virales o secuencias regulatorias, los cuales controlan la expresión constitutiva, específica de tejido o específica del tipo de célula de ARN antisense. El vector de expresión antisense puede presentarse en forma de un plásmido recombinante, fagémido o virus atenuado en el que se produzcan los ácidos nucleicos antisense bajo el control de una región regulatoria altamente efectiva, cuya actividad puede determinarse por el tipo de célula al cual se ha introducido el vector. Véase una explicación de la regulación de la expresión génica por medio de genes antisense en Weintraub, H., y colaboradores, Antisense-RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, volumen 1(1) 1986.

Otro aspecto de la invención se refiere a células huésped a las cuales se ha introducido un vector de expresión recombinante de acuerdo con la invención. Los términos "célula huésped" y "célula huésped recombinante" se usan aquí de manera intercambiable una por otra. Obviamente estos términos se refieren no solo a las células objetivo determinadas sino también a la descendencia o la descendencia potencial de esta célula. Puesto que pueden ocurrir modificaciones específicas en generaciones sucesivas debido a mutación o a efectos ambientales, esta descendencia no es necesariamente idéntica a la célula parental aunque permanece dentro del alcance del término tal como aquí se usa.

Por recombinante o transgénico, por ejemplo vector de expresión recombinante o anfitrión o células huésped recombinantes en el sentido de la invención debe entenderse que los ácidos nucleicos de acuerdo con la invención y/o sus secuencias naturales de regulación en posición 5' y 3' de los ácidos nucleicos no están en su ambiente natural, lo que significa que o bien se ha modificado la locación de las secuencias en el organismo original, o bien han mutado en él las secuencias de ácido nucleico y/o las secuencias regulatorias, o se han transferido las secuencias de ácido nucleico

## ES 2 345 324 T3

según la invención hacia un organismo distinto del organismo original o sus secuencias de regulación. También son posibles combinaciones de estas modificaciones. Por ambiente natural debe entenderse la locación de una secuencia de ácidos nucleicos en un organismo tal como tiene lugar en la naturaleza.

5 Una célula huésped puede ser una célula procariótica o eucariótica. Por ejemplo, una desaturasa puede expresarse en células bacterianas tales como *C. glutamicum*, células de insectos, células de hongos o células de mamíferos (como fósforo de ovario de hámster chino (CHO) o células de COS), algas, ciliados, células vegetales, hongos u otros microorganismos, como *C. glutamicum*. Otras células huésped adecuadas son corrientes para el técnico en la materia.

10 El ADN de vector puede introducirse a células procarióticas o eucarióticas mediante técnicas convencionales de transformación o transfección. Los términos “transformación” y “transfección”, conjugación y transducción, tal como se usan en el presente contexto están dirigidos a abarcar un gran número de métodos conocidos en el campo técnico para introducir ácido nucleico ajeno (por ejemplo, ADN) a una célula huésped, incluyendo coprecipitación por fosfato de calcio o por cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección, competencia natural, transferencia mediada químicamente, electroporación o bombardeo de partículas. Pueden encontrarse métodos adecuados para la transformación o transfección de células huésped, incluyendo células vegetales en Sambrook y colaboradores (Molecular Cloning: A Laboratory Manual., 2. Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) y otros manuales de laboratorio, tales como Methods in Molecular Biology, 1995, volumen 44, Agrobacterium protocols, editores: Gartland y Davey, Humana Press, Totowa, 20 New Jersey.

A través de la transfección estable de células de mamíferos se conoce que dependiendo del vector de expresión usado y de la técnica de transfección usada solo una pequeña parte de la célula integra el ADN ajeno a su genoma. Para identificar y seleccionar estos integrantes usualmente se introduce un gen que codifica un marcador seleccionable (por ejemplo, resistencia frente a antibióticos), junto con el gen de interés a las células huésped. Marcadores seleccionables preferidos comprenden aquellos que confieren resistencia frente a medicamentos tales como G418, higromicina y metotrexato, o en vegetales aquellos que confieren resistencia frente a un herbicida, como glifosato o glufosinato. Otros marcadores adecuados son, por ejemplo, marcadores que codifican genes que participan en rutas de biosíntesis de azúcares o aminoácidos, por ejemplo, tales como  $\beta$ -galactosidasa, *ura3* o *ilv2*. Así mismo son adecuados marcadores que codifican genes como luciferasa, *gfp* u otros genes de fluorescencia. Estos marcadores pueden usarse en mutantes en los que estos genes no son funcionales puesto que éstos hab sido deletados mediante métodos convencionales, por ejemplo. Adicionalmente, en el mismo vector pueden introducirse marcadores que codifican un ácido nucleico a una célula huésped como aquella que codifica una desaturasa, o pueden introducirse en un vector separado. Las células que han sido transfectadas de manera estable con el ácido nucleico introducido pueden identificarse mediante selección de medicamentos (por ejemplo, sobreviven células que tienen integrados los marcadores seleccionables, mientras que las otras mueren).

Para generar un microorganismo recombinado de manera homóloga, se genera un vector que contiene al menos un segmento de un gen de desaturasa al cual se ha introducido una delección, adición o sustitución para modificar el gen desaturasa, por ejemplo para destruirlo funcionalmente. Este gen desaturasa es preferentemente un gen desaturasa de *Phaeodactylum tricornutum*, aunque puede usarse un homólogo o análogo de otros organismos, incluso de una fuente de mamíferos, hongos o insectos. En una forma preferida de realización se configura el vector de tal manera que se destruye funcionalmente la desaturasa endógena durante la recombinación homóloga (es decir, ya no codifica una proteína funcional, también denominada vector knock-out). De manera alterna, el vector puede configurarse de tal manera que el gen desaturasa endógeno muta durante la recombinación homóloga o se modifica de otra manera, pero siempre codifica todavía una proteína funcional (por ejemplo, la región regulatoria ubicada hacia arriba puede modificarse de tal manera que se modifique así la expresión de la desaturasa endógena). Para generar una mutación puntual por recombinación homóloga también pueden usarse híbridos de ADN-ARN conocidos como quimeraplastos, los cuales se conocen de Cole-Strauss y colaboradores, 1999, Nucleic Acids Research 27(5):1323-1330 y Kmiec, Gene therapy, 1999, American Scientist, 87(3): 240-247.

En el vector para la recombinación homóloga, el segmento modificado del gen desaturasa está flanqueado en su extremo 5' y 3' por ácido nucleico adicional del gen desaturasa, de modo que es posible una recombinación homóloga entre el gen desaturasa exógeno, que se encuentra presente en el vector, y un gen desaturasa en un microorganismo o un vegetal. El ácido nucleico de desaturasa que flanquea adicionalmente es suficientemente largo para una recombinación homóloga exitosa con el gen endógeno. Habitualmente en el vector están contenidos ADN que flanquean de varios cientos de pares de bases hasta kilobases (tanto en el extremo 5' como también en el 3') (véase una descripción de vectores para la recombinación homóloga, por ejemplo, en Thomas, K.R., y Capecchi, M.R. (1987) Cell 51:503 o para la recombinación en *Physcomitrella patens* a base de ADNc véase en Strepp y colaboradores, 1998, Proc. Natl. Acad. Sci. USA 95 (8): 4368-4373). El vector se introduce a un microorganismo o a una célula vegetal (por ejemplo por medio de ADN mediado por polietilenglicol), y se seleccionan células en las que el gen desaturasa introducido se recombina con el gen desaturasa endógeno de manera homóloga, usando técnicas conocidas en el campo técnico.

En otra forma de realización, pueden generarse organismos recombinantes, tales como microorganismos, que contienen sistemas seleccionados que hacen posible una expresión regulada del gen introducido. La inclusión de un gen desaturasa en un vector, cuando se ha introducido bajo el control del lac-operon, hace posible por ejemplo la expresión del gen desaturasa solo en presencia de IPTG. Estos sistemas de regulación se conocen en el campo técnico.

Pueden usarse células huésped de acuerdo con la invención, como una célula huésped procariótica o eucariótica, que crecen en cultivo o en un campo, para la producción (es decir, expresión) de una desaturasa. En las plantas puede aplicarse un método alternativo mediante transferencia directa de ADN a flores que se desarrollan mediante electroporación o transferencia génica por medio de agrobacterium. Por lo tanto, la invención proporciona además métodos para la producción de desaturasas usando células huésped de la invención. En una forma de realización, el método comprende hacer crecer la célula huésped de la invención (a la que se ha introducido un vector de expresión recombinante que codifica una desaturasa, o a cuyo genoma se ha introducido un gen que codifica un tipo silvestre o una desaturasa modificada) en un medio adecuado, hasta que se ha producido la desaturasa. El método comprende en otra forma de realización el aislamiento de la desaturasa del medio o de la célula huésped.

Células huésped que son adecuadas para recibir en principio el ácido nucleico de la invención, el producto génico o el vector de acuerdo con la invención, son todos organismos procarióticos o eucarióticos. Los organismos usados de manera ventajosa son organismos como bacterias, hongos, levaduras, células animales o vegetales. Otros organismos ventajosos son animales o preferentemente vegetales o partes de éstos. Hongos, levaduras o vegetales se usan preferentemente, particularmente preferible hongos o vegetales, muy particularmente se prefiere usar vegetales como plantas de frutos oleaginosos, los cuales contienen grandes cantidades de compuestos lipídicos, como colza, onagra, canola, maní o cacahuate, lino, soya, cártamo o cardo, girasol, borraja o plantas como maíz, trigo, centeno, avena, triticale, arroz, cebada, algodón, yuca, pimienta, tagetes, plantas solanáceas como patata, tabaco, berenjena y tomate, variedades de vicia, guisantes, alfalfa, arbustos (café, cacao, té), variedades de salix, árboles (plasma de aceite, coco) así como gramíneas perennes y frutos de campo para forraje o alimento de animales. De acuerdo con la invención se prefieren particularmente plantas de frutos oleaginosos como soya, maní, colza, canola, lino, onagra, girasol, cártamo, árboles (palma de aceite, coco).

#### D. Desaturasas aisladas

Otro aspecto de la invención se refiere a desaturasas aisladas y partes biológicamente activas de las mismas. Una proteína "aislada" o "purificada" o una parte biológicamente activa de la misma está esencialmente libre de material celular, si se produce por técnicas de recombinación de ADN, o de precursores químicos u otros productos químicos, si se sintetiza químicamente. El término "esencialmente libre de material celular" comprende preparaciones de desaturasa en las que se separa la proteína de componentes de las células en las que se produce la desaturasa de manera natural o recombinante. En una forma de realización, la expresión "material esencialmente libre de material celular" comprende preparaciones de desaturasa con menos de aproximadamente 30% (respecto del peso seco) de no desaturasa (denominado aquí también como "proteína contaminante"), más preferido menos de aproximadamente 20% de no desaturasa, aún más preferido menos de aproximadamente 10% de no desaturasa y lo más preferido menos de aproximadamente 5% de no desaturasa. Si la desaturasa o una parte biológicamente activa de la misma se ha producido de manera recombinante, esta también se encuentra esencialmente libre de medio de cultivo si el medio de cultivo constituye menos de aproximadamente 20%, más preferible menos de aproximadamente 10% y lo más preferido menos de aproximadamente 5% del volumen de la preparación de proteína. La expresión "esencialmente libre de precursores químicos u otros productos químicos" comprende preparaciones de desaturasa en las que se separa la proteína de precursores químicos u otros productos químicos que participan en la síntesis de la proteína. En una forma de realización la expresión "esencialmente libre de precursores químicos u otros productos químicos" comprende preparaciones de desaturasa con menos de aproximadamente 30% (respecto del peso seco) de precursores químicos o productos químicos que no son desaturasa, más preferido menos de aproximadamente 20% de precursores químicos o productos químicos que no son desaturasa, aún más preferido menos de aproximadamente 10% de precursores químicos o productos químicos que no son desaturasa y lo más preferido menos de aproximadamente 5% de precursores químicos o productos químicos que no son desaturasa. En formas preferidas de realización, las proteínas aisladas, o las partes biológicamente activas de las mismas, no tienen proteínas contaminantes del mismo organismo del cual proviene la desaturasa. Estas proteínas se producen habitualmente mediante expresión recombinante de desaturasa de *Phaeodactylum tricorutum*, por ejemplo, en vegetales como *Physcomitrella patens* o las arriba mencionadas o microorganismos, por ejemplo bacterias como *E. coli*, *Bacillus subtilis*, *C. glutamicum*, hongos como mortierella, levaduras como saccharomyces, o ciliados como colpidium o algas como *Phaeodactylum*.

Una desaturasa aislada de acuerdo con la invención o una parte de la misma también puede participar en el metabolismo de compuestos necesarios para la síntesis de membranas celulares en *Phaeodactylum tricorutum* o en el transporte de moléculas por estas membranas. En formas preferidas de realización, la proteína, o la parte de la misma, comprende una secuencia de aminoácidos que es lo suficientemente homóloga a una secuencia de aminoácidos de las SEQ ID NO: 2, 4, 6 ó 12 para que la proteína, o parte de la misma, mantenga la habilidad de participar en el metabolismo de compuestos necesarios para la síntesis de membranas celulares en *Phaeodactylum* o en el transporte de moléculas a través de estas membranas. La parte de la proteína es preferentemente una parte biológicamente activa, tal como se describe aquí. En otra forma preferida de realización una desaturasa de acuerdo con la invención tiene una secuencia de aminoácidos mostrada en SEQ ID NO: 2, 4, 6 ó 12. En otra forma preferida de realización, la desaturasa tiene una secuencia de aminoácidos, que se codifica por una secuencia de nucleótidos que se hibrida en una secuencia de nucleótidos de las SEQ ID NO: 1, 3, 5 u 11, por ejemplo en condiciones estrictas. En otra forma preferida más de realización la desaturasa tiene una secuencia de aminoácidos que se codifica por una secuencia de nucleótidos que tiene una homología de al menos aproximadamente 50 hasta 60%, preferentemente de al menos aproximadamente 60 hasta 70%, más preferible de al menos aproximadamente 70 hasta 80%, 80 hasta 90%, 90 hasta 95% y aún más preferible de al menos aproximadamente 96%, 97%, 98%, 99% o aún más homología a una de las secuencias de aminoácidos

## ES 2 345 324 T3

de las SEQ ID NO: 2, 4, 6 ó 18. La desaturasa preferida de acuerdo con la invención también posee preferentemente al menos una de las actividades de desaturasa aquí descritas. Por ejemplo una desaturasa preferida de acuerdo con la invención comprende una secuencia de aminoácidos que se codifica por una secuencia de nucleótidos, que se hibrida a una secuencia de nucleótidos de las SEQ ID NO: 1, 3, 5 u 11 en condiciones estrictas, por ejemplo y puede participar en el metabolismo de compuestos necesarios para la síntesis de membranas celulares en *Phaeodactylum tricorutum* o en el transporte de moléculas por estas membranas o introduce un enlace doble a un ácido graso con uno, dos, tres o cuatro enlaces dobles y una longitud de cadena de C<sub>18</sub>, C<sub>20</sub> ó C<sub>22</sub>.

En otras formas de realización, la desaturasa es esencialmente homóloga a una secuencia de aminoácidos de las SEQ ID NO: 2, 4 ó 6 y mantiene la actividad funcional de la proteína de una de las secuencias de las SEQ ID NO: 2, 4 ó 6, aunque su secuencia de aminoácidos se diferencia debido a la mutación natural o a la mutagénesis, tal como se ha descrito detalladamente en la subsección I de arriba. En otra forma de realización, la desaturasa es por lo tanto una proteína que comprende una secuencia de aminoácidos que tiene una homología de al menos aproximadamente 50 hasta 60%, preferentemente de al menos aproximadamente 60 hasta 70% y más preferido de al menos aproximadamente 70 hasta 80%, 80 hasta 90%, 90 hasta 95% y lo más preferido de al menos aproximadamente 96%, 97%, 98%, 99% o más de homología a una secuencia completa de aminoácidos de las SEQ ID NO: 2, 4 ó 6 y tiene al menos una de las actividades de desaturasa aquí descritas. En otra forma de realización la invención se refiere a una proteína completa de *Phaeodactylum tricorutum* que es esencialmente homóloga a una secuencia completa de aminoácidos de las SEQ ID NO: 2, 4 ó 6.

Las partes biológicamente activas de una desaturasa comprenden péptidos que comprenden secuencias de aminoácidos que se derivan de la secuencia de aminoácidos de una desaturasa, por ejemplo una secuencia de aminoácidos mostrada en SEQ ID NO: 2, 4 ó 6 o la secuencia de aminoácidos de una proteína que es homóloga a una desaturasa, las cuales comprenden menos aminoácidos que la desaturasa de longitud completa o la proteína de longitud completa, la cual es homóloga a una desaturasa, y tienen al menos una actividad de una desaturasa.

Habitualmente, las partes activas biológicamente (péptidos, por ejemplo péptidos, que tienen una longitud, por ejemplo, de 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 o más aminoácidos) comprenden un dominio o un motivo con al menos una actividad de una desaturasa. Además, mediante técnicas recombinantes pueden producirse otras partes biológicamente activas en las que se han deletado otras regiones de la proteína, e investigarse respecto de una o varias de las actividades aquí descritas. Las partes activas biológicamente de una desaturasa comprenden preferiblemente uno o varios dominios/motivos seleccionados o partes de los mismos con actividad biológica.

Las desaturasas se producen preferentemente mediante técnicas de recombinación de ARN. Por ejemplo, se clona una molécula de ácido nucleico que codifica la proteína en un vector de expresión (tal como se describió previamente), el vector de expresión se introduce a una célula huésped (como se describió previamente), y se expresa la desaturasa en la célula huésped. Entonces la desaturasa puede aislarse de las células a través de un esquema de purificación mediante técnicas de purificación estándar. De manera alterna a la expresión recombinante una desaturasa, un polipéptido desaturasa o un péptido desaturasa pueden sintetizarse químicamente por medio de técnicas estándar de síntesis de péptidos. Además, pueden aislarse desaturasas nativas de células (por ejemplo, células endotéllicas) usando, por ejemplo, un anticuerpo anti-desaturasa, que puede producirse mediante técnicas estándar, y se usa una desaturasa de acuerdo con la invención o un fragmento de la misma.

La invención también proporciona proteínas desaturasas quiméricas o proteína de fusión desaturasas. Tal como aquí se usa, una "proteína desaturasa quimérica" o una "proteína de fusión -desaturasa" comprende un polipéptido desaturasa que está funcionalmente enlazado con un polipéptido no desaturasa. Un "polipéptido desaturasa" se refiere a un polipéptido con una secuencia de aminoácidos que corresponde a una desaturasa, mientras que un "polipéptido no desaturasa" se refiere a un polipéptido con una secuencia de aminoácidos que corresponde a una proteína que es esencialmente no homóloga a la desaturasa; por ejemplo, una proteína que se diferencia de la desaturasa y proviene del mismo o de otro organismo. Dentro de la proteína de fusión la expresión "enlazado funcionalmente" debe significar que el polipéptido desaturasa y el polipéptido no desaturasa están fusionados uno con otro de tal modo que ambas secuencias cumplen la función predicha, atribuida a la secuencia usada. El polipéptido no desaturasa puede estar fusionado en el terminal N o en terminal C del polipéptido desaturasa. En una forma de realización, la proteína de fusión es, por ejemplo, una proteína de fusión desaturasa GST, en la que las secuencias desaturasa se fusionan en el terminal C de las secuencias GST. Estas proteínas de fusión pueden facilitar la purificación de las desaturasas recombinantes. En otra forma de realización, la proteína de fusión es una desaturasa que tiene una secuencia de señal heteróloga en su terminal N. En determinadas células huésped (por ejemplo, células huésped de mamíferos) la expresión y/o secreción de una desaturasa puede incrementarse usando una secuencia de señal heteróloga.

Una proteína desaturasa quimérica o proteína de fusión desaturasa de acuerdo con la invención se produce mediante técnicas de recombinación de ADN estándar. Por ejemplo, los fragmentos de ADN que codifican diversas secuencias de polipéptidos, se ligan unos a otros según técnicas convencionales en el marco de lectura empleando, por ejemplo, extremos lisos o que sobresalen colgando para la ligazón, descomposición de enzima de restricción para proporcionar extremos adecuados, llenado de extremos cohesivos, como se requiera, tratamiento con fosfatasa alcalina para impedir enlazamientos indeseados y ligazón enzimática. En otra forma de realización, el gen de fusión puede sintetizarse mediante técnicas convencionales, incluyendo sintetizadores automáticos de ADN. De manera alterna, una amplificación PCR de fragmentos de gen puede realizarse usando cebadores anclas que generan sobresalientes que cuelgan entre fragmentos de gen sucesivos que pueden hibridarse y reamplificarse unos con otros a continuación, de modo

que se genera una secuencia génica quimérica (véase, por ejemplo, *Current Protocols in Molecular Biology*, Editores Ausubel y colaboradores, John Wiley & Sons: 1992). Además, pueden obtenerse comercialmente muchos vectores de expresión que ya codifican una unidad de fusión (por ejemplo, un polipéptido GST). Un ácido nucleico que codifica desaturasa puede clonarse en un vector de expresión así, de modo que la unidad de fusión se enlaza en el marco de lectura con la proteína desaturasa.

Pueden generarse homólogos de la desaturasa mediante mutagénesis, por ejemplo mediante mutación puntual específica o truncación de la desaturasa. El término “homólogos”, tal como se usa aquí, se refiere a una forma variante de la desaturasa que actúa como agonista o antagonista de la actividad desaturasa. Un agonista de la desaturasa puede mantener esencialmente la misma actividad que la parte o una parte de las actividades biológicas de la desaturasa. Un antagonista de la desaturasa puede inhibir una o varias actividades de la forma que existe de manera natural de la desaturasa mediante, por ejemplo, enlace competitivo a un elemento ubicado hacia abajo o hacia arriba de la cascada metabólica para componentes de membrana celular que comprende la desaturasa, o mediante enlace a una desaturasa que facilita el transporte de compuestos por las membranas celulares, por lo cual se inhibe la translocación. En una forma alterna de realización, pueden identificarse homólogos de la desaturasa mediante clasificación de bancos combinatorios de mutantes, por ejemplo mutantes de truncación de la desaturasa con respecto a la actividad agonista o antagonista. En una forma de realización se genera un banco variegado de variantes de desaturasa mediante mutagénesis combinatoria a nivel de ácido nucleico y se codifica por un banco génico variegado. Un banco variegado de variantes desaturadas puede generarse, por ejemplo, mediante ligazón enzimática de una mezcla de oligonucleótidos sintéticos a secuencias génicas de modo que pueda expresarse un juego de secuencias potenciales de desaturasa como polipéptidos individuales o, de manera alterna, como un juego de proteínas de fusión más grandes (por ejemplo para display de fago) las cuales comprenden este juego de secuencias de desaturasa. Existe un gran número de métodos que pueden usarse para generar bancos de homólogos potenciales de desaturasa a partir de una secuencia degenerada de oligonucleótidos. La síntesis química de una secuencia génica degenerada puede realizarse en un sintetizador de ADN y el gen sintético puede luego ligarse a un vector de expresión adecuado. El uso de un juego degenerado de genes hace posible el suministro en una mezcla de todas las secuencias que codifican el juego deseado en secuencias potenciales de desaturasa. En el campo técnico son conocidos métodos para la síntesis de oligonucleótidos degenerados (véase, por ejemplo, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura y colaboradores (1984) *Annu. Rev. Biochem.* 53:323; Itakura y colaboradores, (1984) *Science* 198:1056; Ike y colaboradores (1983) *Nucleic Acids Res.* 11:477).

Adicionalmente, pueden usarse bancos de fragmentos de desaturasa para preparar una población variegada de fragmentos de desaturasa para la clasificación y para la selección enseguida de homólogos de una desaturasa. En una forma de realización, un banco de fragmentos de la secuencia que codifica puede generarse mediante tratamiento de un fragmento PCR bicatenario de una secuencia desaturada codificante con una nucleasa en condiciones en las que se efectúan rupturas bicatenarias solo aproximadamente una vez por molécula, desnaturalización del ADN bicatenario, renaturalización del ADN con la formación de ADN bicatenario que puede comprender pares sense/antisense de diversos productos con rupturas bicatenarias, remoción de secciones monocatenarias de dúplex recién formados mediante tratamiento con nucleasa S1, y ligazón del banco de fragmentos resultante en un vector de expresión. Con este método puede derivarse un banco de expresión que codifica terminales N, terminales C y fragmentos internos de la desaturasa de diversos tamaños.

En el campo técnico se conocen varias técnicas para la clasificación de productos génicos en bancos combinatorios que se han producido por mutaciones puntuales o truncación, y para la clasificación de bancos de ADNc para productos génicos con una propiedad seleccionada. Estas técnicas pueden adaptarse a clasificación rápida de los bancos de genes que se han generado por mutagénesis combinatoria de homólogos de desaturasa. Las técnicas más frecuentemente usadas para clasificar bancos génicos grandes, que pueden someterse a análisis de alto rendimiento, usualmente abarcan la clonación del banco génico en vectores de expresión replicables, la transformación de células adecuadas con el banco de vectores resultante y la expresión de los genes combinatorios en condiciones en las que la detección de la actividad deseada facilita el aislamiento del vector que codifica el gen cuyos productos han sido detectados. Recursive-Ensemble-Mutagenese (REM), una nueva técnica que eleva la frecuencia de mutantes funcionales en los bancos, puede usarse en combinación con los ensayos de clasificación para identificar homólogos de desaturasa. (Arkin y Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89: 7811-7815; Delgrave y colaboradores (1993) *Protein Engineering* 6(3): 327-331).

Otra técnica conocida para la modificación de propiedades catalíticas de enzimas o de sus genes codificantes es “gen-shuffling” (barajar genes) (véase por ejemplo en Stemmer, *PNAS* 1994, 91: 10747-10751, WO9720078 o WO9813487), que representa una combinación de fragmentos de gen y esta nueva combinación puede variarse adicionalmente mediante reacciones en cadena fallidas y de esta manera se crea una alta diversidad de secuencia a ensayarse. Sin embargo, el prerrequisito para usar un planteamiento así es un sistema adecuado de clasificación para ensayar la diversidad de genes resultante para funcionalidad.

En particular para clasificar actividades de desaturasa un prerrequisito para el método de clasificación es que capte la(s) actividad(es) enzimática(s) PUFA-dependientes. Respecto de las actividades de desaturasa con una especificidad por PUFAs, puede aprovecharse la toxicidad del ácido araquidónico en presencia de un metabolito tóxico (aquí: ácido salicílico o derivados de ácido salicílico) en especies *Mucor* que pueden transformarse con constructos deseado de genes mediante métodos conocidos de transformación (Eroshin y colaboradores, *Mikrobiologiya*, Vol. 65, No.1 1996, páginas 31-36), para realizar una clasificación primaria basada en crecimiento. Pueden analizarse entonces clones resultantes para sus componentes lipídicos por medio de cromatografía de gases y espectroscopía de masas para identificar la naturaleza y la cantidad de los materiales iniciales y los productos.

En otra forma de realización pueden aprovecharse ensayos sobre la base celular para analizar un banco variegado de desaturasa usando otros métodos conocidos en el campo técnico.

#### 5 E. Usos y métodos de acuerdo con la invención

Las moléculas de ácido nucleico, las proteínas, los homólogos de proteínas, las proteínas de fusión, los cebadores, los vectores y las células huésped descritos aquí pueden usarse en uno o más de los métodos que siguen: identificación de *Phaeodactylum* y organismos relacionados, mapeo de genoma de organismos que se relacionan con *Phaeodactylum tricorutum*, identificación y localización de *Phaeodactylum tricorutum*, identificación y localización de secuencias de interés, estudios de evolución, determinación de regiones de proteína desaturasa requeridos para la función, modulación de una actividad desaturasa, modulación del metabolismo de uno o más de los componentes de membrana celular, modulación del transporte transmembrana de uno o más compuestos, y modulación de la producción celular de un compuesto deseado tal como un producto químico fino. Las moléculas de ácido nucleico desaturasa según la invención tienen un gran número de usos. Primero, pueden usarse para identificar un organismo como *Phaeodactylum tricorutum* o un derivado cercano del mismo. También pueden usarse para identificar la presencia de *Phaeodactylum tricorutum* o de un pariente del mismo en una población mixta de microorganismos. La invención suministra las secuencias de ácido nucleico de una serie de genes de *Phaeodactylum tricorutum*; la presencia o la ausencia de este organismo puede determinarse mediante clasificación del ADN genómico extraído de un cultivo de una población uniforme o mezclada de microorganismos en condiciones rigurosas con una sonda que cubre una región de un gen de *Phaeodactylum tricorutum* o partes del mismo, el cual es único para este organismo. *Phaeodactylum tricorutum* mismo se usa para la producción comercial de ácidos poliinsaturados y adicionalmente también es adecuado para la producción de PUFAs en otros organismos, en particular cuando se deba lograr que los PUFAs resultantes se incorporen también a la fracción de triacilglicerol.

Adicionalmente, las moléculas de ácido nucleico y proteína según la invención pueden actuar como marcadores para regiones específicas del genoma. Esto es adecuado no solo para mapear el genoma sino también para proteína funcionales de *Phaeodactylum tricorutum*. Para identificar la región de genoma a la que se enlaza una proteína determinada de *Phaeodactylum tricorutum* que enlaza ADN, el genoma de *Phaeodactylum tricorutum* podría escindir, por ejemplo, y los fragmentos se incubarían con la proteína que enlaza ADN. Aquellos que enlazan la proteína pueden sondearse adicionalmente con las moléculas de ácido nucleico según la invención, preferiblemente con marcadores detectables fácilmente; el enlazamiento de una molécula de ácido nucleico así al fragmento de genoma hace posible la localización del fragmento en el mapa de genoma de *Phaeodactylum tricorutum* y, si esto se realiza repetidamente con diferentes enzimas, facilita una determinación rápida de la secuencia de ácido nucleico a la cual se enlaza la proteína. Además, las moléculas de ácido nucleico según la invención pueden tener suficiente homología a las secuencias de especies relacionadas para que estas moléculas de ácido nucleico sean capaces de actuar como marcadores para la construcción de un mapa genómico en hongos o algas relacionados.

Las moléculas de ácido nucleico desaturasa de acuerdo con la invención son adecuadas también para estudios de evolución y estudios de la estructura de proteína. Los procesos metabólicos y de transporte en los que participan las moléculas según la invención se utilizan por muchas células procariótidas y eucariótidas; el grado evolutivo de la familiaridad de los organismos puede determinarse comparando las secuencias de las moléculas de ácido nucleico según la invención con aquellas que codifican enzimas similares de otros organismos. De manera correspondiente, una comparación así hace posible la determinación de cuáles regiones de secuencia se conservan y cuáles no, lo cual puede ser de ayuda en la determinación de regiones de la proteína que son esenciales para la función enzimática. Este tipo de determinación es valioso para estudios de ingeniería de proteína y pueden suministrar una clave de cuánta mutagénesis puede tolerar la proteína sin perder la función.

La manipulación de las moléculas de ácido nucleico desaturasa según la invención puede llevar a la producción de desaturasas con diferencias funcionales frente a las desaturasas de tipo silvestre. La eficiencia o actividad de estas proteínas puede mejorarse, pueden estar presentes en la célula en cantidades más grandes que lo usual, o su eficiencia o actividad puede reducirse. Una eficiencia o mejoradas significa, por ejemplo, que la enzima tenga una selectividad y/o actividad más altas, preferible una actividad que es al menos 10% superior, especialmente preferible una actividad que es al menos 20% superior, muy especialmente preferible una actividad que es al menos 30% superior que aquella de la enzima original.

Existe una serie de mecanismos por los que la modificación de una desaturasa según la invención puede afectar directamente el rendimiento, la producción y/o la eficiencia de producción de un producto químico fino que comprende una proteína modificada así. La obtención de compuestos químicos finos a partir de cultivos de ciliados, algas u hongos a gran escala se mejora significativamente cuando la célula secreta los compuestos deseados, puesto que estos compuestos pueden aislarse fácilmente a partir del medio de cultivo (en contraste con la extracción desde la biomasa de las células cultivadas). Por otro lado, la purificación puede mejorarse cuando la célula almacena compuestos *in vivo* en un compartimento especializado con un tipo de mecanismo de concentración. En plantas que expresan desaturasas, un transporte incrementado puede conducir a una mejor distribución dentro del tejido vegetal y los órganos vegetales. Aumentar el número o la actividad de moléculas transportadoras que exportan productos químicos finos desde la célula puede permitir que se incremente la cantidad de los productos químicos finos producidos que están presentes en el medio extracelular, facilitando así cosechar y purificar o, en el caso de plantas, distribuir más eficientemente. Por lo contrario, para la sobreproducción eficiente de uno o más productos químicos finos se requieren cantidades aumen-

tadas de co-factores, moléculas precursoras e intermediarios para las rutas adecuadas de biosíntesis. Aumentando el número y/o la actividad de proteínas transportadoras involucradas en la importación de nutrientes tales como fuentes de carbono (es decir, azúcares), fuentes de nitrógeno (es decir, aminoácidos, sales de amonio), fosfato y azufre, puede mejorar la producción de un químico fino debido a la eliminación de todas las limitaciones de los nutrientes disponibles en el proceso de biosíntesis. Ácidos grasos, como PUFAs, y lípidos que contienen PUFAs son por sí mismos productos químicos finos deseables; optimizando la actividad o aumentando el número de una o varias desaturasas de acuerdo con la invención que participan en la biosíntesis de estos compuestos, o interfiriendo en la actividad de una o más desaturasas que participan en el catabolismo de estos compuestos, es posible aumentar de esta manera el rendimiento, la producción y/o la eficiencia de la producción de moléculas de ácido graso y lípidos en ciliados, algas, vegetales, hongos, levaduras u otros microorganismos.

La manipulación de uno o varios genes desaturasa de acuerdo con la invención también puede conducir a desaturasas con actividades modificadas que afectan indirectamente la producción de uno o varios productos químicos finos deseados a partir de algas, vegetales, ciliados u hongos. Los procesos metabólicos bioquímicos normales conducen, por ejemplo, a la producción de una gran número de productos residuales (por ejemplo, peróxido de hidrógeno y otras especies reactivas de oxígeno) que pueden interferir activamente en estos procesos metabólicos (por ejemplo, peroxinitrito se conoce por nitrar cadenas laterales de tirosina, desactivando así algunas enzimas que tienen tirosina en el centro activo (Groves, J.T. (1999) Curr. Opin. Chem. Biol. 3(2);226-235)). Estos productos residuales se excretan normalmente, pero las células usadas para la producción fermentativa a gran escala se optimizan para la sobreproducción de uno o varios productos químicos finos y de esta manera pueden producir más productos residuales de lo que es usual para una célula de tipo silvestre. Optimizando la actividad de una o más desaturasas según la invención que participan en la exportación de moléculas residuales, es posible mejorar la viabilidad de la célula y mantener una actividad metabólica eficiente. La presencia de cantidades intracelulares altas de los productos químicos finos también puede ser tóxica en la realidad, de modo que la viabilidad de la célula puede mejorarse incrementando la capacidad de la célula de secretar estos compuestos.

Las desaturasas de acuerdo con la invención pueden además manipularse de tal manera que se modifican las cantidades relativas de diversos lípidos y moléculas de ácido graso. Esto puede tener un efecto decisivo en la composición de lípidos de la membrana celular. Puesto que cada tipo de lípido tiene diferentes propiedades físicas, una modificación de la composición de lípidos puede cambiar significativamente la fluidez de la membrana. Las modificaciones de la fluidez de la membrana pueden afectar el transporte de moléculas a través de la membrana, tal como se explica previamente, lo cual puede modificar la exportación de productos residuales o del químico fino producido o la importación de nutrientes indispensables. Estos cambios en la fluidez de la membrana también pueden tener un efecto decisivo en la integridad de la célula; células con membranas comparativamente más débiles son más susceptibles a condiciones de estrés abióticas o bióticas que pueden dañar o matar la célula. Manipulando desaturasas que participan en la producción de ácidos grasos y lípidos para la síntesis de membrana, de modo que la membrana resultante tenga una composición de membrana que es más susceptible a las condiciones ambientales reinantes en los cultivos que se usan para la producción de químicos finos, debe sobrevivir una mayor fracción de las células y multiplicarse. Cantidades mayores de células productoras deben manifestarse en rendimientos mayores, producción más alta o eficiencia de producción más alta de los químicos finos del cultivo.

Las estrategias de mutagénesis mencionadas previamente para desaturasas dirigidas a conducir hacia un rendimiento elevado de un producto químico fino no deben interpretarse de manera limitante; para el técnico en la materia son fácilmente obvias las variaciones de estas estrategias. Usando estos mecanismos y con la ayuda de los mecanismos allí divulgados, las moléculas de ácido nucleico y proteínas según la invención pueden usarse para generar algas, ciliados, vegetales, animales, hongos u otros microorganismos tales como *C. glutamicum*, los cuales expresan moléculas de ácido nucleico y proteína desaturasas mutadas de modo que se mejore el rendimiento, la producción y/o eficiencia de producción de un compuesto deseado. Este compuesto deseado puede ser cualquier producto natural de algas, ciliados, vegetales, animales, hongos o bacterias que comprenden los productos finales de rutas de biosíntesis e intermedios de rutas metabólicas que se dan de manera natural, y también moléculas que no se producen de manera natural en el metabolismo de estas células pero que se producen por células según la invención.

Otra forma de realización de acuerdo con la invención es un proceso para la producción de PUFAs que comprende cultivar un organismo que contiene un ácido nucleico según la invención, un constructo génico según la invención o un vector según la invención que codifican un polipéptido que alarga ácidos grasos de C<sub>18</sub>, C<sub>20</sub> o C<sub>22</sub> con al menos dos enlaces dobles en el molécula de ácido graso en al menos dos átomos de carbono en condiciones en las que se producen PUFAs en el organismo. Los PUFAs producidos mediante este proceso pueden aislarse cosechando los organismos ya sea desde el cultivo en el que crecen o desde el campo y quebrando y/o extrayendo el material cosechado con un solvente orgánico. De este solvente puede extraerse el aceite que contiene lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerina y/o ácidos grasos libres con contenido superior de PUFAs. Mediante hidrólisis básica o ácida de los lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerina pueden aislarse ácidos grasos con contenido superior de PUFAs. Un contenido superior de PUFAs significa al menos 5%, preferentemente 10%, particularmente preferido 20%, muy particularmente preferido 40% más de PUFAs que el organismo original que no posee ácido nucleico adicional que codifique desaturasa de acuerdo con la invención.

Los PUFAs producidos mediante este proceso son preferentemente moléculas de ácido graso de C<sub>18</sub> o de C<sub>20-22</sub> con al menos dos enlaces dobles en la molécula de ácido graso, preferiblemente tres, cuatro, en combinación con otra elongasa y una desaturasa Δ4, cinco o seis enlaces dobles. Estas moléculas de ácido graso de C<sub>18</sub> o de C<sub>20-22</sub> pueden

## ES 2 345 324 T3

aislarse del organismo en forma de un aceite, lípido o un ácido graso libre. Ejemplos de organismos adecuados son aquellos mencionados arriba. Los organismos preferidos son plantas transgénicas.

Una forma de realización de acuerdo con la invención son aceites, lípidos o ácidos grasos o fracciones de los mismos que se han producido mediante el proceso arriba descrito, particularmente preferible un aceite, lípido o una composición de ácido graso, que comprenden PUFAs y que provienen de plantas transgénicas.

Otra forma de realización de acuerdo con la invención es el uso de aceite, lípido o de la composición de ácido graso en forrajes, productos alimenticios, cosméticos o fármacos.

Otro objeto de la definición es un proceso para la identificación de un antagonista o agonista de desaturasa que comprende

a) poner en contacto las células que expresan el polipéptido de la presente invención con una sustancia candidato;

b) ensayar la actividad de desaturasa;

c) comparar la actividad de desaturasa en ausencia de la sustancia candidata; un incremento en la actividad de desaturasa más allá del estándar indica que el material candidato es un agonista y una reducción en la actividad de desaturasa indica que el material candidato es un antagonista.

La sustancia candidata mencionada puede ser una sustancia que se ha sintetizado químicamente o producida por microbios y puede darse, por ejemplo, en extractos celulares de, por ejemplo, plantas, animales o microorganismos. Además, la sustancia mencionada aunque puede ser conocida en el estado de la técnica, puede no ser conocida hasta ahora como la que incrementa la actividad de las desaturasas. La mezcla de reacción puede ser un extracto desprovisto de células o comprender una célula o cultivo celular. Los métodos adecuados son conocidos por el técnico en la materia y se describen en términos generales, por ejemplo, en Alberts, *Molecular Biology the cell*, 3rd Edition (1994), por ejemplo capítulo 17. Las sustancias mencionadas pueden adicionarse, por ejemplo, a la mezcla de reacción o al medio de cultivo o también inyectadas a las células o aspergidas sobre una planta.

Si se ha identificado una muestra que comprende una sustancia activa según el método de la invención, es posible aislar directamente la sustancia de la muestra original, o bien la muestra puede dividirse en diversos grupos, por ejemplo cuando se compone de un gran número de diversos componentes, para reducir el número de las diversas sustancias por muestra y luego repetir el método según la invención con una "submuestra" de la muestra original. Dependiendo de la complejidad de la muestra, pueden repetirse varias veces los pasos descritos arriba, preferiblemente hasta que la muestra identificada de conformidad con el método de la invención solo contenga un pequeño número de sustancias o solo una sustancia. Preferiblemente, la sustancias identificadas de conformidad con el método de la invención, o derivados de la misma, se siguen formulando de modo que sean adecuadas para usar en la reproducción de vegetales o cultivo de células o tejidos vegetales.

Las sustancias que se han ensayado e identificado de conformidad con el método de la invención pueden ser: bibliotecas de expresión, por ejemplo bibliotecas de expresión de ADNc, péptidos, proteínas, ácidos nucleicos, anticuerpos, pequeñas sustancias orgánicas, PNAs o similares (Milner, *Nature Medicin* 1 (1995), 879-880; Hupp, *Cell*. 83 (1995), 237-245; Gibbs, *Cell*. 79 (1994), 193-198 y referencias allí citadas). Estas sustancias también pueden ser derivados o análogos funcionales de inhibidores o activadores conocidos. Los métodos para preparar derivados o análogos químicos son conocidos para el técnico en la materia. Los derivados y análogos mencionados pueden ensayarse de conformidad con métodos del estado de la técnica. Además, el diseño apoyado en ordenador o la péptido mimética pueden usarse para producir derivados y análogos adecuados. La célula o el tejido que pueden usarse para el método de la invención es preferiblemente una célula huésped, una célula vegetal o un tejido vegetal, de acuerdo con la invención, tal como se describe en las formas de realización arriba mencionadas.

De manera correspondiente la presente invención también se refiere a una sustancia que ha sido identificada de conformidad con los métodos de la invención. La sustancia es, por ejemplo, un homólogo de las desaturasas de la invención. Pueden generarse homólogos de desaturasas mediante mutagénesis, por ejemplo mediante mutación puntual o delección de las desaturasas. El término "homólogo" tal como se usa aquí denota una forma variante de las desaturasas que actúa como agonista o antagonista para la actividad de las desaturasas. Un agonista puede tener esencialmente la misma actividad biológica, o parte de ella, de las desaturasas. Un antagonista de las desaturasas puede inhibir una o más actividades de las formas que se dan naturalmente de las desaturasas, por ejemplo puede enlazarse de manera competitiva a un miembro ubicado hacia abajo o hacia arriba (downstream o upstream) de la ruta metabólica de la síntesis de ácido graso, las cuales incluyen las desaturasas o enlazarse a desaturasas y reducir o inhibir la actividad.

Además, la presente invención también se refiere a un anticuerpo o a un fragmento del mismo, tal como se han descrito aquí, que inhibe la actividad de las desaturasas de la invención.

En otro aspecto, la presente invención se refiere a un anticuerpo que reconoce o enlaza específicamente agonistas o antagonistas de acuerdo con la invención arriba descritos.

Otro aspecto se refiere a una composición que comprende el anticuerpo, el stop identificado según el método de la invención o la molécula antisense.

En otra forma de realización, la presente invención se refiere a un kit que comprende el ácido nucleico de la invención, el constructo génico según la invención, la secuencia de aminoácido según la invención, la molécula de ácido nucleico antisense según la invención, el anticuerpo y/o la composición según la invención, un antagonista o agonista preparados por el método de la invención, y/o aceites, lípidos y/o ácidos grasos de acuerdo con la invención o una fracción de los mismos. Igualmente, el kit puede comprender las células huéspedes, organismos, vegetales según la invención o partes de los mismos, partes cosechables de las plantas según la invención o material de propagación o bien el antagonista o agonista según la invención. Los componentes del kit de la presente invención pueden empacarse en contenedores adecuados, por ejemplo con o en búferes u otras soluciones. Uno o más de los componentes arriba mencionados pueden empacarse en un contenedor o el mismo contenedor. Adicionalmente, o como alternativa, uno o más de los componentes arriba mencionados puede adsorberse sobre una superficie sólida, por ejemplo filtros de nitrocelulosa, láminas de vidrio, chips, membranas de nylon o placas de microtitulación. El kit puede usarse para cualquiera de los métodos y modalidad descritas aquí, por ejemplo para la producción de células huésped, plantas transgénicas, para la detección de secuencias homólogas, para la identificación de antagonistas o agonistas y similares. Además, el kit puede comprender instrucciones para el uso del kit para una de las aplicaciones arriba mencionadas.

Esta invención se ilustra con mayor detalle mediante los ejemplos siguientes, los cuales no deben ser considerados limitantes. El contenido de todas las referencias bibliográficas, solicitudes de patentes, patentes y solicitudes publicadas de patentes, citadas en esta solicitud de patente, debe incorporarse aquí por referencia.

## 25 Sección de ejemplos

### Ejemplo 1

#### *Métodos generales*

30 a) *Métodos generales de clonación:*

Se llevaron a cabo métodos de clonación tales como, por ejemplo, escisiones de restricción, electroforesis en gel de agarosa, purificación de fragmentos de ADN, transferencia de ácidos nucleicos a nitrocelulosa y membranas de nylon, enlazamiento de fragmentos de ADN, transformación de células de *Escherichia coli* y de levadura, cultivo de bacterias y análisis de secuencias de ADN recombinante, tal como se describen en Sambrook y colaboradores (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) o Kaiser, Michaelis y Mitchell (1994) "Methods in Yeast Genetics" (Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3). La transformación y cultivo de algas como *Clorella* o *Phaeodactylum* se realizan tal como se describe por El-Sheekh (1999), *Biologia Plantarum* 42: 209-216; Apt y colaboradores (1996) *Molecular and General Genetics* 252 (5): 872-9.

#### b) *Sustancias químicas*

45 Si no se indica algo diferente en el texto, las sustancias químicas usadas se obtuvieron de calidad analítica de las empresas Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) y Sigma (Deisenhofen). Se prepararon soluciones usando agua pura libre de pirógenos, denominada en el texto que sigue como H<sub>2</sub>O, de una unidad de purificación de agua del sistema de agua Milli-Q (Millipore, Eschborn). Se compraron nucleasas de restricción, enzimas modificadoras de ADN y kits de biología molecular de las empresas AGS (Heidelberg), Amersham (Bruns-  
50 wick), Biometra (Göttingen), Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/ Taunus), Novagen (Madison, Wisconsin, Estados Unidos de America), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) y Stratagene (Amsterdam, Holanda). Si no se indica otra cosa, se han usado según las indicaciones del fabricante.

55 c) *Material celular*

Las secuencias aisladas de ácidos nucleicos de acuerdo con la invención están contenidas en el genoma de un *Phaeodactylum tricornutum* de cepa UTEX646, el cual se encuentra disponible de la colección de algas de la University of Texas, Austin.

*Phaeodactylum tricornutum* se cultivó a 25°C con un ritmo de luz/oscuridad de 14:10 horas a 22°C y 35 micro-Einstein (corresponde a micromol de fotones por metro cuadrado y segundo) en tubos de vidrio, por los cuales se hizo pasar aire.

65

## ES 2 345 324 T3

Como medio de cultivo para *Phaeodactylum tricornerutum* se usó el medio de cultivo f/2 con 10% de medio orgánico según Guillard, R.R.L. (1975; Culture of phytoplankton for feeding marine invertebrates. En: Smith, W.L. and Chanley, M.H. (Eds.) Culture of marine Invertebrate animals, NY Plenum Press, pp. 29-60.): contiene

5 995,5 ml de agua marina (artificial)

1 ml de NaNO<sub>3</sub> (75 g/l), 1 ml de NaH<sub>2</sub>PO<sub>4</sub> (5 g/l), 1 ml de solución de microelementos (o de trazas), 1 ml de tris/Cl pH 8,0, 0,5 ml de solución vitamínica f/2

10 Solución de microelementos: Na<sub>2</sub>EDTA (4,36 g/l), FeCl<sub>3</sub> (3,15 g/l), microelementos primarios: CuSO<sub>4</sub> (10 g/l), ZnSO<sub>4</sub> (22 g/l),

CoCl<sub>2</sub> (10 g/l), MnCl<sub>2</sub> (18 g/l), NaMoO<sub>4</sub> (6,3 g/l)

15 Solución vitamínica f/2: biotina: 10 mg/l, tiamina 200 mg/l, Vit B12 0,1 mg/l

Medio orgánico: Na-acetato (1 g/l), glucosa (6 g/l), Na-succinato (3 g/l), bacto-triptona (4 g/l), extracto de levadura (2 g/l).

20

### Ejemplo 2

#### *Aislamiento de ADN total de Phaeodactylum tricornerutum UTEX646 para experimentos de hibridación*

25 Los detalles del aislamiento de ADN total se refieren al procesamiento de material vegetal con un peso fresco de un gramo.

Búfer de CTAB: 2% (peso/vol.) bromuro de N-acetil-N,N,N-trimetilamonio (CTAB); 100 mM tris-HCl, pH 8,0; 1,4 M NaCl; 20 mM EDTA.

30

Búfer de N-laurilsarcosina: 10% (peso/vol) de N-laurilsarcosina; 100 mM tris-HCl, pH 8,0; 20 mM EDTA.

Material celular de *Phaeodactylum tricornerutum* fue triturado en un mortero bajo nitrógeno líquido de modo que se obtuvo un polvo fino que fue transferido a envases Eppendorf de 2 ml. El material vegetal congelado se cubrió luego con una capa de 1 ml de búfer de desintegración (1 ml de búfer CTAB, 100 ml de búfer N-laurilsarcosina, 20 ml de  $\beta$ -mercaptoetanol y 10 ml de solución K de proteinasa, 10 mg/ml) y se incubó a 60°C por una hora agitando continuamente. El homogeneizado obtenido se distribuyó en dos envases Eppendorf (2 ml) y se extrajo dos veces agitando con un volumen igual de cloroformo/alcohol isoamílico (24:1). Para la separación de fases se realizó una centrifugación a 8000 x g y TA (= temperatura ambiente = ~ 23°C) en cada caso por 15 minutos. Luego se precipitó el ADN por 30 minutos a -70°C usando isopropanol helado. El ADN precipitado se sedimentó por 30 minutos a 10 000 g a 4°C y se resuspendió en 180 ml de búfer TE (Sambrook y colaboradores, 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6). Para purificación adicional se trató el ADN con NaCl (1,2 M de concentración final) y nuevamente se precipita por 30 min usando el doble de volumen de etanol absoluto a -70°C. Después de un paso de lavado con etanol al 70%, se secó el ADN y a continuación se llevó a 50 ml de H<sub>2</sub>O + RNAsa (50 mg/ml de concentración final). El ADN se disolvió por toda una noche a 4°C y se llevó a cabo la escisión de RNAsa a continuación por 1 hora a 37°C. La conservación del ADN se efectuó a 4°C.

### Ejemplo 3

50

#### *Aislamiento del ARN total y poli(A)<sup>+</sup>-ARN de vegetales y Phaeodactylum tricornerutum*

55 El aislamiento de ARN total de vegetales como lino y colza, etc. se efectúa según uno de los métodos descritos por Logemann *et al* (1987, Anal. Biochem. 163, 21). El ARN total de musgo puede obtenerse de tejido de protonema usando el método GTC (Reski y colaboradores, 1994, Mol. Gen. Genet., 244: 352-359).

#### *Aislamiento de ARN de Phaeodactylum tricornerutum*

60 Muestras de algas congeladas (-70°C) se trituraron en un mortero helado bajo nitrógeno líquido hasta un polvo fino. Medio de homogenización de 2 volúmenes (12,024 g de sorbitol, 40,0 ml de 1M tris-HCl, pH 9 (0,2 M); 12,0 ml de 5 M NaCl (0,3 M), 8,0 ml de 250 mM EDTA, 761,0 mg de EGTA, 40,0 ml de SDS al 10% se llenaron a 200 ml con H<sub>2</sub>O y el pH se ajustó a 8,5) y se adicionaron 4 volúmenes de fenol con mercaptoetanol al 0,2% a 40 hasta 50°C mezclando bien hasta un polvo de células congelado. Después se adicionaron 2 volúmenes de cloroformo y se revolviéron vigorosamente por 15 minutos. Se centrifugó por 10 minutos a 10000 g y la fase acuosa se extrajo con fenol/cloroformo (2 volúmenes) y finalmente con cloroformo.

65

## ES 2 345 324 T3

El volumen obtenido de la fase acuosa se mezcló con 1/20 de volumen de acetato de sodio de 4 M (pH 6) y 1 volumen de isopropanol (helado) y los ácidos nucleicos se precipitaron a -20°C durante una noche (= UN). A continuación se centrifugó por 30 min a 10000 g y se succionó el sobrenadante. A esto le siguió un paso de lavado con EtOH al 70% EtOH y nuevamente una centrifugación. El sedimento se llevó a búfer de tris-borato (búfer de tris-borato de 80 mM, EDTA de 10 mM, pH 7,0). Luego se mezcló el sobrenadante con 1/3 de volumen de LiCl de 8 M, mezclado se incubó por 30 min a 4°C. Después de centrifugar nuevamente el sedimento se lavó con etanol al 70%, se centrifugó y se disolvió el sedimento en agua libre de RNAsa.

El aislamiento de poly(A)<sup>+</sup>-ARN se efectuó usando Dyna Beads<sup>®</sup> (Dyna, Oslo, Finland) según las indicaciones en el protocolo del fabricante.

Después de la determinación de la concentración de ARN o poli(A)<sup>+</sup>-ARN, el ARN se precipitó por adición de 1/10 de volumen de acetato de sodio de 3M, pH 4,6, y 2 volúmenes de etanol y se conservó a -70°C.

Para el análisis se separaron respectivamente 20 µg de ARN en un gel de agarosa al 15% que contiene formaldehído y se transfirieron a membranas de nylon (Hybond, Amersham). Se realizó la detección de transcritos específicos tal como se describe por Amasino ((1986) Anal. Biochem. 152, 304)).

### 20 Ejemplo 4

#### *Construcción del banco de ADNc*

Para la construcción del banco de ADNc de *Phaeodactylum tricornutum* la síntesis del primer cordón se llevó a cabo usando transcriptasa inversa del virus de leucemia en ratones (Roche, Mannheim, Alemania) y oligo-d(T)-cebadores, la síntesis del segundo cordón se logró por incubación con ADN-polimerasa I, enzima Klenow y H-escisión ARNasa a 12°C (2 horas), 16°C (1 hora) y 22°C (1 hora). La reacción se detuvo por incubación a 65°C (10 min) y a continuación se transfirió a hielo. Las moléculas de ADN bicatenario se provisionaron con T4-ADN-polimerasa (Roche, Mannheim) a 37°C (30 min) con extremos lisos. Los nucleótidos se retiraron por extracción con fenol/cloroformo y columnas de centrifugación Sefadex-G50. Adaptadores EcoRI/XhoI (Pharmacia, Freiburg, Alemania) se ligaron a los extremos de ADNc por medio de 4-ADNligasa (Roche, 12°C, por una noche) a los extremos de ADNc, se recortó con XhoI y se fosforiló por incubación con polinucleotidocinasa (Roche, 37°C, 30 min). Esta mezcla se sometió a la separación sobre un gel de agarosa con punto bajo de fusión. Las moléculas de ADN con más de 300 pares de bases se eluyeron sobre el gel, se extrajo con fenol se concentró en columnas Elutip-D (Schleicher y Schüll, Dassel, Alemania) y se ligó a brazos de vector y se empacó en fagos lambda-ZAP-Express usando el Gigapack Gold-Kits (Stratagene, Amsterdam, Holanda); se usó el material del fabricante y se siguieron sus indicaciones.

### 40 Ejemplo 5

#### *Secuenciación de ADN y análisis en ordenador*

Bancos de ADNc, como los descritos en el ejemplo 4, se usaron para la secuenciación de ADN según métodos estándar, en especial el método de terminación de cadena usando el ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction-Kit (Perkin-Elmer, Weiterstadt, Alemania). La secuenciación de clones aleatorios individualizados se llevó a cabo a continuación de la obtención preparativa de plásmido a partir de bancos de ADNc por excisión *in vivo* y retransformación de DH10B sobre placas de agar (detalles sobre el material y el protocolo: Stratagene, Amsterdam, Holanda). Se preparó ADN de plásmido a partir de cultivos de *E. coli* reproducidos durante una noche en caldo de Luria con ampicilina (véase Sambrook y colaboradores (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)), en un robot de preparación de ADN de Qiagen (Qiagen, Hilden) según los protocolos del fabricante. Se usaron cebadores de secuencia con las siguientes secuencias de nucleótidos:

5'-CAGGAAACAGCTATGACC-3'

5'-CTAAAGGGAACAAAAGCTG-3'

5'-TGTAACGACGGCCAGT-3'

Las secuencias se procesaron y anotaron usando el paquete de software estándar EST-MAX, que se suministra comercialmente por Bio-Max (Munich, Alemania). Aprovechando algoritmos comparativos usando la secuencia de búsqueda mostrada en SEQ ID NO: 8, se buscaron genes homólogos con ayuda del programa BLAST (Altschul y colaboradores (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25: 3389-3402.). Se caracterizaron detalladamente dos secuencias de *Phaeodactylum tricornutum* con homologías hacia la secuencia de búsqueda de *Physcomitrella patens*.

## ES 2 345 324 T3

### Ejemplo 5a

*Aislamiento de desaturasas de Phaeodactylum tricornutum por reacción en cadena de polimerasa con ayuda de oligonucleótidos degenerados*

5 Con ayuda de desaturasas publicadas pudieron identificarse motivos que son típicos para desaturasas  $\Delta$ -5 y  $\Delta$ -6. En lo sucesivo se representan secuencias de oligonucleótidos con posibles variaciones. Bajo la secuencia de oligonucleótidos, el aminoácido del cual puede derivarse la combinación base se representa en el código de una letra. Por ejemplo, A/G significa que en esta posición durante la síntesis de la unidad estructural, una A o una G se incorporan aleatoriamente en el oligonucleótido, puesto que la tripleta base derivada del aminoácido correspondiente puede ser AAA o AAG. La secuencia de ADN también puede contener una inosina (i) si la determinación de una base en esta posición permite tres o cuatro bases diferentes debido al código genético. Pueden usarse las siguientes secuencias y cebadores (primers):

15

5'-adelante-cebador:

20 F1a: TGG TGG AA A/G TGG AAi CA T/C AA  
 F1b: TGG TGG AA A/G TGG ACi CA T/C AA  
 F1a: W W K W N/T H K/N  
 F1b: W W K W K H K/N

25

F2a: Gi TGG AA A/G GAI A/C Ai CA T/C AA  
 F2b: Gi TGG AA A/G TTG A/C Ai CA T/C AA  
 F2a: G/W W K E/D K/Q/N H K/N  
 F2b: G/W W K W K/Q/N H K/N

30

F3a: T A/T i TTG AAi A/C A A/G C/A G/A i CA  
 F3b: T A/T i TTG AAi A/C A A/G CAi CA  
 F3a: W W K/N H/N R/Q H  
 F3b: Y W K/N H/N R/Q H

35

F4a: GTi TGG A A/T G/A GA A/G CA A/G CA  
 F4b: GTi TGG A A/T G/A A/T A T/C CA A/G CA  
 F4a: V W K/M E Q H  
 F4b: V W K/M N/Y Q H

40

F5a1: CA T/C TA T/C TGG AA A/G AA T/C CA G C  
 F5a1: CA T/C TA T/C TGG AA A/G AA T/C CA A C  
 F5a1: H Y W K N Q H/Q

45

F6a: TTG TTG AAi A/C A A/G AA i CA T/C AA  
 F6a: W W K/N H/N K/N H K/N

50

55

60

65

## ES 2 345 324 T3

3'-reversa cebador

```

5   R1b:      GG  A/G AA  iAG  G/A TG  G/A TG  T/C TC
   R1b:      GG  A/G AA  iAA  G/A TG  G/A TG  T/C TC
   R1a:      P    F    L    H    H    E
   R1b:      P    F    F    H    H    E
10  R2a1:     AA   iAG  A/G TG  A/G TG  iA C/T  iA/G  T/C TG
   R2a2:     AA  T/C AA  A/G TG  A/G TG  iA C/T  iA/G  T/C TG
   R2a1:     F    L    H    H    V/I    V/A    Q
   R3a1:     AT  iTG   iGG  A/G AA  iAA   A/G TG  A/G TG
   R3a2:     AT  A/G TT  iGG  A/G AA  iAA   A/G TG  A/G TG
15  R3a3:     AT  iTG   iGG  A/G AA  iAG   A/G TG  A/G TG
   R3a4:     AT  A/G TT  iGG  A/G AA  iAG   A/G TG  A/G TG
   R3a1:     I/M H/Q   P    F    F    H    H
   R3a2:     I/M N    P    F    L    H    H
20  R4a1:           CT   iGG  A/G AA  iA A/G  A/G TG  A/G TG
   R4a2:           GA   iGG  A/G AA  iA A/G  A/G TG  A/G TG
   R4a3:           GT   iGG  A/G AA  iA A/G  A/G TG  A/G TG
   R4a1:     =  T/R/S  P    F    F/L    H    H
25  R5a1:     AA  iAA   A/G TG  A/G TG  T/C TC  T/A/G AT  T/C TG
   R5a2:     AA  iAG   A/G TG  A/G TG  T/C TC  T/A/G AT  T/C TG
   R5a1:     F  F    H    H    E    I    Q
   R5a2:     F  L    H    H    E    I    Q
30  R6a1:     T    iGG  iA A/G  iAA   A/G TG  A/G TG  iAC
   R6a1:     T    iGG  iA A/G  iAG   A/G TG  A/G TG  iAC
   R6a1:     T/N   P    L    F/L    H    H    V

```

35

Debido a las diversas posibilidades de variaciones son posibles muchos oligonucleótidos derivados, aunque ha sido encontrado de manera sorprendente que los oligonucleótidos representados pueden ser particularmente adecuados para el aislamiento de desaturasa.

40

Los cebadores pueden aplicarse en todas las combinaciones para reacciones en cadena de polimerasa. Con ayuda de combinaciones individuales pudieron aislarse fragmentos de desaturasa cuando se tuvieron en cuenta las siguientes condiciones: Para reacciones PCR se emplearon en cada caso 10 nMol de cebador y 10 ng de un banco de plásmidos obtenidos por excisión *in vivo*. El banco de plásmidos pudo aislarse según protocolos del fabricante (Stratagene) a partir del banco de fagos. La reacción PCR se realizó en un termociclador (Biometra) con la Pfu-ADN-polimerasa (Stratagene) y el siguiente programa de temperaturas: 3 min a 96°C, seguido de 35 ciclos con 30 s a 96°C, 30 s a 55°C y 1 min a 72°C. Después del primer paso a 55°C, la temperatura de adhesión se redujo a pasos en cada caso de 3°C y después del quinto ciclo, se mantuvo una temperatura de adhesión de 40°C. Finalmente, se llevó a cabo un ciclo de 10 minutos a 72°C y la reacción se detuvo enfriando a 4°C.

50

Las combinaciones de cebadores F6a y R4a2 se resaltan en el texto subrayados y pueden aprovecharse exitosamente para el aislamiento de un fragmento de desaturasa. El fragmento resultante pudo verificarse por secuenciación y mostró homologías con una desaturasa con el No. de acceso al banco genético T36617 de *Streptomyces coelicolor*. La homología se obtuvo con ayuda del programa BLASTP. La comparación se representa en la Figura 4. Resultaron identidades de 34% y una homología de 43% con la secuencia T36617. El fragmento de ADN se empleó de acuerdo con la invención como se muestra en el ejemplo 7 en un experimento de hibridación para el aislamiento de un gen de longitud completa en condiciones estándar.

55

60

La región de codificación de una secuencia de ADN aislada de esta manera se obtuvo traduciendo el código genético a una secuencia de polipéptidos. En SEQ ID NO: 3 se representa una secuencia de 1434 pares de bases de longitud que pudo aislarse por el método descrito. La secuencia posee un codón de inicio en posición 1 a 3 y un codón stop en posición 1432-1434 y pudo traducirse a un polipéptido de 477 aminoácidos de longitud. Comparando con una secuencia génica descrita en WO 98 46763 se ha encontrado que un fragmento no idéntico pero homólogo de *Phaeodactylum tricorutum* que codifica para 87 aminoácidos ya había sido designado. Sin embargo, WO 98/46763 no divulga una desaturasa completa, funcionalmente activa ni especificidad por posición o sustrato. Esto también se ha hecho claro por el hecho de que tanto la homología con desaturasa  $\Delta$ -5 como con  $\Delta$ -6 de *Mortierella alpina* se reportan sin indicar una función específica. La secuencia según la invención, por lo contrario, codifica una desaturasa de  $\Delta$ -6 acil lípido funcionalmente activa.

65

## ES 2 345 324 T3

### Ejemplo 6

#### *Identificación de secuencias de ADN que codifican para desaturasas de Phaeodactylum tricorutum*

5 La secuencia de longitud completa de la desaturasa  $\Delta$ -6-acil lípido Pp\_des6 AJ222980 (NCBI No. de acceso al banco genómico) del musgo *Physcomitrella patens* (véase también Tabla 1) así como la secuencia de desaturasa  $\Delta$ -12-acil lípido (Tabla 1, véase Ma\_des12) de *Mortierella alpina* AF110509 (AF110509 NCBI No de acceso al banco genómico) se emplearon para alineamiento de secuencia con la ayuda del algoritmo de búsqueda TBLASTN.

10 Las secuencias EST PT0010070010R, PT001072031R y PT001078032R se consideraron primero como gen objetivo entre otros genes candidatos debido a homología débiles con las secuencias de búsqueda de *Physcomitrella* y *Mortierella*. En las figuras 1 y 2 así como en la figura 2a se muestra el resultado de las dos secuencias EST encontradas. Las secuencias encontradas son parte de los ácidos nucleicos de la invención de SEQ ID NO: 1 (nombre de gen: Pt\_des5, No. del banco de datos propio del inventor PT001078032R), SEQ ID NO: 5. (nombre del gen: Pt\_des12, No. del banco de datos propio del inventor PT0010070010R) y SEQ ID NO: 11 (nombre del gen: Pt\_des12.2, banco de datos propio del inventor PT001072031R). las letras indican aminoácidos idéntico mientras que el signo más significa un aminoácido químicamente similar. Las identidades u homologías de todas las secuencias encontradas de acuerdo con la invención se infieren de la recopilación de la tabla 2.

20 Las desaturasas pueden presentar dominios citochrom b5 que también se dan en otros genes que no codifican desaturasas. Dominios citochrom b5 muestran así altas homologías, aunque se trate de funciones génicas diferentes. Dentro de regiones conservadas débilmente, las desaturasas pueden identificarse solo como genes candidatos putativos y deben ensayarse para la actividad enzimática y la especificidad para la posición de la función enzimática. Por ejemplo, diversas hidroxilasas, acetilenasas y epoxigenasas, como desaturasas, muestran también motivos de histidín-box, de modo que debe probarse experimentalmente una función específica y solo la verificación adicional del enlace doble hace posible una actividad enzimática garantizada y una especificidad para posición de una desaturasa. Sorprendentemente se ha encontrado que desaturasas  $\Delta$ 6 y  $\Delta$ 5 según la invención tienen especificidades de sustrato particularmente adecuadas y son particularmente adecuadas para aprovecharlas, en combinación con una elongasa  $\Delta$ 6 de *Physcomitrella*, para producir ácidos grasos poliinsaturados tales como ácido araquidónico, ácido eicosapentanoico y ácido docosahexanoico.

35 La secuenciación del fragmento completo de ADN del clon PT001078032R dio lugar a una secuencia de 1652 pares de bases de longitud. La secuencia codifica para un polipéptido de 469 aminoácidos representada en SEQ ID NO: 2. Esta se obtuvo traduciendo el código genético de SEQ ID NO: 1 con un codón de inicio en posición de par de base 115-117 y con un codón stop en posición de par de base 1522-1524. El clon comprende un polipéptido desaturasa completo, como puede verse de la alineación de secuencia en la figura 3. Las rayas denotan aminoácidos idénticos, mientras que los dos puntos y los puntos representan aminoácidos químicamente intercambiables, es decir químicamente equivalentes. La comparación se realizó usando matriz de intercambio BLOSUM62 para aminoácidos según Henikoff & Henikoff: ((1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919). Parámetros usados:

Gap Weight: 8; Average Match: 2.912, Length Weight: 2, Average Mismatch: -2.003.

45 En la figura 6 y figura 7 se representa la comparación de la secuencia de péptidos MA\_des12 con las secuencias encontradas.

50 La secuenciación del fragmento completo de ADNc del clon PT0010070010R dio lugar a una secuencia representada en SEQ ID NO: 5 con 1651 pares de bases de longitud, con un codón de inicio en posición 67-69 y un codón de pare en posición 1552-1554. La secuencia de polipéptidos de acuerdo con la invención se representa en SEQ ID NO: 6.

55 La secuenciación del fragmento completo de ADNc del clon PT0010072031R dio lugar a una secuencia representada en SEQ ID NO: 11 con 1526 pares de bases de longitud, con un codón de inicio en posición 92-94 y un codón de pare en posición 1400-1402. La secuencia de polipéptidos de acuerdo con la invención se representa en SEQ ID NO: 12.

60 En la Tabla 2 se representan las identidades y homologías de las desaturasas de acuerdo con la invención unas con otras y con la desaturasa de *Physcomitrella patens* y *Mortierella alpina*. Los datos se obtuvieron con la ayuda del programa Bestfit bajo los parámetros dados tal como se define abajo como un programa parcial del siguiente software: Wisconsin Package Version 10.0 (Genetics Computer Group (GCG), Madison, Wisc., USA). Henikoff, S. and Henikoff, J.G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.

65 Adicionalmente en la figura 5 se representa la comparación de la desaturasa  $\Delta$ -6-acil lípido de *Physcomitrella patens* con la secuencia de polipéptido del clon Pt\_des6.

## ES 2 345 324 T3

TABLA 2

Homología/identidad en %	Secuencia de búsqueda Pp_des6	Secuencia de búsqueda Ma_des12
Pt_des5	34.92/26.37	n.r.
Pt_des6	50.69/41.06	n.r.
Pt_des12	n.r.	48.58/38.92
Pt_des12.2	n.r.	48.37/41.60
n.r. = no realizado		

Con la ayuda del algoritmo TBLASTN 2.0.10: Altschul *et al* 1997, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25: 3389-3402 se identificaron secuencias con homología, o identidad, superior de secuencia por medio de una comparación local de banco de datos. Los resultados se representan en la siguiente Tabla 2A.

TABLA 2A

*Homólogos con las homología o identidades de secuencia más altas a las secuencias de polipéptidos según la invención de SEQ NO. 2, 4, 6 ó 12*

Homología / identidad (%)	Secuencia de búsqueda PT001070010R	Secuencia de búsqueda PT001072031R	Secuencia de búsqueda PT001078032R	Secuencia de búsqueda Pt_ des6
L26296: Fad2 A. thaliana	50 % / 37 %	n.r.	n.r.	n.r.
U86072 Petroselinum crispum Fad2	n.r.	51/40	n.r.	n.r.
AL358652 L. major putative desaturase	n.r.	n.r.	45/30	n.r.
AB020032 M. alpina delta 6 desaturase	n.r.	n.r.	n.r.	53/38

### Ejemplo 7

#### *Identificación de genes por medio de hibridación*

Pueden usarse secuencias génicas para la identificación de genes homólogos o heterólogos de bancos de ADNc o genómicos.

Genes homólogos (es decir, clones de ADNc de longitud completa que son homólogos, u homólogos) pueden aislarse mediante hibridación de ácido nucleico, usando por ejemplo bancos de ADNc: en particular el método puede usarse para el aislamiento de genes de longitud completa funcionalmente activos de las que se muestran en la SEQ ID NO: 3. Dependiendo de la frecuencia del gen de interés, se colocan sobre la placa 100000 hasta 1000000 de bacteriofagos recombinantes y se trasladan a una membrana de nylon. Después de la desnaturalización con álcali, se inmovilizó el ADN sobre la membrana, por ejemplo mediante reticulación UV. La hibridación se realiza en condiciones

## ES 2 345 324 T3

altamente severas. En solución acuosa se realizan la hibridación y los pasos de lavado a una fuerza iónica de NaCl de 1 M y una temperatura de 68°C. Se han producido sondas de hibridación, por ejemplo, mediante marcación por medio de nick-transcripción radioactiva (<sup>32</sup>P-) (High Prime, Roche, Mannheim, Alemania). Las señales se detectan por medio de autoradiografía.

Genes parcialmente homólogos o heterólogos que están relacionados, pero no son idénticos, pueden identificarse de manera análoga al método descrito arriba usando condiciones de hibridación y de lavado de baja severidad. Para la hibridación acuosa se mantuvo la fuerza iónica habitualmente a NaCl de 1 M y la temperatura se disminuyó gradualmente de 68 a 42°C.

El aislamiento de secuencias génicas que solo exhiben homología con un dominio individual de, por ejemplo, 10 a 20 aminoácidos puede realizarse usando sondas de oligonucleótidos sintéticas, marcadas radioactivamente. Oligonucleótidos marcados radioactivamente se producen por medio de fosforilación del extremo 5' de dos oligonucleótidos complementarios con T4-polinucleótidocinasa. Los oligonucleótidos complementarios se hibridan y se ligan unos a otros de modo que surgen concatémeros. Los concatémeros bicatenarios se marcan radiactivamente por nicktranscripción. La hibridación se efectúa habitualmente en condiciones de baja severidad usando concentraciones altas de oligonucleótidos.

Solución de hibridación de oligonucleótidos:

6 x SSC

0,01 M fosfato de sodio

1 mM EDTA (pH 8)

0,5% SDS

100 microg/ml de ADN de esperma desnaturalizado de salmón

0,1% leche seca magra

Durante la hibridación la temperatura baja a pasos a 5 hasta 10°C por debajo de la T<sub>m</sub> calculada de oligonucleótido o hasta temperatura ambiente (a menos que se especifique otra cosa T<sub>A</sub> = ~ 23°C en todos los experimentos), seguido de pasos de lavado y autoradiografía. El lavado se efectúa a severidad extremadamente baja, por ejemplo tres pasos de lavado usando 4xSSC. Otros detalles son tal como se describe por Sambrook, J., y colaboradores (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, o Ausubel, F.M., y colaboradores (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.

### Ejemplo 8

*Identificación de genes objetivo mediante clasificación de bancos de expresión con anticuerpos*

Para generar proteína recombinante, por ejemplo *E. coli*, se usaron secuencias de ADNc (por ejemplo Qiagen QIAexpress pQE-System). Las proteínas recombinantes se purificaron por afinidad habitualmente mediante cromatografía de afinidad Ni-NTA (Qiagen). Las proteínas recombinantes se usaron entonces para la producción de anticuerpos específicos, usando por ejemplo técnicas estándar para la inmunización de conejos. A continuación, los anticuerpos se purificaron por afinidad usando una columna Ni-NTA que se desatura con antígeno recombinante, tal como se describe por Gu y colaboradores, (1994) *BioTechniques* 17: 257-262. El anticuerpo puede usarse entonces para expresión de cribado de los bancos de ADNc de expresión por medio de clasificación inmunológica (Sambrook, J., y colaboradores (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, o Ausubel, F.M., y colaboradores (1994) "Current Protocols in Molecular Biology", John Wiley & Sons).

### Ejemplo 9

*Transformación de Agrobacterium*

La transformación vegetal mediada por *Agrobacterium* puede realizarse, por ejemplo, usando la cepa *Agrobacterium tumefaciens* GV3101-(pMP90-) (Koncz y Schell, *Mol. Gen. Genet.* 204 (1986) 383-396) o LBA4404 (Clontech) o C58C1 pGV2260 (Deblaere *et al* 1984, *Nucl. Acids Res.* 13, 4777-4788). LA transformación puede realizarse mediante técnicas estándar de transformación (también Deblaere y colaboradores 1984).

## Ejemplo 10

*Transformación vegetal*

5 La transformación vegetal mediada por *Agrobacterium* puede realizarse usando técnicas estándar de transformación y de regeneración (Gelvin, Stanton B., Schilperoort, Robert A., *Plant Molecular Biology Manual*, 2. Edición, Dordrecht: Kluwer Academic Publ., 1995, en Sect., Ringbuc Zentrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton: CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

10 Por ejemplo, colza puede transformarse mediante transformación de cotiledóneas o hipocotiledóneas (Moloney y colaboradores, *Plant Cell* 8 (1989) 238-242; De Block y colaboradores, *Plant Physiol.* 91 (1989) 694-701). El uso de antibióticos para la selección de *Agrobacterium* y de vegetales depende del vector binario usado para la transformación y de la cepa de *Agrobacterium*. La selección de colza se realiza habitualmente usando canamicina como marcador vegetal seleccionable.

15 La transferencia de genes mediada por *Agrobacterium* en lino (*Linum usitatissimum*) puede realizarse usando, por ejemplo, una técnica descrita por Mlynarova y colaboradores (1994) *Plant Cell Report* 13: 282-285.

20 La transformación de soya puede realizarse usando, por ejemplo, una técnica descrita en EP-A-0 0424 047 (Pioneer Hi-Bred International) o en EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo).

25 La transformación de vegetales usando bombardeo de partículas, absorción de ADN mediante polietilenglicol o mediante la técnica de fibra de carbonato de silicio se describe, por ejemplo por Freeling y Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

## Ejemplo 11

30 *Plásmidos para la transformación de vegetales*

Para la transformación de vegetales pueden usarse vectores binarios, como pBinAR (Höfgen y Willmitzer, *Plant Science* 66 (1990) 221-230) o pGPTV (Becker *et al* 1992, *Plant Mol. Biol.* 20: 1195-1197) o derivados de los mismos. La construcción de los vectores binarios puede efectuarse mediante lizagón del ADNc en orientación sense o antisense en T-ADN. 5' del ADNc, un promotor vegetal activa la transcripción del ADNc. Una secuencia de poliadenilación se encuentra 3' del ADNc. Los vectores binarios pueden tener diferentes genes marcadores. En particular, el gen marcador nptII que codifica para resistencia de canamicina, mediada por neomicina fosfotransferasa puede intercambiarse por la forma resistente a herbicidas de un gen sintasa acetolactato (abreviatura: AHAS o ALS). El gen ALS se describe en Ott y colaboradores, *J. Mol. Biol.* 1996, 263: 359-360. El promotor v-ATPasa-c1 puede clonarse al plásmido pBin19 o pGPTV y se aprovecha para expresión génica de marcador mediante clonación antes de la región codificantes ALS. El promotor nombrado corresponde a un fragmento de pares de bases 1153 de beta-Vulgaris (*Plant Mol Biol*, 1999, 39: 463-475). Tanto sulfonilureas e imidazolinonas tales como imazetapir o sulfonilureas pueden usarse como antimetabolitos para selección.

45 La expresión específica para tejido puede lograrse usando un promotor específico para tejido. Por ejemplo, la expresión específica de semillas puede alcanzarse clonando el promotor DC3 o el LeB4 o el USP o el promotor 5' de faseolina del ADNc. También pueden usarse cada otro elemento promotor específico para semillas, como por ejemplo el promotor aapina o arcelina (Goossens y colaboradores 1999, *Plant Phys.* 120(4): 1095-1103 y Gerhardt y colaboradores 2000, *Biochimica et Biophysica Acta* 1490(1-2): 87-98). Para la expresión constitutiva en las plantas completas puede usarse el promotor CaMV-35S o un promotor de v-ATPasa C1.

50 En particular pueden clonarse a un vector binario genes que codifican desaturasas y elongasas, mediante construcción de varios casetes de expresión unos tras otros, con el fin de imitar la ruta metabólica en vegetales.

55 Dentro de un casete de expresión la proteína a expresarse puede dirigirse hacia un compartimiento celular usando un péptido de señal, por ejemplo para plástidos, mitocondrias o el retículo endoplasmático (Kermode, *Crit. Rev. Plant Sci.* 15, 4 (1996) 285-423). La señal de péptido se clona a 5' en el marco de lectura con el ADNc con el fin de alcanzar la localización subcelular de la proteína de fusión.

60 *Ejemplos de casetes de multiexpresión se dan a continuación*

I.) *Casetes promotor-terminador*

65 Los casetes de expresión están compuestos de al menos dos unidades funcionales, tales como un promotor y un terminador. Entre el promotor y el terminador pueden insertarse otras secuencias génicas deseadas como secuencias targeting (direccionadoras), regiones codificadoras de genes o partes de los mismos, etc. Para la estructuración de casetes de expresión, se aíslan promotores y terminadores (Promotor de USP: Baumlein y colaboradores, *Mol Gen*

## ES 2 345 324 T3

Genet, 1991, 225 (3): 459-67); Terminator de OCS: Gielen y colaboradores EMBO J. 3 (1984) 835 y siguientes) con la ayuda de la reacción en cadena de polimerasa y se corta a la medida con secuencias flanqueadoras a base de oligonucleótidos sintéticos.

5 Pueden usarse los siguientes oligonucleótidos, por ejemplo:

USP1 frente: CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

USP2 frente: CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

10

USP3 frente: CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

USP1 atrás: AAAACTGCAGGCGGCCGCCACCGCGGTGGGCTGGCTATGAAGAAATT

15

USP2 atrás: CGCGGATCCGCTGGCTATGAAGAAATT

USP3 atrás: TCCCCCGGGATCGATGCCGGCAGATCTGCTGGCTATGAAGAAATT

20

OCS1 frente: AAAACTGCAGTCTAGAAGGCCTCCTGCTTTAATGAGATAT

OCS2 frente: CGCGGATCCGATATCGGGCCCCGCTAGCGTTAACCCCTGCTTTAATGAGATAT

OCS3 frente: TCCCCCGGGCCATGGCCTGCTTTAATGAGATAT

25

OCS1 atrás: CCCAAGCTTGGCGCGCCGAGCTCGAATTCGTCGACGGACAATCAGTAAATTGA

OCS2 atrás: CCCAAGCTTGGCGCGCCGAGCTCGAATTCGTCGACGGACAATCAGTAAATTGA

30

OCS3 atrás: CCCAAGCTTGGCGCGCCGAGCTCGTCGACGGACAATCAGTAAATTGA

Los métodos son conocidos para el técnico en la materia en en general son conocidos en la literatura.

35 En un primer paso se amplifican un promotor y un terminador mediante PCR. Luego se clona el terminador en un plásmido receptor y en un segundo paso se inserta el promotor antes del terminador. Con esto se obtiene un casete de expresión sobre un plásmido soporte. Los plásmidos pUT1, pUT2 y pUT3 se generan sobre la base de plásmido pUC19.

40 Los constructos se definen de acuerdo con la invención en SEQ ID NO: 13, 14 y 15. Con base en pUC19 contienen el promotor USP y el terminador OCS. Con base en estos plásmidos se genera el constructo pUT12 cortando pUT1 mediante SalI/ScaI y cortando pUT2 mediante XhoI/ScaI. Los fragmentos que contienen casetes de expresión se ligan y se transforman en *E. coli* XLI blue MRF. Después de separar colonias resistentes a ampicilina se prepara ADN y mediante análisis de restricción se identifican aquellos clones que comprenden dos casetes de expresión. La ligazón de extremos compatibles ha eliminado los dos sitios de escisión XhoI/SalI. Esto da lugar al plásmido pUT12 que se define en SEQ ID NO: 16. A continuación, a su vez se corta pUT12 por medio de Sal/ScaI y se corta pUT3 por medio de XhoI/ScaI. Los fragmentos que contienen casetes de expresión se ligan y se transforman en *E. coli* XLI blue MRF. Después de separar colonias resistentes a ampicilina se prepara ADN y mediante análisis de restricción se identifican aquellos clones que contienen tres casetes de expresión. De esta manera se logra un conjunto de casetes multiexpresión los cuales pueden aprovecharse para la inserción de ADN deseado y se describe en la tabla 3 y adicionalmente puede incorporar otros casetes de expresión más.

50 Estos contienen los siguientes elementos:

55

60

65

## ES 2 345 324 T3

TABLA 3

Derivado de pUC19	Sitios de escisión antes del promotor USP	Sitios de escisión clonación múltiples	Sitios de escisión dentro del terminador OCS
pUT1	EcoRI/AscI/SacI/XhoI	BstXI/NotI/PstI/XbaI/StuI	SaII/EcoRI/SacI/AscI/ HindIII
pUT2	EcoRI/AscI/SacI/XhoI	BamHI/EcoRV/ApaI/NheI/ HpaI	SaII/EcoRI/SacI/AscI/ HindIII
pUT3	EcoRI/AscI/SacI/XhoI	BglII/NaeI/ClaI/SmaI/NcoI	SaII/SacI/AscI/HindIII
pUT12 Doble casete de expresión	EcoRI/AscI/ SacI/XhoI BstXI	BstXI/NotI/PstI/XbaI/StuI y BamHI/EcoRV/ApaI/ NheI/HpaI	SaII/EcoRI/SacI/AscI/ HindIII
pUT123 Triple casete de expresión	EcoRI/AscI/SacI/XhoI	1. BstXI/NotI/ PstI/XbaI/ StuI y 2. BamHI/EcoRV/ ApaI/ NheI/ HpaI y 3. BglII/NaeI/ ClaI/SmaI/ NcoI	SaII/SacI/AscI/HindIII

Adicionalmente, otros casetes de multiexpresión pueden describirse y especificarse más detalladamente como en la tabla 4 con la ayuda de

i) el promotor USP o con la ayuda de

ii) el fragmento 3' con alrededor de 700 del promotor LeB4 o generarse con la ayuda de

iii) el promotor DC3 y emplearse para expresión génica específica de semilla.

El promotor DC3-P se describe en Thomas, Plant Cell 1996, 263: 359-368 y se compone solo de la región -117 hasta +27 por lo que con esto representa uno de los promotores más pequeños conocidos. Los casetes de expresión pueden contener varias veces el mismo promotor o estar contruidos por tres promotores diferentes.

TABLA 4

*Casetes de expresión múltiples*

Nombre de plásmido del derivado pUC19	Sitios de escisión antes del promotor respectivo	Sitios de escisión clonación múltiples	Sitios de escisión detrás del terminador OCS
pUTI (pUC19 con USP-OCS1)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/PstI/XbaI/ StuI	SaII/EcoRI/SacI/AscI/ HindIII
pDCT (pUC19 con DC3-OCS)	EcoRI/AscI/SacI/XhoI	(2) BamHI/EcoRV/ApaI/ NheI/HpaI	SaII/EcoRI/SacI/AscI/ HindIII

## ES 2 345 324 T3

Nombre de plásmido derivado pUC19	Sitios de escisión antes del promotor respectivo	Sitios de escisión clonación múltiples	Sitios de escisión detrás del terminador OCS
pLeBT (pUC19-mit LeB4 (700)-OCS)	EcoRI/AscI/SacI/XhoI	(3) BglII/NaeI/ClaI/SmaI/ NcoI	Sall/SacI/AscI/HindIII
pUD12 (pUC19 mit con USPOCS1 y con DC3-OCS)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/PstI/XbaI/ StuI y (2) BamHI/EcoRV/ApaI/ NheI/HpaI	Sall/EcoRI/SacI/AscI/ HindIII
pUDL123 Casete de expresión triple (pUC19 con USP/DC3 y LeB4-700)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/PstI/XbaI/ StuI y (2) BamHI/ (EcoRV*)/ApaI/ NheI/HpaI y (3) BglII/NaeI/ClaI/SmaI/ NcoI	Sall/SacI/AscI/HindIII
* Sitio de escisión EcoRV corta en el fragmento de 700 pares de bases del promotor LeB4 (LeB4-700)			

De manera análoga otros promotores pueden generarse para constructos multigénicos, especialmente usando

- a) el fragmento de 2,7 kB del promotor LeB4-P con la ayuda del
- b) promotor faseolina o con la ayuda del
- c) Promotor constitutivo v-ATPasa c1.

Puede ser deseable en particular usar otros promotores particularmente adecuados para la estructuración casetes de expresión específicos de semillas, como por ejemplo el promotor napina o el promotor arcelina-5.

ii) *Generación de constructos de expresión en derivados pUC19 o pGPTV que obtienen promotor y terminador y contienen en combinación con secuencias génicas deseadas para la expresión génica de PUFA en casetes de expresión vegetales*

Los casetes de multiexpresión pueden insertarse por medio de AscI directamente de derivados pUC19 de la tabla 3 al vector pGPTV+AscI (véase iii.) por la posición de corte AscI y se encuentran s disposición para inserción de genes objetivos. Los constructos génicos correspondientes (pBUT1 está representado en SEQUENZ ID NO: 20, pBUT2 en SEQUENZ ID NO: 21, pBUT 3 en SEQUENZ ID NO: 22, pBUT12 en SEQUENZ ID NO: 22 y pBUT123 en SEQUENZ ID NO: 24) están disponibles según la invención como kits. De manera alterna pueden insertarse secuencias génicas a los casetes de expresión basados en pUC19 y pueden emplearse como fragmento AscI en pGPTV+AscI.

En pUT12 primero se inserta a través de BstXI y XbaI la  $\Delta$ -6-elongasa Pp\_PSE1 al primer casete. Luego la  $\Delta$ -6-desaturasa de musgo (Pp\_des6) se inserta a través de BamHI/NaeI al segundo casete. Se genera el constructo pUT-ED. El fragmento AscI del plásmido pUT-ED se inserta al vector pGPTV+AscI cortado con AscI y se determina la orientación del fragmento insertado mediante restricción o secuenciación. Se genera el plásmido pB-DHGLA, cuya secuencia completa se representa en SEQUENZ ID NO. 25. La región codificadora de la Physcomitrella delta 6 elongasa se representa en SEQUENZ ID NO. 26, la de la la delta 6 desaturasa de Physcomitrella en SEQUENZ ID NO: 27.

## ES 2 345 324 T3

En pUT123 primero se inserta a través de BstXI y XbaI la  $\Delta$ -6-elongasa Pp\_PSE1 al primer casete. Luego se inserta la  $\Delta$ -6-desaturasa de musgo (Pp\_des6) a través de BamHI/NaeI al segundo casete y finalmente se inserta la  $\Delta$ -5-desaturasa de *Phaeodactylum* (Pt\_des5) a través de BglII al tercer casete. El constructo triple obtiene el nombre pARA1. Tomando en cuenta posiciones de escisión de restricción específicos de secuencia, pueden generarse otros casetes de expresión según como se representan en la tabla 5 con la denominación pARA2, pARA3 y pARA4.

El fragmento AscI del plásmido pARA1 se inserta al vector pGPTV+AscI cortado con AscI y se determina la orientación del fragmento insertado por medio de restricción o secuenciación. La secuencia total del plásmido resultante pBARA1 se representa en SEQUENZ ID NO. 28. La región codificadora de la *Physcomitrella* delta 6 alongasa se representa en SEQUENZ ID NO. 29, la de la delta 6 desaturasa de *Physcomitrella* en SEQUENZ ID NO: 30 y la de la delta-5 desaturasa de *Phaeodactylum tricorutum* en SEQUENZ ID NO: 31.

TABLA 5

	Plásmido de gen	Desaturasa- $\Delta$ -6	Desaturasa- $\Delta$ -5	Elongasa- $\Delta$ -6
1	PUT-ED	Pp_des6	---	Pp_PES1
2	pARA1	Pt_des6	Pt_des5	Pp_PES1
3	pARA2	Pt_des6	Ce_des5	Pp_PES1
4	pARA3	Pt_des6	Ce_des5	Pp_PES1
5	pARA4	Ce_des6	Ce_des5	Ce_PES1
6	PBDHGLA	Pt_des6	---	Pp_PES1
7	PBARAI	Pt_des6	Pt_des5	Pp_PES1

Plásmidos 1 hasta 5 son derivados de pUC, plásmidos 6 hasta 7 con vectores binarios de transformación vegetal

Pp = *Physcomitrella patens*, Pt = *Phaeodactylum tricorutum* Pp\_PSE1 corresponde a la secuencia de SEQ ID NO: 9.

PSE = PUFA  $\Delta$ -6-elongasa Ce\_des5 =  $\Delta$ -5-desaturasa de *Caenorhabditis elegans* (No de acceso de banco genético AF078796).

Ce\_des6 =  $\Delta$ -6-desaturasa de *Caenorhabditis elegans elegans* (No de acceso a banco genético AF031477, bases 11-1342).

Ce\_PSE1 =  $\Delta$ -6-elongasa de *Caenorhabditis elegans* (No de acceso a banco genético AF244356, bases 1-867).

Otras secuencias de desaturasas o de elongasas también pueden insertarse a casetes de expresión del tipo descrito, como por ejemplo No de acceso a banco genético AF231981, NM\_013402, AF206662, AF268031, AF226273, AF110510 o AF110509.

iii) *Transferencia de casetes de expresión a vectores para la transformación de Agrobacterium tumefaciens y para la transformación de vegetales*

Constructos génicos quiméricos a base de aquellos descritos en pUC19 pueden insertarse al vector binario pGPTV por medio de AscI. Para este propósito se extiende la secuencia de clonación múltiple en una posición de escisión AscI. Para este propósito recién se sintetiza el polylinker (polienlazador) como oligonucleótidos bicatenarios y se agrega adicionalmente una secuencia ADN AscI. El oligonucleótido se inserta al vector pGPTV por medio de EcoRI y HindIII. Se genera el plásmido pGPTV+AscI. Las técnicas de clonación requeridas son conocidas por el técnico en la materia y pueden consultarse simplemente como se describe en el ejemplo 1.

### Ejemplo 12

#### *Mutagénesis in vivo*

La mutagénesis *in vivo* de microorganismos puede efectuarse por medio del paso del ADN del plásmido (u de otro vector) por *E. coli* u otros microorganismos (por ejemplo *Bacillus* spp. o levaduras como *Saccharomyces cerevisiae*), en los que se hayan interrumpido las capacidades de mantener la integridad de la información genética. Las

## ES 2 345 324 T3

cepas usuales mutadoras tienen mutaciones en los genes para el sistema de reparación de ADN (por ejemplo mutHLS, mutD, mutT etc.; como referencia bibliográfica véase Rupp, W.D. (1996) DNA repair mechanisms, en: *Escherichia coli* and *Salmonella*, páginas 2277-2294, ASM: Washington). Estas cepas son conocidas para el técnico en la materia. El uso de estas cepas es, por ejemplo, ilustrado en Greener, A., y Callahan, M. (1994) *Strategies* 7: 32-34. La transferencia de moléculas mutadas de ADN a vegetales se efectúa preferiblemente después de seleccionar y ensayar los microorganismos. Se generan plantas transgénicas según diversos ejemplos en la sección de ejemplos de este documento.

### 10 Ejemplo 13

#### *Estudio de la expresión de un producto génico recombinante en un organismo transformado*

15 La actividad de un producto génico recombinante en el organismo huésped transformado puede medirse al nivel de transcripción y/o traducción.

Un método adecuado para determinar la cantidad de transcripción del gen (que indica la cantidad de ARN disponible para traducción del producto génico) es llevar a cabo un Northern blot tal como se especifica aquí abajo (véase Ausubel y colaboradores (1988) *Current Protocols in Molecular Biology*, Wiley: New York como referencia o la sección de ejemplos arriba mencionada), en el que un cebador que se configura de tal modo que enlace al gen de interés se marca con una marcación detectable (habitualmente radioactiva o quemoluminiscentemente), de tal modo que cuando el ARN total de un cultivo del organismo se extrae, se separa sobre un gel, se transfiere a una matriz estable y se incuba con esta sonda, el enlazamiento y la medida del enlazamiento de la sonda indican la presencia y también la cantidad del ARNm para este gen. Esta información indica el grado de transcripción del gen transformado. El ARN celular total puede prepararse a partir de células, tejidos u órganos mediante un gran número de métodos, todos los cuales son conocidos en el estado de la técnica, tales como por ejemplo el método de Bormann, E.R., y colaboradores (1992) *Mol. Microbiol.* 6: 317-326.

### 30 *Hibridación Northern*

Para la hibridación de ARN se separaron 20  $\mu\text{g}$  de ARN total o 1  $\mu\text{g}$  de poli(A)<sup>+</sup>ARN fueron separados mediante electroforesis en gel de agarosa con una fuerza del 1,25% usando formaldehído como se describe por Amasino (1986, *Anal. Biochem.* 152, 304), transferidos a membranas de nylon cargadas positivamente (Hybond N+, Amersham, Braunschweig) por medio de atracción capilar usando 10 x SSC, inmovilizados mediante luz UV y prehibridados por 3 horas a 68°C usando búfer de hibridación (sulfato de dextran al 10% peso/vol, NaCl de 1 M, SDS al 1%, 100 mg de ADN de esperma de arenque). La marcación de la sonda de ADN con el kit de etiquetado Highprime DNA (Roche, Mannheim, Alemania) se efectuó durante la pre-hibridación usando alfa-<sup>32</sup>P-dCTP (Amersham, Braunschweig, Alemania). La hibridación se realizó por toda una noche después de la adición de la sonda de ADN marcada en el mismo búfer a 68°C. Los pasos de lavado se realizaron dos veces por 15 min usando 2 X SSC y dos veces por 30 min usando 1 X SSC, SDS al 1%, a 68°C. La exposición de los filtros cerrados se realizó a -70°C por un lapso de tiempo de 1 hasta 14 días.

Para el estudio de la presencia o de la cantidad relativa de proteína traducida desde este ARNm pueden emplearse técnicas estándar tales como un Western blot (véase, por ejemplo, Ausubel y colaboradores (1988) *Current Protocols in Molecular Biology*, Wiley: New York). En este método se extrae el total de proteína celular, se separa por medio de electroforesis en gel, se transfiere a una matriz como nitrocelulosa y se incuba con una sonda tal como un anticuerpo que se enlaza específicamente a la proteína deseada. Esta sonda se suministra usualmente con una etiqueta luminiscente o colorimétrica que puedan detectarse fácilmente. La presencia y la cantidad de la etiqueta observadas indican la presencia y la cantidad de la proteína mutada deseada que está presente en la célula.

### Ejemplo 14

#### 55 *Análisis del efecto de la proteína recombinante sobre la producción del producto deseado*

El efecto de la modificación genética en plantas, hongos, algas, ciliados o en la producción de un compuesto deseado (tal como un ácido graso) puede determinarse haciendo crecer los microorganismos deseados o la planta modificada en condiciones adecuadas (tales como aquellas descritas arriba) y analizando el medio y/o los componentes de la célula para la producción incrementada del producto deseado (es decir, de lípidos o un ácido graso). Estas técnicas analíticas son conocidas para la persona técnica en la materia y comprenden espectroscopía, cromatografía de capa delgada, diversos métodos de tinte, métodos enzimáticos y microbiológicos y cromatografía analítica tal como cromatografía líquida de alto rendimiento (véase por ejemplo Ullman, *Encyclopedia of Industrial Chemistry*, Bd. A2, páginas 89-90 y páginas 443-613, VCH: Weinheim (1985); Fallon, A., y colaboradores, (1987) "Applications of HPLC in Biochemistry" en: *Laboratory Techniques in Biochemistry and Molecular Biology*, volumen 17; Rehm y colaboradores (1993) *Biotechnology*, volumen 3, Capítulo III: "Product recovery and purification", páginas 469-714, VCH: Weinheim; Belter, P.A., y colaboradores (1988) *Bioseparations; downstream processing for Biotechnology*, John Wiley and Sons; Kennedy, J.F., y Cabral, J.M.S. (1992) *Recovery processes for biological Materials*, John

## ES 2 345 324 T3

Wiley and Sons; Shaeiwitz, J.A., y Henry, J.D. (1988) Biochemical Separations, en: Ullmann's Encyclopedia of Industrial Chemistry, volumen B3; Capítulo 11, páginas 1-27, VCH: Weinheim; y Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

5 Además de los métodos arriba mencionados, los lípidos vegetales se extraen a partir de materiales vegetales tal como se describe por Cahoon y colaboradores (1999) Proc. Natl. Acad. Sci. USA 96 (22): 12935-12940, y Browse y colaboradores (1986) Analytic Biochemistry 152: 141-145. El análisis cualitativo y cuantitativo de lípidos y ácidos grasos se describe en Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 10 1989, Repr. 1992, IX, 307 pág. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford": Pergamon Press, 1 (1952)-16 (1977) bajo el título: Progress in the Chemistry of Fats and Other Lipids CODEN.

Adicionalmente a la medición del producto final de la fermentación, para determinar la eficiencia total de producción del compuesto también es posible analizar otros componentes de las rutas metabólicas que se usan para producir el compuesto deseado, tales como intermedios o productos secundarios. Los métodos analíticos abarcan mediciones de las cantidades de nutrientes en el medio (por ejemplo, azúcares, hidrocarburos, fuentes de nitrógeno, fosfato y otros 15 ionnes), mediciones de concentración y crecimiento de biomasa, análisis de la producción de metabolitos usuales de las rutas biosintéticas y mediciones de gases que se generan durante la fermentación. Se describen métodos estándar para estas mediciones en Applied Microbial Physiology; A Practical Approach, P.M. Rhodes y P.F. Stanbury, Editores, 20 IRL Press, S. 103-129; 131-163 y 165-192 (ISBN: 0199635773) y las citas bibliográficas allí indicadas.

Un ejemplo es el análisis de ácidos grasos (abreviaturas: FAME, éster metílico de ácido graso; GC-MS, cromatografía de gas-líquido-espectroscopía de masas; TAG, triacilglicerina; TLC, cromatografía de capa delgada).

25 La detección no ambigua de la presencia de productos de ácido graso puede obtenerse analizando organismos recombinantes mediante métodos analíticos estándar: GC, GC-MS o TLC, tal como se describen en diversas ocasiones por Christie y las referencias bibliográficas de allí (1997, en: Advances on Lipid Methodology, 4ª. Edición: Christie, Oily Press, Dyce, 119-169; 1998, Métodos de cromatografía de gas-espectroscopía de masas, Lípidos 33: 343-353).

30 El material a analizar puede quebrarse por ultrasonido, molienda en un molino de vidrio, nitrógeno líquido y molienda o mediante otros métodos aplicables. Después de quebrar, el material debe centrifugarse. El sedimento se resuspende en agua destilada, se calienta por 10 minutos a 100°C, se enfría con hielo y se recentrifuga, le sigue una extracción en ácido sulfúrico de 0,5 M en metanol con dimetoxipropano al 2% durante 1 hora a 90°C, lo cual conduce a 35 aceite hidrolizado y a compuestos lipídicos lo cual da lípidos transmetilados. Estos ésteres metílicos de ácido graso se extraen en éter de petróleo y se someten finalmente a análisis de GC usando una columna capilar (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 micrómetros, 0,32 mm) con un gradiente de temperatura entre 170°C y 240°C por 20 min y 5 min a 240°C. La identidad del éster metílico del ácido graso obtenido debe definirse usando estándares que pueden conseguirse de fuentes comerciales (es decir Sigma).

40 En el caso de ácidos grasos para los que no hay estándares disponibles la identidad debe demostrarse por derivatización seguida de análisis de GC/MS. Por ejemplo, la localización de ácidos grasos con enlace triple deben demostrarse mediante GC/MS seguido de derivatización con derivado de 4,4-dimetoxioxazolina (Christie, 1998, véase arriba).

### Constructos de expresión en cepas de sistemas microbianos heterólogos, condiciones de lavado y plásmidos

45 La cepa de *Escherichia coli* XL1 Blue MRF<sup>r</sup> kan (Stratagene) se usó para subclonar las nuevas desaturasas pPDe-saturasa1 de *Physcomitrella patens*. Para la expresión funcional de este gen nosotros usamos la cepa de *Saccharomyces cerevisiae* INVSc 1 (Invitrogen Co.). Se cultivó *E. coli* en caldo de Luria-Bertini (LB, Duchefa, Haarlem, Holanda) a 37°C. Cuando fue necesario se adicionó ampicilina (100 mg/litro), y se adicionó agar al 1,5% (peso/ volumen) área medios sólidos LB. *S. cerevisiae* se hizo crecer a 30°C, o bien en medio de YPG, o en medio mínimo completo sin uracilo (CMdum; véase en: Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Albright, L.B., Coen, D.M., y Varki, A. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, 55 New York) con rafinosa o glucosa al 2% (peso/vol). Para medios sólidos se agregaron 2% (peso/vol) de agar Bacto™ (Difco). Los plásmidos usados para la clonación y la expresión son pUC18 (Pharmacia) y pYES2 (Invitrogen Co.).

### 60 Ejemplo 16

#### Clonación y expresión de desaturasas específicas de PUFA de *Phaeodactylum tricornerutum*

65 Para la expresión en levaduras primero se modificaron los clones de ADNc de *Phaeodactylum tricornerutum* de las Seq ID NO: 1, 3, 5 ó 11 o las secuencias de SEQ ID NO: 7 ó 9 u otras secuencias deseadas de tal manera que solo se amplifican las regiones codificadoras por medio de reacción en cadena de polimerasa con la ayuda de dos oligonucleótidos. En este caso se prestó cuidado de que se mantuviera una secuencia de consenso antes del codón de inicio para la traducción (translation) eficiente. Para este fin se seleccionó la serie de bases ATA o AAA y se

## ES 2 345 324 T3

inserta a la secuencia antes de la ATG (Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-292). Adicionalmente antes de la tripleta de consenso se introdujo un sitio de escisión de restricción que debe ser compatible con el sitio de escisión del vector objetivo al cual debe clonarse el fragmento y con su ayuda debe efectuarse la expresión génica en microorganismos o vegetales.

La reacción PCR se realizó con ADN de plásmido como matriz en un termociclador (Biometra) usando polimerasa de ADN Pfu (Stratagene) y el siguiente programa de temperatura: 3 min a 96°C, seguido de 30 ciclos con 30 s a 96°C, 30 s a 55°C y 2 min a 72°C, 1 ciclo con 10 min a 72°C y parada a 4°C. La temperatura de adhesión se varió dependiendo de los oligonucleósidos elegidos. Puede tomarse un tiempo de síntesis de aproximadamente de un minuto por pares de kilobase. Otros parámetros que tienen un efecto sobre la PCR, tales como por ejemplo iones de Mg, sal, polimerasa de ADN y similares son corrientes para el técnico en la materia y pueden variarse según la necesidad.

El tamaño correcto del fragmento de ADN amplificado se confirmó por medio de electroforesis en gel TBE de agarosa. El ADN amplificado se extrajo del gel usando el kit de extracción de gel QIAquick (QIAGEN) y se ligó al sitio de restricción SmaI del vector defosforilado pUC18 usando el kit Sure Clone Ligations Kit (Pharmacia), y se obtuvieron los derivados pUC. Después de la transformación *E. coli* XL1 Blue MRF' kan se realizó una minipreparación de ADN (Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid minipreparation. BioTechniques 4, 310-313) en transformantes resistentes a ampicilina y se identificaron clones positivos mediante análisis de restricción de BamHI. Se confirmó la secuencia del producto PCR clonado por medio de resecuenciación usando el kit ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt).

Δ5 Acil Lipid desaturasa, Pt\_des5

Cebador 1 GAG CTC ACA TAA TGG CTC CGG ATG CGG ATA AGC

Cebador 2 CTC GAG TTA CGC CCG TCC GGT CAA GGG

El fragmento PCR (1428bp) se clonó con la ayuda del Sure Clone Kit (Pharmacia) a pUC 18, el fragmento insertado se digirió con SacI/XhoI y el fragmento se insertó a pYES2 o pYES6 con la ayuda de sitios de escisión de restricción adecuados.

Δ6 Acil Lipid desaturasa, Pt\_des6

Cebador 3 GGA TCC ACA TAA TGG GCA AAG GAG GGG ACG CTC GGG

Cebador 4 CTC GAG TTA CAT GGC GGG TCC ATC GGG

El fragmento PCR (1451 bp) se clonó a pUC18 con la ayuda del Sure Clone Kit (Pharmacia), el fragmento insertado BamHI/XhoI se digirió y se insertó el fragmento a pYES2 o pYES6 con la ayuda de sitios de escisión de restricción correspondientes.

Δ12 Acil Lipid desaturasa, Pt\_des12

Cebador 5 GGA TCC ACA TAA TGG TTC GCT TTT CAA CAG CC

Cebador 6 CTC GAG TTA TTC GCT CGA TAA TTT GC

Δ12 Acil Lipid desaturasa, Pt\_des12.2

Cebador 7 GGA TCC ACA TAA TGG GTA AGG GAG GTC AAC G

Cebador 8 CTC GAG TCA TGC GGC TTT GTT TCG C

El fragmento PCR (1505bp) se clonó a pUC 18 con la ayuda del Sure Clone Kit (Pharmacia), el fragmento insertado se digirió con BamHI/XhoI y el fragmento se insertó a pYES2 o pYES6 con la ayuda de sitios de escisión de restricción correspondientes.

El ADN de plásmido se escindió con enzima(s) de restricción para hacer coincidir el sitio de escisión introducido de la secuencia de cebador, y el fragmento obtenido se ligó a sitios de restricción compatibles del vector shuttle de

## ES 2 345 324 T3

levadura defosforilada- *E. coli* pYES2 o pYES6, en lo cual se obtienen derivados de pYES. Después de la transformación de *E. coli* y de la minipreparación de ADN de los transformandos se verificó la orientación del fragmento de ADN en el vector por medio de escisión de restricción o secuenciación. Para la maxipreparación de ADN se usó un clon con el kit de extracción de ADN de plásmido Nucleobon® AX 500 (Macherey-Nagel, Düren).

5 *Saccharomyces cerevisiae* INVSc1 se transformó con los derivados pYES y el vector vacío pYES por medio de un protocolo de PEG/acetato de litio (Ausubel y colaboradores, 1995). Después de la selección sobre placas de agar CMDum con glucosa al 2% se seleccionó transformandos derivados de pYES y un transformando pYES2 para seguir cultivando y expresar funcionalmente. En derivados pYES6 se usó blasticidina como antimetabolito. En el caso de  
10 coexpresión a base de pYES2 y pYES6 se seleccionó con blasticidina en medio mínimo.

### *Expresión funcional de una actividad desaturasa en levaduras*

#### 15 *Precultivo*

20 20 ml de Cmdum líquido sin uracilo pero con rafinosa al 2% (peso/vol) se inocularon con clones transgénicos de levadura (pYES2) y se hicieron crecer por 3 días a 30°C, 200 rpm hasta lograr una densidad óptica a 600 nm (OD<sub>600</sub>) de 1,5 hasta 2. Si se usó como vector pYES6, entonces se seleccionó adicionalmente sobre blasticidina como antimetabolito.

#### *Cultivo principal*

25 Para la expresión se suplementaron 20 ml de medio líquido CMDum sin uracilo pero con 2% de rafinosa y 1% (vol./vol.) de Tergitol NP-40 con sustratos de ácido graso a una concentración final de 0,003% (peso/vol). Los medios se inocularon con los precultivos hasta una OD<sub>600</sub> de 0,05. La expresión se indujo por 16 horas a una OD<sub>600</sub> de 0,2 con galactosa al 2% (peso/vol) por 16 horas, después de lo cual se cosecharon los cultivos a una OD<sub>600</sub> de 0,8-1,2.

#### 30 *Análisis de ácido graso*

35 Se extrajo el total de los ácidos grasos de los cultivos de levadura y se analizaron por medio de cromatografía de gases. De esto se cosecharon células de 5 ml de cultivo por medio de centrifugación (1000 x g, 10 min, 4°C) y se lavó una vez con NaHCO<sub>3</sub> de 100 mM, pH 8,0 con el fin de retirar el medio residual y los ácidos grasos. Para producir el éster metílico de ácido graso (FAMES o, en singular, FAME) se trataron los sedimentos celulares con H<sub>2</sub>SO<sub>4</sub> de 1 M en metanol y dimetoxipropano al 2% (vol./vol.) durante 1 hora a 80°C. Los FAMES se extrajeron dos veces con 2 ml de éter de petróleo, una vez con NaHCO<sub>3</sub> de 100 mM, pH 8,0, y se lavó una vez con agua destilada y se secó con Na<sub>2</sub>SO<sub>4</sub>.  
40 El solvente orgánico se evaporó bajo una corriente de argón y los FAMES se disolvieron en 50 µL de éter de petróleo. Las muestras se separaron en una columna capilar Wax ZEBRON-ZB (30 m, 0,32 mm, 0,25 µm; Fenomenex) en un cromatógrafo de gases Hewlett Packard-6850 con un detector de ionización por llama. La temperatura del horno se programó desde 70°C (1 min sostenida) hasta 200°C con una velocidad de 20°C/min, luego a 250°C (5 min sostenida) con una velocidad de 5°C/min y finalmente a 260°C con una velocidad de 5°C/min. Se usó nitrógeno como gas de soporte (4,5 ml/min a 70°C). Se identificaron los ácidos grasos mediante comparación con tiempos de retención de de  
45 los FAMES estándares.

#### *Análisis de expresión*

50 Se determinaron las proporciones de los sustratos de ácido graso adicionados y recuperados y de esta manera se registran la cantidad y la calidad de la reacción de desaturasa según las Tablas 6 a 8.

55

60

65

## ES 2 345 324 T3

El resultado de la expresión de una desaturasa  $\Delta$ -6-acil lípido de *Phaeodactylum tricornutum* en levadura:

TABLA 6

5

	Ácido graso	pYes2 alimentado con			
		pYes2	-	+18:2	+18:3
10	16:0	13,3	18,9	28,4	16,7
	16:1D9	45,4	44,7	12,5	16,9
	16:2D6,9	-	4,3	-	-
	18:0	4,9	6,3	10,4	9,1
15	18:1D9	36,4	24,1	6,8	11,8
	18:2D6,9	-	1,8	-	-
	18:2D9,12	-	-	33,4	-
20	18:3D6,12,15	-	-	4,9	-
	18:3D9,12,15	-	-	-	43,1
	18:4D6,9,12,15	-	-	-	2,3

25

Los datos representan % molar de los ácidos grasos cis correspondientes.

El resultado de la expresión de una desaturasa  $\Delta$ -5-acil lípido de *Phaeodactylum tricornutum* en levadura:

30

TABLA 7

35

40

45

50

55

60

65

Acido graso	pYES2		Constructo Pyes_PtD5 alimentado con						
	Vacío	Control	18:2	18:3	20:1	-20:1	20:2	20:3	20:3
					$\Delta$ 8	$\Delta$ 11	$\Delta$ 11,14	$\Omega$ 3	$\Omega$ 6
16:0 $\Delta$	16,9	20,4	27,7	24,4	16,2	21	17,6	19,5	22,8
16:1 $\Delta$ 9	44,7	44,1	13,2	9,6	37,4	39,4	38,3	36,9	30,7
18:0	6,1	6,9	10,54	9,8	4,7	7,9	6,3	6,8	8,2
18:1 $\Delta$ 9	31,72	28,1	8,77	6	15	26	29,5	25,6	21,1
18:2 $\Delta$ 5,9		0,17	0	0	0	0,09	0,21	0,09	9
18:2 $\Delta$ 9,12		-	39,7	-	-	-	-	-	-
18:3 $\Delta$ 9,12,15		-		49,9	--	-	-	-	-
20:1 $\Delta$ 8		-		-	25,5	-	-	-	-
20:1 $\Delta$ 11		-		-	-	5,41	-	-	-
20: 2 $\Delta$ 5,11		-		-	-	0,21	-	-	-
20:2 $\Delta$ 11,14		-	-	-	-	-	6,48	-	-
20:3 $\Delta$ 5,11,4		-					0,76	-	-
20:3 $\Delta$ 11,14,17		-	-	-	-	-	-	9,83	-
20:3 $\Delta$ 8,11,14		-	-	-	-	-	-	-	13,69
20:4 $\Delta$ 5,11,14,17		-	-	-	-	-	-	1,26	-
20:4 $\Delta$ 5,8,11,14		-	-	-	-	-	-	-	3,08

Los datos representan el % molar de ácidos grasos de ácidos grasos cis.

## ES 2 345 324 T3

A partir de más experimentos de alimentación se encontró que C18:1 $\Delta$ 9 no se desatura en la presencia de ácidos grasos de C18:2 $\Delta$ 9,11 o C18: $\Delta$ 9,12,15 o C20:1 $\Delta$ 8 mientras que en presencia de C20:1 $\Delta$ 11, C20:2  $\Delta$ 11, 14 y C20:3  $\Delta$ 8,11,14 también se desatura C18:1. Así mismo, no se efectuó una desaturación en presencia de C20:3 $\Delta$ 8,11,14.

5 Al usar la cepa de levadura deficiente de proteasa C13BYS86 (Kunze I. y colaboradores, Biochemica et Biophysica Acta (1999) 1410: 287-298) para la expresión de la  $\Delta$ -5-desaturasa de *Phaeodactylum tricornutum* sobre medio completo con blasticidina se encontró que C20:4  $\Delta$ 8,11,14,17 como sustrato de la  $\Delta$ -5-desaturasa dio una rata de conversión de 20% e igualmente bien se convirtió como C20:3  $\Delta$ 8,11,14. De manera alterna, para expresión génica pueden usarse los marcadores de auxotrofismo leu2, ura3 o his.

10 En otro experimento de coexpresión de  $\Delta$ -5 desaturasa de *Phaeodactylum* y  $\Delta$ -6 elongasa de *Physcomitrella* se utilizó la cepa UTL7A (Warnecke y colaboradores, J. Biol. Chem. (1999) 274(19): 13048-13059), y la  $\Delta$ -5 desaturasa se convirtió aproximadamente 10% C20:3  $\Delta$ 8, 11, 14 a C20:4  $\Delta$ 5,8,11,14.

15 Pueden realizarse otros experimentos de alimentación con ácidos grasos muy diversos, solos o en combinación (por ejemplo ácido linoleico, 20:3  $\Delta$ -5,11,14-ácido graso, ácido alfa- o gamma linolénico, ácido estearidónicosäure, ácido araquidónico, ácido eicosapentaenoico etc.) para la confirmación detallado de la selectividad y especificidad por sustrato de estas desaturasas.

TABLA 8

*Resultado de la co-expresión de una desaturasa de  $\Delta$ -5-acilo lípido de *Phaeodactylum tricornutum* y de una elongasa  $\Delta$ -6 de musgo en levadura sobre la base de vectores de expresión pYes2 y pYes6*

	pYes2-Elo		pYes2-Elo y pYes6-Ptd5	
	+18:3	+18:4	+18:3	+18:4
16:0	15,0	14,8	15,6	15,1
16:1 $\Delta$ 9	27,7	29,2	27,5	29,0
18:0	5,6	6,3	5,7	6,4
18:1 $\Delta$	17,1	30,8	27,4	31,6
18:3 $\Delta$	7,60	-	7,8	-
18:4 $\Delta$	-	6,71	-	6,4
20:3 $\Delta$	15,92	-	13,55	-
20:4 $\Delta$	-	-	1,31	-
20:4 $\Delta$	-	11,4	-	10,31
20:5 $\Delta$	-	-	-	0,53

45 De la conversión de sustrato se infiere que la  $\Delta$ -5-desaturasa de *Phaeodactylum* y la  $\Delta$ -6-elongasa de *Physcomitrella patens* respecto de la actividad de sustrato y especialmente de la especificidad de sustrato son adecuadas para producir ácido araquidónico o ácido eicosapentaenoico con la ayuda de las secuencias de acuerdo con la invención.

50 Los patrones de fragmentación y los espectros de masas de los derivados de AMOX de estándares y también las fracciones de pico de los ácidos grasos identificados mediante GC que se muestran en las tablas 6, 7 y 8, muestran resultados idénticos comparativamente, por lo cual se aseguró la posición respectiva del enlace doble más allá de la simple detección con GC.

### 55 Ejemplo 17

#### *Purificación del producto deseado de organismos transformados*

60 La recuperación del producto deseado del material vegetal o de hongos, algas, ciliados, células animales o del sobrenadante de los cultivos que se describieron antes puede efectuarse por diferentes métodos conocidos en el campo técnico. Si el producto deseado no se secreta por las células, pueden cosecharse las células del cultivo mediante centrifugación de baja velocidad, pueden lisarse las células mediante técnicas estándar, tales como fuerza mecánica o por ultrasonido. Pueden separarse órganos de plantas de manera mecánica de otros tejidos u otros órganos. Después de homogenizar, se retiran los detritos celulares mediante centrifugación y se conserva la fracción sobrenadante que contiene las proteínas disueltas para seguir purificando el compuesto deseado. Si el producto es secretado de las células deseadas, las células se retiran mediante centrifugación lenta y se mantiene la fracción sobrenadante para seguir purificando.

## ES 2 345 324 T3

La fracción sobrenadante de cada método de purificación se somete a una cromatografía con una resina adecuada y la molécula deseada, o bien se retiene sobre la resina de cromatografía, aunque muchas impurezas en la muestra no, o bien se quedan las impurezas en la resina mientras que la muestra no. Si es necesario pueden repetirse estos pasos de cromatografía y se usan resinas de cromatografía iguales o diferentes. La persona técnica en la materia está familiarizada en la selección de resinas adecuadas de cromatografía y su aplicación más efectiva para una determinada molécula a purificarse. El producto purificado puede concentrarse mediante filtración o ultrafiltración y mantenerse a una temperatura a la que la estabilidad del producto es máxima.

En el campo técnico se conoce un amplio espectro de métodos de purificación y el método de purificación previo no debe limitarse. Estos métodos de purificación se describen, por ejemplo, en Bailey, J.E., & Ollis, D.F., *Biochemical Engineering Fundamentals*, McGraw-Hill: New York (1986).

La identidad y la pureza de los compuestos aislados puede determinarse mediante técnicas estándar del campo técnico. Estos incluyen cromatografía líquida de alto rendimiento (HPLC), métodos espectroscópicos, métodos de tinturación, cromatografía de capa delgada, en particular cromatografía de capa delgada y detección por ionización de llama (IATROSCAN, Iatron, Tokio, Japón), NIRS, ensayo por ensayo enzimático o microbiológicamente. Véase una recopilación de estos métodos de análisis en: Patek y colaboradores (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova y colaboradores (1996) *Biotekhnologiya* 11: 27-32; y Schmidt y colaboradores (1998) *Bioprocess Engineer.* 19: 67-70. *Ulmann's Encyclopedia of Industrial Chemistry* (1996) vol. A27, VCH: Weinheim, páginas 89-90, páginas 521-540, páginas 540-547, páginas 559-566, 575-581 y páginas 581-587; Michal, G (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A., y colaboradores (1987) *Applications of HPLC in Biochemistry en: Laboratory Techniques in Biochemistry and Molecular Biology*, volumen 17.

*Equivalentes*

La persona técnica en la materia reconoce o puede establecer muchos equivalentes de las modalidades específicas de acuerdo con la invención aquí descritas usando solo experimentos rutinarios. Estos equivalentes deben abarcarse por las reivindicaciones de la patente.

REIVINDICACIONES

5 1. Método para producir ésteres de ácido graso con un contenido elevado de ácidos grasos poliinsaturados con al menos dos enlaces dobles **caracterizado** porque se introduce a un organismo no humano, productor de ésteres de ácido graso, al menos un ácido nucleico seleccionado del grupo de

10 a) Un ácido nucleico con la secuencia representada en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11,

10 b) Ácidos nucleicos que se obtienen debido al código degenerado mediante retraducción de las secuencias de aminoácidos representadas en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 o SEQ ID NO: 12,

15 c) Derivados de los ácidos nucleicos representados en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11 que codifican para partes biológicamente activas de una desaturasa y presentan al menos 50% de identidad a la secuencia de aminoácidos de la SEQ ID NO: 2, 4, 6 ó 12,

se hace crecer el organismo y se aíslan los ésteres de ácido graso obtenidos del organismo.

20 2. Método según la reivindicación 1, en el cual los ésteres de ácidos grasos producidos mediante el método contienen moléculas de de ácido graso poliinsaturadas de C<sub>19</sub>, C<sub>20</sub> o C<sub>22</sub> con al menos dos enlaces dobles en el éster de ácido graso.

25 3. Método según la reivindicación 1 ó 2, en el que las moléculas de ácido graso de C<sub>18</sub>, C<sub>20</sub> o C<sub>22</sub> se aíslan del organismo en forma de un aceite o lípido.

30 4. Método según las reivindicaciones 1 hasta 3, en el que el organismo es un microorganismo, un animal no humano o un vegetal.

30 5. Método según las reivindicaciones 1 hasta 4, en el que el organismo es una planta transgénica.

35 6. Método según las reivindicaciones 1 hasta 5, en el que los ésteres de ácido graso contienen ácidos grasos de C<sub>18</sub>, C<sub>20</sub> o C<sub>22</sub> con tres, cuatro o cinco enlaces dobles en el éster de ácido graso.

35 7. Método según las reivindicaciones 1 hasta 6, **caracterizado** porque se liberan los ácidos grasos poliinsaturados contenidos en los ésteres de ácido graso.

40 8. Ácido nucleico aislado que codifica para un polipéptido con actividad de desaturasa y que se selecciona del grupo de:

45 a) Un ácido nucleico con la secuencia representada en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11,

45 b) Ácidos nucleicos que se obtienen debido al código genético degenerado mediante retraducción de las secuencias de aminoácidos representadas en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 o SEQ ID NO: 12,

50 c) Derivados de los ácidos nucleicos representados en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11,

que codifican para partes biológicamente activas de una desaturasa y al menos 50% de identidad a la secuencia de aminoácidos de la SEQ ID NO: 2, 4, 6 o 12.

55 9. Polipéptido codificado por una secuencia de ácidos nucleicos según la reivindicación 8.

60 10. Polipéptido según la reivindicación 9 codificado por la secuencia representada en SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 o SEQ ID NO: 11.

60 11. Constructo de ácido nucleico que contiene un ácido nucleico según la reivindicación 8, y el ácido nucleico está conectado con una o varias señales de regulación.

65 12. Constructo de ácido nucleico según la reivindicación 11, y en el constructo de ácido nucleico se encuentran contenidos genes adicionales de biosíntesis del metabolismo de ácido graso o lípido.

## ES 2 345 324 T3

13. Constructo de ácido nucleico según la reivindicación 11 ó 12, y como gen de biosíntesis del metabolismo de ácido graso o lípido que está contenido en el constructo de ácido nucleico se selecciona un gen del grupo acil-CoA-dehidrogenasa(s), acil-ACP[= acil carrier protein]-desaturasa(s), acil-ACP-tioesterasa(s), ácido graso-acil-transferasa(s), ácido graso-suntasa(s), ácido graso-hidroxilasa(s), acetil-coenzima A-carboxilasa(s), acil-coenzima A-oxidasa(s), ácido graso-desaturasa(s), ácido graso-acetilenasas, lipoxigenasas, triacilglicerol-lipasas, alenoxid-sintasas, hidropéroxido-liasas o ácido graso-elongasa(s).

14. Vector que contiene un ácido nucleico según la reivindicación 8 o un constructo de ácido nucleico según la reivindicación 11.

15. Planta transgénica que contiene un ácido nucleico según la reivindicación 8, un constructo de ácido nucleico según la reivindicación 11 o un vector según la reivindicación 14.

16. Uso de un ácido nucleico según la reivindicación 8 o de un constructo de ácido nucleico según la reivindicación 11 para la producción de plantas transgénicas.

17. Anticuerpo que enlaza específicamente un polipéptido que se codifica por uno de los ácidos nucleicos según la reivindicación 8 a) o b).

18. Molécula de ácido nucleico antisense que comprende la secuencia complementario del ácido nucleico según la reivindicación 8.

19. Kit que comprende el ácido nucleico según la reivindicación 8, el constructo de ácido nucleico según las reivindicaciones 11 hasta 13, el anticuerpo según la reivindicación 17, o la molécula de ácido nucleico antisense según la reivindicación 18.

20. Composición que contiene el anticuerpo según la reivindicación 17 o constructo de ácido nucleico antisense según la reivindicación 18.

## ES 2 345 324 T3

Figura 1: Comparación de polipéptidos de la región codificadora de Pp\_des 6 (serie superior) con la secuencia EST de PT001078032R

```

398 WKPLVWMAVTELMSCMLLGFVFLSHNGMEVYNSKKEFVSAQI-----VSTR 444
    W+   + + +   + L +F LSHN   + S+   +A +               V T
430 WRVFGNIMLMGVAESLALAVLPSLSHN----FESADRDPTAPLKKTGEPVDWFKTQVETS 263
445 DIKGNIFNDWFTGGLNRQIEHHLFFPTMPRHNLNKIAPRVEVFCKKHGLVY 494
    G   + FTGGLN Q+EHHLFF M           IAP+V   C KHG+ Y
262 CTYGGFLSGCFTGGLNFQVEHHLFFRMSSAWYXYIAPKRVREICAKHGVMHY 113
  
```

Figura 2: Comparación de polipéptidos de la región codificadora de Ma\_des12 (secuencia superior) con PT001070010R

```

105 GVVWLAHECGHQSFSTSKTLNN 126
    G WVLAHECGH +FS .+++L +
533 GFWVLAHECGHGAFSKNRSLOD 598
  
```

Figura 2a: Comparación de polipéptidos de la región codificadora de Ma\_des12 (secuencia superior) con PT001072031R

```

117 SFSTSKTLNNMTVGWILHSMLLVPYHSWRISHSKHH 151
    ++S S+T N+ VG+I+H LLVPY +W+ +H+KHH
465 AYSDSQTFNDVVGFIVHQALLVPYFAWQYTHAKHH 569
  
```

Figura 3: Comparación de polipéptidos de la región codificadora de un producto de PCR a partir de un par de cebadores F6a y R4a2 que codifica para un fragmento de desaturasa (serie superior) de *Phaodactylum* con la secuencia T36617 de *Streptomyces coelicolor* (serie inferior)

```

1   WWKMKHNGHHA VPNLHCSSAVAQDCDPDIDT M PLLAWSVQQAQSYRELQADGKDSGLVKF 60
    WW++KH  HHA PN           +D DPDI  LL WS  QA++           +GL +
114 WWQDKHTRH HANPN-----TEDLDPDIGP-DLLVWSPDQARAA-----TGLPRL 156
61  MIRNQSYFYFPILLARLSWLNESFKCAFGLGAASENAALELKAKGLQYPLLEKAGILLH 120
    + R Q++ +FP+L L           E F           G A N L+ +A           L+ A +L H
157 LGRWQAFLLFFPLLT-----EGFNLHVASGRAMANRRLKRR-----LDGALLAH 202
121 YANMLTVSSGFGRXXXXXXXXXXXXXXXXXXCGFLLAIVFGLGHNGMATYNADARPDFWKLQ 180
    A LT   F           G L   F   H GM   AD RPDF + Q
203 CAVYLTAL--FWVLPFGMAIAFLAVHQCLFGVYLGSAFAPNHKGMFILTADDRPDFLRRQ 260
181 VTTTRNVITGGHGFQAPVDWFCGGLQYQVDHHLFPS 216
    V T+RNV GG           F D   GGL +Q++HHLFPS
261 VLTSRNVNCG-----LPTDLALQGLNHQIEHHLFPS 291
  
```

Figura 4: Comparación de polipéptidos de la región codificadora de Pp\_des6 (serie superior) comparada con Pt\_des6 (serie inferior)

ES 2 345 324 T3

```

51 KRLTSKKRVSESAAVQCISAEVQRNSSTQGTAEALAESVVKPTRRRSSQW 100
      . |   |-.
1  .....MGKGGDARASRG 12
      .
101 KKSTHPLS..EVAVHNKPSDCWIVVKNEVYDVSNFADEHPGGSVISTYFG 148
      . :| || | | ||: ||||| |. ||||. || |: |
13 STAARKISWQEVKTHASPEDAWIHSNKVYDVSNW.HEKPGGAVIFTHAG 61
      .
149 RDGTDVFSFPHAASTWKILQDFYIGDV..ERVEPTPELL...KDPREMRA 193
      | ||:|..|| | :. |||:|. |   |: : | :|:|.
62 DDMTDIFAAFHAPGSQSLMKKIFYIGELLPETTGKEPQQIAFEKGYRDLRS 111
      .
194 LFLREQLFKSSKLYVVKLLTNVAIFAASIAIICWSKTISAVLASACNMA 243
      : :|||. | :| | |. |. ||. ||. |:: :|   ||| |:
112 KLIMMGMFXSNKWVYVKCLSNMAIWAAACALVFYSDRFWVHLASAVMLG 161
      .
244 LCFQQCGWLSHDFLHNQVFETRNLNEVVGIVIGNAVLGFSTGWWKEKGNL 293
      ||| |||. ||||. ||| | :. |   || . |:| ||| |||
162 TFFQQSGNLAHDFLHHQVPTKRXHGDLDGLFWGNLMQGYSVQWVKNKHNG 211
      .
294 HHAAPN.ECDQTY.QPIDEDIDTFLPIAWS.....KDILATVENKTFL 334
      ||| || | | | |||:| |:| ||| ::: | ..
212 HHAVPNLHCSSAVAQDGDPDFIDTMTPLLAWSVQQAQSYRELQADGKDSGLV 261
      .
335 R.ILOYQHLLFFMGLLPFARGSWLFWSWR.....YTSTAVLSPVDR... 373
      : :. | |: :| || ||| |:   . | | :
262 KFMIRNQSYFYFPILLARLSWLNESFKCAFGLGAASENAALELKAKGLQ 311
      .
374 ..LLEKGTVLFHYFWVGTAC.YLLPGWKPLVWMAVTELMG.GMLLGFVF 419
      ||| | :| || | - - : - - . | | || || |
312 YPLLEKAGILLHYAWMLTVSSGGRFSPAYTAFYFLTATASCGFLLAIVF 361
      .
420 VLSHNGMEVYNSS..KEFVSAQIVSTRDIKG.....NIFNDWFTGGLNRQ 462
      | |||| ||. :| |: :||. | | ||| ||| |
362 GLGHNGMATYNADARPDFWKLQVTTTSTNVTGGHGFPQAFVDWFCGGLQYQ 411
      .
463 IEHHLFPTMPRHNLNKIAPRVEVPCKKHGLVYEDVSIATGTCKVLKALKE 512
      ::||| |. :||| | | ||| |. | | : : || . || |
412 VDHHLFPSLPRHNLAKTHALVESFCKEWSVQYHEADLVDGTMEVLHHLGS 461
      .
513 VAERAAAEQHATTS.... 525
      || |
462 VAGEFVVDFVRDGPAM. 477

```

Figura 5: Comparación de regiones codificadoras de Pp\_des6 (serie superior) comparadas con Pt\_des5 (serie inferior)

ES 2 345 324 T3

```

51 KRLTSKKRVSESAAVQCISAQVQRNSSTQGTAEALAESVVKPTERRRSSQW 200
      . . . . . : | | . | - - -
1 . . . . . MAPDADKLRQRYTAV 16
      . . . . .
101 KKSTHPLSEVAVHNEKPSDC . . . . . WIVVKNKVYDVSNFADSHPPGGSVIS 144
      | | . : : : : : | } : | } } } } }
17 AK . . HNAATISTQERLCSLSSLKGEEVCIDGIYDLQSF . . DHPGGETIK 62
      . . . . .
145 TYFGRDGYDVFSSPHAASTWKILQDF . YIGDVERVEFPPELLKDF . REM . 191
      : | } | : | | | : : | | . : | } | .
63 MFGGNDVTVQYKMIHPYHTEKHLEKMKRVGKVTDFVCEYKFDTEPEREIK 112
      . . . . .
192 RALFLREQLPKS . SKLYYVMKLLTNVAIFAASIALICNSKT . ISAVLASA 239
      | . | - | | : : : | | | | | | | | | | | | |
113 REVFKIVRRGKDFGTLGWFFRAFCYZAIF . . FYLQYHWVVTIGTSWLLAVA 160
      . . . . .
240 CMMALCFQQCGWLSHDFLHNQVFPETRWLNQVVGVIQNAVLGFSTYGWKKE 289
      . . . . . | | | | . | . | : . . | : | : | | | | . |
161 YGISQAMIGGN . VQHDANHGATSKRPPWVNDMLG . . LGADFIGGSKWLWQE 207
      . . . . .
290 KHNLHHAAPNECDQTYQPIDEDIDTLPPLIAWSKDILATVENKTFPLRILQY 339
      . | | | | | : | : | : : . | : | . :
208 QHWTHHAYTNHAE . . DP . . DSFGAEPMLLFN . DYPLDHPARTWLH . . RF 250
      . . . . .
340 QHLFFMGLEFFARGSWLFWWR . . . . . YTSTAVLS . . PVDRLLEKGTVL 381
      | | : | | | | | | | . | | | | . |
251 QAFFYMPVL . . AGYWLSAVFNBQILDLOQRGALSVGIRLDNAFIHSRRK 297
      . . . . .
382 FHYFW . . FVG . . . TACYLLPG . . . WKPLVWMAVTELMSGMLLGEVWV 420
      : | | : : | | | | : . . . : | . |
298 YAVFWRAVVIAVNVIAPFYTNSGLEWSWRVFGNIMLGVVAESLALAVLPS 347
      . . . . .
421 LSHN . . . . . GMEVYNSSKEFVSAQIVSTRDIKGNIFNDWPTGGLN 460
      | | | | . . . . | | | | | . | | | | |
348 LSHNFESADRDPTAPLAKTGEFVDWFKTQ . VETSCTYGGFLSGCFPGGLN 396
      . . . . .
461 RQIEHHLFPMPRHNLNKIAPRVEVFCRKHGLVYEDVS . IATGTCKVLKA 509
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
397 FQVEHHLFPRMSSAWYPYIAPKRVREICAKHGVHYAYYPWIHQNFLSTVRY 446
      . . . . .
510 LKEVAEAAA . EQHATTS . . . . . 525
      : | | | | . |
447 MHAAGTGANWRQMARENPLTGRA . 469

```

ES 2 345 324 T3

Figura 6: Comparación de polipéptidos de las regiones codificadoras de la desaturasa Δ-12 de *Mortierella alpina* (Ma\_des12) serie superior con la secuencia homóloga de *Phaeodactylum tricornutum* (Pt\_des12) en la serie inferior

```

40 KEIRECIPAHCFERSGLRGLCHVAIDLTWASLL..FLAATQIDKFE..NP 85
   |::| || ||| : | :..: | :| . | : : ||
107 KDLRAVTPKDCFEPDTAKSLGYLSVS.TMGTILCSVVGANLLSVLDPSNP 155
   . . . . .
86 LIRYLAWPVYWIMQGI VCTGVWVLAHECCGQSPSTSKTLNNTVGVWILHSM 135
   | : | | . | | }-||| ||||| }-|| .:.) . ||:|:|:|.
156 L.TWPLWAAYGAVTGTVAMGLWVLAHECCGHCAPSKNRSLODAVGYIHSI 204
   . . . . .
136 LLVPYHSWRISHSKHHKATGHMTKDQVFPKTRSQVGLPPKENAAAAVQE 185
   :||| ||- ||. ||. | || : || . | | . | | .
205 MLVPEYFSWQRSHAVHHQYTNHMELGETHVDPDRADKEG....EKSLALRQF 250
   . . . . .
186 EDMSVHLDEEAPIVTLFWYVIQFLFGWPAYLIMNASGQDYGRWTSHFHTY 235
   | [ . : : ||||| |: . }-| | .| |:
251 MLDSFGKDKGMKAYGGLOSFLHLIVGNPAYLLIGATGGPDRGMTNHFYP. 299
   . . . . .
236 SPIFEP....RNFF.....DIIISDLGVLAALGALYASMQLSLLTVTK 275
   .|: | : | : ||:|: | .||| . | |
300 NPLSTFTQPKKELFPGNWKEKVYQSDIGIAAVVGALIAWTATSGLAPVMA 349
   . . . . .
276 YYIVPYL FVNFWLVLITFLQHTDPKLPHYREGANNFQRGALCTVDRSFGK 325
   | | : :| |||| |.||||| }-||: || :||| |:| | : |
350 LYGGPLIVINAWLVLVLTWLTQHTDFTDVPHFSSDNHNFVKGALHTIDRPFYDK 399
   . . . . .
326 .....FLDHMFHGI VHTVVAHHLFSQMPFYHAREATYHLKLLGEYVYVD 370
   :| : | ] ||||| | .| | |: || :| | | .| |
400 LDPWGIIDFLHHKIGTTHVAHHPDSTIPHYKAQIATDAIKAKFPEVYLYD 449
   . . . . .
371 PSPIVVAVWRSFREC RFVEDQGDVVFVK 398
   |.|| |.|| : | || }-|| .|
450 PTFIPQAMWRVAKGCTAVEQRGDVWVWK 477

```

Figura 7: Comparación de polipéptidos de regiones codificadoras de la desaturasa D-12 de *Mortierella alpina* (Ma\_des12) serie superior con la secuencia homóloga de *Phaeodactylum tricornutum* clon PT001012031R (Pt\_des12.2) en la serie inferior



# ES 2 345 324 T3

## LISTA DE SECUENCIAS

- <110> BASF Plant Science GmbH
- 5 <120> Métodos para producir ácidos grasos poliinsaturados, nuevos genes de biosíntesis y nuevos constructos de expresión vegetales
- <130> 2000\_873
- <140> 2000\_873
- 10 <141> 2000-12-22
- <160> 31
- <170> PatentIn Vers. 2.0
- <210> 1
- 15 <211> 1652
- <212> DNA
- <213> *Phaeodactylum tricornutum*
- 20 <220>
- <221> CDS
- <222> (115)..(1524)

25 <400> 1

```

gaccaacaaa acccaacaat cccaacaatc ccatcaacag gaattgggtt tcgttgagtc 60
aataattgct agaatccaaa cagacagaca gagaccaacc gcatctatta caga atg 117
Met
1
gct ccg gat gcg gat aag ctt cga caa cgc cag acg act gcg gta gcg 165
Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val Ala
5 10 15
aag cac aat gct gct acc ata tcg acg cag gaa cgc ctt tgc agt ctg 213
Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser Leu
20 25 30
tct tcg ctc aaa ggc gaa gaa gtc tgc atc gac gga atc atc tat gac 261
Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr Asp
35 40 45
ctc caa tca ttc gat cat ccc ggg ggt gaa acg atc aaa atg ttt ggt 309
Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe Gly
50 55 60 65
ggc aac gat gtc act gta cag tac aag atg att cac ccg tac cat acc 357
Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His Thr
70 75 80
gag aag cat ttg gaa aag atg aag cgt gtc ggc aag gtg acg gat ttc 405
Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp Phe
85 90 95

```

60

65

ES 2 345 324 T3

gtc tgc gag tac aag ttc gat acc gaa ttt gaa cgc gaa atc aaa cga 453

5 Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys Arg  
100 105 110

gaa gtc ttc aag att gtg cga cga ggc aag gat ttc ggt act ttg gga 501  
Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu Gly  
10 115 120 125

tgg ttc ttc cgt gcg ttt tgc tac att gcc att ttc ttc tac ctg cag 549  
Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu Gln  
15 130 135 140 145

tac cat tgg gtc acc acg gga acc tct tgg ctg ctg gcc gtg gcc tac 597  
Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala Tyr  
150 155 160

20 gga atc tcc caa gcg atg att ggc atg aat gtc cag cac gat gcc aac 645  
Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala Asn  
165 170 175

25 cac ggg gcc acc tcc aag cgt ccc tgg gtc aac gac atg cta ggc ctc 693  
His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly Leu  
180 185 190

30 ggt gcg gat ttt att ggt ggt tcc aag tgg ctc tgg cag gaa caa cac 741  
Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln His  
195 200 205

tgg acc cac cac gct tac acc aat cac gcc gag atg gat ccc gat agc 789  
Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp Ser  
35 210 215 220 225

ttt ggt gcc gaa cca atg ctc cta ttc aac gac tat ccc ttg gat cat 837  
Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp His  
230 235 240

40 ccc gct cgt acc tgg cta cat cgc ttt caa gca ttc ttt tac atg ccc 885  
Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met Pro  
245 250 255

45 gtc ttg gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att ctt 933  
Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile Leu  
260 265 270

50 gac ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac aac 981  
Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp Asn  
275 280 285

55 gct ttc att cac tcg oga cgc aag tat gcg gtt ttc tgg cgg gct gtg 1029  
Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala Val  
290 295 300 305

tac att gcg gtg aac gtg att gct cgg ttt tac aca aac tcc ggc ctc 1077  
Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly Leu  
310 315 320

60 gaa tgg tcc tgg cgt gtc ttt gga aac atc atg ctc atg ggt gtg gcg 1125  
Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val Ala  
325 330 335

65

# ES 2 345 324 T3

```

gaa tgg ctc gcg ctg gcg gtc ctg ttt tgg ttg tgg cac aat ttc gaa -1173
Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe Glu
      340                      345                      350

5
tcc ggg gat cgc gat ccg acc gcc cca ctg aaa aag acg gga gaa cca 1221
Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu Pro
      355                      360                      365

10
gtc gac tgg ttc aag aca cag gtc gaa act tcc tgc act tac ggt gga 1269
Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly Gly
      370                      375                      380                      385

15
ttc ctt tcc ggt tgc ttc acg gga ggt ctc aac ttt cag gtt gaa cac 1317
Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu His
      390                      395                      400

20
cac ttg ttc cca cgc atg agc agc gct tgg tat ccc tac att gcc ccc 1365
His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro
      405                      410                      415

25
aag gtc cgc gaa att tgc gcc aaa cac ggc gtc cac tac gcc tac tac 1413
Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr Tyr
      420                      425                      430

30
ccg tgg atc cac caa aac ttt ctc tcc acc gtc cgc tac atg cac gcg 1461
Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His Ala
      435                      440                      445

35
gcc ggg acc ggt gcc aac tgg cgc cag atg gcc aga gaa aat ccc ttg 1509
Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro Leu
      450                      455                      460                      465

40
acc gga cgg gcg taa aagtacacga cagcaccaaa ggtggcgtat ggtgatctct 1564
Thr Gly Arg Ala
      470

agaaaacaga catagcctac tggaaatata gacgtccaaa caataatttt aaagactatt 1624
40
tttctgcgta aaaaaaaaaa aaaaaaaaaa 1652

<210> 2
45 <211> 469
<212> PRT
<213> Phaeodactylum tricornutum

50 <400> 2

Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
55      1                      5                      10                      15

Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
      20                      25                      30

60
Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
      35                      40                      45

65
Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe

```

ES 2 345 324 T3

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

ES 2 345 324 T3

	260					265					270					
5	Leu	Asp	Leu	Gln	Gln	Arg	Gly	Ala	Leu	Ser	Val	Gly	Ile	Arg	Leu	Asp
			275					280					285			
	Asn	Ala	Phe	Ile	His	Ser	Arg	Arg	Lys	Tyr	Ala	Val	Phe	Trp	Arg	Ala
		290					295					300				
10	Val	Tyr	Ile	Ala	Val	Asn	Val	Ile	Ala	Pro	Phe	Tyr	Thr	Asn	Ser	Gly
	305					310					315					320
	Leu	Glu	Trp	Ser	Trp	Arg	Val	Phe	Gly	Asn	Ile	Met	Leu	Met	Gly	Val
15					325					330					335	
	Ala	Glu	Ser	Leu	Ala	Leu	Ala	Val	Leu	Phe	Ser	Leu	Ser	His	Asn	Phe
20				340					345					350		
	Glu	Ser	Ala	Asp	Arg	Asp	Pro	Thr	Ala	Pro	Leu	Lys	Lys	Thr	Gly	Glu
			355					360					365			
	Pro	Val	Asp	Trp	Phe	Lys	Thr	Gln	Val	Glu	Thr	Ser	Cys	Thr	Tyr	Gly
25		370					375					380				
	Gly	Phe	Leu	Ser	Gly	Cys	Phe	Thr	Gly	Gly	Leu	Asn	Phe	Gln	Val	Glu
30		385				390					395					400
	His	His	Leu	Phe	Pro	Arg	Met	Ser	Ser	Ala	Trp	Tyr	Pro	Tyr	Ile	Ala
					405					410					415	
	Pro	Lys	Val	Arg	Glu	Ile	Cys	Ala	Lys	His	Gly	Val	His	Tyr	Ala	Tyr
35				420					425					430		
	Tyr	Pro	Trp	Ile	His	Gln	Asn	Phe	Leu	Ser	Thr	Val	Arg	Tyr	Met	His
40			435					440					445			
	Ala	Ala	Gly	Thr	Gly	Ala	Asn	Trp	Arg	Gln	Met	Ala	Arg	Glu	Asn	Pro
		450					455					460				
45	Leu	Thr	Gly	Arg	Ala											
	465															

50 <210> 3  
 <211> 1434  
 <212> DNA  
 <213> *Phaeodactylum tricornutum*  
 55 <220>  
 <221> CDS  
 <222> (1)..(1434)

60

65



# ES 2 345 324 T3

	Pro	Gln	Gln	Ile	Ala	Phe	Glu	Lys	Gly	Tyr	Arg	Asp	Leu	Arg	Ser	Lys	
				100					105					110			
5	ctc	atc	atg	atg	ggc	atg	ttc	aag	tcc	aac	aag	tgg	ttc	tac	gtc	tac	384
	Leu	Ile	Met	Met	Gly	Met	Phe	Lys	Ser	Asn	Lys	Trp	Phe	Tyr	Val	Tyr	
			115					120					125				
10	aag	tgc	ctc	agc	aac	atg	gcc	att	tgg	gcc	gcc	gcc	tgt	gct	ctc	gtc	432
	Lys	Cys	Leu	Ser	Asn	Met	Ala	Ile	Trp	Ala	Ala	Ala	Cys	Ala	Leu	Val	
		130					135					140					
15	ttt	tac	tcg	gac	cgc	ttc	tgg	gta	cac	ctg	gcc	agc	gcc	gtc	atg	ctg	480
	Phe	Tyr	Ser	Asp	Arg	Phe	Trp	Val	His	Leu	Ala	Ser	Ala	Val	Met	Leu	
	145					150					155					160	
20	gga	aca	ttc	ttt	cag	cag	tcg	gga	tgg	ttg	gca	cac	gac	ttt	ctg	cac	528
	Gly	Thr	Phe	Phe	Gln	Gln	Ser	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His	
					165					170					175		
25	cac	cag	gtc	ttc	acc	aag	cgc	aag	cac	ggg	gat	ctc	gga	gga	ctc	ttt	576
	His	Gln	Val	Phe	Thr	Lys	Arg	Lys	His	Gly	Asp	Leu	Gly	Gly	Leu	Phe	
				180					185						190		
30	tgg	ggg	aac	ctc	atg	cag	ggt	tac	tcc	gta	cag	tgg	tgg	aaa	aac	aag	624
	Trp	Gly	Asn	Leu	Met	Gln	Gly	Tyr	Ser	Val	Gln	Trp	Trp	Lys	Asn	Lys	
			195					200					205				
35	cac	aac	gga	cac	cac	gcc	gtc	ccc	aac	ctc	cac	tgc	tcc	tcc	gca	gtc	672
	His	Asn	Gly	His	His	Ala	Val	Pro	Asn	Leu	His	Cys	Ser	Ser	Ala	Val	
			210				215					220					
40	gcg	caa	gat	ggg	gac	ccg	gac	atc	gat	acc	atg	ccc	ctt	ctc	gcc	tgg	720
	Ala	Gln	Asp	Gly	Asp	Pro	Asp	Ile	Asp	Thr	Met	Pro	Leu	Leu	Ala	Trp	
			225			230					235					240	
45	tcc	gtc	cag	caa	gcc	cag	tct	tac	cgg	gaa	ctc	caa	gcc	gac	gga	aag	768
	Ser	Val	Gln	Gln	Ala	Gln	Ser	Tyr	Arg	Glu	Leu	Gln	Ala	Asp	Gly	Lys	
					245					250					255		
50	gat	tcg	ggt	ttg	gtc	aag	ttc	atg	atc	cgt	aac	caa	tcc	tac	ttt	tac	816
	Asp	Ser	Gly	Leu	Val	Lys	Phe	Met	Ile	Arg	Asn	Gln	Ser	Tyr	Phe	Tyr	
				260					265						270		
55	ttt	ccc	atc	ttg	ttg	ctc	gcc	cgc	ctg	tcg	tgg	ttg	aac	gag	tcc	ttc	864
	Phe	Pro	Ile	Leu	Leu	Leu	Ala	Arg	Leu	Ser	Trp	Leu	Asn	Glu	Ser	Phe	
				275				280					285				
60	aag	tgc	gcc	ttt	ggg	ctt	gga	gct	gcg	tcg	gag	aac	gct	gct	ctc	gaa	912
	Lys	Cys	Ala	Phe	Gly	Leu	Gly	Ala	Ala	Ser	Glu	Asn	Ala	Ala	Leu	Glu	
			290				295					300					
65	ctc	aag	gcc	aag	ggt	ctt	cag	tac	ccc	ctt	ttg	gaa	aag	gct	ggc	atc	960
	Leu	Lys	Ala	Lys	Gly	Leu	Gln	Tyr	Pro	Leu	Leu	Glu	Lys	Ala	Gly	Ile	
						305		310			315					320	
70	ctg	ctg	cac	tac	gct	tgg	atg	ctt	aca	ggt	tcg	tcc	ggc	ttt	gga	cgc	1008
	Leu	Leu	His	Tyr	Ala	Trp	Met	Leu	Thr	Val	Ser	Ser	Gly	Phe	Gly	Arg	
					325					330					335		

ES 2 345 324 T3

ttc tgc ttc gcg tac acc gca ttt tac ttt cta acc gcg acc gcg tcc 1056  
 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser  
 340 345 350  
 5  
 tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc ggc cac aac ggc atg 1104  
 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met  
 355 360 365  
 10  
 gcc acc tac aat gcc gac gcc cgt cgg gac ttc tgg aag ctc caa gtc 1152  
 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val  
 370 375 380  
 15  
 acc acg act cgc aac gtc acg ggc gga cac ggt ttc ccc caa gcc ttt 1200  
 Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe  
 385 390 395 400  
 20  
 gtc gac tgg ttc tgt ggt ggc ctc cag tac caa gtc gac cac cac tta 1248  
 Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu  
 405 410 415  
 25  
 ttc ccc agc ctg ccc cga cac aat ctg gcc aag aca cac gca ctg gtc 1296  
 Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val  
 420 425 430  
 30  
 gaa tgc ttc tgc aag gag tgg ggt gtc cag tac cac gaa gcc gac ctt 1344  
 Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu  
 435 440 445  
 35  
 gtg gac ggg acc atg gaa gtc ttg cac cat ttg ggc agc gtg gcc ggc 1392  
 Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly  
 450 455 460  
 40  
 gaa ttc gtc gtg gat ttt gta cgc gat gga ccc gcc atg taa 1434  
 Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met  
 465 470 475

40 <210> 4  
 <211> 477  
 <212> PRT  
 45 <213> *Phaeodactylum tricornutum*  
 <400> 4

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

ES 2 345 324 T3

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

45

50

55

60

65

ES 2 345 324 T3

Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe  
 275 280 285  
 5 Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu  
 290 295 300  
 10 Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile  
 305 310 315 320  
 Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg  
 325 330 335  
 15 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser  
 340 345 350  
 20 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met  
 355 360 365  
 25 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val  
 370 375 380  
 Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe  
 385 390 395 400  
 30 Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu  
 405 410 415  
 35 Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val  
 420 425 430  
 40 Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu  
 435 440 445  
 Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly  
 450 455 460  
 45 Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met  
 465 470 475

<210> 5

50 <211> 1651

<212> DNA

<213> *Phaeodactylum tricornutum*

<220>

55 <221> CDS

<222> (67)..(1554)

60

65

# ES 2 345 324 T3

<400> 5

```

gaagaaggaa catataaaag taagccatct cctcggcacc atctaaagac ctaatatcta 60
5
ctcgtc atg gtt cgc ttt tca aca gcc gct cta ctt tct ctg tcg aca      108
    Met Val Arg Phe Ser Thr Ala Ala Leu Leu Ser Leu Ser Thr
        1                5                10

ttg aca act tca tgt att ggt gcc ttc cag ctg tct tcg cca gca caa      156
Leu Thr Thr Ser Cys Ile Gly Ala Phe Gln Leu Ser Ser Pro Ala Gln
    15                20                25                30

ctt ccg aca agt agg ctt cgt cgg cat acg aac acg gcg ccg ctt tcg      204
Leu Pro Thr Ser Arg Leu Arg Arg His Thr Asn Thr Ala Pro Leu Ser
        35                40                45

gcc gtg gcc gtc gac tcc ggt tct tcc gat ccg gcc ttg gta ggc aac      252
Ala Val Ala Val Asp Ser Gly Ser Ser Asp Pro Ala Leu Val Gly Asn
        50                55                60

ctc ccc ctt ccc aac aac aat gat aat gag gac aag aac cgt aga atg      300
Leu Pro Leu Pro Asn Asn Asn Asp Asn Glu Asp Lys Asn Arg Arg Met
        65                70                75

cca atg atg gac ttg aaa ggt att gct ctg tct ggt ctc aaa ggg caa      348
Pro Met Met Asp Leu Lys Gly Ile Ala Leu Ser Gly Leu Lys Gly Gln
    80                85                90

gct ctt tcc gtc cga gcg gaa gat ttt cct cag gcg aaa gac ttg cgt      396
Ala Leu Ser Val Arg Ala Glu Asp Phe Pro Gln Ala Lys Asp Leu Arg
    95                100                105                110

35

40

45

50

55

60

65

```

# ES 2 345 324 T3

	gcc gtc att ccg aaa gat tgc ttc gaa ccc gac acg gcc aaa tcg ttg	444
	Ala Val Ile Pro Lys Asp Cys Phe Glu Pro Asp Thr Ala Lys Ser Leu	
	115 120 125	
5	gga tat ctt tcc gtt tca act atg ggg aca att ctc tgc tcc gtc gtc	492
	Gly Tyr Leu Ser Val Ser Thr Met Gly Thr Ile Leu Cys Ser Val Val	
	130 135 140	
10	ggc gcg aac ctc ctt agt gtg ctc gat ccc tcc aat cca tta acc tgg	540
	Gly Ala Asn Leu Leu Ser Val Leu Asp Pro Ser Asn Pro Leu Thr Trp	
	145 150 155	
15	cct ctc tgg gcg gcc tac ggt gcc gtc acg ggg acg gtc gcc atg ggg	588
	Pro Leu Trp Ala Ala Tyr Gly Ala Val Thr Gly Thr Val Ala Met Gly	
	160 165 170	
20	ctt tgg gtg ctg gcc cac gaa tgc gga cac ggc gcc ttt tcc aaa aac	636
	Leu Trp Val Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Lys Asn	
	175 180 185 190	
25	cga tcc ctc cag gat gcc gtg ggg tac att atc cat tcc atc atg ctg	684
	Arg Ser Leu Gln Asp Ala Val Gly Tyr Ile Ile His Ser Ile Met Leu	
	195 200 205	
30	gtg cca tac ttt agt tgg cag cga tgc cat gcc gtg cat cac cag tat	732
	Val Pro Tyr Phe Ser Trp Gln Arg Ser His Ala Val His His Gln Tyr	
	210 215 220	
35	acc aat cat atg gaa ctg ggg gaa aca cac gtt cct gat cga gcc gat	780
	Thr Asn His Met Glu Leu Gly Glu Thr His Val Pro Asp Arg Ala Asp	
	225 230 235	
40	aag gag ggc gag aag agc ctg gcg ctc cgc cag ttc atg ttg gat tcc	828
	Lys Glu Gly Glu Lys Ser Leu Ala Leu Arg Gln Phe Met Leu Asp Ser	
	240 245 250	
45	ttt ggt aaa gac aag ggc atg aaa gca tac gga ggc ctc cag tgc ttt	876
	Phe Gly Lys Asp Lys Gly Met Lys Ala Tyr Gly Gly Leu Gln Ser Phe	
	255 260 265 270	
50	ttg cat ctc atc gtg gga tgg cca gcc tac ctc ctg atc ggt gcg acc	924
	Leu His Leu Ile Val Gly Trp Pro Ala Tyr Leu Leu Ile Gly Ala Thr	
	275 280 285	
55	ggt gga ccc gac cgt ggt atg acc aac cat ttt tat ccc aac cct ttg	972
	Gly Gly Pro Asp Arg Gly Met Thr Asn His Phe Tyr Pro Asn Pro Leu	
	290 295 300	
60	tcg acg cca aca cag ccc aag aaa gaa ctt ttc cct ggg aac tgg aaa	1020
	Ser Thr Pro Thr Gln Pro Lys Lys Glu Leu Phe Pro Gly Asn Trp Lys	
	305 310 315	
65	gaa aag gtc tac cag tca gat att gga atc gcc gcc gtt gtc ggc gcc	1068
	Glu Lys Val Tyr Gln Ser Asp Ile Gly Ile Ala Ala Val Val Gly Ala	
	320 325 330	
70	ctc att gct tgg acc gcc act tgc ggt cta gcc ccc gtc atg gcc ttg	1116

# ES 2 345 324 T3

```

Leu Ile Ala Trp Thr Ala Thr Ser Gly Leu Ala Pro Val Met Ala Leu
335                               340                               345                               350

5   tac ggt ggt ccc ttg atc gtc att aat gcc tgg ctg gta ctg tac acg 1164
    Tyr Gly Gly Pro Leu Ile Val Ile Asn Ala Trp Leu Val Leu Tyr Thr
                                355                               360                               365

10  tgg ttg caa cat aca gat acc gat gtt ccg cac ttt tcc tcc gac aac 1212
    Trp Leu Gln His Thr Asp Thr Asp Val Pro His Phe Ser Ser Asp Asn
                                370                               375                               380

15  cac aac ttt gtc aag ggc gca ctg cat acg atc gat cgt ccc tac gac 1260
    His Asn Phe Val Lys Gly Ala Leu His Thr Ile Asp Arg Pro Tyr Asp
                                385                               390                               395

20  aaa ctt gat ccc tgg gga atc ata gac ttt ctg cac cac aag att gga 1308
    Lys Leu Asp Pro Trp Gly Ile Ile Asp Phe Leu His His Lys Ile Gly
    400                               405                               410

25  aca acg cat gtg gca cac cat ttt gac agt act atc ccc cac tat aag 1356
    Thr Thr His Val Ala His His Phe Asp Ser Thr Ile Pro His Tyr Lys
    415                               420                               425                               430

30  gct cag att gct acc gat gcc atc aaa gcc aag ttt cca gaa gtg tac 1404
    Ala Gln Ile Ala Thr Asp Ala Ile Lys Ala Lys Phe Pro Glu Val Tyr
                                435                               440                               445

35  ctc tat gac ccg aca cca att cca caa gcc atg tgg cgc gtc gcc aag 1452
    Leu Tyr Asp Pro Thr Pro Ile Pro Gln Ala Met Trp Arg Val Ala Lys
                                450                               455                               460

40  gga tgt act gca gta gag caa cgc ggt gac gcc tgg gtg tgg aaa aac 1500
    Gly Cys Thr Ala Val Glu Gln Arg Gly Asp Ala Trp Val Trp Lys Asn
                                465                               470                               475

45  gaa gga ata gaa gat ttg gtg gaa cat cgt caa agc aaa tta tcg agc 1548
    Glu Gly Ile Glu Asp Leu Val Glu His Arg Gln Ser Lys Leu Ser Ser
    480                               485                               490

50  gaa taa agcaacatat cgctttatgg aagaacaaac gtccattgtg taaaaccctg 1604
    Glu
    495

55  ataatttcaa tattgtgttt tgttttaaaa aaaaaaaaaa aaaaaaa 1651

```

<210> 6

<211> 495

50 <212> PRT

<213> *Phaeodactylum tricornutum*

55 <400> 6

```

Met Val Arg Phe Ser Thr Ala Ala Leu Leu Ser Leu Ser Thr Leu Thr
  1           5           10           15

60  Thr Ser Cys Ile Gly Ala Phe Gln Leu Ser Ser Pro Ala Gln Leu Pro
    20           25           30

```

65

ES 2 345 324 T3

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

# ES 2 345 324 T3

	340	345	350
5	Gly Pro Leu Ile Val Ile Asn Ala Trp Leu Val Leu Tyr Thr Trp Leu 355                                  360                                  365		
	Gln His Thr Asp Thr Asp Val Pro His Phe Ser Ser Asp Asn His Asn 370                                  375                                  380		
10	Phe Val Lys Gly Ala Leu His Thr Ile Asp Arg Pro Tyr Asp Lys Leu 385                                  390                                  395                                  400		
15	Asp Pro Trp Gly Ile Ile Asp Phe Leu His His Lys Ile Gly Thr Thr 405                                  410                                  415		
	His Val Ala His His Phe Asp Ser Thr Ile Pro His Tyr Lys Ala Gln 420                                  425                                  430		
20	Ile Ala Thr Asp Ala Ile Lys Ala Lys Phe Pro Glu Val Tyr Leu Tyr 435                                  440                                  445		
25	Asp Pro Thr Pro Ile Pro Gln Ala Met Trp Arg Val Ala Lys Gly Cys 450                                  455                                  460		
	Thr Ala Val Glu Gln Arg Gly Asp Ala Trp Val Trp Lys Asn Glu Gly 465                                  470                                  475                                  480		
30	Ile Glu Asp Leu Val Glu His Arg Gln Ser Lys Leu Ser Ser Glu 485                                  490                                  495		

35 <210> 7  
 <211> 1578  
 <212> DNA  
 <213> *Physcomitrella patens*  
 40 <220>  
 <221> CDS  
 <222> (1)..(1578)  
 45 <400> 7

50	atg gta ttc gcg ggc ggt gga ctt cag cag ggc tct ctc gaa gaa aac 48 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn 1                                  5                                  10                                  15
55	atc gac gtc gag cac att gcc agt atg tct ctc ttc agc gac ttc ttc 96 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe 20                                  25                                  30
60	agt tat gtg tct tca act gtt ggt tcg tgg agc gta cac agt ata caa 144 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln 35                                  40                                  45
65	cct ttg aag cgc ctg acg agt aag aag cgt gtt tcg gaa agc gct gcc 192 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala 50                                  55                                  60
70	gtg caa tgt ata tca gct gaa gtt cag aga aat tcg agt acc cag gga 240 Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly

# ES 2 345 324 T3

	65		70		75		80	
5	act gcg gag gca ctc		gca gaa tca gtc gtg	aag ccc acg aga cga agg				288
	Thr Ala Glu Ala		Leu Ala Glu Ser Val Val Lys	Pro Thr Arg Arg				
		85		90		95		
10	tca tct cag tgg aag aag		tcg aca cac ccc cta tca	gaa gta gca gta				336
	Ser Ser Gln Trp Lys Lys		Ser Thr His Pro Leu Ser	Glu Val Ala Val				
		100		105		110		
15	cac aac aag cca agc gat		tgc tgg att gtt gta aaa	aac aag gtg tat				384
	His Asn Lys Pro Ser Asp		Cys Trp Ile Val Val Lys	Asn Lys Val Tyr				
		115		120		125		
20	gat gtt tcc aat ttt gcg		gac gag cat ccc gga gga	tca gtt att agt				432
	Asp Val Ser Asn Phe Ala		Asp Glu His Pro Gly Gly	Ser Val Ile Ser				
		130		135		140		
25	act tat ttt gga cga gac		ggc aca gat gtt ttc tct	agt ttt cat gca				480
	Thr Tyr Phe Gly Arg Asp		Gly Thr Asp Val Phe Ser	Ser Ser Phe His Ala				
		145		150		155		160
30	gct tct aca tgg aaa att		ctt caa gac ttt tac att	ggt gac gtg gag				528
	Ala Ser Thr Trp Lys Ile		Leu Gln Asp Phe Tyr Ile	Gly Asp Val Glu				
		165		170		175		
35	agg gtg gag ccg act cca		gag ctg ctg aaa gat ttc	cga gaa atg aga				576
	Arg Val Glu Pro Thr Pro		Glu Leu Lys Asp Phe Arg	Glu Met Arg				
		180		185		190		
40	gct ctt ttc ctg agg gag		caa ctt ttc aaa agt tcg	aaa ttg tac tat				624
	Ala Leu Phe Leu Arg Glu		Gln Leu Phe Lys Ser Ser	Lys Leu Tyr Tyr				
		195		200		205		
45	ggt atg aag ctg ctc acg		aat gtt gct att ttt gct	gcg agc att gca				672
	Val Met Lys Leu Leu Thr		Asn Val Ala Ile Phe Ala	Ala Ala Ser Ile Ala				
		210		215		220		
50	ata ata tgt tgg agc aag		act att tca gcg gtt ttg	gct tca gct tgt				720
	Ile Ile Cys Trp Ser Lys		Thr Ile Ser Ala Val Leu	Ala Ser Ala Cys				
		225		230		235		240
55	atg atg gct ctg tgt ttc		caa cag tgc gga tgg cta	tcc cat gat ttt				768
	Met Met Ala Leu Cys Phe		Gln Gln Cys Gly Trp Leu	Ser His Asp Phe				
		245		250		255		
60	ctc cac aat cag gtg ttt		gag aca cgc tgg ctt aat	gaa gtt gtc ggg				816
	Leu His Asn Gln Val Phe		Glu Thr Arg Trp Leu Asn	Glu Val Val Gly				
		260		265		270		
65	tat gtg atc ggc aac gcc		ggt ctg ggg ttt agt aca	ggg tgg tgg aag				864
	Tyr Val Ile Gly Asn Ala		Val Leu Gly Phe Ser Thr	Gly Trp Trp Lys				
		275		280		285		
70	gag aag cat aac ctt cat		cat gct gct cca aat gaa	tgc gat cag act				912
	Glu Lys His Asn Leu His		His Ala Ala Pro Asn Glu	Cys Asp Gln Thr				
		290		295		300		



ES 2 345 324 T3

<400> 8

5 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn  
1 5 10 15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe  
20 25 30

10 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln  
35 40 45

15 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala  
50 55 60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly  
65 70 75 80

20 Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg  
85 90 95

25 Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val  
100 105 110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr  
115 120 125

30 Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser  
130 135 140

35 Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala  
145 150 155 160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu  
165 170 175

40 Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg  
180 185 190

45 Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr  
195 200 205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala  
210 215 220

50 Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys  
225 230 235 240

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe  
245 250 255

55 Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly  
260 265 270

60 Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys  
275 280 285

65

ES 2 345 324 T3

Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr  
 290 295 300  
 5 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp  
 305 310 315 320  
 10 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile  
 325 330 335  
 Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg  
 340 345 350  
 15 Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu  
 355 360 365  
 20 Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr  
 370 375 380  
 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro  
 385 390 395 400  
 25 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly  
 405 410 415  
 30 Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser  
 420 425 430  
 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly  
 435 440 445  
 35 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu  
 450 455 460  
 40 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala  
 465 470 475 480  
 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp  
 485 490 495  
 45 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu  
 500 505 510  
 50 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser  
 515 520 525

<210> 9

<211> 873

55 <212> DNA

<213> *Physcomitrella patens*

<220>

60 <221> CDS

<222> (1)..(873)

65

# ES 2 345 324 T3

<400> 9

```

5      atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg      48
      Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
      1          5          10          15

10     cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg gat      96
      Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
      20          25          30

15     acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc atc      144
      Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
      35          40          45

20     gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt ttg      192
      Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
      50          55          60

25     tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt ttg      240
      Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
      65          70          75          80

30     ctc caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc agt      288
      Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
      85          90          95

35     ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg tac      336
      Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
      100          105          110

40     tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg att      384
      Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
      115          120          125

45     ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat acc      432
      Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
      130          135          140

50     gtt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc cac      480
      Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
      145          150          155          160

55     gtt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct cat      528
      Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
      165          170          175

60     cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca gga      576
      His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
      180          185          190

65     gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt cga      624
      Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
      195          200          205

70     agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac ttg      672
      Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
      210          215          220

75     aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct tac      720
      Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
      225          230          235          240

```

# ES 2 345 324 T3

```

tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att   768
Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
                245                      250                      255

5   ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt tac   816
Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
                260                      265                      270

10  gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa   864
Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
                275                      280                      285

15  act gag tga                                                    873
Thr Glu
    290

<210> 10
<211> 290
<212> PRT
<213> Physcomitrella patens

25  <400> 10

30  Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
     1          5          10          15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
     20          25          30

35  Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
     35          40          45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
     50          55          60

40  Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
     65          70          75          80

45  Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
     85          90          95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
     100         105         110

50  Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
     115         120         125

55  Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
     130         135         140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
     145         150         155         160

60  Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
     165         170         175

65  His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly

```

ES 2 345 324 T3

			180					185				190				
5	Val	His	Val	Leu	Met	Tyr	Ala	Tyr	Tyr	Phe	Leu	Ala	Ala	Cys	Leu	Arg
			195					200				205				
	Ser	Ser	Pro	Lys	Leu	Lys	Asn	Lys	Tyr	Leu	Phe	Trp	Gly	Arg	Tyr	Leu
		210					215					220				
10	Thr	Gln	Phe	Gln	Met	Phe	Gln	Phe	Met	Leu	Asn	Leu	Val	Gln	Ala	Tyr
	225					230					235					240
	Tyr	Asp	Met	Lys	Thr	Asn	Ala	Pro	Tyr	Pro	Gln	Trp	Leu	Ile	Lys	Ile
15					245					250					255	
	Leu	Phe	Tyr	Tyr	Met	Ile	Ser	Leu	Leu	Phe	Leu	Phe	Gly	Asn	Phe	Tyr
20				260					265					270		
	Val	Gln	Lys	Tyr	Ile	Lys	Pro	Ser	Asp	Gly	Lys	Gln	Lys	Gly	Ala	Lys
			275					280					285			
25	Thr	Glu														
		290														

30 <210> 11  
 <211> 1526  
 <212> DNA  
 <213> *Phaeodactylum tricornutum*  
 35 <220>  
 <221> CDS  
 <222> (92)..(1402)  
 40 <400> 11

	gcttccggtta gcgtcccata gtttggttaca cttggctgtg aaacgaatac gttcttggtc															60
45	tacttactac	aacgaagcaa	ccaccagcag	c	atg	ggg	aag	gga	ggg	caa	cga					112
					Met	Gly	Lys	Gly	Gly	Gln	Arg					
					1					5						
50	gct gta gct	ccc aag agt	gcc acc agc	tct act ggc	agt gct acc	ctt										160
	Ala Val Ala	Pro Lys Ser	Ala Thr Ser	Ser Thr Gly	Ser Ala Thr	Leu										
		10		15				20								
55	agc caa agc	aag gaa cag	gta tgg act	tcg tcg tac	aac cct ctg	gcg										208
	Ser Gln Ser	Lys Glu Gln	Val Trp Thr	Ser Ser Tyr	Asn Pro Leu	Ala										
		25		30				35								
60	aag gat tcc	ccg gag ctg	cca acc aaa	ggc caa atc	aag gcc gtc	att										256
	Lys Asp Ser	Pro Glu Leu	Pro Thr Lys	Gly Gln Ile	Lys Ala Val	Ile										
		40		45		50		55								
65	ccg aag gaa	tgt ttc caa	cgc tca gcc	ttt tgg tct	acc ttc tac	ctg										304
	Pro Lys Glu	Cys Phe Gln	Arg Ser Ala	Phe Trp Ser	Thr Phe Tyr	Leu										
			60		65			70								
65	atg cgc gat	ctc gcc atg	gct gcc gcc	ttt tgc tac	gga acc tca	cag										352
	Met Arg Asp	Leu Ala Met	Ala Ala Ala	Phe Cys Tyr	Gly Thr Ser	Gln										

# ES 2 345 324 T3

	75	80	85	
5	gtc ctc tcc acc gac ctt ccc caa gac gcc acg ctc att ctg ccc tgg Val Leu Ser Thr Asp Leu Pro Gln Asp Ala Thr Leu Ile Leu Pro Trp 90 95 100			400
10	gct ctc ggc tgg ggc gtc tac gcc ttt tgg atg gga acc att ctc acc Ala Leu Gly Trp Gly Val Tyr Ala Phe Trp Met Gly Thr Ile Leu Thr 105 110 115			448
15	ggg cct tgg gta gtt gcg cac gaa tgt gga cac gcc gct tac tcc gac Gly Pro Trp Val Val Ala His Glu Cys Gly His Gly Ala Tyr Ser Asp 120 125 130 135			496
20	tcc cag acg ttc aat gac gtg gtc gcc ttt atc gtc cac caa gct ttg Ser Gln Thr Phe Asn Asp Val Val Gly Phe Ile Val His Gln Ala Leu 140 145 150			544
25	ctc gtc ccc tac ttt gcc tgg cag tac acc cac gcg aaa cac cac cgt Leu Val Pro Tyr Phe Ala Trp Gln Tyr Thr His Ala Lys His His Arg 155 160 165			592
30	cga acc aac cat ctg gtg gac gcc gag tcc cac gtc cct tct acc gcc Arg Thr Asn His Leu Val Asp Gly Glu Ser His Val Pro Ser Thr Ala 170 175 180			640
35	aag gat aac ggc ctc ggg ccg cac aac gag cga aac tcc ttc tac gcc Lys Asp Asn Gly Leu Gly Pro His Asn Glu Arg Asn Ser Phe Tyr Ala 185 190 195			688
40	gcg tgg cac gag gcc atg gga gac gcc gcc ttt gcc gtc ttt caa gtc Ala Trp His Glu Ala Met Gly Asp Gly Ala Phe Ala Val Phe Gln Val 200 205 210 215			736
45	tgg tcg cac ttg ttc gtc ggc tgg cct ctc tac ttg gcc ggt ctg gcc Trp Ser His Leu Phe Val Gly Trp Pro Leu Tyr Leu Ala Gly Leu Ala 220 225 230			784
50	agt acc gga aag ctt gcg cac gaa ggt tgg tgg ctg gaa gaa cgg aac Ser Thr Gly Lys Leu Ala His Glu Gly Trp Trp Leu Glu Glu Arg Asn 235 240 245			832
55	gcg att gcg gat cac ttt cga ccc agc tct ccc atg ttc ccc gcc aag Ala Ile Ala Asp His Phe Arg Pro Ser Ser Pro Met Phe Pro Ala Lys 250 255 260			880
60	atc cgt gcc aag att gcc ctt tcc agc gcg acg gaa ctc gcc gtg ctc Ile Arg Ala Lys Ile Ala Leu Ser Ser Ala Thr Glu Leu Ala Val Leu 265 270 275			928
65	gct gga ctc ttg tat gtc ggt aca cag gtc gga cac ctt ccc gtc ctg Ala Gly Leu Leu Tyr Val Gly Thr Gln Val Gly His Leu Pro Val Leu 280 285 290 295			976
70	ctg tgg tac tgg gga ccg tac acc ttt gtc aac gct tgg ctt gta ctc Leu Trp Tyr Trp Gly Pro Tyr Thr Phe Val Asn Ala Trp Leu Val Leu 300 305 310			1024

# ES 2 345 324 T3

```

    tac acg tgg ctg cag cat acg gac ccg tcc atc ccg cac tac ggt gaa 1072
    Tyr Thr Trp Leu Gln His Thr Asp Pro Ser Ile Pro His Tyr Gly Glu
                315                      320                      325

5      ggc gag tgg acc tgg gtc aag ggc gcg ctc tct acc att gat cga gac 1120
    Gly Glu Trp Thr Trp Val Lys Gly Ala Leu Ser Thr Ile Asp Arg Asp
                330                      335                      340

    tac ggc atc ttc gat ttc ttt cac cac acc atc ggt tcc acg cac gtg 1168
    Tyr Gly Ile Phe Asp Phe Phe His His Thr Ile Gly Ser Thr His Val
                345                      350                      355

    gta cac cat ttg ttc cac gaa atg ccc tgg tac aat gcc ggc att gcc 1216
    Val His His Leu Phe His Glu Met Pro Trp Tyr Asn Ala Gly Ile Ala
    360                      365                      370                      375

    acg caa aag gtc aag gaa ttt ttg gaa ccc cag ggc ttg tac aat tac 1264
    Thr Gln Lys Val Lys Glu Phe Leu Glu Pro Gln Gly Leu Tyr Asn Tyr
                380                      385                      390

    gat ccg acc ccc tgg tac aag gcc atg tgg cgc att gcc egg acc tgt 1312
    Asp Pro Thr Pro Trp Tyr Lys Ala Met Trp Arg Ile Ala Arg Thr Cys
                395                      400                      405

25     cac tat gtg gag tca aac gag ggt gtg cag tat ttc aag agt atg gaa 1360
    His Tyr Val Glu Ser Asn Glu Gly Val Gln Tyr Phe Lys Ser Met Glu
                410                      415                      420

    aac gtg ccg ctg act aag gat gtg cga aac aaa gcc gca tga 1402
    Asn Val Pro Leu Thr Lys Asp Val Arg Asn Lys Ala Ala
                425                      430                      435

35     gaaaaagtgc caccgacgca taattttaca atcctaccaa caagaccaac attatatggt 1462

    tttcgcttaa aagatagttt tttctaccat ctgtgtagtc ggcacaaaaa aaaaaaaaaa 1522

    aaaa 1526

40     <210> 12
        <211> 436
        <212> PRT
45     <213> Phaeodactylum tricornutum

        <400> 12

50     Met Gly Lys Gly Gly Gln Arg Ala Val Ala Pro Lys Ser Ala Thr Ser
           1           5           10           15

55     Ser Thr Gly Ser Ala Thr Leu Ser Gln Ser Lys Glu Gln Val Trp Thr
           20           25           30

        Ser Ser Tyr Asn Pro Leu Ala Lys Asp Ser Pro Glu Leu Pro Thr Lys
           35           40           45

60     Gly Gln Ile Lys Ala Val Ile Pro Lys Glu Cys Phe Gln Arg Ser Ala
           50           55           60

65     Phe Trp Ser Thr Phe Tyr Leu Met Arg Asp Leu Ala Met Ala Ala Ala

```



## ES 2 345 324 T3

	Pro	Gln	Gly	Leu	Tyr	Asn	Tyr	Asp	Pro	Thr	Pro	Trp	Tyr	Lys	Ala	Met
	385					390					395					400
5	Trp	Arg	Ile	Ala	Arg	Thr	Cys	His	Tyr	Val	Glu	Ser	Asn	Glu	Gly	Val
					405					410					415	
10	Gln	Tyr	Phe	Lys	Ser	Met	Glu	Asn	Val	Pro	Leu	Thr	Lys	Asp	Val	Arg
				420					425					430		
	Asn	Lys	Ala	Ala												
				435												

15

<210> 13

<211> 3598

<212> DNA

20

<213> Desconocido

<220>

<223> Secuencia representa un casete de expresión vegetal promotor-terminador en vector pUC19

25

<400> 13

	tcgcgcgcttt	cggatgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacgggtca	60
30	cagcttgtct	gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	120
	ttggcgggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
35	accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcaggcgcc	240
	attcgccatt	caggctgccc	aactggtggg	aagggcgatc	ggtgcccggc	tcttcgctat	300
40	tacgccagct	ggcgaaaggg	ggatgtgctg	caagggcatt	aagttgggta	acgccagggg	360
	tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgaatt	cggcggcgccg	agctcctcga	420
	gcaaatttac	acattgccac	taaacgtcta	aacccttgta	atttgttttt	gttttactat	480
45	gtgtgttatg	tatttgattt	gcgataaatt	tttatatttg	gtactaaatt	tataaacacct	540
	tttatgctaa	cgtttgccaa	cacttagcaa	tttgcaagtt	gattaattga	ttctaaatta	600
50	ttttgtctt	ctaaatacat	atactaataca	actggaaatg	taaatatttg	ctaataatttc	660
	tactatagga	gaattaaagt	gagtgaatat	ggtaccacaa	ggtttggaga	tttaattggt	720
	gcaatgctgc	atggatggca	tatacaccaa	acattcaata	attccttgagg	ataataatgg	780
55	taccacacaa	gatttgaggt	gcatgaacgt	cacgtggaca	aaaggtttag	taatttttca	840
	agacaacaat	gttaccacac	acaagttttg	aggtgcatgc	atggatgccc	tgtggaaagt	900
60	ttaaaaatat	tttgaaatg	atttgcatgg	aagccatgtg	taaaaccatg	acatccactt	960
	ggaggatgca	ataatgaaga	aaactacaaa	tttacatgca	actagttagt	catgtagtct	1020

65

## ES 2 345 324 T3

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080  
 taatttcttc atagccagcc caccgcggtg ggcggcgccc tgcagtctag aaggcctcct 1140  
 5 gctttaatga gatatgcgag acgectatga tcgcatgata tttgctttca attctgttgt 1200  
 gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgcccgt ttcggttcat 1260  
 10 tctaataaat atatcacccg ttaactatcg atttttatga ataataattct ccgttcaatt 1320  
 tactgattgt ccgtcgacga attcgagctc ggcgcgcaa gcttggcgta atcatggtea 1380  
 tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga 1440  
 15 agcataaagt gtaaagcctg gggtgccctaa tgagtgcgct aactcacatt aattgcgttg 1500  
 cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 1560  
 20 caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac 1620  
 tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata 1680  
 cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa 1740  
 25 aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccct 1800  
 gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgcac aggactataa 1860  
 30 agataccagg cgtttcccc tggaaactcc ctcgctgcgt ctctgtttcc gacctgccc 1920  
 cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca 1980  
 cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 2040  
 35 cccccgctc agcccagccg ctgcgcctta tccggtaact atcgtcttga gtccaaccgc 2100  
 gtaagacacg acttatcgc actggcagca gccactggta acaggattag cagagcgagg 2160  
 40 tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta cactagaagg 2220  
 acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 2280  
 tcttgatccg gcaaaaaaac caccgctggt agcgggtggt tttttgttg caagcagcag 2340  
 45 attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 2400  
 gctcagtgga acgaaaactc acgttaaggg attttggtea tgagattatc aaaaaggatc 2460  
 50 ttcacctaga tccccctaaa ttaaaaatga agtttttaat caatctaaag tatatatgag 2520  
 taaacttggg ctgacagtta ccaatgctta atcagtgcg cactatctc agcgatctgt 2580  
 55 ctatttcggt catccatagt tgcctgactc cccgtcgtgt agataactac gataccgggag 2640  
 ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca 2700  
 gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcttgcaact 2760  
 60

65

## ES 2 345 324 T3

ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttegcca 2820  
gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg 2880  
5 tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 2940  
atgttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg 3000  
10 gccgcagtgt tatcactcac ggttatggca gcactgcata attctcttac tgatcatgcca 3060  
tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg agaatagttg 3120  
atgcgggcag cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc 3180  
15 agaactttaa aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc 3240  
ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca 3300  
20 tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa 3360  
aagggaataa ggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat 3420  
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 3480  
25 aataa caaaa taggggttcc ggcacattt ccccgaaaag tgccacctga cgtctaagaa 3540  
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctctcgtc 3598

30

<210> 14

<211> 3590

<212> DNA

35

<213> Desconocido

<220>

<223> Secuencia representa un casete de expresión vegetal promotor-terminador en vector pUC19

40

<400> 14

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
45 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
50 accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcggc 240  
attcgcatt caggctgcgc aactgttggg aagggcagtc ggtgcgggccc tcttcgctat 300  
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360  
55 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccc agctcctcga 420  
gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat 480  
60 gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540

65

## ES 2 345 324 T3

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600  
 tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660  
 5 tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgaga tttaatgtt 720  
 gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780  
 10 taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca 840  
 agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt 900  
 ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt 960  
 15 ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct 1020  
 atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080  
 20 taatttcttc atagccagcg gatccgatat cgggccgct agcgtaacc ctgctttaat 1140  
 gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg 1200  
 taaaaaacct gagcatgtgt agctcagatc cttaccgccc gtttcggttc attetaatga 1260  
 25 atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa tttactgatt 1320  
 gtccgtcgac gaattcgagc tcggcgcgcc aagcttggcg taatcatggt catagctgtt 1380  
 30 tcctgtgtga aattgttate cgctcacaat tccacacaac atacgagccc gaagcataaa 1440  
 gtgtaaagcc tggggtgrect aatgagtgag otaactcaca ttaattgcgt tgcgctcact 1500  
 gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 1560  
 35 ggggagagggc ggtttgctga ttgggcgctc ttccgcttcc tcgctcactg actcgtgcg 1620  
 ctccgctggt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc 1680  
 40 cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 1740  
 gaaccgtaaa aaggccgcgt tgctggcgtt ttccatagg ctccgcccc ctgacgagca 1800  
 tcacaaaaat cgacgctcaa gtcagagtg gcgaaaccgg acaggactat aaagatacca 1860  
 45 ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgacctgc cgcttaccgg 1920  
 atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct cagcgtgtag 1980  
 50 gtatctcagt tcgggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt 2040  
 tcagcccagc cgctgcgect tatccgtaa ctatcgtctt gagtccaacc cggtaagaca 2100  
 cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 2160  
 55 cgggtctaca gatttctga agtggtggcc taactacggc tacactagaa ggacagtatt 2220  
 tggtatctgc gctctgctga agccagttac ctccgaaaa agagttggta gctcttgatc 2280  
 60

65

## ES 2 345 324 T3

```

cggcaaaaca accaccgctg gtagcggtag tttttttggt tgcaagcagc agattacgcg 2340
cagaaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg 2400
5 gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 2460
gatcctttta aattaataat gaagttttaa atcaatctaa agtatatatg agtaaacttg 2520
10 gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatctcg 2580
ttcatccata gttgcoctgac tccccgctgt gtagataact acgatacggg agggcttacc 2640
atctggcccc agtgctgcaa tgataccgcg agaccacgc tcaccggctc cagatctatc 2700
15 agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 2760
ctccatccag tctattaatt gttgcccggga agctagagta agtagttcgc cagttaatag 2820
20 tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttgggat 2880
ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttggtg 2940
caaaaaagcg gttagctcct teggtcctcc gatcgttgtc agaagtaagt tggccgcagt 3000
25 gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 3060
atgcttttct gtgactgggt agtactcaac caagtcattc tgagaatagt gtatgccccg 3120
30 accgagttgc tcttgccccg cgtaatacag ggataatacc gcgccacata gcagaacttt 3180
aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 3240
35 gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag catcttttac 3300
tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaaggggat 3360
aagggcgaca cggaaatggt gaatactcat actcttcctt tttcaatatt attgaagcat 3420
40 ttatcagggt tattgtotca tgagcggata catatctgaa tgtatctaga aaaataaaca 3480
aataggggtt ccgcgcacat tccccgaaa agtgcacact gacgtctaag aaaccattat 3540
45 tatcatgaca ttaacctata aaaataggcg tatcacgagg ccttttcgtc 3590

```

<210> 15

<211> 3584

50 <212> DNA

<213> Desconocido

<220>

55 <223> Secuencia representa un casete de expresión vegetal promotor-terminador en vector pUC19

60

65

# ES 2 345 324 T3

<400> 15

```

tgcgcgcttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
5   cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
   ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
   accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
10  attcgccatt caggctgcgc aactgttggg aaggcgcatc ggtgcgggcc tcttcgctat 300
   tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggtg acgccagggt 360
   tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccc agctcctcga 420
15  gcaaatttac acattgccc taaacgtcta aacccttgta atttgttttt gttttactat 480
   gtgtgttatg tttttgatlt gcgataaatt tttatatttg gtactaaatt tataacacct 540
   tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600
20  tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaaatattc 660
   tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga ttttaattgtt 720
   gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780
25  taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca 840
   agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900
   ttaaaaaat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt 960
30  ggaggatgca ataataaga aactacaaa tttacatgca actagttatg catgtagtct 1020
   atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080
   taattttctc atagccagca gatctgccgg catcgatccc gggccatggc ctgctttaat 1140
35  gagatatgcg agacgcctat gatcgcatga tttttgcttt caattctgtt gtgcacgttg 1200
   taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attetaatga 1260
   atatatcacc cgttactatc gtattttttat gaataatatt ctccgttcaa tttactgatt 1320
   gtccgctcgc gagctcggcg cgccaagctt ggcgtaatca tggtcatagc tgtttcctgt 1380
   gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa 1440
45  agcctggggg gcctaataag tgagctaact cacattaatt gcgttgcgct cactgccccg 1500
   tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag 1560
   aggcgggttg cgtattgggc gctcttcgcg ttctctgctc actgactcgc tgcgctcggg 1620
50  cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 1680
   atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 1740
   taaaaggccc gcgttgctgg cgtttttcca taggctcgcg cccctgacg agcatcacia 1800

```

55

60

65

## ES 2 345 324 T3

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65

```

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accagggcgtt 1860
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta cgggatacct 1920
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040
cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtaa gacacgactt 2100
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg tagggcgtgc 2160
tacagagttc ttgaagtggg ggctaacta cggctacact agaaggacag tatttggtat 2220
ctgcgctctg ctgaagccag ttacctcgg aaaaagagtt ggtagctctt gatccggcaa 2280
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 2340
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 2400
aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct 2460
tttaaattaa aatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttgggtctga 2520
cagttaccaa tgcttaata gtgaggcacc tatctcagcg atctgtctat ttcgttcctc 2580
catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg 2640
ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat 2700
aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat 2760
ccagtctatt aattggtgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg 2820
caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 2880
attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 2940
agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttctc 3000
actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt 3060
ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 3120
ttgctcttgc cggcgctcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 3180
gctcatcatt ggaaaacggt cttcggggcg aaaactotca aggatcttac cgctgttgag 3240
atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac 3300
cagcgtttct ggggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 3360
gacacggaaa tgttgaatac tcatactctt ctttttcaa tattattgaa gcatttatca 3420
gggttattgt ctcatgagcg gatcacatatt tgaatgtatt tagaaaaata aacaaatagg 3480
ggttccgcgc acatttccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 3540
gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc 3584
  
```

## ES 2 345 324 T3

<210> 16

<211> 4507

<212> DNA

5 <213> Desconocido

<220>

<223> Secuencia representa un casete de expresión vegetal promotor-terminador en vector pUC19

10 <400> 16

```
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
15 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
20 accatattcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactggtggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
25 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccc agctcctcga 420
gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat 480
30 gtgtgttatg tatttgattc gcgataaatt tttatatttg gtactaaatt tataacacct 540
tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600
ttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatctc 660
35 tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgagga ttttaattgtt 720
gcaatgctgc atggatggca tatacaccaa acattcaata attcctgagg ataataatgg 780
40 taccacacaa gatttgaggt gcatgaacgt cagtgagaca aaaggtttag taatttttca 840
agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt 900
ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt 960
45 ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct 1020
atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080
50 taattttctc atagccagcc caccgcggtg ggcggccgcc tgcagtctag aaggcctcct 1140
gctttaatga gatattgcgag acgcctatga tcgcatgata tttgctttca attctgttgt 1200
gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgcccgt ttcggttcat 1260
55 tctaataaat ataccaccg ttactatcgt atttttatga ataataattct ccgttcaatt 1320
```

60

65

## ES 2 345 324 T3

tactgattgt ccgtcgagca aatttacaca ttgccactaa acgtctaaac ctttgaatt 1380  
 5 tgtttttggt ttactatgtg tgttatgtat ttgatctgcg ataaattttt atatttggtta 1440  
 ctaaatttat aacacctttt atgctaacgt ttgccaacac ttagcaattt gcaagttgat 1500  
 taattgattc taaattattt ttgtcttcta aatacatata ctaatcaact ggaaatgtaa 1560  
 10 atatttgcta atatttctac tataggagaa ttaaagtgag tgaatatggt accacaaggt 1620  
 ttggagattt aattggtgca atgctgcatg gatggcatat acaccaaca ttcaataatt 1680  
 cttgaggata ataatggtac cacacaagat ttgaggtgca tgaacgtcac gtggacaaaa 1740  
 15 ggtttagtaa tttttcaaga caacaatggt accacacaca agttttgagg tgcattgcatg 1800  
 gatgccctgt ggaaagttaa aaaatatttt ggaaatgatt tgcattggaag ccatgtgtaa 1860  
 20 aaccatgaca tccacttgga ggatgcaata atgaagaaaa ctacaaattt acatgcaact 1920  
 agttatgcat gtagtctata taatgaggat ttgcaatac tttcattcat acacactcac 1980  
 taagttttac acgattataa tttcttcata gccagcggat ccgatatcgg gcccgctagc 2040  
 25 gttaaccctg ctttaatgag atatgcgaga cgccatgat cgcatgatat ttgctttcaa 2100  
 ttctgttggt cacgttgtaa aaaacctgag catgtgtgac tcagatcctt accgccggtt 2160  
 tcggttcatt ctaatgaata tatcacccgt tactatcgta tttttatgaa taatattctc 2220  
 cgttcaattt actgattgtc cgtcgacgaa ttcgagctcg gcgcgccaag cttggcgtaa 2280  
 tcatggatcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata 2340  
 35 cgagccggaa gcataaagtg taaagcctgg ggtgcctaata gagtgagcta actcacatta 2400  
 attgctgtgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa 2460  
 40 tgaatcggcc aacgcgcggg gagagggcgg ttgctgattg ggctctcttc cgttctctcg 2520  
 ctcaactgact cgtgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag 2580  
 gcggtaatac ggttatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa 2640  
 45 ggccagcaaa aggccaggaa ccgtaaaaag gccgcggttc tggcgttttt ccataggttc 2700  
 cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca 2760  
 50 ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc tctgttccg 2820  
 accctgccgc ttaocggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 2880  
 catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 2940  
 55 gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggttaacta tcgtcttgag 3000  
 tccaaccgg taagacacga cttatcgcca ctggcagcag ccaactggtta caggattagc 3060

60

65

## ES 2 345 324 T3

```

agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 3120
actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 3180
5 gttggtagct cttgatccgg caaacaacc accgctggta gcggtggttt ttttgtttgc 3240
aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 3300
10 gggctctgacg ctcagtggaa cgaaaactca cgtaaggga ttttggtcat gagattatca 3360
aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt 3420
15 atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 3480
gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg 3540
atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca 3600
20 ccgctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc 3660
cctgcaactt tatccgcctc catccagtct attaattggt gccgggaagc tagagtaagt 3720
agttcgccag ttaatagttt gcgcaacgct gttgccattg ctacaggcat cgtgggtgca 3780
25 cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca 3840
tgatccccca tgttggtgcaa aaaagcgggt agctccttcg gtccctccgat cgttggtcaga 3900
30 agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact 3960
gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga 4020
gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg 4080
35 ccacatagca gaactttaa agtgcctcctc attggaaaac gttcttcggg gcgaaaactc 4140
tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccaactcgtgc acccaactga 4200
40 tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat 4260
gccgcaaaaa agggaataag ggcgacacgg aaatggtgaa tactcatact cttccttttt 4320
caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 4380
45 attagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 4440
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat caccaggccc 4500
50 tttcgtc 4507

```

<210> 17

55 <211> 5410

<212> DNA

<213> Desconocido

<220>

60 <223> Secuencia representa un casete de expresión vegetal promotor-terminador en vector pUC19

65

# ES 2 345 324 T3

<400> 17

5            ttttggaaat gatttgcacg gaagccatgt gtaaaacocat gacatccact tggaggatgc 60  
             aataatgaag aaaactacaa atttacatgc aactagttat gcatgtagtc tatataatga 120  
             ggattttgca atactttcat tcatacacac tcaactaagt ttacacgatt ataatttctt 180  
 10           catagccagc ggatccgata tcgggcccgc tagcgtaaac cctgcttaa tgagatatgc 240  
             gagacgccta tgatcgcacg atatttgctt tcaattctgt tgtgcacggt gtaaaaaacc 300  
             tgagcatgtg tagctcagat ccttaccgcc ggtttcgggt cattctaata aatataatcac 360  
 15           cegttactat cgtatTTTTA tgaataatat tctccgttca atttactgat tgtccgctga 420  
             gcaaatttac acattgccac taaacgtcta aacccttgta atttgTTTTT gttttactat 480  
 20           gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540  
             tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600  
 25           tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660  
             tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga ttttaattgtt 720  
             gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780  
 30           taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca 840  
             agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900  
 35           ttaaaaatat tttggaaatg atttgcacg aagccatgtg taaaaccatg acatccactt 960  
             ggaggatgca ataataaga aaactacaaa ttacatgca actagttatg catgtagtct 1020  
             atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080  
 40           taatttcttc atagccagca gatctgccgg catcgatccc gggccatggc ctgctttaat 1140  
             gagatatgcg agacgcctat gatcgcacga tatttgcttt caattctggt gtgcacggtg 1200  
 45           taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcgggtc attctaataga 1260  
             atatatcacc cgttactatc gtatTTTTAT gaataatatt ctccgttcaa tttactgatt 1320  
             gtccgctgac gagctcggcg cgccaagctt ggcgtaataca tggatcatagc tgtttcctgt 1380  
 50           gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa 1440  
             agcctggggg gcctaataag tgagctaact cacattaatt gcgttgcgct cactgcccgc 1500  
 55           tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag 1560  
             aggcgggttg cgtattgggc gctcttccgc ttctcgcctc actgactcgc tgcgctcggg 1620  
 60  
 65

## ES 2 345 324 T3

cgttcggctg cggcgagcgg tadcagctca ctcaaaggcg gtaatacggg tatccacaga 1680  
 atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 1740  
 5 taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia 1800  
 aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 1860  
 10 tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta cgggatacct 1920  
 gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980  
 cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040  
 15 cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt 2100  
 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtagt taggcggtgc 2160  
 20 tacagagttc ttgaagtggg ggccctaacta cggctacact agaaggacag tatttggtat 2220  
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagt ggtagctctt gatccggcaa 2280  
 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 2340  
 25 aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 2400  
 aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct 2460  
 30 tttaaattaa aaatgaagt ttaaatcaat ctaaagtata tatgagtaa cttggtctga 2520  
 cagttaccaa tgcttaatea gtgaggcacc tatctcagcg atctgtctat ttcgctcacc 2580  
 catagttgcc tgactccccg tctgttagat aactacgata cgggagggct taccatctgg 2640  
 35 ccccagtctg gcaatgatac cgcgagaccc acgctcaccg gctccagatt tadcagcaat 2700  
 aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgctccat 2760  
 ccagtctatt aattgttgc gggagctag agtaagtagt tcgccagtta atagtttgcg 2820  
 40 caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 2880  
 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 2940  
 45 agcggtagc tccttcggtc ctccgatcgt tgtcagaagt aagtggccg cagtgttacc 3000  
 actcatggtt atggcagcac tgcataatc tcttactgtc atgccatccg taagatgctt 3060  
 50 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 3120  
 ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 3180  
 gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag 3240  
 55 atccagttcg atgtaaccca ctctgtcacc caactgatct tcagcatctt ttactttcac 3300  
 cagcgtttct gggtagcaaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 3360  
 60  
 65

## ES 2 345 324 T3

gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca 3420  
 gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 3480  
 5 ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 3540  
 gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtctcgcgc gtttcgggtga 3600  
 tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gctctgtaagc 3660  
 10 ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg 3720  
 ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga 3780  
 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgcattcgc cattcaggct 3840  
 15 gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa 3900  
 aggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 3960  
 ttgtaaaaacg acggccagtg aattcggcgc gccgagctcc tcgagcaaat ttacacattg 4020  
 20 ccactaaaacg tctaaacct tgtaatttgt tttgtttta ctatgtgtgt tatgtatttg 4080  
 atttgcgata aatthttata tttgttacta aatthataac acctthtatg ctaacgtttg 4140  
 ccaacactta gcaatttga agttgattaa ttgattctaa attatthttg tcttctaaat 4200  
 25 acatatacta atcaactgga aatgtaaata tttgctaata tttctactat aggagaatta 4260  
 aagtgagtga atatggtacc acaaggtttg gagatttaat tgttgcaatg ctgcatggat 4320  
 ggcatataca ccaaacattc aataattctt gaggataata atggtaccac acaagatttg 4380  
 30 aggtgcatga acgtcacgtg gacaaaagg ttagtaattt tcaagacaa caatgttacc 4440  
 acacacaagt tttgaggtgc atgcatggat gccctgtgga aagtttaaaa atatthttgga 4500  
 aatgatttgc atggaagcca tgtgtaaac catgacatcc acttgaggga tgcaataatg 4560  
 35 aagaaaacta caaatttaca tgcaactagt tatgcatgta gtctatataa tgaggatttt 4620  
 gcaatacttt cattcataca cactcactaa gttttacacg attataattt ctcatagcc 4680  
 40 agcccaccgc ggtgggcccgc cgcctgcagt ctagaaggcc tccctgttta atgagatag 4740  
 cgagacgcct atgatcgcac gatatttgc tccaattctg ttgtgcacgt tgtaaaaaac 4800  
 ctgagcatgt gtagctcaga tccttaccgc cggtttcggg tcattctaat gaatatatca 4860  
 45 cccgttacta tcgtattttt atgaataata ttctccgttc aatthactga ttgtccgctc 4920  
 agcaaattha cacattgcca ctaaactctt aaaccttct aatthgtttt tgtthtacta 4980  
 50 tgtgtgttat gtatthgatt tgcgataaat ttttatattt ggtactaaat ttataacacc 5040  
 ttttatgcta acgtthtcca acacttagca atthtgaagt tgattaattg attctaaatt 5100  
 atthttgtct tctaaataca tataactaat aactggaaat gtaaatattt gctaataatt 5160  
 55 ctactatagg agaattaaag tgagtgaata tggtaaccaca aggtthtgag atttaattgt 5220  
 tgcaatgctg catggatggc atatacacca aacattcaat aatthctgag gataataatg 5280  
 60 gtaccacaca agatthgagg tgcatgaacg tcacgtggac aaaaggthta gtaatthttc 5340  
 aagacaacaa tgttaccaca cacaagthtt gaggtgcatg catggatgcc ctgtggaaag 5400  
 tttaaaaata 5410

65

# ES 2 345 324 T3

<210> 18  
 <211> 648  
 <212> DNA  
 5 <213> *Phaeodactylum tricornutum*  
 <220>  
 <221> CDS  
 10 <222> (1)..(648)  
 <220>  
 <223>  
 15 <400> 18

		tgg tgg aaa aac aag cac aac gga cac cac gcc gtc ccc aac ctc cac			48
	Trp Trp Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His				
20	1	5	10	15	
	tgc tcc tcc gca gtc gcg caa gat ggg gac ccg gac atc gat acc atg				96
	Cys Ser Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met	20	25	30	
25	ccc ctt ctc gcc tgg tcc gtc cag caa gcc cag tct tac cgg gaa ctc				144
	Pro Leu Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu	35	40	45	
30	caa gcc gac gga aag gat tgc ggt ttg gtc aag ttc atg atc cgt aac				192
	Gln Ala Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn	50	55	60	
35	caa tcc tac ttt tac ttt ccc atc ttg ttg ctc gcc cgc ctg tgc tgg				240
	Gln Ser Tyr Phe Tyr Phe Pro Ile Leu Leu Ala Arg Leu Ser Trp	65	70	75	80
40	ttg aac gag tcc ttc aag tgc gcc ttt ggg ctt gga gct gcg tgc gag				288
	Leu Asn Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu	85	90	95	
45	aac gct gct ctc gaa ctc aag gcc aag ggt ctt cag tac ccc ctt ttg				336
	Asn Ala Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu	100	105	110	
50	gaa aag gct ggc atc ctg ctg cac tac gct tgg atg ctt aca gtt tgc				384
	Glu Lys Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser	115	120	125	
55	tcc ggc ttt gga cgc ttc tgc ttc gcg tac acc gca ttt tac ttt cta				432
	Ser Gly Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu	130	135	140	
60	acc gcg acc gcg tcc tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc				480
	Thr Ala Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu	145	150	155	160
65	ggc cac aac ggc atg gcc acc tac aat gcc gac gcc cgt ccg gac ttc				528
	Gly His Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe	165	170	175	
70	tgg aag ctc caa gtc acc acg act cgc aac gtc acg ggc gga cac ggt				576
	Trp Lys Leu Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly	180	185	190	
75	ttc ccc caa gcc ttt gtc gac tgg ttc tgt ggt ggc ctc cag tac caa				624
	Phe Pro Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln	195	200	205	
80	gtc gac cac cac tta ttc ccc agc				648
	Val Asp His His Leu Phe Pro Ser	210	215		

ES 2 345 324 T3

<210> 19

<211> 216

<212> PRT

5 <213> *Phaeodactylum tricornutum*

<400> 19

10 Trp Trp Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His  
1 5 10 15

15 Cys Ser Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met  
20 25 30

Pro Leu Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu  
35 40 45

20 Gln Ala Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn  
50 55 60

25 Gln Ser Tyr Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp  
65 70 75 80

Leu Asn Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu  
85 90 95

30 Asn Ala Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu  
100 105 110

35 Glu Lys Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser  
115 120 125

Ser Gly Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu  
130 135 140

40 Thr Ala Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu  
145 150 155 160

45 Gly His Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe  
165 170 175

50 Trp Lys Leu Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly  
180 185 190

Phe Pro Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln  
195 200 205

55 Val Asp His His Leu Phe Pro Ser  
210 215

<210> 20

60 <211> 12093

<212> DNA

<213> Desconocido

<220>

65 <223> Vector de expresión vegetal con un casete de expresión promotor-terminador

# ES 2 345 324 T3

<400> 20

5 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca 120  
tagtgggccc tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc 180  
10 ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
15 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420  
gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag cagccggcgc 480  
20 tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg 540  
cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
25 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
ccggcacgcg accggggcgca ccgacagatgg aaacggccga cgcgcagctt cgcttcctct 720  
gcgagggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
30 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca 840  
ccgttgaaca ggctccgctc tcgcccgtgt tcgggggccg gatagacgcc ttcgacgaag 900  
35 ccggtccgga cgcagcgttc gagcaggac tcgcggtgat tgcgatgga ttggcgaaaa 960

40

45

50

55

60

65

## ES 2 345 324 T3

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatecc ctcccccttt 1080  
 5 ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcattgcct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct ggcgcccttc tggcgctctt ccgcttcctc 1200  
 10 gctcaactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcaactcaa 1260  
 ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt tccataggct 1380  
 15 ccgccccct gacgagcctc acaaaaatcg acgctcaagt cagaggtggc gaaaccgcac 1440  
 aggactataa agataccagg cgtttcccc tggaaagctc ctctgctgct ctctgttcc 1500  
 gaccctgccg ctaccggat acctgtccgc cttctcctc tcgggaagcg tggcgctttt 1560  
 20 ccgctgcata accctgcttc ggggtcatta tagcgatctt ttcggtatat ccatcctttt 1620  
 tcgcacgata tacaggatct tgccaaaggg ttcgtgtaga cttccttgg tgtatccaac 1680  
 25 ggcgtcagcc ggcagcata ggtgaagtag gccaccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttcgacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaccaag ccaaccagga 1860  
 30 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 aggcggcggc ggcggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980  
 35 aaatcacggg cgtcgtggac tatgagcagc tccgcgagct ggcccgcatc aatggcgacc 2040  
 tgggccgctt ggcggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt 2100  
 tcggtgatgc cagatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 40 gcaaggctcat gatggcgctg gtccgccga ggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg gggctgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctggtg ccctcggcgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 50 cgccagcaa acgctcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata 2460  
 cctcgggaa aacttgccc tcaactgacag atgaggggag gacgctgaca cttgaggggc 2520  
 cgactcccc ggcggcgct tgacagatga gggcaggct cgatttcggc cggcgacgtg 2580  
 55 gagctggcca gcctcgaaa tcggcgaaaa cgcctgattt tacgaggtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcct gcggtattga cacttgagg gcgcgactac 2700

60

65

## ES 2 345 324 T3

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60

```

tgacagatga ggggcgcat ccttgacact tgaggggag agtgctgaca gatgaggggc 2760
gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820
ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880
aaaccttggt ttttaaccagg gctgcccctt gtgcccgtga ccgcccagcg cgaagggggg 2940
tgccccctt tctcgaacct tcccggcccg ctaacggcgg cctcccatcc ccccaggggc 3000
tgccccttc ggccgcgaac ggccctcacc caaaaatggc agcgcctggca gtccttgcca 3060
ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120
ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180
gcggcctggg tggcggcctg cccttcaact cggccgtcgg ggcattcacg gacttcatgg 3240
cggggccggc aatttttaac ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300
tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360
gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420
ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccctgaat atattgacaa 3480
tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540
ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600
tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660
attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720
tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
gctgcctcag attcaggta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt 3840
gcagctttcc cttcaggcgg gattcataca gcccagcc atccgctatc catatcacca 3900
cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960
atacgtgccc aacaaccgtc ttccggagac tgtcatagc gtaaaacagc cagcgcctggc 4020
gcgatttagc cccgacatag cccactggt cgtccatttc cgcgcagacg atgacgtcac 4080
tgcccggctg tatgcccgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140
cgtggtgagg ccaacgcccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200
catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260
ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtg ttttgccgtt 4320
acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380
agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
  
```

65

## ES 2 345 324 T3

tggtttcaaa atcggetccg togatactat gttatacgcc aactttgaaa acaactttga 4500  
 5 aaaagctggt ttctggattt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 10 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctgggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaagggg ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 15 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcctc 4980  
 20 aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag 5160  
 25 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280  
 30 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340  
 ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 35 caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggccacg gcaagtattt 5520  
 gggcaagggg tcgctgggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 40 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc ggggtcaaate aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgcaagga ggggtgaatga atcggacggt tgaccggaag gcatacaggc aagaactgat 5760  
 45 cgacgcgggg ttttccgccc aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttccagt ccgtcggctc gatgggtccag caagctacgg ccaagatcga 5880  
 50 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940  
 ttgcgctcgt ctcgaaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 55 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aatgcagct 6120  
 60 ttecttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc 6180

65

## ES 2 345 324 T3

ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccc cgcgaggcgc tgcaaaacaa 6240  
 5 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 10 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcggtgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct 6540  
 15 ggaccgtggc aagaaaacgt cccggttgcca ggtcctgac gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgcggac 6660  
 ggccccgacg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 20 aaccttccgc ctcatgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgcaagagt tgcgaggcag cggcctggtg gaacacgcct gggtaaatga 6840  
 25 tgacctggtg cattgcaaac gctagggcct tgtgggggtca gttccggctg ggggttcagc 6900  
 agccagcgct ttactggcat ttcaggaaca agcgggcact gtcgacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020  
 30 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 35 cacgaggaga aaaagcccat ggaggcgttc gctgaacggg tgcgagatgc cgtggcattc 7200  
 gggcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 40 ggggtcgcgg gtatgctgct gcgggcggtg ccggcgggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaattatt 7440  
 45 tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccggggcg ggtcggggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc 7560  
 ccgatacgat tgatggcggg cctgggggct atttgcgaa ctgcggggcgt ggcgctgttg 7620  
 50 gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggcg 7680  
 gttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc 7740  
 55 acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 ttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actogaacct 7920  
 60  
 65

## ES 2 345 324 T3

acagttgttt ccttactggg cttctcagc cccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa ctccgggtcc ccgacctgta ccattcgggtg agcaatggat 8040  
 5 aggggagttg atatcgtaaa cgttcacttc taaagaaata gcgccactca gcttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 10 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgagag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgagaga 8280  
 tcatccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 15 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctggt ggctggctgg 8460  
 20 tggcaggata tattgtgggtg taaacaaatt gacgcttaga caacttaata acacattgag 8520  
 gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgccctcac cgccctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 25 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggg tgagtgttgt tccagtttgg aacaagagtc cactatataa 8760  
 30 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacy tggcgagaaa 8940  
 35 ggaaggggag aaagcgaaag gagcggggcg cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 40 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaataatt attgataaaa taacaagtca 9180  
 45 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatc cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgctgaagct 9360  
 50 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggctag ccatttcgcc gccaaactct 9480  
 55 tcagcaatat cacgggtagc caacgctatg tctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 60 tcgccatggg tcacgacgag atcctcgcgg tcgggcatgc gcgccttgag cctggcgaac 9660

65

## ES 2 345 324 T3

agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgac gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 5 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
 10 tcccttcccg cttcagtgac aacgtcgagc acagctgccc aaggaacgcc cgctcgtggcc 9960  
 agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc 10020  
 15 ttgacaaaaa gaaccgggag cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcccc gtcatagccg aatagcctct ccacccaagc ggccggagaa 10140  
 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 20 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aatatattgc tagctgatag tgacctagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 25 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccg 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gctcatgac ttgatccct 10500  
 30 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 cttaccagag ggcgccccag ctggcaatc cggttcgctt gctgtccata aaaccgcca 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcgtt 10680  
 35 ttcccttctc cagatagccc agtagctgac attcacccg ggtcagcacc gtttctgcgg 10740  
 actggcttcc tacgtgttcc gcttccctta gcagcccttg cgccctgagt gcttgcggca 10800  
 40 gcgtgaagct tgcatgcctg caggtcgagc gcgcgccgag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 45 tttgatttgc gataaatctt tatatttggg actaaattta taacacctt tatgctaacg 10980  
 tttgccaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100  
 50 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 55 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta attttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 60 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400

65

## ES 2 345 324 T3

```

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460
ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat 11520
5 agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga 11580
tatgcgagac gcctatgata gcatgatatt tgctttcaat tctggtgtgc acgttgtaaa 11640
10 aaacctgagc atgtgtagct cagatcctta ccgcccgttt cggttcattc taatgaatat 11700
atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760
gtcgcgcaat tcgagctcgg cgcgcctcta gaggatcgat gaattcagat cggctgagtg 11820
15 gctccttcaa cgttgctggt ctgtcagttc caaacgtaaa acggcttgtc ccgcgtcacc 11880
ggcggggggtc ataacgtgac tcccttaatt ctccgctcat gatcagattg tcgtttcccg 11940
20 ccttcagttt aaactatcag tgtttgacag gatataattgg cgggtaaacc taagagaaaa 12000
gagcgtttat tagaataatc ggatatttaa aagggcgtga aaaggtttat ccttcgtcca 12060
tttgtatgtg catgccaacc acagggttcc cca 12093

```

```

<210> 21
<211> 12085
30 <212> DNA
<213> Desconocido
<220>
<223> Vector de expresión vegetal con un casete de expresión promotor-terminador
35 <400> 21

```

```

gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60
gcccagcga caggtgcgca ggcaaattgc accaacgcat acagcgcag cagaatgcca 120
tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180
45 ataatcaggc cgatgccgac agcgtcagc gcgacagtgc tcagaattac gatcaggggt 240
atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300
50 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360
tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420
gcgtagacgg tctgacgaca cgcaaactgg cggaacgggt gggggttcag cagccggcgc 480
55 tttactggca cttcaggaac aagcgggccc tgctcagcgc actggccgaa gccatgctgg 540
cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgtca tttctgatcg 600
60 ggaatgcccc cagcttcagg cagggcgtgc tcgcctaacc cgatggcgcg cgcatecatg 660

```

65

## ES 2 345 324 T3

ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720  
 gcgaggcggg tttttcggcc ggggacgcg tcaatgcgct gatgacaatc agctacttca 780  
 5 ctgltggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca 840  
 ccgttgaaca ggctccgctc tcgcccgtgt tgccggccgc gatagacgcc ttcgacgaag 900  
 10 ccggtccgga cgcagcgttc gagcaggac tcgcggtgat tgctgatgga ttggcgaaaa 960  
 ggaggctcgt tgctaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccctc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 15 ccaccgcgtc agacgcccg agcagccgc tacgggcttt ttcatgccct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct ggccgecttc tggcgctctt ccgcttcctc 1200  
 20 gctcactgac tcgctgcgct cggtcgttcg gctgcccga gcggtatcag ctcaactcaa 1260  
 ggcgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt tccataggct 1380  
 25 ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagagggtggc gaaaccgcac 1440  
 aggactataa agataccagg cgtttcccc tggaagctcc ctcgctgcgct ctccgttcc 1500  
 30 gaccctgccg cttaccgat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatectttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac 1680  
 35 ggcgtcagcc gggcaggata ggtgaagtag gccaccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg 1800  
 40 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtaactgc ttccagacga acgaagagcg attgaggaaa 1920  
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgctggc cagggtaca 1980  
 45 aaatcacggg cgtcgtggac tatgagcag tccgcgagct ggcccgcac aatggcgacc 2040  
 tgggcccctt gggcggcctg ctgaaactct ggctaccga cgaccgcgc acggcgcggt 2100  
 50 tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggcat gatggcgctg gtccgccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg gggcgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 55 gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcttttgc 2340  
 gacgctcacc gggctggtt cctcgcgcg tgggctggcg gccgtctatg gccctgcaa 2400  
 60  
 65

## ES 2 345 324 T3

cgcgccagaa acgcccgtoga agccgtgtgc gagacaccgc ggcccggcgc gttgtggata 2460  
 cctcgcggaa aacttggccc tcaactgacag atgaggggcg gacgttgaca cttgaggggc 2520  
 5 cgactcaccg ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcyagtt tcccacagat 2640  
 10 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcyactac 2700  
 tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 15 ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 aaaccttggt ttttaaccagg gctgcgcctt gtgcgcgtga ccgcycaagc cgaagggggg 2940  
 20 tgccccctt tctcgaacct tcccggcccg ctaacgcggg cctcccatcc ccccaggggc 3000  
 tgccccctc ggccgcgaac ggcctcacc caaaaatggc agcgtggca gtccttgcca 3060  
 ttgcccggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 25 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 ggggcctggg tgggggcctg cccttcaact cggccgtcgg ggcattcagc gacttcatgg 3240  
 30 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaga ttataccgag 3360  
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgcctt atttaaaaag 3420  
 35 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 40 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720  
 45 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcaggtta tgccgtcaa ttcgctcgtt atatcgcttg ctgattacgt 3840  
 50 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgcgc aacaaccgtc ttccggagac tgcatacgc gtaaacagc cagcgtggc 4020  
 55 gcgatttagc cccgacatag cccaactggt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgccccgctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140

60

65

## ES 2 345 324 T3

5 cgtggtgagg ccaacgcccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 10 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
 tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctggatatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 15 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 20 gtaaaagata cggaaggaat gtctctgct aaggtatata agctgggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggcagccgg tataaagga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgctg ttccaaaggt cctgcacttt 4860  
 25 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 30 aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctgccc atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
 35 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcggg caagtggat 5280  
 40 gacattgctt tctgctccg gtcgatcagg gaggatatcg ggaagaaca gtatgtcgag 5340  
 ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 45 caccgacttc ttccgatca agtggttttg ctctcaggcc gagggccacg gcaagtattt 5520  
 gggcaagggg tcgtgggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 50 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc ggggtcaaac aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgaagga gggatgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 55 cgacggggg ttttccgccc aggatgccga aaccatcgca agccgcaccg tcatgctgctc 5820  
 gccccgggaa acctccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga 5880

60

65

## ES 2 345 324 T3

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcgggccg ccgtggagcg 5940  
 ttcgcgtcgt ctgcaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 5 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aatgcagct 6120  
 10 ttccttgctc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccc cgcgaggcgc tgcaaaaaca 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 15 cgacgatgac gaactgggtg ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 20 ccggtattac acgaaggccg aggaatgcct gtcgccccta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcgttgggc acctggaatc ggtgtcgcctg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 25 gtttgcctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgcggac 6660  
 ggcccgcagg atgttcgact atttcagctc gcaccgggag ccgtaccgcg tcaagctgga 6720  
 30 aaccttcgcg ctcatgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tcgcaagagt tgcgagcag cggcctggtg gaacacgcct gggtcfaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc 6900  
 35 agccagcgct ttactggcat ttcaggaaca agcgggcaact gctcgacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgctc tacgaaactgc cgataaacag aggattaaaa 7020  
 40 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc 7200  
 45 ggcgcctaca tcgacggcga gatcattggg ctgtcgttct tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 50 ggggtcgcgg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatat 7440  
 55 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggccc ggtcgcggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc 7560  
 ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcggcgtt ggcgctgttg 7620

60

65

## ES 2 345 324 T3

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggect ggcgggggag 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctccgt gcctctgtc 7740  
 5 acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagt 7800  
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 10 ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcggc actcgaacct 7920  
 acagtgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat 8040  
 15 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 20 cggttaagcg agaaatgaat aagaaggctg ataattcggg tctctcggag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280  
 25 tcatccgtgt ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg 8460  
 30 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat 8580  
 35 tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcceca 8640  
 gcaggcgaaa atcctgtttg atgggtggtc cgaaatcggc aaaatccct ataaatcaaa 8700  
 agaatagccc gagataggtg tgagtgttgt tccagtttgg aacaagagtc cactatataa 8760  
 40 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaacatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 45 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 ggaaggggag aaagcgaaag gagcggggcg cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 50 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaacg acggccagtg 9120  
 aattaattcc catcttgaag gaaatatagt taaatattt attgataaaa taacaagtca 9180  
 55 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 60 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360

65

## ES 2 345 324 T3

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattegcc gccaaagctct 9480  
 5 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcca tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 10 tcgccatggg tcacgacgag atcctcgcog tcgggcatgc gcgccttgag cctggcgaac 9660  
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 15 gtagccggat caagcgtatg cagccgcgcg attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
 20 tcccttcccc cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
 agccacgata gccgcgctgc ctgcctctgc agttcattca gggcaccgga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg ccctcgcgct gacagccgga acacggcggc atcagagcag 10080  
 25 ccgattgtct gttgtgcccc gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 cctgcgtgca atccatcttg ttcaatocaa gctcccattg gccctcgact agagtcgaga 10200  
 30 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 35 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacatgac ttgatccct 10500  
 40 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctcttg cgcttgctt 10680  
 45 ttccttctg cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
 actggcttcc taegtgttcc gcttcttcta gcagccctg cgccctgagt gcttggggca 10800  
 50 gcgtgaagct tgcatgcctg caggtcgacg gcgcgcccag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacacctt tatgctaacg 10980  
 55 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgct aatattteta ctataggaga 11100

60

65

## ES 2 345 324 T3

attaaagtga gtgaatatgg taccacaagg tttggagatt taattggtgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 5 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atctttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 10 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cagcattata atttcttcat 11520  
 15 agccagcgga tccgatatcg ggcccgctag cgttaaccct gctttaatga gatatgagag 11580  
 acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga 11640  
 20 gcatgtgtag ctcagatcct taccgcccgt ttcggttcat tctaatgaat atatcaccgg 11700  
 ttactatcgt atttttatga ataataattct ccgttcaatt tactgattgt ccgtcgacga 11760  
 attcgagctc ggcgcgcctc tagaggatcg atgaattcag atcggctgag tggctccttc 11820  
 25 aacgttgccg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 11880  
 tcataacgtg actcccttaa ttctccgctc atgatecagat tgtcgtttcc cgccttcagt 11940  
 30 ttaaactatc agtgtttgac aggatatatt ggccgggtaaa cctaagagaa aagagcgttt 12000  
 attagaataa tcggatattt aaaagggcgt gaaaaggttt atccttcgtc catttgatg 12060  
 tgcattgcca ccacaggggt cccca 12085

35

<210> 22

<211> 12079

40 <212> DNA

<213> Desconocido

<220>

<223> Vector de expresión vegetal con un casete de expresión promotor-terminador

45

<400> 22

gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaacg atccgacagc 60  
 50 ggcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgcag cagaatgcca 120  
 tagtgggccc tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc 180  
 55 ataatcaggc cgatgccgac agcgtcagac gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 60 ttctttatca ctgataagt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaagt gcagccgaat acagtgatec gtgccccct ggacctgtg aacgaggtcg 420

65

## ES 2 345 324 T3

gcgtagacgg tctgacgaca cgcaaactgg cggaacgggtt gggggttcag cagccggcgc 480  
 tttactggca cttcaggaac aagcggggcgc tgctcgacgc actggccgaa gccatgctgg 540  
 5 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 10 ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720  
 gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gaggcggccg gcgacagcga tgccggcgag cgcggcggca 840  
 15 ccgttgaaca ggctccgctc tcgccgctgt tcggggccgc gatagacgcc ttcgacgaag 900  
 ccggtccgga cgcagcgctc gaggcgggac tcgccgtgat tgtcgatgga ttggcgaaaa 960  
 20 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccctc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgctc agaccccgt agcagcccgc tacgggcttt ttcatgcct gccctagcgt 1140  
 25 ccaagcctca cggcccgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc 1200  
 gctcactgac tcgtcgctc cggctcgtcg gctgcggcga gcggtatcag ctactcaaa 1260  
 30 ggcgtaata cggttatcca cagaatcagg ggataacgca ggaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaa gcccgcgttg ctggcgcttt tccataggct 1380  
 ccgccccct gacgagcacc acaaaaatcg acgctcaagt cagagggtgc gaaaccgac 1440  
 35 aggactataa agataccagg cgtttcccc tggaagctcc ctctgctcct ctctgttcc 1500  
 gaccctgccc cttaccgat acctgtccgc ctttctccc tcgggaagcg tggcgctttt 1560  
 40 ccgctgcata acctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttcttgg tgtatccaac 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccaccgc gagcgggtgt tccttcttca 1740  
 45 ctgtccctta ttcgacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 50 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980  
 aaatcacggg cgtcgtggac tatgagcag tccgcgagct ggcccgcac aatggcgacc 2040  
 55 tgggccgctt gggcggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt 2100  
 tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160

60

65

## ES 2 345 324 T3

5 gcaaggtcat gatgggctg gtccgccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 10 gacttcgctg agctgggtgaa gtacatcacc gacgagcaag gcaagaccga ggccttttgc 2340  
 gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata 2460  
 cctcgccgaa aacttggccc tcaactgacag atgagggggc gacgttgaca cttgaggggc 2520  
 15 cgactcacc ggcgcccgtg tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgagagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg ggcgactac 2700  
 20 tgacagatga ggggctgat ccttgacact tgaggggagc agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 ccgcccgttt ttggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 25 aaaccttgtt ttaaccagg gctgcccctt gtgcgctga ccgcgacgc cgaagggggg 2940  
 tgccccctt tctcgaacc tcccggccc ctaacgccc gctcccacc ccccaggggc 3000  
 30 tgccccctt ggccgcgaa ggcctcacc caaaaatggc agcgtggca gtccttgcca 3060  
 ttgcccggat cgggctgta acgggatggc cgatcagccc gagcgcgacg cccggaagca 3120  
 35 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 ggggctggg tggcggcctg cccttcaact cggccgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 40 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420  
 45 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 50 tgcattgact aatgctttaa acccaggaca ataacctat agcttgtaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720  
 55 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcaggtta tgcgctcaa ttcgctcgt atatcgcttg ctgattacgt 3840  
 60 gcagctttcc cttcagggcg gattcataka gggccagcc atccgtcctc catatcacca 3900

65

## ES 2 345 324 T3

5           cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga 3960  
           atacgtgcmc aacaaccgtc ttccggagac tgtcatacgc gtaaacacgc cagcgctggc 4020  
 10           gcgatttagc cccgacatag ccccactggt cgtccatttc cgcgcagacg atgacgtcac 4080  
           tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 15           cgtgttgagg ccaacgcca taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200  
           catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 20           ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
           acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 25           agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
           tggtttcaaa atcggctccg tcgatactat gttatacggc aactttgaaa acaactttga 4500  
 30           aaaagctggt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
           cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 35           taaattggcta aaatgagaat atcacgggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
           gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa 4740  
 40           aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
           gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 45           gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
           gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 50           aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
           ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 55           gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
           cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 60           gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280  
           gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340  
 65           ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
           ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 70           caccgacttc ttccgcatca agtgTTTTTg ctctcaggcc gaggccacg gcaagtattt 5520  
           gggcaagggg tcgctggat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 75           cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640

## ES 2 345 324 T3

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60

```

ggcaccaggc gggTcaaate aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700
cccgcaagga gggTgaatga atcggacgTt tgaccgggaag gcatacaggc aagaactgat 5760
cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgctgTc 5820
gccccgcgaa accttccagt ccgTcggctc gatggTccag caagctacgg ccaagatcga 5880
gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatacggccg ccgTggagcg 5940
ttcgcgTcgt ctcgaaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000
aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggTcag 6060
cgaggccaag caggccgctg tgctgaaaca cacgaagcag cagatcaagg aatgcagct 6120
ttccttgTtc gatattgTc cgtggccgga cacgatgTca gcgatgcca acgacacggc 6180
ccgctctgcc ctgTtcacca cgcgcaaaa gaaaatcccg cgcgaggcgc tgcaaaaaca 6240
ggTcattttc cacgTcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300
cgacgatgac gaaactggTt ggcagcaggt gTtgagTac gcgaagcga cccctatcgg 6360
cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggT cgatcaatgg 6420
ccggtattac acgaaggccg aggaatgcct gTcgcgccta caggcgacgg cgatgggctt 6480
cacgTccgac cgcgTtgggc acctggaatc ggtgTcgtg ctgcaccgct tccgctcct 6540
ggaccgtggc aagaaaacgt cccgTtgcca ggtcctgatc gacgaggaaa tcgTcgtgct 6600
gTttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgTcggccgac 6660
ggcccgacgg atgTtcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720
aaccttccgc ctatgTgTc gatcggattc caccgcgTg aagaagtggc gcgagcaggt 6780
cggcgaagcc tgcgaaagTt tgcgaggcag cggcctggTg gaacacgcct gggTcaatga 6840
tgacctggTg cattgcaaac gctagggcct tgtggggTca gTtccggctg ggggTtcagc 6900
agccagcgt tTactggcat tTcaggaaca agcgggcact gTcgcacgca cTtGctTcgc 6960
tcagTatcgc tcgggacgca cggcgcgTc tacgaaactgc cgataaacag aggattaaaa 7020
ttgacaattg tgattaaggc tcagattcga cggcTtgag cggccgacgt gcaggatttc 7080
cgcgagatcc gattgTcggc cctgaagaaa gTcTcagaga tgTtccggTc cgtttacgag 7140
cacgaggaga aaaagcccat ggaggcgtTc gTgaaacggT tcgagatgTc cgtggcattc 7200
ggcgcctaca tcgacggcga gatcattggg ctgTcggTct tcaaacagga ggacggcccc 7260
aaggacgTc acaaggcga tctgTccggc gTtttctgTg agcccgaaca gcgaggccga 7320
gggTcgcgcg gTatgctgct gcgggcgTt cggcgggTt tattgctgT gatgatcTc 7380
  
```

65

## ES 2 345 324 T3

cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaattatt 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccggggcg ggctcggggc 7500  
 5 acggtaggcg ctgtgcagcc gctgatggtc gtgttcattc ctgccgctct gctaggtagc 7560  
 ccgatacgat tgatgggggt cctgggggct atttgcggaa ctgccggcgt ggcgctgttg 7620  
 10 gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggggc 7680  
 gtttccatgg cgttcggaa cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc 7740  
 acctttaccg cctggcaact gggggccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 15 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct 7920  
 acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980  
 20 catcaggccg acagtcggaa ctccgggtcc ccgacctgta ccattcggtg agcaatggat 8040  
 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttctcag 8100  
 25 cggctttatc cagcgatttc ctattatgtc ggcattagtc tcaagatca cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 30 tcattccgtg ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccggg ctgatgggct 8400  
 35 gcctgtatcg agtggtgatt ttgtccgag ctgccggctg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcy 8520  
 gacgttttta atgtactggg gtggtttttc ttttaccag tgagacgggc aacagctgat 8580  
 40 tgccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 45 agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aaggcgaaa aaccgctctat cagggcgatg gccactacg 8820  
 tgaacatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 50 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 ggaaggggag aaagcgaaag gagcggggcg cattcaggct gcgcaactgt tgggaagggc 9000  
 55 gatcggtgcy ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120

60

65

## ES 2 345 324 T3

aattaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca 9180  
 5 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 10 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctcgcaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tccatgatagc ggtcgccac acccagccgg 9540  
 15 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac 9660  
 20 agttcggctg gcgagagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 25 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgccccaa tagcagccag 9900  
 tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgctcgtggc 9960  
 30 agccacgata gccgcgctgc ctgcctctgc agttcattca gggcaccgga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcccc gtcatagccc aatagcctct ccaccaagc ggccggagaa 10140  
 35 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 40 agtggagcat ttttgacaag aaatatttgc tagctgatag tgacctagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaacgg cttgtcccgc 10440  
 45 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcctatgatc ttgatcccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 50 cttaccagag ggcgccccag ctggcaattc cggctcgctt gctgtccata aaaccgcccc 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctcttg cgcttgctt 10680  
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
 55 actggcttct tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcgcca 10800  
 gcgtgaagct tgcattgctg caggctgacg gcgagccgag ctccctgagc aaatttacac 10860  
 60  
 65

## ES 2 345 324 T3

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
5 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980  
tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100  
10 attaaagtga gtgaatatgg taccacaagg tttggagatt taattggtgc aatgctgcat 11160  
ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
15 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
taccacacac aagttttgag gtgcatgcat ggatgcctg tggaaagttt aaaaatattt 11340  
tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400  
20 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat 11520  
agccagcaga tctgccggca tcgatcccgg gccatggcct gctttaatga gatatgagag 11580  
25 acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga 11640  
gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaataat atatcaccgg 11700  
30 ttactatcgt atttttatga ataataattct ccgttcaatt tactgattgt ccgtcgacga 11760  
gctcggcgcg cctctagagg atcgatgaat tcagatcggc tgagtggctc cttcaacgtt 11820  
35 gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc gtcacggcg ggggtcataa 11880  
cgtgactccc ttaattctcc gctcatgatc agattgtcgt tteccgcctt cagtttaaac 11940  
tatacagtgt tgacaggata tattggcggg taaacctaaag agaaaagagc gtttattaga 12000  
40 ataatcggat atttaaaagg gcgtgaaaag gtttatcctt cgtccatttg tatgtgcatg 12060  
ccaaccacag ggttcccca 12079

45 <210> 23  
<211> 13002  
<212> DNA  
50 <213> Desconocido  
<220>  
<223> Vector de expresión vegetal con dos casetes de expresión promotor-terminador

55

60

65

# ES 2 345 324 T3

<400> 23

```

5      gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60
      gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcc 120

      tagtggggcg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcacggc 180
10     ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240
      atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgggga 300
15     ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360
      tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgtg aacgaggtcg 420
      gcgtagacgg tetgacgaca cgcaaactgg cggaacggtt gggggttcag cagccggcgc 480
20     tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg 540
      cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600
25     ggaatgcccg cagcttcagg caggecgtgc tcgcctaccg cgatggcgcg cgcattcatg 660
      ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttccctt 720
      gcgaggcggg ttttcggcc ggggacgcc tcaatgcgct gatgacaatc agctacttca 780
30     ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga 840
      ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc ttcgacgaag 900
35     ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaa 960
      ggaggtcgtt tgtcaggaaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020
      tgccggagcg caaccctc actacagcag agccatgtag acaacatccc ctccccctt 1080
40     ccaccgctc agacgcccgt agcagcccgc tacgggcttt ttcattgccct gccctagcgt 1140
      ccaagctca cggccgct cggcctctct gggggccttc tggcgctctt ccgcttctc 1200
45     gctcaactgac tcgctgcgct cggtcgttct gctgcggcga gcggtatcag ctcactcaa 1260
      ggcgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaa 1320
      aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggt 1380
50     ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac 1440
      aggactataa agataccagg cgtttcccc tggaagctcc ctcgctgcgt ctctgttcc 1500
55     gaccctgccg ctaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttt 1560
      ccgctgcata acctgcttc ggggtcatta tagcgatttt ttcggtatat ccatccttt 1620
      tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttcttgg tgtatccaac 1680
60     ggcgtcagcc gggcaggata ggtgaagtag gccaccgc gagcgggtgt tcttcttca 1740
      ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg 1800
65     ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860

```

## ES 2 345 324 T3

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980  
 5 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggccccatc aatggcgacc 2040  
 tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt 2100  
 10 tcggtgatgc cacgatectc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaaag cacgtcccca tgcgctccat caagaagagc 2280  
 15 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctggttg ccctcgcgcg tgggctggcg gccgtctatg gccctgcaaa 2400  
 20 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460  
 cctcgcggaa aacttggccc tcaactgacag atgaggggcg gacgttgaca cttgaggggc 2520  
 cgactcacc. ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 25 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac 2700  
 30 tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 ccgccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 35 aaacctgtt ttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctcgaacct tcccggcccg ctaacgcggg cctcccatcc ccccaggggc 3000  
 40 tgccccctc ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 45 gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcattg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 50 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaga ttataccgag 3360  
 gtatgaaaac gagaattgga cttttacaga attactctat gaagcgcctat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 55 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 60  
 65

## ES 2 345 324 T3

5      tgc atggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc cggatgactt 3720  
 10     tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcaggtta tgcgctcaa ttcgctgctg atatcgcttg ctgattacgt 3840  
 15     gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcacc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga 3960  
 20     atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020  
 gcgatttagc cccgacatag ccccaactgtt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgccccgtg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 25     cgtgttgagg ccaacgcccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 30     ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgcggtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atactactaa tcagtaagtt ggcagcatca cccataattg 4440  
 35     tggtttcaaa atcggctccg tcgatactat gttatacggc aactttgaaa acaactttga 4500  
 aaaagctggt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagtccgt 4560  
 40     cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aagggtatata agctgggtggg agaaaatgaa 4740  
 45     aacctatatt taaaaatgac ggacagccgg tataaagggg ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 50     gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcacc 4980  
 aggcctcttc actccatoga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 55     ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
 60     cccgaagagg aacttgctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcggg caagtgggat 5280  
 65     gacattgcct tctgcgtccg gtcgatcagg gaggatatcg ggaagaaca gtatgtcgag 5340

## ES 2 345 324 T3

ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaaata ttatatttta 5400  
 ctggatgaat tgTTTTagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 5 caccgacttc ttccgcatca agtgTTTTgg ctctcaggcc gagggcccacg gcaagtattt 5520  
 gggcaagggg tcgctgggat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 10 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc gggTcaaate aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 15 cccgcaagga gggTgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg tttccgcgcg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttccagt ccgtcggctc gatggTccag caagctacgg ccaagatcga 5880  
 20 gcgcgacagc gtgcaactgg ctccccctgc cctgccccgc ccatcggccg ccgtggagcg 5940  
 ttcgcgtcgt ctcgaaacagg aggcggcagg tttggcgaa tcgatgacca tcgacacgcg 6000  
 25 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aatgcagct 6120  
 ttccttgTtc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc 6180  
 30 ccgctctgcc ctgttcacca cgcgcaaca gaaaatcccg cgcgaggcgc tgcaaaaca 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 35 cgacgatgac gaactggTgt ggcagcaggT gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 40 cacgtccgac cgcgttgggc acctggaate ggtgtcgtg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 45 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660  
 ggccccagcg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 aaccttcgc ctcattgtgc gatcggattc caccgcgtg aagaagtggc gcgagcaggT 6780  
 50 cggcgaagcc tgcgaaagat tgcgaggcag cggcctggTg gaacacgcct gggTcaatga 6840  
 tgacctggTg cattgcaaac gctagggcct tgtgggTca gttccggctg ggggttcagc 6900  
 55 agccagcgt tTactggcat ttcaggaaca agcgggact gctogacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgtc tacgaaactgc cgataaacag aggattaaa 7020  
 60 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080

65

## ES 2 345 324 T3

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 5 cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc 7200  
 ggcgectaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 10 ggggtcgcgc gtatgctgct gcgggcgttg ccggcggggt tattgctcgt gatgatcgtc 7380  
 cgacagatc caacgggaat ctgggtgatg cgcattctca tctcgggcgc acttaatat 7440  
 15 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggagg ggtcgcggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc 7560  
 ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggctg ggcgctgttg 7620  
 20 gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggctt ggcgggggcg 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aaacctccgt gcctctgctc 7740  
 25 acctttaccg cctggcaact ggcgcccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttccct tggttccggg ggatctcgcg actcgaacct 7920  
 30 acagttgttc cttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa ctccgggtcc ccgacctgta ccattcgggt agcaatggat 8040  
 35 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttccctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcggg tctctgcgag ggagatgata 8220  
 40 tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt tcaaaccccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 45 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctggt ggctggctgg 8460  
 tggcaggata tattgtggtg taacaaatt gacgcttaga caacttaata acacattgcg 8520  
 50 gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgctcggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 55 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataatcaaa 8700  
 agaatagccc gagatagggt tgagtgtgt tccagtttg aacaagagtc cactatataa 8760  
 60 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gccactacg 8820

65

## ES 2 345 324 T3

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcgga 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 5 ggaaggggaag aaagcgaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtgcg ggccctctcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 10 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca 9180  
 15 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 20 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataaccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tctgatagc ggtccgccac acccagccgg 9540  
 25 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcgatgc gcgccttgag cctggcgaac 9660  
 30 agttcggctg gcgagagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccc attgcatcag ccatgatgga tactttctcg 9840  
 35 gcaggagcaa ggtgagatga caggagatcc tgccccgga cttcgcccaa tagcagccag 9900  
 tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
 40 agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgocca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 45 cctgcgtgca atccatcttg ttcaatocaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataaccg aggggaattt atggaacgtc 10260  
 50 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc 10380  
 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaacgg cttgtcccgc 10440  
 55 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacgatc ttgatcccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560

60

65

## ES 2 345 324 T3

ettaccagag ggcgccccag ctggcaattc cggttegett gctgtccata aaaccgcccc 10620  
 5 gcttagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgettgcggt 10680  
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
 actggccttc tacgtgttcc gcttccttla gcagcccttg cgccttgagt gcttgccgca 10800  
 10 gcgtgaagct tgcattgctg caggtcgacg gcgcgcccag ctctctgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 15 tttgatttgc gataaatttt tatatttggg actaaattta taacacctt tatgctaacy 10980  
 tttgccaca cttagcaatt tgcaagtga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgc aatatttcta ctataggaga 11100  
 20 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatgta ccacacaaga 11220  
 tttgaggtgc atgaacgca cgtggacaaa aggttttagta attttcaag acaacaatgt 11280  
 25 taccacacac aagttttgag gtgcatgcat ggatgcctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatgga gccatgtgta aaacctgac atccacttg aggatgcaat 11400  
 30 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttla cagattata atttcttcat 11520  
 agccagcccc ccgcggtggg cggccgctg cagtctagaa ggctcctgc tttaatgaga 11580  
 35 tatgagagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaa 11640  
 aaacctgagc atgtgtagct cagatcctta ccgcccgtt cggttcattc taatgaatat 11700  
 40 atcacccgtt actatcgat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760  
 gtcgagcaaa tttacacatt gccactaac gtctaaacc ttgtaatttg tttttgtttt 11820  
 actatgtgtg ttatgtattt gatttgcgat aaattttat atttggact aaattataa 11880  
 45 caccttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 11940  
 aattattttt gtcttctaaa tacatatact aatcaactgg aatgtaaat atttgtaat 12000  
 50 attctacta taggagaatt aaagtgagtg aatattgtac cacaaggtt ggagatttaa 12060  
 ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 12120  
 aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt 12180  
 55 tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgccctgtgg 12240  
 aaagtttaa aatatttttg aatgatttg catggaagcc atgtgtaaaa ccatgacatc 12300

60

65

## ES 2 345 324 T3

cacttggagg atgcaataat gaagaaaact acaaatttac atgcaactag ttatgcatgt 12360  
agtctatata atgaggattt tgcaataactt tcattcatac aactcacta agttttacac 12420  
5 gattataatt tcttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct 12480  
ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctggttgca 12540  
10 cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccggtttc ggttcattct 12600  
aatgaatata tcaccggtta ctatcgtatt tttatgaata atattctccg ttcaatttac 12660  
tgattgtccg tcgacgaatt cgagctcggc gcgcctctag aggatcgatg aattcagatc 12720  
15 ggctgagtgg ctccttcaac gttgcggttc tgtcagttcc aaacgtaaaa cggttgtcc 12780  
cgcgtcatcg gcggggggtca taacgtgact ccttaatte tccgctcatg atcagattgt 12840  
20 cgtttccgc cttcagttta aactatcagt gtttgacagg atatattggc gggtaaacct 12900  
aagagaaaag agcgtttatt agaataatcg gatatttaaa agggcgtgaa aaggtttacc 12960  
cttcgtccat ttgtatgtgc atgccaacca cagggttccc ca 13002

25

<210> 24

<211> 13905

30 <212> DNA

<213> Desconocido

<220>

35 <223> Vector de expresión vegetal con tres casetes de expresión promotor-terminador

<400> 24

40 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
gcccagcga cagggtgcga ggcaaattgc accaacgcat acagcggcag cagaatgccca 120  
tagtgggccc tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc 180  
45 ataatcaggc cgatgccgac agcgtcagc gcgacagtgc tcagaattac gatcaggggt 240  
atggtggggt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
50 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420  
55 gcgtagacgg tctgacgaca cgcaaactgg cggaacgggt gggggttcag cagccggcgc 480  
ttactggca cttcaggaac aagcgggccc tgctcgacgc actggccgaa gccatgctgg 540  
cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
60 ggaatgcccg cagcttcagg caggcgtgc tcgcctaccg cgatggcgcg cgcattccatg 660  
cgggcacgcg accgggcccga ccgagatgg aaacggccga cgcgcagctt cgcttctct 720

65

## ES 2 345 324 T3

gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca 840  
 5 ccgltgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag 900  
 ccggtccgga cgcagcgctc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa 960  
 10 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccctc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgctc agacgcccgt agcagcccgc tacgggcttt ttcattgccct gccctagcgt 1140  
 15 ccaagcctca cggccgctc cggcctctct ggccgcttc tggcgctctt ccgcttctct 1200  
 gctcactgac tcgctgctc cggctgctc gctgcccga gcggtatcag ctcaactcaa 1260  
 20 ggccgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aagccagga accgtaaaaa ggccgcttg ctggcgcttt tccataggct 1380  
 ccgccccct gacgagcctc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac 1440  
 25 aggactataa agataccagg cgtttcccc tggaagctcc ctctgctcct ctctgttcc 1500  
 gaccctgccg ctaccggat acctgtccgc ctttctcct tcgggaagcg tggcgctttt 1560  
 30 ccgctgcata acctgcttc ggggtcatta tagcgatttt ttccgtatat ccatcctttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttcttgg tgtatccaac 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccaccgc gagcgggtgt tccttcttca 1740  
 35 ctgtccctta ttccacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg 1800  
 ctaccgcccg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 40 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgctcgc cagggctaca 1980  
 aatcacggg cgtcgtggac tatgagcacg tccgagct ggcccgcctc aatggcgacc 2040  
 45 tggccgccc gggcggcctg ctgaaactct ggctcaccga cgaccgctc acggcgcggt 2100  
 tcggtgatgc cagcatctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 50 gcaaggteat gatggcgctg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg gggctgctg gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgagg agctgggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcttttg 2340  
 55 gacgctcacc gggctggtg cctcggcgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgcccgtga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460

60

65

## ES 2 345 324 T3

cctcgcggaa aacttggccc tcaactgacag atgaggggcg gacgttgaca cttgaggggc 2520  
 cgactcaccc ggcgcgcgct tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 5 gagctggcca gcctcgcaaa tcggcgaaaa cgcttgattt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac 2700  
 10 tgacagatga ggggcgcgat ccttgacact tgaggggag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggtt 2820  
 cgcgccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 15 aaaccttggt ttaaccagg gctgcgccct gtgcgctga ccgcgcacgc cgaagggggg 2940  
 tgccccctt tctcgaacct tcccggcccg ctaacgcggg cctcccatcc ccccaggggc 3000  
 20 tgcgccccct ggccggaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgcgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 25 gcggcctggg tggcggcctg cccttcaact cggccgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 30 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 35 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 40 tgcattggact aatgcttga acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 45 gctgcctcag attcaggtta tgccgetcaa ttcgctgctg atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca ggggccagcc atccgtcacc catatcacca 3900  
 50 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgcctatagt cgttcaccga 3960  
 atacgtgctc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020  
 55 gcgatttagc ccgacatag cccactggt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgccccgctg tatgcgcgag gttaccgact gggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200  
 60  
 65

## ES 2 345 324 T3

5      catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
 10     tggtttcaaa atcggctccg tcgatactat gttatagcc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 15     taaatggcta aatgagaat atcaccgga ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctgggtggg agaaaatgaa 4740  
 20     aacctatatt taaaaatgac ggacagccgg tataaagga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 25     gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagcctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtctttc actccatoga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 30     ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
 35     cccgaagagg aacttgctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcggg caagtgggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340  
 40     ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 45     caccgacttc ttccgcatca agtgTTTTg ctctcaggcc gaggcccacg gcaagtattt 5520  
 gggcaagggg tcgctgggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 50     ggcaccaggc gggTcaaTc aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgcaagga gggTgaatga atcggacgTt tgaccggaag gcatacaggc aagaactgat 5760  
 55     cgacgcgggg ttttccgccc aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttccagt ccgtcggctc gatggTccag caagctacgg ccaagatcga 5880  
 60     gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatacggccg ccgtggagcg 5940

65

## ES 2 345 324 T3

ttcgcgtcgt ctgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 5 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aatgcagct 6120  
 ttccttgttc gatattgctc cgtggccgga cacgatgca gcgatgcaa acgacacggc 6180  
 10 ccgctctgcc ctggtcacca cgcgcaācaa gaaaatcccg cgcgaggcgc tgcaaaācaa 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 15 cgacgatgac gaactgggtg ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 20 cacgtccgac cgcggttggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt ccggttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 25 gtttgcctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgctgcggac 6660  
 ggcccgcagg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 aacctccgc ctcatgtgct gatcggatc caccgcgtg aagaagtggc gcgagcaggt 6780  
 30 cggcgaagcc tgcaagagt tgcgaggcag cggcctggtg gaacacgcct gggtaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtgggtca gttccggtg ggggttcagc 6900  
 35 agccagcgtt ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattāaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cgcttgagg cgcccgacgt gcaggatttc 7080  
 40 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgaacggg tgcgagatgc cgtggcattc 7200  
 45 ggcgcctaca tcgacggcga gatcattggg ctgtcggctt tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 ggggtcgcgg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc 7380  
 50 cgacagatcc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatatt 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccggggcg ggtcgcggcg 7500  
 55 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatt ctgccgctct gctaggtagc 7560  
 ccgatacgat tgatggcggc cctgggggct atttgccgaa ctgcgggcgt ggcgctgttg 7620  
 60 gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggcg 7680

65

## ES 2 345 324 T3

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc 7740  
 acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 5 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct 7920  
 10 acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat 8040  
 15 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 cggtaagcg agaaatgaat aagaaggctg ataattcggg tctctgcgag ggagatgata 8220  
 20 tttgatcaca ggcagcaacg ctctgtcctc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400  
 25 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtggtg taacaaaatt gacgcttaga caacttaata acacattgcg 8520  
 30 gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atgggtggtc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 35 agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 40 tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaaag ccggcgaacg tggcgagaaa 8940  
 ggaaggggaag aaagcgaag gagcggggcg cattcaggct gcgcaactgt tgggaagggc 9000  
 45 gatcggtgcg ggctctctcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgccg gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 50 aattaattcc catcttgaaa gaaatatagt ttaaattttt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 55 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctcggaa 9420

60

65

## ES 2 345 324 T3

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tectgatagc ggcccgccac acccagccgg 9540  
 5 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgceatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac 9660  
 10 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgac gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccc attgcatcag ccatgatgga tactttctcg 9840  
 15 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgccca tagcagccag 9900  
 tccottcccg cttcagtgac aacgtcgagc acagctgccc aaggaacgcc cgtcgtggcc 9960  
 20 agccacgata gccgcgctgc ctgctcctgc agttcattca gggcacccga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg ccctgcccgt gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 25 cctgcgtgca atccatcttg ttcaatcaa gtcceatgg gccctcgact agagtogaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 30 agtggagcat ttttgacaag aaatatttgc tagctgatag tgacctagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaacgg cttgtcccgc 10440  
 35 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacatgac ttgatccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 40 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620  
 gtctagctat cgccatgtaa gcccaactgca agctacctgc tttctctttg cgcttgcgtt 10680  
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgagg 10740  
 45 actggctttc tacgtgttcc gtttcttta gcagccctg gcacctgagt gcttggggca 10800  
 gcgtgaagct tgcacgcctg caggtcgacg gcgcgcccag ctctcgagc aaatttacac 10860  
 50 attgccacta aacgtctaaa cccttgaat ttgtttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacacctt tatgctaacg 10980  
 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
 55 aaatacatat actaatcaac tggaaatgta aatatttgc aatatttcta ctataggaga 11100  
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160

60

65

## ES 2 345 324 T3

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65

```

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
tttgagggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280
taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340
tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400
aatgaagaaa actacaaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460
ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat 11520
agccagccca ccgcggtggg cggccgctg cagtctagaa ggctcctgc tttaatgaga 11580
tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 11640
aaacctgagc atgtgtagct cagatcctta ccgccggtt cggttcattc taatgaatat 11700
atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760
gtcgagcaaa tttacacatt gccactaac gtctaaacct ttgtaatttg tttttgtttt 11820
actatgtgtg ttatgtattt gatttgcat aaattttat atttggact aaattataa 11880
caccttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 11940
aattattttt gtcttctaaa tacatatact aatcaactgg aaatgtaaat atttgcta 12000
atcttacta taggagaatt aaagtgagtg aatatggtac cacaaggtt ggagatttaa 12060
ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 12120
aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg ttagtaatt 12180
tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgcctgtgg 12240
aaagtttaaa aatattttgg aaatgatttg catggaagcc atgtgtaaaa ccatgacatc 12300
cacttgaggg atgcaataat gaagaaaact acaatttac atgcaactag ttatgcatgt 12360
agtctatata atgaggattt tgcaataactt tcattcatac aactcacta agttttacac 12420
gattataatt tttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct 12480
ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttgtgca 12540
cgttgtaaaa aacctgagca tgtgtagctc agatccttac ccgcggttcc ggttcattct 12600
aatgaatata tcaccggtta ctatcgtatt tttatgaata atattctccg ttcaatttac 12660
tgattgtccg tcgagcaaat ttacacattg ccactaaacg tctaaacct tgtaatttgt 12720
ttttgtttta ctatgtgtgt tatgtatttg atttgcgata aattttata tttggtacta 12780
aatttataac accttttatg ctaacgtttg ccaaacctta gcaatttgca agttgattaa 12840
ttgattctaa attatttttg tcttctaaat acatatacta atcaactgga aatgtaata 12900
  
```

## ES 2 345 324 T3

tttgctaata tttctactat aggagaatta aagtgagtga atatggtacc acaaggtttg 12960  
gagatttaat tgttgcaatg ctgcatggat ggcataata ccaaaccattc aataattctt 13020  
5 gaggataata atggtaccac acaagatttg aggtgcatga acgtcacgtg gacaaaagggt 13080  
ttagtaattt tccaagacaa caatgttacc acacacaagt tttgaggtgc atgcatggat 13140  
10 gccctgtgga aagtttaaaa atattttgga aatgatttgc atggaagcca tgtgtaaaac 13200  
catgacatcc acttggagga tgcaataatg aagaaaacta caaatttaca tgcaactagt 13260  
tatgcatgta gtctatataa tgaggatttt gcaatacttt cattcatata cactcactaa 13320  
15 gttttacacg attataattt cttcatagcc agcagatctg ccggcatcga tcccgggcca 13380  
tggcctgctt taatgagata tgcgagacgc ctatgatcgc atgatatttg ctttcaattc 13440  
20 tgttggtgac gttgtaaaaa acctgagcat gtgtagctca gatccttacc gccggtttcg 13500  
gttcattcta atgaatatat caccggttac tatcgtattt ttatgaataa tattctccgt 13560  
tcaatttact gattgtccgt cgacgagctc ggcgcgcctc tagaggatcg atgaattcag 13620  
25 atcggctgag tggctccttc aacgttgcgg ttctgtcagt tccaaacgta aaacggcttg 13680  
tcccgcgtca tcggcggggg tcataacgtg actcccttaa ttctccgctc atgatcagat 13740  
30 tgtcgtttcc cgccttcagt ttaaactatc agtgtttgac aggatatatt ggcgggtaaa 13800  
cctaagagaa aagagcgttt attagaataa tcggatattt aaaagggcgt gaaaagggtt 13860  
atccttcgctc catttgatg tgcgatgcaa ccacaggggt cccca 13905

35  
<210> 25  
<211> 15430  
40 <212> DNA  
<213> Desconocido  
<220>  
45 <223> Vector de expresión vegetal con tres casetes de expresión promotor-terminador insertado es elongasa y desaturasa de *Physcomitrella patens*  
<220>  
<221> CDS  
50 <222> (11543)..(12415)  
<220>  
<221> CDS  
<222> (13313)..(14890)

55

60

65

# ES 2 345 324 T3

<400> 25

5           gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
           gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgccca 120  
 10           tagtggggcg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180  
           ataatcaggc cgatgccgac agcgtcgagc ggcacagtgc tcagaattac gatcaggggt 240  
           atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 15           ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
           tgacaaagtt gcagccgaat acagtgatec gtgccgccct ggacctgttg aacgaggtcg 420  
           gcgtagacgg tctgacgaca cgaaaactgg cggaaacggtt gggggttcag cagccggcgc 480  
 20           tttactggca cttcaggaac aagcgggccc tgctcgacgc actggccgaa gccatgctgg 540  
           cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 25           ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
           ccggcacgcg accggggcga ccgcagatgg aaacggccga cgcgcagctt cgettcctct 720  
           gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
 30           ctgttggggc cgtgcttgag gaggaggccc ggcacagcga tgccggcgag cgcggcggca 840  
           ccgttgaaca ggetccgctc tcgcccgtgt tgcgggcccgc gatagacgcc ttcgacgaag 900  
 35           ccggtccgga cgcagcgttc gaggaggac tcgcggtgat tgtcgatgga ttggcgaaaa 960  
           ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
           tgccggagcg caaccctc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 40           ccaccgcgtc agacgcccg agcagcccgc tacgggcttt tcatgcctt gccttagcgt 1140  
           ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt cgettcctc 1200  
 45           gctcaactgac tcgctgcgct cggtcgctgc gctgcccga gcggtatcag ctactcaaa 1260  
           ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 50           aggccagcaa aaggccagga accgtaaaaa ggcgcggtt ctggcgtttt tccataggct 1380  
           ccgccccctc gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgcac 1440  
           aggactataa agataccagg cgtttcccc tggaagctcc ctcgctgcgt ctctgttcc 1500  
 55           gaccctgccg ctaccggat acctgtccgc ctttctcct tcgggaagcg tggcgctttt 1560  
           ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt 1620  
 60           tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tglatccaac 1680  
           ggcgtcagcc gggcaggata ggtgaagtag gccaccgcg gagcgggtgt tccttcttca 1740  
           ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 65           ctaccgcccg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860

## ES 2 345 324 T3

agggcagccc acctatcaag gtgtactgcc tccagacga acgaagagcg attgaggaaa 1920  
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980  
 5 aaatcacggg cgctcgtggac tatgagcacg tccgcgagct ggcccgcacg aatggcgacc 2040  
 tgggccgctt gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt 2100  
 10 tcggtgatgc caccatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaaag cacgtcccca tgcgctccat caagaagagc 2280  
 15 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 20 cgccccagaa acgcccgcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460  
 cctcgccgaa aacttgcccc tcaactgacag atgaggggcg gacgttgaca cttgaggggc 2520  
 cgactcacc ggccggcgtg tgacagatga ggggcaggct cgatttcggc eggcgacgtg 2580  
 25 gagctggcca gcctcgcaaa tcggcgaaaa cgccgtgatt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac 2700  
 30 tgacagatga ggggcgcgat ccttgacact tgaggggagc agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 35 aaaccttgtt ttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccctt tctcgaacce tcccgcccg ctaacgcggg cctcccatcc ccccggggc 3000  
 40 tgcgccctc ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgcccggc agtgagggcg 3180  
 45 gcggcctggg tggcggcctg cccttcaact cggccgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 50 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 55 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcggc cttatcaata tatctataga atgggcaaag cataaaaact 3600

60

65

## ES 2 345 324 T3

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65

```

    tgc atggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660
    attgggtaat gactccaact tattgatagt gttttatggt cagataatgc cegatgactt 3720
    tgtcatgcag ctccaccgat tttgagaacg acagcgactt cegtcccagc cgtgccaggt 3780
    gctgcctcag attcaggtta tgccgctcaa ttcgctgctg atatcgcttg ctgattacgt 3840
    gcagctttcc cttcaggcgg gattcataca gcgccagcc atccgtcatc catatcacca 3900
    cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960
    atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
    gcgathtagc cccgacatag cccactggt cgtccatttc cgcgcagacg atgacgtcac 4080
    tgccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140
    cgtgttgagg ccaacgcccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200
    catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260
    ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320
    acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380
    agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
    tggtttcaaa atcggctccg tcgatactat gttatacggc aactttgaaa acaactttga 4500
    aaaagctggt ttctgggtatt taaggtttta gaatgcaagg aacagtgaat tggagtctgt 4560
    cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaataa 4620
    taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
    gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa 4740
    aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800
    gaaaaggaca tgatgctatg gctggaagga aagctgctg ttccaaaggt cctgcacttt 4860
    gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920
    gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
    aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040
    ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100
    gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160
    cccgaagagg aacttgtctt tcccacggc gacctgggag acagcaacat ctttgtgaaa 5220
    gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280
    gacattgctt tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340
  
```

# ES 2 345 324 T3

ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatTTTta 5400  
 ctgggatgaat tgTTTTtagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 5 caccgacttc ttccgcatca agtgTTTTtg ctctcaggcc gaggcccacg gcaagtattt 5520  
 gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 10 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc gggTcaaTc aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 15 cccgcaagga gggTgaatga atcggacgTt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgCGTgc 5820  
 gccccgcgaa accttccagt ccgtcggctc gatggTccag caagctacgg ccaagatcga 5880  
 20 gcgCGacagc gtgcaactgg ctccccctgc cctgccccgcg ccatacggccg ccgtggagcg 5940  
 ttCGctcgt ctCGaacagg aggcgGcagg tttggCGaag tcgatgacca tcgacacgCG 6000  
 25 aggaactatg acgaccaaga agcGaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccCGt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttcttGttc gatattgCGc cgtggccgga cacgatgCGa gcgatgcaa acgacacgCG 6180  
 30 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatccccg cgcgaggCGc tgcaaaacaa 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtCG agctgcgggc 6300  
 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagCGca ccctatcgg 6360  
 35 cgagccgata accttcaCGt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgCGccta caggcgacgg cgatgggctt 6480  
 40 cacgtccgac cgcgttggc acctggaatc ggtgtCGctg ctgcaccgct tccgctcct 6540  
 ggaccgtggc aagaaaacgt cccgTtgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 gtttGctggc gaccactaca cGaaattcat atgggagaag taccgcaagc tgtcgCCgac 6660  
 45 ggcccgaCGg atgttcgact atttcagctc gcaccgggag ccgtaccCGc tcaagctgga 6720  
 aaccttccgc ctcatgtgCG gatcggatc caccCGcgtg aagaagtggc gcgagcaggt 6780  
 50 cggcgaagcc tgcgaagagt tgCGaggcag cggcctggTg gaacacgcct gggTcaatga 6840  
 tgacctggTg cattgcaaac gctagggcct tgtggggtca gttccggtg ggggttcagc 6900  
 55 agccagcGct ttactggcat ttcaggaaca agcgggcact gctCGacgca cttgcttcGc 6960  
 tcagtatCGc tcgggacgca cggcCGctc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattCGa cggcTtgag cggccgacgt gcaggatttc 7080

60

65

## ES 2 345 324 T3

5           cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
           caccaggaga aaaagcccat ggaggcggtc gctgaacggt tgcgagatgc cgtggcattc 7200  
           ggcgcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggaaggcccc 7260  
           aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 10          ggggtcgccc gtatgctgct gcgggcggtg cgggcgggtt tattgctcgt gatgategtc 7380  
           cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatat 7440  
           tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccggggcg ggtcggggcg 7500  
           acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc 7560  
           ccgatacgat tgatggcggt cctgggggct atttgcgga ctgcccggct ggcgctggtg 7620  
 20          gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggggc 7680  
           gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aaacctccgt gcctctgctc 7740  
           acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg 7800  
           ttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
           ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct 7920  
 30          acagtgtttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980  
           catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat 8040  
           aggggagtgt atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag 8100  
           cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
           cggttaagcg agaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 40          ttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
           tcacccgtgt ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
           gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400  
           gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctggt ggctggctgg 8460  
           tggcaggata tattgtgggt taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 50          gacgttttta atgtactggg gtggtttttc tttcaccag tgagacgggc aacagctgat 8580  
           tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
           gcaggcgaaa atcctgtttg atggtgggtc cgaaatcggc aaaatccctt ataaatcaaa 8700  
           agaatagccc gagatagggt tgagtgttgt tccagtttg aacaagagtc cactattaa 8760  
 60          gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820

65

## ES 2 345 324 T3

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag cggcggaacg tggcgagaaa 8940  
 5 ggaaggggaag aaagcgaaag gagcggggcg cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtgcg ggccctctcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 10 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca 9180  
 15 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaat cagaaatatt 9240  
 tcaataactg atttatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 20 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctcggaa 9420  
 tcggggagcgg cgataaccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 25 ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac 9660  
 30 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccattgatgga tactttctcg 9840  
 35 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgccca tagcagccag 9900  
 tcccttcccc cttcagtgac aacgtcgagc acagctgccc aaggaacgcc cgtcgtggcc 9960  
 40 agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggtcggtc 10020  
 ttgacaaaaa gaaccggggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatagccg aatagcctct ccacccaagc ggccggagaa 10140  
 45 cctgcgtgca atccatcttg ttcaatccaa gctccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 50 agtggagcat ttttgacaag aatatattgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 55 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaacgg cttgtccccg 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560

60

65







# ES 2 345 324 T3

	Asp Glu His Pro Gly Gly Ser Val Ile Ser Thr Tyr Phe Gly Arg Asp	
	430 435 440	
5	ggc aca gat gtt ttc tct agt ttt cat gca gct tct aca tgg aaa att Gly Thr Asp Val Phe Ser Ser Phe His Ala Ala Ser Thr Trp Lys Ile	13810
	445 450 455	
10	ctt caa gac ttt tac att ggt gac gtg gag agg gtg gag ccg act cca Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu Arg Val Glu Pro Thr Pro	13858
	460 465 470	
15	gag ctg ctg aaa gat ttc cga gaa atg aga gct ctt ttc ctg agg gag Glu Leu Leu Lys Asp Phe Arg Glu Met Arg Ala Leu Phe Leu Arg Glu	13906
	475 480 485	
20	caa ctt ttc aaa agt tct gaa ttg tac tat gtt atg aag ctg ctc acg Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr Val Met Lys Leu Leu Thr	13954
	490 495 500 505	
25	aat gtt gct att ttt gct gcg agc att gca ata ata tgt tgg agc aag Asn Val Ala Ile Phe Ala Ala Ser Ile Ala Ile Ile Cys Trp Ser Lys	14002
	510 515 520	
30	act att tca gcg gtt ttg gct tca gct tgt atg atg gct ctg tgt ttc Thr Ile Ser Ala Val Leu Ala Ser Ala Cys Met Met Ala Leu Cys Phe	14050
	525 530 535	
35	caa cag tgc gga tgg cta tcc cat gat ttt ctc cac aat cag gtg ttt Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His Asn Gln Val Phe	14098
	540 545 550	
40	gag aca cgc tgg ctt aat gaa gtt gtc ggg tat gtg atc ggc aac gcc Glu Thr Arg Trp Leu Asn Glu Val Val Gly Tyr Val Ile Gly Asn Ala	14146
	555 560 565	
45	gtt ctg ggg ttt agt aca ggg tgg tgg aag gag aag cat aac ctt cat Val Leu Gly Phe Ser Thr Gly Trp Trp Lys Glu Lys His Asn Leu His	14194
	570 575 580 585	
50	cat gct gct cca aat gaa tgc gat cag act tac caa cca att gat gaa His Ala Ala Pro Asn Glu Cys Asp Gln Thr Tyr Gln Pro Ile Asp Glu	14242
	590 595 600	
55	gat att gat act ctc ccc ctc att gcc tgg agc aag gac ata ctg gcc Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp Ser Lys Asp Ile Leu Ala	14290
	605 610 615	
60	aca gtt gag aat aag aca ttc ttg cga atc ctc caa tac cag cat ctg Thr Val Glu Asn Lys Thr Phe Leu Arg Ile Leu Gln Tyr Gln His Leu	14338
	620 625 630	
65	ttc ttc atg ggt ctg tta ttt ttc gcc cgt ggt agt tgg ctc ttt tgg Phe Phe Met Gly Leu Leu Phe Phe Ala Arg Gly Ser Trp Leu Phe Trp	14386
	635 640 645	
70	agc tgg aga tat acc tct aca gca gtg ctc tca cct gtc gac agg ttg Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu Ser Pro Val Asp Arg Leu	14434
	650 655 660 665	



ES 2 345 324 T3

<400> 26

5 Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser  
1 5 10 15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp  
20 25 30

10 Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
35 40 45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu  
50 55 60

15 Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu  
65 70 75 80

20 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser  
85 90 95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr  
100 105 110

25 Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile  
115 120 125

30 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr  
130 135 140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His  
145 150 155 160

35 Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His  
165 170 175

40 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly  
180 185 190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg  
195 200 205

45 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu  
210 215 220

50 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr  
225 230 235 240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile  
245 250 255

55 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr  
260 265 270

60 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys  
275 280 285

65 Thr Glu  
290

ES 2 345 324 T3

<210> 27

<211> 525

<212> PRT

5 <213> Desconocido

<400> 27

10 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn  
1 5 10 15

15 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe  
20 25 30

20 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln  
35 40 45

25 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala  
50 55 60

30 Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly  
65 70 75 80

35 Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg  
85 90 95

40 Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val  
100 105 110

45 His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr  
115 120 125

50 Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser  
130 135 140

55 Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala  
145 150 155 160

60 Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu  
165 170 175

65 Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg  
180 185 190

70 Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr  
195 200 205

75 Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala  
210 215 220

80 Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys  
225 230 235 240

ES 2 345 324 T3

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe  
 245 250 255

5 Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly  
 260 265 270

10 Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys  
 275 280 285

Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr  
 290 295 300

15 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp  
 305 310 315 320

20 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile  
 325 330 335

Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg  
 340 345 350

25 Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu  
 355 360 365

Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr  
 370 375 380

30 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro  
 385 390 395 400

35 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly  
 405 410 415

Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser  
 420 425 430

40 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly  
 435 440 445

45 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu  
 450 455 460

His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala  
 465 470 475 480

50 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp  
 485 490 495

55 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu  
 500 505 510

60 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser  
 515 520 525

65

## ES 2 345 324 T3

<210> 28  
 <211> 17752  
 <212> DNA  
 5 <213> Desconocido  
 <220>  
 <223> vector de expresión vegetal con 3 casetes de expresión promotor-terminador insertado con elongasa + desaturasa de Physcomitrella + desaturasa de Phaeodactylum  
 10 <220>  
 <221> CDS  
 <222> (11543)..(12415)  
 15 <220>  
 <221> CDS  
 <222> (13313)..(14890)  
 <220>  
 20 <221> CDS  
 <222> (15791)..(17200)  
  
 <400> 28  
 25  
  
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
 gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgccca 120  
 30 tagtggggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc 180  
 ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 35 atgttggggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctggtg aacgaggtcg 420  
 40 gcgtagacgg tctgacgaca cgcaaactgg cggaaacggtt gggggttcag cagccggcgc 480  
 tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540  
 45 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 50 ccggcacgcg accgggcgca ccgcagatgg aaacggcoga cgcgcagctt cgettcctct 720  
 gcgaggcggg tttttcggcc ggggacgcgc tcaatgcgct gatgacaatc agctacttca 780  
 ctggtggggc cgtgcttgag gaggaggccg gcgacagcga tgccggcgag cgcggcggca 840  
 55 ccggtgaaca ggctccgctc tcgccgctgt tgcgggcccgc gatagacgcc ttcgacgaag 900  
 ccggtccgga cgcagcgttc gaggaggac tcgcggtgat tgtcgatgga ttggcgaaaa 960  
 60 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080

65

## ES 2 345 324 T3

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcattgcct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct gggggccttc tggcgctctt ccgcttcttc 1200  
 5 gctcactgac tcgctgcgct cggtcgcttc gctgcccgcg gcggtatcag ctactcaaaa 1260  
 ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 10 aggccagcaa aaggccagga accgtaaaaa gggcgcgctt ctggcgcttt tccatagcgt 1380  
 ccgccccctc gacgagcctc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac 1440  
 aggactataa agataccagg cgtttccccc tggaaagctc ctctgctcct ctctgttcc 1500  
 15 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata acctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt 1620  
 20 tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttcttggc tgtatccaac 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccaccgcg gagcgggtgt tccttcttca 1740  
 ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg 1800  
 25 ctaccgcccg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 30 agggggcggc ggccggcatg agcctgtcgg cctacctgct ggccgctcgg cagggttaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcctc aatggcgacc 2040  
 tgggcccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgcg acggcgcggt 2100  
 35 tcggtgatgc caogacctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggatc gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 40 aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctgggtg ccctcgcgcg tgggctggcg gccgtctatg gcctgcaaaa 2400  
 45 cgccagcagaa acgcccgcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460  
 cctcgcggaa aacttggccc tctactgacag atgagggggc gacgttgaca cttgaggggc 2520  
 50 cgactcacc ggcgcccgtg tgacagatga gggcagagct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgagagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgagactac 2700  
 55 tgacagatga ggggcgagc ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820

60

65

## ES 2 345 324 T3

cgcgccgttt ttgggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 5   aaccttggt ttaaccagg gctgcgcct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgcccccoct tctcgaacce tcccggcccg ctaacgcggg cctcccatcc ccccaggggc 3000  
 tgcgccccctc ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 10   ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcacgcacat tcagcgacca ggtgcggggc agtgagggcg 3180  
 15   gcggcctggg tggcggcctg cccttactt cggccgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcatc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgagtg ataggtaaga ttataccgag 3360  
 20   gtatgaaaac gagaattgga cttttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 25   tactgataag ataatatatc ttttatatag aagatcgc cgatagtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataacctat agcttgtaa ttctatcata 3660  
 30   attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccatgactt 3720  
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 35   gctgcctcag attcaggta tgcgcgtcaa ttcgctcgt atatcgctt ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca gggccagcc atccgtcacc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgcctatagt cgttcaccga 3960  
 40   atacgtgccc aacaaccgtc ttccggagac tgcatacgc gtaaacacgc cagcgtggc 4020  
 gcgatttagc cccgacatag ccccactggt cgtccattc cgcgcagacg atgacgtcac 4080  
 45   tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtggtgagg ccaacgccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 50   ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtc ttttgcggtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 55   agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
 tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctggt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 60

65

## ES 2 345 324 T3

5           cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
           taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
           gtaaaagata cggaaggaat gtctcctgct aaggatatata agctggtggg agaaaatgaa 4740  
 10           aacctatatt taaaaatgac ggacagccgg tataaagggg ccacctatga tgtggaacgg 4800  
           gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
           gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 15           gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcadc 4980  
           aggctctttc actocatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
           ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 20           gaagaagaca ctocatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
           cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 25           gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280  
           gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340  
           ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 30           ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
           caccgacttc ttccgatca agtgTTTTg ctctcaggcc gaggcccacg gcaagtattt 5520  
 35           gggcaagggg tcgctgggat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
           cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
           ggcaccaggc ggggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 40           cccgcaagga ggggtgaatga atcggacggt tgaccggaag gcatacaggc aagaactgat 5760  
           cgacgcgggg ttttccgccg aggatgccga aaccatcga agccgcaccg tcatgcgtgc 5820  
 45           gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga 5880  
           gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940  
           ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag togatgacca tcgacacgcg 6000  
 50           aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
           cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 55           ttccttgttc gatattgctc cgtggccgga cacgatcga gcgatgcaa acgacacggc 6180  
           ccgctctgcc ctgttcacca cgcgcaaaa gaaaatcccg cgcgagggcg tgcaaaaaca 6240  
 60           ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300

65

## ES 2 345 324 T3

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca ccctatcgg 6360  
 5 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgcta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcgttgggc acctggaatc ggtgtoctg ctgcaccgct tccgctcct 6540  
 10 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgac gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660  
 15 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgac tcaagctgga 6720  
 aaccttccgc ctcatgtgcg gatcggatc caccgcgctg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgcgaagagt tgcgaggcag cggcctgggt gaacacgcct gggtaatga 6840  
 20 tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc 6900  
 agccagcgct ttactggcat ttcaggaaca agcgggact gctcgacgca ctgcttcgc 6960  
 25 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 30 caccgaggaga aaaagcccat ggaggcgctc gctgaacggt tgcgagatgc cgtggcattc 7200  
 ggcgcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc 7260  
 35 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agccogaaca gcgaggccga 7320  
 ggggtcgcgg gtatgctgct gcggcgctt cggcggggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctgggtgatg cgcattctca tctcggcgc acttaattt 7440  
 40 tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggagg ggtcgcggcg 7500  
 acggtaggag ctgtgcagcc gctgatggtc gtgttcattc ctgccgctct gctaggtagc 7560  
 45 ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggctt ggcgctgttg 7620  
 gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggctt ggcggggcg 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctccgt gcctctgctc 7740  
 50 acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg 7800  
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 55 ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct 7920  
 acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980  
 60 catcaggccg acagtcgga cttcgggtcc ccgacctgta ccattcgggtg agcaatggat 8040

65

## ES 2 345 324 T3

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttctctcag 8100  
 5 cggcctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcggg tctctgcgag ggagatgata 8220  
 tttgatcaca ggcaagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280  
 10 tcatccgtgt ttcaaaccog gcagcttagt tgccggttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgcgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400  
 15 gcctgtatcg agtggtgatt ttgtgccgag ctgccggctcg gggagctggt ggctggctgg 8460  
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat 8580  
 20 tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtcccagctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 25 agaatagccc gagatagggt tgagtgttgt tccagtttg aacaagagtc cactatataa 8760  
 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 30 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 ggaaggggaag aaagcgaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 35 gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaataatt attgataaaa taacaagtca 9180  
 40 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaat cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 45 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggctcag cccattcgcg gccaaagctct 9480  
 50 tcagcaatat cacgggtagc caacgctatg tcttgatagc ggtccgccac acccagccgg 9540  
 ccacagtcga tgaatccaga aaagcggcca ttttcacca tgatattcgg caagcaggca 9600  
 55 tcgccatggg tcacgacgag atcctgcgag tcgggcatgc gcgccttgag cctggcgaac 9660  
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 60 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780

65

## ES 2 345 324 T3

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttegcccaa tagcagccag 9900  
 5 tcccttcccc cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgctcgtggcc 9960  
 agccacgata gccgcgctgc ctcgctctgc agttcattca gggcaccgga caggtcggtc 10020  
 10 ttgacaaaaa gaaccgggcg ccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgccca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 15 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgct 10260  
 agtggagcat ttttgacaag aatatattgc tagctgatag tgaccttagg cgacttttga 10320  
 20 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaaacgg cttgtccccg 10440  
 25 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcctatgac ttgatcccc 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620  
 30 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgctt 10680  
 ttcccttgtc cagatagccc agtagctgac attcacccg ggtcagcacc gtttctgcgg 10740  
 actggctttc tacgtgttcc gtttcttta gcagcccttg cgcctgagt gcttgccgca 10800  
 35 gcgtgaagct tgcatgctg caggtcgacg gcgcgccgag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt ttactatgt gtgttatgta 10920  
 40 tttgatttgc gataaatttt tatatttggg actaaattta taacacctt tatgctaacg 10980  
 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgc aatattteta ctataggaga 11100  
 45 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 50 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 55 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgragtctat ataatgagga 11460  
 60 ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat 11520

65

ES 2 345 324 T3

agccagccca ccgcggtgga aa atg gag gtc gtg gag aga ttc tac ggt gag 11572  
Met Glu Val Val Glu Arg Phe Tyr Gly Glu  
1 5 10

5  
ttg gat ggg aag gtc tcg cag ggc gtg aat gca ttg ctg ggt agt ttt 11620  
Leu Asp Gly Lys Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe  
15 20 25

10  
ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctg gtt 11668  
Gly Val Glu Leu Thr Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val  
30 35 40

15  
gac agt ccc aca ccc atc gtc ctc ggt gtt tct gta tac ttg act att 11716  
Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile  
45 50 55

20  
gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc 11764  
Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg  
60 65 70

25  
gcc tcg gag cca ttt ttg ctc caa gct ttg gtg ctt gtg cac aac ctg 11812  
Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu  
75 80 85 90

30  
ttc tgt ttt gcg ctc agt ctg tat atg tgc gtg ggc atc gct tat cag 11860  
Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln  
95 100 105

35  
gct att acc tgg cgg tac tct ctc tgg ggc aat gca tac aat cct aaa 11908  
Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys  
110 115 120

40  
cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac 11956  
His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr  
125 130 135

45  
gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg 12004  
Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg  
140 145 150

50  
caa ata agc ttc ctc cac gtt tat cat cat tct tca att tcc ctc att 12052  
Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Ser Leu Ile  
155 160 165 170

55  
tgg tgg gct att gct cat cac gct cct ggc ggt gaa gca tat tgg tct 12100  
Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser  
175 180 185

60  
gcg gct ctg aac tca gga gtg cat gtt ctc atg tat gcg tat tac ttc 12148  
Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe  
190 195 200

65  
ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt 12196  
Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu  
205 210 215

70  
ttt tgg ggc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg 12244  
Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu

# ES 2 345 324 T3

	220	225	230	
5	aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro 235	240	245	12292
10	caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe 255	260	265	12340
15	ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly 270	275	280	12388
20	aag caa aag gga gct aaa act gag tga tctagaaggc ctccctgcttt Lys Gln Lys Gly Ala Lys Thr Glu 285	290		12435
25	aatgagatat gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa cctgagcatg tgtagctcag atccttaccg cgggtttcgg ttcattctaa tgaatatatc accggttact atcgtatttt tatgaataat attctccggt caatttactg attgtccgtc gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt ttgttttact atgtgtgta tgtatttgat ttgcgataaa tttttatatt tgggtactaaa			12495 12555 12615 12675 12735
30	tttataacac cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat ttttttgtc ttctaaatac atatactaata caactggaaa tgtaaatatt tgctaataatt tctactatag gagaattaa gtgagtgaat atggtaccac aaggtttggga gatttaattg ttgcaatgct gcatggatgg catatacac aaacattcaa taattcttga ggataataat ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaaggttt			12795 12855 12915 12975 13035
35	agtaattttt caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa gtttaaaaat attttgaaa tgatttgcgcat ggaagccatg tgtaaaacca tgacatccac ttggaggatg caataatgaa gaaaactaca aatttcatg caactagtta tgcattgtagt ctatataatg aggattttgc aatactttca ttcatacaca ctcaactaagt tttacacgat tataattttct tcatagccag oggatcc atg gta ttc gcg ggc ggt Met Val Phe Ala Gly Gly 295			13095 13155 13215 13275 13330
40	gga ctt cag cag ggc tct ctc gaa gaa aac atc gac gtc gag cac att Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn Ile Asp Val Glu His Ile 300	305	310	13378
45	gcc agt atg tct ctc ttc agc gac ttc ttc agt tat gtg tct tca act Ala Ser Met Ser Leu Phe Ser Asp Phe Phe Ser Tyr Val Ser Ser Thr 315	320	325	13426
50				
55				
60				
65				



# ES 2 345 324 T3

	555		560		565		
5	gtt ctg ggg ttt agt Val Leu Gly Phe Ser 570		aca ggg tgg tgg aag Thr Gly Trp Trp Lys 575		gag aag cat aac ctt cat Glu Lys His Asn Leu His 580 585	14194	
10	cat gct gct cca aat His Ala Ala Pro Asn 590		gaa tgc gat cag act Glu Cys Asp Gln Thr 595		tac caa cca att gat gaa Tyr Gln Pro Ile Asp Glu 600	14242	
15	gat att gat act ctc Asp Ile Asp Thr Leu 605		ccc ctc att gcc tgg Pro Leu Ile Ala Trp 610		agc aag gac ata ctg gcc Ser Lys Asp Ile Leu Ala 615	14290	
20	aca gtt gag aat aag Thr Val Glu Asn Lys 620		aca ttc ttg cga atc Thr Phe Leu Arg Ile 625		ctc caa tac cag cat ctg Leu Gln Tyr Gln His Leu 630	14338	
25	ttc ttc atg ggt ctg Phe Phe Met Gly Leu 635		tta ttt ttc gcc cgt Leu Leu Phe Phe Ala 640		ggt agt tgg ctc ttt tgg Gly Ser Trp Leu Phe Trp 645	14386	
30	agc tgg aga tat acc Ser Trp Arg Tyr Thr 650		tct aca gca gtg ctc Ser Thr Ala Val Leu 655		tca cct gtc gac agg ttg Ser Pro Val Asp Arg Leu 660 665	14434	
35	ttg gag aag gga act Leu Glu Lys Gly Thr 670		gtt ctg ttt cac tac Val Leu Phe His Tyr 675		ttt tgg ttc gtc ggg aca Phe Trp Phe Val Gly Thr 680	14482	
40	gcg tgc tat ctt ctc Ala Cys Tyr Leu Leu 685		cct ggt tgg aag cca Pro Gly Trp Lys Pro 690		tta gta tgg atg gcg gtg Leu Val Trp Met Ala Val 695	14530	
45	act gag ctc atg tcc Thr Glu Leu Met Ser 700		ggc atg ctg ctg ggc Gly Met Leu Leu Gly 705		ttt gta ttt gta ctt agc Phe Val Phe Val Leu Ser 710	14578	
50	cac aat ggg atg gag His Asn Gly Met Glu 715		gtt tat aat tcg tct Val Tyr Asn Ser Ser 720		aaa gaa ttc gtg agt gca Lys Glu Phe Val Ser Ala 725	14626	
55	cag atc gta tcc aca Gln Ile Val Ser Thr 730		cggt gat atc aaa gga Arg Asp Ile Lys Gly 735		aac ata ttc aac gac tgg Asn Ile Phe Asn Asp Trp 740 745	14674	
60	ttc act ggt ggc ctt Phe Thr Gly Gly Leu 750		aac agg caa ata gag Asn Arg Gln Ile Glu 755		cat cat ctt ttc cca aca His His Leu Phe Pro Thr 760	14722	
65	atg ccc agg cat aat Met Pro Arg His Asn 765		tta aac aaa ata gca Leu Asn Lys Ile Ala 770		cct aga gtg gag gtg ttc Pro Arg Val Glu Val Phe 775	14770	
70	tgt aag aaa cac ggt Cys Lys Lys His Gly 780		ctg gtg tac gaa gac Leu Val Tyr Glu Asp 785		gta tct att gct acc ggc Val Ser Ile Ala Thr Gly 790	14818	

# ES 2 345 324 T3

act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca 14866  
 Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala Ala  
 795 800 805

5

gag cag cat gct acc acc agt taa gctagcgtaa accctgcttt aatgagatat 14920  
 Glu Gln His Ala Thr Thr Ser  
 810 815

10

gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa 14980  
 cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc 15040  
 acccgttact atcgatatttt tatgaataat attctccggt caatttactg attgtccgtc 15100

15

gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt ttgttttact 15160  
 atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tggactaaa tttataacac 15220

20

cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat 15280  
 tatttttgtc ttctaaatac atatactaata caactggaaa tgtaaatatt tgctaataat 15340  
 tctactatag gagaattaa gtgagtgaat atggtaccac aaggtttggga gatttaattg 15400

25

ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga ggataataat 15460  
 ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaaggttt agtaattttt 15520

30

caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa 15580  
 gtttaaaaat attttggaaa tgatttgcat ggaagccatg tgtaaaacca tgacatccac 15640  
 ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta tgcatgtagt 15700

35

ctatataatg aggattttgc aatactttca ttcatcacaca ctactaagt tttacacgat 15760  
 tataattttct tcatagccag cagatctaaa atg gct ccg gat gcg gat aag ctt 15814  
 Met Ala Pro Asp Ala Asp Lys Leu  
 820 825

40

cga caa cgc cag acg act gcg gta gcg aag cac aat gct gct acc ata 15862  
 Arg Gln Arg Gln Thr Thr Ala Val Ala Lys His Asn Ala Ala Thr Ile  
 830 835 840

45

tcg acg cag gaa cgc ctt tgc agt ctg tct tcg ctc aaa ggc gaa gaa 15910  
 Ser Thr Gln Glu Arg Leu Cys Ser Leu Ser Ser Leu Lys Gly Glu Glu  
 845 850 855

50

gtc tgc atc gac gga atc atc tat gac ctc caa tca ttc gat cat ccc 15958  
 Val Cys Ile Asp Gly Ile Ile Tyr Asp Leu Gln Ser Phe Asp His Pro  
 860 865 870

55

ggg ggt gaa acg atc aaa atg ttt ggt ggc aac gat gtc act gta cag 16006  
 Gly Gly Glu Thr Ile Lys Met Phe Gly Gly Asn Asp Val Thr Val Gln  
 875 880 885

60

tac aag atg att cac ccg tac cat acc gag aag cat ttg gaa aag atg 16054  
 Tyr Lys Met Ile His Pro Tyr His Thr Glu Lys His Leu Glu Lys Met

65

# ES 2 345 324 T3

	890	895	900	905	
5	aag cgt gtc ggc aag Lys Arg Val Gly	aag gtg acg gat ttc gtc Lys Val Thr Asp Phe Val	gag tac aag ttc gat Glu Tyr Lys Phe Asp		16102
		910	915	920	
10	acc gaa ttt gaa cgc gaa atc aaa cga gaa gtc ttc aag att gtg cga Thr Glu Phe Glu Arg Glu Ile Lys Arg Glu Val Phe Lys Ile Val Arg				16150
		925	930	935	
15	cga ggc aag gat ttc ggt act ttg gga tgg ttc ttc cgt gcg ttt tgc Arg Gly Lys Asp Phe Gly Thr Leu Gly Trp Phe Phe Arg Ala Phe Cys				16198
		940	945	950	
20	tac att gcc att ttc ttc tac ctg cag tac cat tgg gtc acc acg gga Tyr Ile Ala Ile Phe Phe Tyr Leu Gln Tyr His Trp Val Thr Thr Gly				16246
		955	960	965	
25	acc tct tgg ctg ctg gcc gtg gcc tac gga atc tcc caa gcg atg att Thr Ser Trp Leu Leu Ala Val Ala Tyr Gly Ile Ser Gln Ala Met Ile				16294
		970	975	980	985
30	ggc atg aat gtc cag cac gat gcc aac cac ggg gcc acc tcc aag cgt Gly Met Asn Val Gln His Asp Ala Asn His Gly Ala Thr Ser Lys Arg				16342
		990	995	1000	
35	ccc tgg gtc aac gac atg cta ggc ctc ggt gcg gat ttt att ggt ggt Pro Trp Val Asn Asp Met Leu Gly Leu Gly Ala Asp Phe Ile Gly Gly				16390
		1005	1010	1015	
40	tcc aag tgg ctc tgg cag gaa caa cac tgg acc cac cac gct tac acc Ser Lys Trp Leu Trp Gln Glu Gln His Trp Thr His His Ala Tyr Thr				16438
		1020	1025	1030	
45	aat cac gcc gag atg gat ccc gat agc ttt ggt gcc gaa cca atg ctc Asn His Ala Glu Met Asp Pro Asp Ser Phe Gly Ala Glu Pro Met Leu				16486
		1035	1040	1045	
50	cta ttc aac gac tat ccc ttg gat cat ccc gct cgt acc tgg cta cat Leu Phe Asn Asp Tyr Pro Leu Asp His Pro Ala Arg Thr Trp Leu His				16534
		1050	1055	1060	1065
55	cgc ttt caa gca ttc ttt tac atg ccc gtc ttg gct gga tac tgg ttg Arg Phe Gln Ala Phe Phe Tyr Met Pro Val Leu Ala Gly Tyr Trp Leu				16582
		1070	1075	1080	
60	tcc gct gtc ttc aat cca caa att ctt gac ctc cag caa cgc gcc gca Ser Ala Val Phe Asn Pro Gln Ile Leu Asp Leu Gln Gln Arg Gly Ala				16630
		1085	1090	1095	
65	ctt tcc gtc ggt atc cgt ctc gac aac gct ttc att cac tcg cga cgc Leu Ser Val Gly Ile Arg Leu Asp Asn Ala Phe Ile His Ser Arg Arg				16678
		1100	1105	1110	
70	aag tat gcg gtt ttc tgg cgg gct gtg tac att gcg gtg aac gtg att Lys Tyr Ala Val Phe Trp Arg Ala Val Tyr Ile Ala Val Asn Val Ile				16726
		1115	1120	1125	

# ES 2 345 324 T3

```

gct ccg ttt tac aca aac tcc ggc ctc gaa tgg tcc tgg cgt gtc ttt 16774
Ala Pro Phe Tyr Thr Asn Ser Gly Leu Glu Trp Ser Trp Arg Val Phe
1130                1135                1140                1145

5   gga aac atc atg ctc atg ggt gtg gcg gaa tcg ctc gcg ctg gcg gtc 16822
Gly Asn Ile Met Leu Met Gly Val Ala Glu Ser Leu Ala Leu Ala Val
                1150                1155                1160

10  ctg ttt tcg ttg tcg cac aat ttc gaa tcc gcg gat cgc gat ccg acc 16870
Leu Phe Ser Leu Ser His Asn Phe Glu Ser Ala Asp Arg Asp Pro Thr
                1165                1170                1175

15  gcc cca ctg aaa aag acg gga gaa cca gtc gac tgg ttc aag aca cag 16918
Ala Pro Leu Lys Lys Thr Gly Glu Pro Val Asp Trp Phe Lys Thr Gln
                1180                1185                1190

20  gtc gaa act tcc tgc act tac ggt gga ttc ctt tcc ggt tgc ttc acg 16966
Val Glu Thr Ser Cys Thr Tyr Gly Gly Phe Leu Ser Gly Cys Phe Thr
                1195                1200                1205

25  gga ggt ctc aac ttt cag gtt gaa cac cac ttg ttc cca cgc atg agc 17014
Gly Gly Leu Asn Phe Gln Val Glu His His Leu Phe Pro Arg Met Ser
1210                1215                1220                1225

30  agc gct tgg tat ccc tac att gcc ccc aag gtc cgc gaa att tgc gcc 17062
Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Lys Val Arg Glu Ile Cys Ala
                1230                1235                1240

35  aaa cac ggc gtc cac tac gcc tac tac ccg tgg atc cac caa aac ttt 17110
Lys His Gly Val His Tyr Ala Tyr Tyr Pro Trp Ile His Gln Asn Phe
                1245                1250                1255

40  ctc tcc acc gtc cgc tac atg cac gcg gcc ggg acc ggt gcc aac tgg 17158
Leu Ser Thr Val Arg Tyr Met His Ala Ala Gly Thr Gly Ala Asn Trp
                1260                1265                1270

45  cgc cag atg gcc aga gaa aat ccc ttg acc gga cgg gcg taa 17200
Arg Gln Met Ala Arg Glu Asn Pro Leu Thr Gly Arg Ala
                1275                1280                1285

50  agatctgccg gcatcgatcc cgggccatgg cctgctttaa tgagatatgc gagacgccta 17260
tgatcgcatg atatttgctt tcaattctgt tgtgcacggt gtaaaaaacc tgagcatgtg 17320

55  tagctcagat ccttaccgcc ggtttcgggtt cattotaatg aatatatcac ccgttactat 17380
cgtattttta tgaataatat tctccgttca atttactgat tgtccgtcga cgagctcggc 17440
gcgcctctag aggatcgatg aattcagatc ggctgagtgg ctcttcaac gttgcgggtc 17500

60  tgtcagttcc aaacgtaaaa cggcttgtcc cgcgtcatcg gcgggggtca taacgtgact 17560
cccttaattc tccgctcatg atcagattgt cgtttcccgc cttcagttta aactatcagt 17620

65  gtttgacagg atatatggc gggtaaacct aagagaaaag agcgtttatt agaataatcg 17680
gatatttaaa agggcgtgaa aaggtttata cttcgtccat ttgtatgtgc atgccaacca 17740

cagggttccc ca 17752

<210> 29
<211> 290
65 <212> PRT
<213> Desconocido

```

ES 2 345 324 T3

<400> 29

5 Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser  
1 5 10 15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp  
20 25 30

10 Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
35 40 45

15 Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu  
50 55 60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu  
65 70 75 80

20 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser  
85 90 95

25 Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr  
100 105 110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile  
115 120 125

30 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr  
130 135 140

35 Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His  
145 150 155 160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His  
165 170 175

40 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly  
180 185 190

45 Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg  
195 200 205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu  
210 215 220

50 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr  
225 230 235 240

55 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile  
245 250 255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr  
260 265 270

60 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys  
275 280 285

65 Thr Glu  
290

ES 2 345 324 T3

<210> 30

<211> 525

<212> PRT

5 <213> Desconocido

<400> 30

10 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn  
1 5 10 15

15 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe  
20 25 30

20 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln  
35 40 45

25 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala  
50 55 60

30 Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly  
65 70 75 80

35 Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg  
85 90 95

40 Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val  
100 105 110

45 His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr  
115 120 125

50 Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser  
130 135 140

55 Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala  
145 150 155 160

60 Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu  
165 170 175

65 Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg  
180 185 190

70 Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr  
195 200 205

75 Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala  
210 215 220

80 Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys

ES 2 345 324 T3

	225					230						235				240
5	Met	Met	Ala	Leu	Cys	Phe	Gln	Gln	Cys	Gly	Trp	Leu	Ser	His	Asp	Phe
					245					250					255	
	Leu	His	Asn	Gln	Val	Phe	Glu	Thr	Arg	Trp	Leu	Asn	Glu	Val	Val	Gly
				260					265					270		
10	Tyr	Val	Ile	Gly	Asn	Ala	Val	Leu	Gly	Phe	Ser	Thr	Gly	Trp	Trp	Lys
			275					280					285			
15	Glu	Lys	His	Asn	Leu	His	His	Ala	Ala	Pro	Asn	Glu	Cys	Asp	Gln	Thr
		290					295					300				
	Tyr	Gln	Pro	Ile	Asp	Glu	Asp	Ile	Asp	Thr	Leu	Pro	Leu	Ile	Ala	Trp
	305					310					315					320
20	Ser	Lys	Asp	Ile	Leu	Ala	Thr	Val	Glu	Asn	Lys	Thr	Phe	Leu	Arg	Ile
					325					330					335	
25	Leu	Gln	Tyr	Gln	His	Leu	Phe	Phe	Met	Gly	Leu	Leu	Phe	Phe	Ala	Arg
				340					345					350		
	Gly	Ser	Trp	Leu	Phe	Trp	Ser	Trp	Arg	Tyr	Thr	Ser	Thr	Ala	Val	Leu
			355					360					365			
30	Ser	Pro	Val	Asp	Arg	Leu	Leu	Glu	Lys	Gly	Thr	Val	Leu	Phe	His	Tyr
		370					375					380				
35	Phe	Trp	Phe	Val	Gly	Thr	Ala	Cys	Tyr	Leu	Leu	Pro	Gly	Trp	Lys	Pro
	385					390					395					400
	Leu	Val	Trp	Met	Ala	Val	Thr	Glu	Leu	Met	Ser	Gly	Met	Leu	Leu	Gly
				405						410					415	
40	Phe	Val	Phe	Val	Leu	Ser	His	Asn	Gly	Met	Glu	Val	Tyr	Asn	Ser	Ser
				420					425					430		
	Lys	Glu	Phe	Val	Ser	Ala	Gln	Ile	Val	Ser	Thr	Arg	Asp	Ile	Lys	Gly
			435					440					445			
45	Asn	Ile	Phe	Asn	Asp	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Arg	Gln	Ile	Glu
	450						455					460				
50	His	His	Leu	Phe	Pro	Thr	Met	Pro	Arg	His	Asn	Leu	Asn	Lys	Ile	Ala
	465					470					475					480
	Pro	Arg	Val	Glu	Val	Phe	Cys	Lys	Lys	His	Gly	Leu	Val	Tyr	Glu	Asp
				485						490					495	
55	Val	Ser	Ile	Ala	Thr	Gly	Thr	Cys	Lys	Val	Leu	Lys	Ala	Leu	Lys	Glu
				500					505					510		
60	Val	Ala	Glu	Ala	Ala	Ala	Glu	Gln	His	Ala	Thr	Thr	Ser			
			515					520					525			

<210> 31

<211> 469

65 <212> PRT

<213> Desconocido

ES 2 345 324 T3

<400> 31

5 Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val  
1 5 10 15

Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser  
20 25 30

10 Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr  
35 40 45

15 Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe  
50 55 60

20 Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His  
65 70 75 80

Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp  
85 90 95

25 Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys  
100 105 110

Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu  
115 120 125

30 Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu  
130 135 140

35 Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala  
145 150 155 160

Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala  
165 170 175

40 Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly  
180 185 190

45 Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln  
195 200 205

His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp  
210 215 220

50 Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp  
225 230 235 240

55 His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met  
245 250 255

Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile  
260 265 270

60 Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp  
275 280 285

65

ES 2 345 324 T3

Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala  
 290 295 300  
 5 Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly  
 305 310 315 320  
 Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val  
 325 330 335  
 10 Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe  
 340 345 350  
 15 Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu  
 355 360 365  
 Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly  
 370 375 380  
 20 Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu  
 385 390 395 400  
 25 His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala  
 405 410 415  
 Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr  
 420 425 430  
 30 Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His  
 435 440 445  
 Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro  
 450 455 460  
 35 Leu Thr Gly Arg Ala  
 465  
 40  
 45  
 50  
 55  
 60  
 65